Mathematical Induction

A powerful and elegant method of proof

Mathematical Induction

A powerful and elegant method of proof

Titu Andreescu

Vlad Crişan

Library of Congress Control Number: 2017932999

ISBN-10: 0-9968745-9-3
ISBN-13: 978-0-9968745-9-5

9 8 7 6 5 4 3 2 1

www.awesomemath.org

Cover design by Iury Ulzutuev

Foreword

Mathematical induction is a method of proof that has been known to mathematicians for hundreds of years, with one of the earliest proofs using Induction dating back to Francesco Maurolico (1575), when he showed that the sum of first n odd positive integers is n^2. Nevertheless, it was only formalized as an axiom of the nonnegative integers in the 19-th century due to the work of Peano and other mathematicians on Set Theory. Nowadays, it has become one of the basic tools that any mathematician is familiar with. Induction is an important technique used in competitions and its applications permeate almost every area of mathematics. Due to its elegance and popularity, we have decided to write this book whose main goal is to offer a detailed exposition of the method, its subtleties and beautiful applications. The book is designed as follows:

First chapter, entitled *A brief overview of Induction* offers an introduction to the subject. We begin by describing Induction in the context of Set Theory, as one of Peano's axioms and present a few examples of how different properties of the nonnegative integers can be derived from it. We then look at some classical examples which are solved using Induction and illustrate in detail the approach one should follow when writing their proof. After discussing some variants of Induction, we move on to presenting one of the most intriguing aspects of the method, called the *Paradox of Induction.* There we look at various examples, fully motivating all the steps that lie behind their elegant proofs, where we strengthen the original statement to make the solution easier. The chapter ends with a section dedicated to Transfinite Induction.

Second chapter, *Sums, products, and identities* is mainly aimed at those who want to familiarize themselves with the basics of applying Induction. The nature of the questions presented is similar to the ones that originally motivated the use of Induction as an algebraic tool. The chapter also illustrates the power of the method, showing how easy it is to use Induction to prove identities, rather than using other techniques.

From the third chapter onwards, we follow a problem solving based approach by discussing Induction in various areas of mathematics. Each chapter is divided into two sections: Theory with examples and Proposed problems. The book is designed so that it is as self-contained as possible. Therefore, each chapter starts by introducing the reader to all the notions that are required for understanding the examples and tackling the proposed problems. Various beautiful examples are then discussed in full detail, explaining the main themes that occur in each specific field (see for example Chapters 6 and 7 on Number theory and Combinatorics). With the aim of being as comprehensive as possible, we explore a total of 10 different areas of mathematics, including topics that are not usually discussed in an Olympiad-oriented book on Induction (such as Chapter 3 on Functions and functional equations). The second part of each chapter consists of a carefully chosen list of proposed problems. These add up to more than 200 elegant questions from over 20 worldwide renowned Olympiads and mathematical magazines, as well as original problems designed by the authors of this book and their collaborators. Fully detailed solutions are provided for each problem at the end of the book.

We truly believe that the book can serve as a very good resource and teaching material for anyone who wants to explore the beauty of Induction and its applications, from novice mathematicians to Olympiad-driven students and professors teaching undergraduate courses. We hope that at the end of the journey, every reader will agree with our belief formulated in the title of the book, and find that Induction is indeed a powerful and elegant method of proof.

Titu Andreescu Vlad Crişan

Contents

Chapter 1

A Brief Overview of Induction

1.1 Theory and Examples

1.1.1 Introductory Notions

What is induction? Is it an axiom or a theorem? What kind of questions does it address? What is the largest context where it holds and can be applied? These are some of the questions that we will address throughout this chapter and in general, throughout this book.

Let us begin with the following informal discussion. Assume that a certain country decides at some point to become a kingdom and they choose their first royal family, call it generation 0. For each generation, the first offspring of the current royal family will be the successor to the throne. Unfortunately, all members of generation 0 suffer from a degenerative disease which is certain to be passed to the next generation. How can we prove that all the royal families that will rule that country will suffer from the same degenerative disease?

This seems like an intuitively clear result. Let us denote by $P(n)$ the proposition that the n-th generation suffers from the degenerative disease. We know that $P(0)$ is true and also that if $P(n)$ is true, then so is $P(n+1)$. We want to prove that $P(n)$ holds for any non-negative integer n. In fact, this is what the Principle of Induction says in general:

Principle of Induction. If $P(n)$ is a proposition that depends on a non-negative integer n such that $P(0)$ is true and $P(n)$ holds implies that $P(n+1)$ holds, then $P(n)$ is true for all non-negative integers n.

From now on, $\mathbb{N}$ denotes the set of natural numbers, i.e. non-negative integers. Here is one way we could try to prove the Principle of Induction:

Assume by contradiction that there is a non-negative integer for which the proposition P is false. Let us denote by A the non-empty set of all non-negative integers for which the proposition does not hold and let a be the smallest element of A, i.e. $P(a)$ is false and $P(0)$, $P(1)$, ..., $P(a-1)$ are all true. This implies in particular that $a > 0$ and $P(a-1)$ is true; but then $P(a)$ is true, which gives the desired contradiction. The conclusion follows.

However, the above argument has a gap. The reason is the following: in the proof we constructed a non-empty subset A of $\mathbb{N}$. Then we picked a to be the smallest element of A. But how do we know that such an a exists? Given that there may be very complicated and hard-to-understand subsets of $\mathbb{N}$, how do we know that **any** non-empty subset of $\mathbb{N}$ has a least element? This might seem like a silly question given the facts that we know about the natural numbers. But the convention in mathematics is that we should take only the simplest of these "facts" as our axioms. We should use this small set of axioms for constructing $\mathbb{N}$. Anything more complicated should be theorems, i.e. facts that can be deduced from our axioms.

The Principle of Induction is usually regarded as the simplest way to say something about all natural numbers. Thus it has been chosen as an axiom itself and it cannot be proved from other axioms. The fact that any non-empty subset of the natural numbers has a least element is seen as a little more complicated. We will prove that it follows from the Principle of Induction and we will think about the implication in this order and not the other way round. The Italian mathematician Giuseppe Peano was the first one to postulate rigorously a set of axioms on the natural numbers $\mathbb{N}$. We present them at the end of this section as an appendix, together with the proof that the Principle of Induction implies that every non-empty subset of the natural numbers has a least element.

Let us now look at a typical example which uses the Principle of Induction:

Show that for all positive integers n, the following identity holds:

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}.$$

Proof. Let $P(n)$ be the statement

$$P(n) : 1 + 2 + \ldots + n = \frac{n(n+1)}{2}.$$

We are required to prove that $P(n)$ is true for all positive integers n. We start by proving that $P(1)$ holds. This is the *base case*, i.e. the smallest value for which we have to show that our property holds.

$P(1)$ simply says that $1 = \frac{1 \cdot 2}{2}$, which is clear.

We now prove that assuming $P(n)$ is true for some $n \geq 1$, we obtain that $P(n+1)$ is true. This is called the *induction step* and the assumption that $P(n)$ is true is called the *induction hypothesis.*

From our induction hypothesis we know that

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}.$$

So to prove $1 + 2 + \ldots + n + n + 1 = \frac{(n+1)(n+2)}{2}$, it suffices to show that

$$\frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2},$$

which follows easily after bringing the left hand side to a common denominator. This shows that $P(n) \Rightarrow P(n+1)$, completing our proof.

Remark. Notice that the Principle of Induction as originally stated above uses the base case $n = 0$, while we started from $n = 1$ in the above proof. Nevertheless, it follows from the Principle of Induction that if $P(n_0)$ holds for some $n_0 \in \mathbb{N}$ and $P(n) \Rightarrow P(n+1)$, then $P(n)$ holds for any $n \geq n_0$ (hint: let $Q(n) = P(n + n_0)$). This observation regarding the base case will apply to all the variants we are going to present later.

The solution we have given to the above example illustrates the outline that a proof by induction should have. The crucial aspect one has to keep in mind is to always check both the base case(s) and the induction step. To see how important this is, consider the following two examples, both of which are false:

a) For any non-negative integer n, $n^2 + n + 41$ is prime.

b) For any non-negative integer n, $3n + 1$ is divisible by 3.

For statement a), one can show that the base cases are true, as in fact $n^2 + n + 41$ is prime for all $n \in \{0, 1, \ldots, 39\}$. However, we cannot prove that $P(n) \Rightarrow P(n + 1)$.

In the example b), assuming $P(n)$ was true i.e. $3 \mid (3n + 1)$ then

$$3(n + 1) + 1 = (3n + 1) + 3,$$

so $P(n + 1)$ is also true. But in this situation, we cannot prove that the base case holds, as in fact $P(0)$ is false.

Appendix: Peano's axioms for $\mathbb{N}$

Before we state the axioms introduced by Giuseppe Peano, we need to recall the following definitions: a **function** f between two sets A and B (written $f : A \to B$) is a correspondence which associates to each element $a \in A$ precisely one element $f(a) \in B$; the **image** of f, denoted $Im(f)$, is the set $Im(f) = \{f(a) : a \in A\}$ (note that $Im(f) \subseteq B$ and this inclusion can be strict, depending on the function f); a function is called **injective** if no two distinct elements of A are sent to the same element of B. For more details and examples, the reader can refer to Section 3.1 of this book.

The axioms introduced by Peano are the following:

1. There exists a special element $0 \in \mathbb{N}$;

2. There exists an injective function (called the *successor function*) $S : \mathbb{N} \to \mathbb{N}$ such that 0 does not lie in the image of S;

3. If $K \subseteq \mathbb{N}$ is a set such that $0 \in K$ and for every natural number $n \in K$ implies $S(n) \in K$, then $K = \mathbb{N}$ (Axiom of Induction).

At first glance, the Axiom of Induction as stated here does not look quite the same as the version we stated informally earlier. Before, we referred to a family of propositions $P(n)$. Now we have just a set K. However, they are easily seen to be equivalent if we interpret K as being the set of all n such that $P(n)$ is true.

Peano originally included several other axioms which are now commonly regarded as first-order logic axioms and in modern treatments are not included as axioms of the natural numbers. It is remarkable how complex a system can become even when it is founded only on a handful of axioms! The universe we live in seems to be governed by just four laws, namely the gravitational force, the electromagnetic force, the strong and the weak nuclear forces. And just look what is out there! Look at the complexity, the intricacy, the splendour! Also, the above three laws for the set of natural numbers which were postulated by Peano govern the whole universe of Number Theory and explain anything from operations such as addition and multiplication to modern results regarding bounded gaps between primes and arbitrary long arithmetic progressions of primes!

To get a small flavour of these, let us see how one can define the usual addition on $\mathbb{N}$ from Peano's axioms and how we can prove some standard properties of it. We define a function $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ inductively by

a) For all $a \in \mathbb{N}$, we define $a + 0 = a$;

b) Having defined $a + b$ for some $b \in \mathbb{N}$, we define $a + S(b) = S(a + b)$.

Lemma. The operation $+$ defined above is commutative.

Proof. We will prove the result in three steps. We start by showing that $n + 0 = 0 + n$ for any $n \in \mathbb{N}$. We do this by induction on n. Let $P(n)$ be the proposition

$$P(n) : n + 0 = 0 + n.$$

By a), we have $0 + 0 = 0$, so $P(0)$ holds.

Assume now that $P(n)$ is true for some $n \in \mathbb{N}$ and we deduce that $P(S(n))$ is true:

$$0 + S(n) \overset{b)}{=} S(0 + n) \overset{P(n)}{=} S(n + 0) \overset{a)}{=} S(n) \overset{a)}{=} S(n) + 0.$$

Therefore, by the Axiom of Induction, $P(n)$ holds for all $n \in \mathbb{N}$.

We now prove that $S(a) + n = S(a + n)$, for any $a, n \in \mathbb{N}$. For a fixed $a \in \mathbb{N}$, let $P(n)$ be the proposition

$$P(n) : S(a) + n = S(a + n).$$

When $n = 0$, we have $S(a) + 0 \overset{a)}{=} S(a) \overset{a)}{=} S(a + 0)$, so $P(0)$ is true.

Assume now that $P(n)$ holds for some $n \in \mathbb{N}$, and we prove that $P(S(n))$ also holds. We have

$$S(a) + S(n) \overset{b)}{=} S(S(a) + n) \overset{P(n)}{=} S(S(a + n)) \overset{b)}{=} S(a + S(n)).$$

Thus $P(S(n))$ is true. Hence, by the Axiom of Induction, we have that $P(n)$ is true for all $n \in \mathbb{N}$.

As a was arbitrary, the property $S(a) + n = S(a + n)$ holds for any a and n in $\mathbb{N}$.

Finally, we prove that $a + n = n + a$ for any $a, n \in \mathbb{N}$. We do so by induction on a, for some fixed n. As before, we let $P(a)$ be the corresponding Proposition.

We have $0 + n = n + 0$ from the first step, so $P(0)$ holds.

Assuming $P(a)$, we have from the second step

$$S(a) + n = S(a + n) \overset{P(a)}{=} S(n + a) \overset{b)}{=} n + S(a).$$

This completes the proof of our last step. Since n was arbitrary, we obtain that $a + n = n + a$, for any $a, n \in \mathbb{N}$, showing that $+$ is commutative, as we wanted.

From now on, we write $S(0) = 1$ and $S(n) = n + 1 = 1 + n$. The other standard properties of addition are proved in a similar manner. With the aid of addition, we can define the usual ordering "$\leq$" between two elements of $\mathbb{N}$ as: $a \leq b$ if there exists $c \in \mathbb{N}$ such that $a + c = b$. The properties of $\leq$ are now deduced from those of addition.

We have established so far what the Principle of Induction is and that it is an axiom. Before we move on to discussing other applications, variations and generalizations, we give the promised proof that the Principle of Induction implies that every non-empty subset $A \subset \mathbb{N}$ has a least element with respect to $\leq$:

Let $A \subset \mathbb{N}$ be a subset that has no minimal element with respect to $\leq$. We would like to prove that A is empty in this case. We set $P(n)$ to be the proposition:

$$P(n): \quad k \notin A, \quad \text{for all} \quad 0 \leq k \leq n.$$

Certainly $P(0)$ is true, as otherwise we would have $0 \in A$ and this would be the minimal element of A (as there is no non-negative integer smaller than 0). Let us now show that $P(n)$ true for some $n \geq 0$ implies $P(n+1)$ true as well:

If $P(n)$ is true and $P(n+1)$ is false, then $0, 1, 2, \ldots n \notin A$, but $n+1 \in A$, and then $n+1$ is the least element of A. This is a contradiction and we conclude that $P(n+1)$ is true.

By the Principle of Induction, $P(n)$ is true for every $n \in \mathbb{N}$ and hence A is empty.

The above argument shows that the only subset of $\mathbb{N}$ with no least element is the empty set. Therefore, every non-empty subset of $\mathbb{N}$ has a least element, which proves what we wanted.

1.1.2 Variants of Induction

We have seen above what the axiom of Induction is and how a typical proof by induction looks. We have also seen that in some applications we could have the base case bigger than 0. The next topic to look at is what happens when we modify our induction step. We prove the following:

Theorem. Let k be some fixed positive integer and $P(n)$ be a mathematical statement that satisfies the following properties:

1. All of $P(0), P(1), \ldots, P(k-1)$ are true;
2. $P(n) \Rightarrow P(n+k)$, for any $n \geq 0$.

Then $P(n)$ is true for every $n \in \mathbb{N}$.

Proof. When $k = 1$, the above statement is the Axiom of Induction, so we have nothing to prove. So let $k \geq 2$ and assume by contradiction that there is some $n \in \mathbb{N}$ for which $P(n)$ does not hold. Then the set $S = \{n \in \mathbb{N} : P(n) \text{ does not hold}\}$ is non-empty and so it has a least element, call it m.

Let r be the remainder when we divide m by k, so that we have $m = q \cdot k + r$, for some $q \in \mathbb{N}$ and $0 \leq r \leq k-1$. From the given hypotheses we know that $P(r)$ holds, thus in particular we cannot have $q = 0$. This implies $q \geq 1$, hence $0 \leq (q-1) \cdot k + r < m$. Because m is the least element of S, we have that $P((q-1) \cdot k + r)$ is true and then using the second hypothesis, $P((q-1) \cdot k + r + k) = P(m)$ is true, giving a contradiction.

Remark. What is important to note when using this variant is that if we prove $P(n) \Rightarrow P(n+k)$, we need to check k base cases instead of just one; for example, if $k = 2$, then checking just $P(0)$, together with $P(n) \Rightarrow P(n+2)$ would only imply that $P(n)$ holds when n is an even positive integer! The above variant (and hence also the Axiom of Induction obtained for $k = 1$) bears the name of the *Weak Principle of Induction.*

Example 1.1. Prove that any square can be divided in n (not necessarily congruent) squares ($n \geq 6$).

Solution. The key observation is that given any square, we can partition it into four smaller congruent squares, which shows that $P(n) \Rightarrow P(n+3)$. Hence all we need is to check that the statement holds for $n = 6, 7, 8$.

For $n = 6$, we can divide the square into 9 congruent squares and then join four of them into a larger one.

For $n = 7$, we first divide the square into 4 equal squares and then divide one of those further into 4 equal squares.

Finally, for $n = 8$, divide the square into 16 congruent squares and then join 9 of them into a bigger one.

The next variant is known as the *Strong Principle of Induction*:

Theorem. Let $P(n)$ be a statement about $n \in \mathbb{N}$. Suppose that

1. $P(0)$ is true;
2. $\forall n \in \mathbb{N}$, if $P(k)$ is true $\forall k < n$ then $P(n)$ is true.

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof. Suppose that $P(0)$ is true and for all $n \in \mathbb{N}$, if $P(0)$, $P(1)$, ..., $P(n-1)$ are all true, then $P(n)$ is true. We want to show that $P(n)$ is true for all n using the weak principle.

Let $Q(n)$ be the statement "$P(k)$ is true $\forall k \leq n$". Then $Q(0)$ is true from the hypothesis. Suppose that $Q(n)$ is true. Then all of $P(0), P(1), \ldots, P(n)$ are true. So $P(n+1)$ is true. Hence $Q(n+1)$ is true. By the Weak Principle of Induction, $Q(n)$ is true for all n. So $P(n)$ is true for all n.

Example 1.2. Show that if $x + \dfrac{1}{x} \in \mathbb{Z}$, then $x^n + \dfrac{1}{x^n} \in \mathbb{Z}$ for all positive integers n.

Solution. We begin with the base cases $P(1)$ and $P(2)$. $P(1)$ is true from the hypothesis. Since

$$x^2 + \frac{1}{x^2} = \left(x + \frac{1}{x}\right)^2 - 2,$$

$P(2)$ is also true.

Assume now that $P(n-1)$ and $P(n)$ are both true for some $n \geq 2$. We now observe that

$$x^{n+1} + \frac{1}{x^{n+1}} = \left(x + \frac{1}{x}\right)\left(x^n + \frac{1}{x^n}\right) - \left(x^{n-1} + \frac{1}{x^{n-1}}\right) \in \mathbb{Z},$$

because all of $P(1)$, $P(n-1)$, $P(n)$ hold. Therefore, by the Strong Principle of Induction, $P(n)$ is true for all positive integers n, completing the proof.

Remark. More generally, we can consider the following variant: We have a proposition depending on $n \in \mathbb{N}$, a non-empty set of natural numbers A for which $P(a)$ is true $\forall a \in A$ and a sequence of operations σ which applied repeatedly to the elements of A give all elements of a non-empty set $K \subset \mathbb{N}$. If for any such operation σ, $P(a)$ true implies $P(\sigma(a))$ true, then using the Principle of Induction, one can show that $P(n)$ is true for all $n \in K$. A good example of such a variant is the *Cauchy Induction*, where usually $A = \{1\}$ or $A = \{2\}$ and we show $P(n) \Rightarrow P(2n)$ and $P(n) \Rightarrow P(n-1)$. Another popular example is when we have $A = \{1\}$ and $K = \{1, \ldots, n\}$, for some fixed $n \geq 1$, so we prove that $P(k)$ holds for $1 \leq k \leq n$. These two examples shall be discussed in greater detail in the chapter on Inequalities, where we will make intensive use of this technique.

1.1.3 Paradox of Induction

We are about to illustrate one of the most intriguing facts about the Principle of Induction, the so called *Paradox of Induction*.

For example, say we want to prove for every positive integer n the proposition

$$1 + \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^n} < 2,$$

denoted by $P(n)$. We easily check that $P(1)$ is true. Then we assume $P(n)$ is true and we try to prove $P(n+1)$ i.e. assume

$$1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^n}<2 \tag{1}$$

and try to prove

$$1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^{n+1}}<2.$$

But the only thing we can deduce directly from (1) is that

$$1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^{n+1}}<2+\frac{1}{2^{n+1}},$$

which is not good enough for what we want. At first glance, it may seem that this statement is too hard to be proved by induction. Actually, it is too weak!

The paradox of induction is that sometimes it is easier to prove a stronger statement by induction. Let us try to use induction to prove for every natural number n the proposition

$$1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^n}=2-\frac{1}{2^n}, \tag{2}$$

denoted by $Q(n)$. Clearly $Q(n)$ is stronger than $P(n)$. Again, we easily check that $Q(1)$ is true. We then assume $Q(n)$ is true and try to prove $Q(n+1)$ i.e. assume (2) and try to prove

$$1+\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^{n+1}}=2-\frac{1}{2^{n+1}},$$

which is straightforward by adding $\frac{1}{2^{n+1}}$ to both sides in (2). The conclusion follows by the Weak Principle of Induction.

The paradox of induction is exactly what makes induction so interesting to us. Say we want to prove that for every natural number n the proposition $P(n)$ holds, but induction does not work. Then we consider for every natural number n another proposition $Q(n)$ stronger than $P(n)$ and use induction to prove that for every natural number n, the proposition $Q(n)$ is true. The idea behind the paradox of induction is quite simple to explain, but there is one key question: how do we choose our stronger statement $Q(n)$? In most situations, the stronger statement is suggested by the induction step that we have to perform. Let us look at some further examples which show how we apply this idea.

Example 1.3. Show that for any $n \geq 1$, there exist integers a_n, b_n such that

$$\left(\frac{1+\sqrt{5}}{2}\right)^n = \frac{a_n + b_n\sqrt{5}}{2}.$$

Solution. We try to proceed by proving the statement using induction on n. For $n = 1$, the statement is clear by taking $a_1 = 1$ and $b_1 = 1$.

Assume now that the result holds for some $n \geq 1$, that is

$$\left(\frac{1+\sqrt{5}}{2}\right)^n = \frac{a_n + b_n\sqrt{5}}{2},$$

for some integers a_n and b_n. Then

$$\begin{aligned}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} &= \frac{a_n + b_n\sqrt{5}}{2} \cdot \frac{1+\sqrt{5}}{2} \\ &= \frac{a_n + b_n\sqrt{5} + a_n\sqrt{5} + 5b_n}{4} \\ &= \frac{\frac{a_n+5b_n}{2} + \frac{a_n+b_n}{2}\sqrt{5}}{2}.\end{aligned}$$

We would now like to say that we take $a_{n+1} = \dfrac{a_n + 5b_n}{2}$ and $b_{n+1} = \dfrac{a_n + b_n}{2}$. But for these to be integers, we need to know that a_n and b_n have the same parity. This suggests that we could probably prove the following stronger statement:

"For any $n \geq 1$, there exist integers a_n, b_n of same parity such that

$$\left(\frac{1+\sqrt{5}}{2}\right)^n = \frac{a_n + b_n\sqrt{5}}{2}."$$

Notice that if we manage to prove this assertion, we clearly prove the one the question asks as well.

For $n = 1$ we had $a_1 = b_1 = 1$ above, so the base case is verified. Now for the induction step, we obtain as before that

$$\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} = \frac{\frac{a_n+5b_n}{2} + \frac{a_n+b_n}{2}\sqrt{5}}{2}.$$

Since a_n and b_n have the same parity from the inductive hypothesis, we have that $a_{n+1} = \dfrac{a_n + 5b_n}{2}$ and $b_{n+1} = \dfrac{a_n + b_n}{2}$ are both integers. Also, we have that $a_{n+1} = b_{n+1} + 2b_n$, so a_{n+1} and b_{n+1} have the same parity, proving the inductive step. By the Weak Principle of Induction, the proof is complete.

Example 1.4. Show that for any $n \geq 3$, $n!$ can be written as the sum of n distinct divisors of itself.

Solution. We try to prove our statement by induction on n. The base case is $n = 3$, for which we have $3! = 1 + 2 + 3$.

Assume now that the statement holds for some n, that is

$$n! = d_1 + d_2 + \ldots + d_n,$$

with $d_1 < d_2 < \ldots < d_n$ distinct divisors of $n!$. Using $(n+1)! = n!(n+1)$ we have that

$$(n+1)! = (n+1)d_1 + (n+1)d_2 + \ldots + (n+1)d_n,$$

where $(n+1)d_1 < (n+1)d_2 < \ldots < (n+1)d_n$.

To make the inductive step work, we need to come up with a sum of $n+1$ divisors of $(n+1)!$ from the n divisors $(n+1)d_1 < (n+1)d_2 < \ldots < (n+1)d_n$. One of the most natural things that one could try is to write $(n+1)d_1$ as $nd_1 + d_1$. Then this would create an extra divisor and definitely $d_1 \mid (n+1)!$, because from the inductive hypothesis we know that $d_1 \mid n!$. The problem with this idea is that it might be that even though $d_1 \mid n!$, we might not have that $nd_1 \mid (n+1)!$ (for example, if $d_1 = n$ and n is a prime number). However, this would not be an issue if we could get some control over d_1 in the inductive assumption. For example, if we could always set $d_1 = 1$, then clearly nd_1 is always a divisor of $(n+1)!$. Moreover, when we write $(n+1)d_1 = d_1 + nd_1$, we also get for the inductive step that $d_1 = 1$. So let us try to prove the following, stronger statement:

"For any $n \geq 3$, $n!$ can be written as the sum of n distinct divisors of itself, one of them being equal to 1."

We prove this statement by induction on n. We saw above that $3! = 1+2+3$, so the assertion holds for $n = 3$. For the inductive step, assume that

$$n! = 1 + d_2 + \ldots + d_n,$$

where $d_2 < d_3 < \ldots < d_n$ are distinct divisors of $n!$, all bigger than 1. We have that

$$\begin{aligned}(n+1)! &= (n+1) + (n+1)d_2 + \ldots + (n+1)d_n \\ &= 1 + n + (n+1)d_2 + \ldots + (n+1)d_n.\end{aligned}$$

Clearly $1 \mid (n+1)!$, $n \mid (n+1)!$ and since $d_k \mid n!$ for $k = 2, \ldots n$, we also have that $(n+1)d_k \mid (n+1)!$. Also, since $n \geq 3$ and $d_2, \ldots, d_n$ are bigger than 1, we have that the numbers 1, n, $(n+1)d_2, \ldots, (n+1)d_n$ are all distinct. This completes the inductive step and the proof of our statement.

Example 1.5. Prove that for all positive integers $n \geq 1$, one has

$$\frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} < \frac{7}{10}.$$

Solution. First of all, let us see why we should strengthen the inequality. Let us denote

$$x_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n}.$$

Ideally, we would like to set $P(n)$ to be the mathematical statement $x_n < \frac{7}{10}$ and prove it by induction. Unfortunately, we have that

$$x_{n+1} - x_n = \frac{1}{2n+1} + \frac{1}{2n+2} - \frac{1}{n+1} = \frac{1}{(2n+1)(2n+2)} > 0.$$

So our sequence is increasing and therefore, from $x_n < \frac{7}{10}$ we cannot deduce that $x_{n+1} < \frac{7}{10}$, as we would like for the induction step. So let us sharpen the condition to something of the form

$$x_n + f(n) < \frac{7}{10},$$

where $f : \mathbb{Z}_{>0} \to [0, \infty)$. This statement clearly implies the one we are interested in, since f takes non-negative values and therefore $x_n + f(n) < \frac{7}{10}$ implies $x_n < \frac{7}{10}$.

We now discuss two ideas which are useful in determining what our choice for f should be:

Firstly, notice that the obstruction we had above for applying induction directly was that our original sequence was increasing. So if the sequence $y_n = x_n + f(n)$ was non-increasing, then clearly $y_n < \frac{7}{10}$ implies $y_{n+1} < \frac{7}{10}$, as we have

$$y_{n+1} \leq y_n < \frac{7}{10}.$$

So to make the induction step work, we impose the condition $y_n - y_{n+1} \geq 0$, which gives

$$f(n+1) \leq f(n) - \frac{1}{(2n+1)(2n+2)}.$$

In particular, this shows that f should be a decreasing function.

This is the key point for this type of questions: we strengthen the result, so that we get rid of the obstructions we had in applying induction in the first place. On a similar note, if we had an inequality of the form $a_n > c$, for some constant c, and $(a_n)_{n \geq 0}$ was a decreasing sequence, then we would define a sequence b_n such that b_n was non-decreasing and $a_n \geq b_n > c$. This is a powerful tool which is also often used in Analysis when dealing with limits of sequences.

The second idea we are about to discuss is rather technical and may or may not occur in this type of problems:

We know that the other crucial part of a proof by induction is the validity of the base case. If we want the result to hold for $n = 1$, then this gives

$$\frac{1}{2} + f(1) < \frac{7}{10} \Rightarrow f(1) < \frac{1}{5}.$$

However, here we have a technical observation which gives us some flexibility in the following sense:

We know that $x_n < \frac{7}{10}$ holds for the first few values of n by hand computations (for $n = 1$ we get $\frac{1}{2} < \frac{7}{10}$, for $n = 2$ we get $\frac{1}{3} + \frac{1}{4} < \frac{7}{10}$, etc). We also know that $y_n < \frac{7}{10}$ implies $x_n < \frac{7}{10}$. Therefore, since y_n is non-increasing, it could be that even if we let $y_1 > \frac{7}{10}$, we obtain $y_{n_0} < \frac{7}{10}$, for some small positive integer n_0, provided the sequence y_n decreases fast enough.

So instead of asking for the result to hold from $n = 1$, we could start our induction from $n = n_0$, where n_0 is some positive integer (in our case it will turn out that $n_0 = 4$) and verify by hand the result for the values from 1 to n_0. To illustrate how this could actually become useful, we return to our question:

We established so far that f should be a function taking non-negative values and which satisfies

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

for all positive integers n. Notice that the above inequality could be rewritten as

$$f(n) - f(n+1) > \frac{1}{2n+1} - \frac{1}{2n+2}.$$

Now, what is a better candidate for a function that takes non-negative values and decreases at a rate comparable to $\frac{1}{n}$ than something of the form $\frac{k}{n}$?

So we let $f(n) = \frac{k}{n}$ and we determine the suitable value for k. We have

$$f(n+1) < f(n) - \frac{1}{(2n+1)(2n+2)} \Leftrightarrow \frac{(1-4k)n - 2k}{2n(n+1)(2n+1)} < 0.$$

Since the second inequality must hold for any $n \geq 1$, we obtain that $k \geq \frac{1}{4}$. Notice that if we choose $k = \frac{1}{4}$, we actually have

$$y_1 = \frac{1}{2} + \frac{1}{4} > \frac{7}{10}.$$

But

$$y_4 = \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{16} < \frac{7}{10},$$

and since y_n is decreasing, the same will be true for any y_n with $n \geq 4$. Hence, we basically provided all the results required for a proof by induction that $y_n < \frac{7}{10}$ for all $n \geq 4$ and therefore $x_n < \frac{7}{10}$ for all $n \geq 4$. As x_n is increasing, we clearly have that the result also holds for $n = 1, 2, 3$, completing the proof.

We conclude this section by discussing the following example:

Example 1.6. Let F_n denote the Fibonacci numbers defined by $F_0 = F_1 = 1$ and $F_{n+1} = F_n + F_{n-1}$ for $n \geq 1$. Prove that for any non-negative integer m, one has

$$F_{2m+1} = F_{m+1}^2 + F_m^2.$$

Solution. It is fair to start by saying that one can actually use Induction to prove the question in the given form, though the computations are lengthy and messy. We present below a way in which we can generalize the statement and make the proof easier. So let us try to prove instead:

"For any $m, n \in \mathbb{N}$, we have $F_{m+n} = F_{m+1}F_n + F_mF_{n-1}$."

Notice that the above statement implies the original one by taking $n = m+1$. For the second statement, we have the following proof:
For a fixed m, we let $P(n)$ represent the statement that

$$F_{m+n+1} = F_{m+1}F_{n+1} + F_mF_n.$$

Now,

$$P(0) : F_{m+1} = F_{m+1};$$

$$P(1) : F_{m+2} = F_{m+1} + F_m,$$

which are trivially true. Now, if $P(n)$ is true for $n \leq k$, then:

$$\begin{aligned} F_{m+k+1} &= F_{m+k} + F_{m+k-1} \\ &= F_{m+1}F_k + F_mF_{k-1} + F_{m+1}F_{k-1} + F_mF_{k-2} \\ &= F_{m+1}F_{k+1} + F_mF_k, \end{aligned}$$

so $P(k+1)$ is true as well. Thus, the given relation is always true for any m, n. (Notice that our proof of the induction step used $P(k)$ and $P(k-1)$, so it was crucial that we checked two base cases $P(0)$ and $P(1)$.)

Remark. We will see a different proof of the previous example in the next chapter, where we also study several other identities regarding the Fibonacci numbers.

1.1.4 Well-Ordering and Transfinite Induction

For a set X, a **partial order** $\preceq$ on X is a binary relation (i.e. involving two elements) satisfying the following properties:

1. $x \preceq x$, $\forall x \in X$ (reflexivity);

2. For any $x, y \in X$, if $x \preceq y$ and $y \preceq x$, then $x = y$ (antisymmetry);

3. For any $x, y, z \in X$, if $x \preceq y$ and $y \preceq z$, then $x \preceq z$ (transitivity).

Example. As mentioned in the first section, we can prove starting from Peano's axioms that the usual ordering $\leq$ on the natural numbers is a partial order. Another example of partial order on $\mathbb{N}$ is $a \preceq b$ if $a \mid b$ (check!).

A partial order under which every pair of elements is comparable is called a **total order**. In the above examples, one can easily check that the usual ordering on $\mathbb{N}$ is a total order while the partial ordering given by $a \preceq b$ if $a \mid b$ is not.

A subset $Y \subset X$ for which every pair of elements is comparable is called a **chain**. A total order under which every non-empty subset $Y \subset X$ has a least element is called a **well-order**.

Example. $\mathbb{N}$ with the usual order is well-ordered, while $\mathbb{Z}$ is not well-ordered, since the set of even integers has no minimum. The positive rationals are also not well-ordered under the usual order.

We have seen in the first section that once we have defined the three axioms as postulated by Peano, we can prove that the natural numbers $\mathbb{N}$ satisfy the nice property of being a well-ordered set (the well-ordering principle). However, we saw that the problem when we tried to prove that the Principle of Induction holds was that we needed to assume that $\mathbb{N}$ with the usual order relation was a well-ordered set. So the natural question one can ask is why do we not simply take the well-ordering principle as our axiom and deduce the other axioms, including the principle of induction from it?

It turns out that assuming the well-ordering of $\mathbb{N}$ we can deduce the first two axioms that Peano postulated, but not the Axiom of Induction. Part of the reason for this is that we actually used the Axiom of Induction in the process of defining the usual order $\leq$ on $\mathbb{N}$. We can prove that taking the well-ordering principle as our axiom is not good enough as follows: let $\mathbb{N}$ now be a general set and $\leq$ be a well-order on $\mathbb{N}$, such that $\mathbb{N}$ does not have an upper bound with respect to $\leq$ (i.e. there is no $x \in \mathbb{N}$ such that $y \in \mathbb{N} \Rightarrow y \leq x$). The special element $0 \in \mathbb{N}$ can be defined as the least element of the non-empty subset $\mathbb{N} \subset \mathbb{N}$; for any $n \in \mathbb{N}$, the successor of n, $S(n)$, can be defined as the least element of the non-empty subset $A = \{y > n\} \subset \mathbb{N}$; Indeed, the successor function $S : \mathbb{N} \to \mathbb{N}$ is injective and $0 \notin Im(S)$. But how do we know that if $K \subset \mathbb{N}$ is any set such that $0 \in K$ and for every natural number n, if $n \in K$ then $S(n) \in K$, then $K = \mathbb{N}$? We do not know! But how could we guarantee that we do not know?

Surprisingly, the answer is in any dictionary or phone book! Consider the set $\mathbb{N}$ with the usual well-order $\leq$ and consider the set $X = \mathbb{N} \times \mathbb{N}$ with the lexicographic well-order $\preceq$ i.e. $(a_1, b_1) \preceq (a_2, b_2)$ if $a_1 \leq a_2$ or $a_1 = a_2$ and $b_1 \leq b_2$. Then $(0,0)$ is the smallest element of X with respect to $\preceq$ and the successor function $S : X \to X$ satisfies $S((a,b)) = (a, b+1)$. Let $K = \{(0,a) : \ a \in \mathbb{N}\} \subset X$. Then $(0,0) \in K$ and for every $x \in X$, if $x = (0,a) \in K$, then $S(x) = (0, a+1) \in K$. However, K is a proper subset of X. Notice that actually we cheated a bit as we assumed the existence of the set of natural numbers as postulated by Peano. So actually we know for sure we do not know, given that we still know something... Therefore, if we take "$\mathbb{N}$ with usual order $\leq$ is well-ordered" as an axiom, this axiom does NOT define the natural numbers.

Up to this stage, we have only discussed induction in the context of the natural numbers. But what happens if we consider a general set X? Is there an equivalent notion of induction as the one we defined for $\mathbb{N}$?

Notice that if we look at a very large set, like $\mathbb{R}$ (the reals), we do not have an equivalent notion of successor function as we had for $\mathbb{N}$, as between any two reals, there are infinitely many others. Thus, we cannot simply take the three axioms of Peano and generalize them to any set. However, we saw that once we have established what the natural numbers are, we can define the notion of induction on $\mathbb{N}$ if we assumed that $\mathbb{N}$ is well-ordered. So what happens if we postulate the axiom that any set X admits a well-order $\preceq$? Then a universe is born. Strangely enough, this axiom is equivalent to the following axiom:

given an index set X and a family of sets $F = \{A_i : i \in X\}$, then there exists a choice function $f : X \to \bigcup_{i \in X} A_i$, such that for any $i \in X$ we have $f(i) \in A_i$. Usually, the latter is called *The Axiom of Choice* and the former is called *The well-ordering theorem.*

How can we imagine a well-ordered set X, especially a big one? We have to imagine fractals; a fractal is an object with the property that no matter how much we zoom in or out we still see the same thing. Imagine the elements of X as points on a line. If we sit very close to X at the beginning of X, we see a copy of $\mathbb{N}$ of points. From far away, the copy of $\mathbb{N}$ appears as a single point and we see a copy of $\mathbb{N}$ of these single points each of which is actually a copy of $\mathbb{N}$. From really far away, this copy of $\mathbb{N}$ appears as a single point and we see a copy of $\mathbb{N}$ of these single points each of which is actually a copy of $\mathbb{N}$ copies of $\mathbb{N}$. The trip to a fractal is the most boring trip ever as the landscape never changes.

Theorem. (**Transfinite Induction**) Let X be a set with a well-order $\preceq$, and $P(x)$ a proposition for every element $x \in X$ which satisfies the following property:

(*) If $x \in X$ and for every $y \prec x$, $P(y)$ is true, then $P(x)$ is true ($y \prec x$ means $y \preceq x$ and $y \neq x$).

Then $P(x)$ is true for all $x \in X$.

Proof. Assume for the sake of contradiction that the set $K = \{x \in X : P(x) \text{ is false}\}$ is non-empty and let x be the least element of K. Then for every $y \prec x$, $P(y)$ is true. It follows from the hypothesis that $P(x)$ is true, a contradiction. Therefore, we conclude that $P(x)$ is true for all x.

Given a well-ordered set S of indices, a transfinite construction on S means to construct a family of sets $K(i)$ for each $i \in S$ as follows: say that for all $i \in S$, given all the sets $K(j)$ with $j \prec i$, we can construct a specific set $K(i)$. Then by the above theorem, we know that we can construct $K(i)$ for all $i \in S$. Let us look at a very beautiful example which uses this ideas:

Example 1.7. (**Sierpinski**) Is there a set X of points in the plane such that every possible line in the plane contains exactly two points from our set?

The answer to the question is, quite remarkably, affirmative. Before we present the solution to the problem, we need to present some facts from Set Theory. The reader who is not familiarized with some the terms we are about to use or would like to see some further details can consult any standard text book on Set Theory.

We say two sets X, Y have the same cardinality and we write $|X| = |Y|$ if there is a bijection between them. If there is an injective function $f : X \to Y$, then we define $|X| \leq |Y|$. The Cantor-Schröder-Bernstein theorem says that if $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$, so cardinality behaves like the usual notion of size. This allows us to extend the notion of two sets being the same size or one being larger than the other to infinite sets. A set with the same cardinality as a subset of $\mathbb{N}$ is said to be countable and a set with the same cardinality as $\mathbb{N}$ is said to be countably infinite. Examples of countably infinite sets include $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$. There are many more countable sets than you might at first suppose. If A is a countable set then so is $A \times A$. In fact any union of countably many countable sets is still countable. A set with the same cardinality as $\mathbb{R}$ is said to have the cardinality of the continuum. A famous theorem of Cantor shows that $\mathbb{R}$ is not countable (is uncountable), and hence $\mathbb{R}$ is "bigger" than $\mathbb{N}$ in this sense.

The **Continuum hypothesis** asserts that there is no cardinal bigger than $\mathbb{N}$ and smaller than $\mathbb{R}$. For the solution which we present below we shall assume the Continuum hypothesis, though there are solutions which do not require this. More precisely, we will need the following statement, which is equivalent to the Continuum hypothesis: "We can well-order $\mathbb{R}$ such that for every real $r \in \mathbb{R}$, if we look at the set of points $a < r$, this set is countable."

Let us quickly see that the two statements are indeed equivalent. Assume first the Continuum hypothesis. Choose an arbitrary well-order $\preceq$ on $\mathbb{R}$. If this satisfies what we want, we are done. Otherwise, for each $r \in \mathbb{R}$, we look at the set of elements less than r with respect to $\preceq$. This set is either countable or has the same cardinality as $\mathbb{R}$. Choose the minimal r with the property that the set $S = \{x \in \mathbb{R} : x \prec r\}$ is uncountable. So there must be a bijection $f : \mathbb{R} \to S$. Notice that S is well-ordered with respect to $\preceq$. Now we can reorder $\mathbb{R}$ via $x \leq y$ if $f(x) \preceq f(y)$. One can check immediately that this defines a well-order on $\mathbb{R}$ which satisfies the condition we want. Proving the converse statement is immediate and we leave it for the reader.

The second fact which we will need is that $\mathbb{R}^2$ and $\mathbb{R}$ have the same cardinality. To see this, notice that we have an injection $(0,1) \times (0,1) \to (0,1)$ given by $(0.a_1a_2a_3\ldots, 0.b_1b_2b_3\ldots) \to 0.a_1b_1a_2b_2\ldots$ (we discard here the infinite sequences of 9's, e.g. $0.099999\ldots = 0.1$) and $(0,1)$ is in bijection with $\mathbb{R}$ (check!).

We are now ready to give the solution to our problem. Let us denote the set of all possible lines in the plane by L. Clearly $|\mathbb{R}| \leq |L| \leq |\mathbb{R}^2|$, so by the remarks we made above, L has the same cardinality as $\mathbb{R}$. Assuming the Continuum hypothesis, we can choose a well-order $\preceq$ on L so that for each $l \in L$ we have that the set $\{x \in L : x \prec l\}$ is countable. We now use a transfinite construction on L to construct our set X. Namely, for each $l_i \in L$, we construct a set of points $K(l_i) \subset \mathbb{R}^2$ such that: $K(l_i)$ is countable, no three points of $K(l_i)$ are collinear, each line $l \in L$ with $l \preceq l_i$ contains exactly two points from $K(l_i)$, and if $l_a \prec l_b$, then $K(l_a) \subseteq K(l_b)$. Let l_1 be the least element of L with respect to $\preceq$. We construct $K(l_1)$ by choosing two random points on l_1.

Let now l_i be some element in L and assume we have constructed the sets $K(l_j)$ for all $l_j \prec l_i$. Consider the set $S = \bigcup_{l_j \prec l_i} K(l_j)$. By construction, there are no three collinear points in any $K(l_j)$ for any $l_j \prec l_i$, hence there are no three collinear points in S. Also by construction each of the sets $K(l_j)$ is countable, hence so is S. If l_i already contains two points from S, we take $K(l_i) = S$ and this satisfies all our conditions. Otherwise we need to add one or two points to S (depending on whether l_i already passes through one or zero points of S). These points must lie on l_i and must not be collinear with any two other points of S. Since S is countable, so is the set of lines they determined by points in S (since A countable implies $A \times A$ countable). On the other hand, the set of points on l_i is uncountable, so we can pick either one or two points on l_i with the required properties. Then we let $K(l_i)$ be S union with these points.

We now let $X = \bigcup_{l \in L} K(l)$ and we are done.

1.2 Proposed Problems

Problem 1.1. Prove that for all $n \geq 1$:

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} < \frac{1}{\sqrt{3n}}.$$

Problem 1.2. Show that for every positive integer $n \geq 2$ one has

$$\frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{n^2} < 1.$$

Problem 1.3. (**China 2004**) Prove that every positive integer n, except a finite number of them, can be represented as a sum of 2004 positive integers: $n = a_1 + a_2 + \cdots + a_{2004}$, where $1 \leq a_1 < a_2 < \cdots < a_{2004}$, and $a_i \mid a_{i+1}$ for all $1 \leq i \leq 2003$.

Problem 1.4. The sequence $(a_n)_{n \geq 1}$ is defined by

$$a_1 = \frac{1}{2},\ a_{n+1} = \frac{2n-1}{2n+2} a_n.$$

Prove that $a_1 + a_2 + \ldots + a_n < 1$ for all $n \geq 1$.

Problem 1.5. Consider S the set of all binary sequences of length n. It is partitioned into two sets A and B, each having 2^{n-1} elements. Two sequences, one from A and one from B, form a *tentacle* if they differ in only one position. Prove that there are at least 2^{n-1} tentacles.

Problem 1.6. Let S be a set of points in the plane (not necessarily finite) and consider a complete graph G which has the points in S as its nodes. We color the edges of this graph with two colors. It is known (for example from Ramsey Theory) that if S has the same cardinality as $\mathbb{N}$, then we will be able to find a monochromatic complete subgraph whose vertex set also has the cardinality of $\mathbb{N}$. Is it true that if S has same cardinality as $\mathbb{R}$ we will find a monochromatic subgraph whose vertex set also has the cardinality of $\mathbb{R}$?

Chapter 2

Sums, Products, and Identities

2.1 Theory and Examples

The first mathematical statements proved by induction were related to sum identities, to the extent where nowadays it is very common to write $S(n)$ (where S stands for sum) for the mathematical statement depending on n that we would like to prove. These are perhaps the easiest applications of induction, since in order to prove such an identity we usually employ only basic algebraic manipulations such as expanding the parentheses or taking common denominators. We also mainly require only the Weak Principle of Induction for such problems.

We recall the following standard facts which we shall require later in this chapter:

For a positive integer n and an integer $0 \leq k \leq n$, we denote by $\binom{n}{k}$ the number of ways in which one can choose a set of k elements from a given set with n elements. We have the well-known formula:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}.$$

When $k < 0$ or $k > n$, we set $\binom{n}{k} = 0$. Using the above formula, one can prove the following relation, called Pascal's identity:

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}, \quad \text{for} \quad 1 \le k \le n.$$

Let us look at some examples:

Example 2.1. Prove that for any positive integer n, one has

$$1 \cdot 2 \cdot 3 + 2 \cdot 3 \cdot 4 + \ldots + n \cdot (n+1) \cdot (n+2) = \frac{n(n+1)(n+2)(n+3)}{4}.$$

Solution. We begin our solution by setting the propositional statement which we will prove by induction:

Let

$$P(n) : 1 \cdot 2 \cdot 3 + 2 \cdot 3 \cdot 4 + \ldots + n \cdot (n+1) \cdot (n+2) = \frac{n(n+1)(n+2)(n+3)}{4}.$$

The base case is the least value of n which the question addresses and since we are asked to prove that the identity holds for all positive integers n, the smallest value is $n = 1$.

For $n = 1$ we need to check that

$$1 \cdot 2 \cdot 3 = \frac{1 \cdot 2 \cdot 3 \cdot 4}{4},$$

which is true. Therefore $P(1)$ holds.

Assume now that $P(n)$ holds for some $n \ge 1$, that is

$$1 \cdot 2 \cdot 3 + 2 \cdot 3 \cdot 4 + \ldots + n \cdot (n+1) \cdot (n+2) = \frac{n(n+1)(n+2)(n+3)}{4}. \quad (1)$$

We need to prove that $P(n+1)$ is true, namely

$$1 \cdot 2 \cdot 3 + 2 \cdot 3 \cdot 4 + \ldots + (n+1) \cdot (n+2) \cdot (n+3) = \frac{(n+1)(n+2)(n+3)(n+4)}{4}.$$

To derive this, we add $(n+1)\cdot(n+2)\cdot(n+3)$ to both sides of (1) and we obtain

$$\begin{aligned}
&1\cdot 2\cdot 3+\ldots+(n+1)\cdot(n+2)\cdot(n+3)\\
&=\frac{n(n+1)(n+2)(n+3)}{4}+(n+1)\cdot(n+2)\cdot(n+3)\\
&=(n+1)(n+2)(n+3)\left(\frac{n}{4}+1\right)\\
&=\frac{(n+1)(n+2)(n+3)(n+4)}{4}.
\end{aligned}$$

This shows that $P(n+1)$ is true and by the Weak Principle of Induction, $P(n)$ is true for all $n\geq 1$, which is what we wanted to show.

Example 2.2. Show that for any positive integer n, we have

$$\left(1-\frac{1}{n^2}\right)\left(1-\frac{1}{(n+1)^2}\right)\cdots\left(1-\frac{1}{(2n-1)^2}\right)=1-\frac{1}{2n-1}.$$

Solution. We prove the result by induction on n. For $n=1$, we get $0=0$, and for $n=2$, we get $\frac{3}{4}\cdot\frac{8}{9}=\frac{2}{3}$. Thus the base cases are verified.

Assume now that the identity holds for some $n\geq 2$, i.e.

$$\left(1-\frac{1}{n^2}\right)\left(1-\frac{1}{(n+1)^2}\right)\cdots\left(1-\frac{1}{(2n-1)^2}\right)=1-\frac{1}{2n-1}. \tag{1}$$

We need to prove it for $n+1$, that is

$$\left(1-\frac{1}{(n+1)^2}\right)\left(1-\frac{1}{(n+1)^2}\right)\cdots\left(1-\frac{1}{(2n+1)^2}\right)=1-\frac{1}{2n+1}.$$

To derive this, we multiply both sides of (1) by

$$\left(1-\frac{1}{(2n)^2}\right)\left(1-\frac{1}{(2n+1)^2}\right)\left(1-\frac{1}{n^2}\right)^{-1}.$$

Note that to avoid dividing by zero, we must have $n > 1$ here, and this is why we included $n = 2$ as a base case. We obtain

$$\begin{aligned}
&\left(1-\frac{1}{(n+1)^2}\right)\cdots\left(1-\frac{1}{(2n+1)^2}\right)\\
&=\left(1-\frac{1}{2n-1}\right)\left(1-\frac{1}{(2n)^2}\right)\left(1-\frac{1}{(2n+1)^2}\right)\left(1-\frac{1}{n^2}\right)^{-1}\\
&=\frac{2(n-1)}{2n-1}\cdot\frac{(2n-1)(2n+1)}{4n^2}\cdot\frac{4n(n+1)}{(2n+1)^2}\cdot\frac{n^2}{(n-1)(n+1)}\\
&=\frac{2n}{2n+1}\\
&=1-\frac{1}{2n+1},
\end{aligned}$$

as we wanted.

Example 2.3. Prove that for any positive integer n we have

$$1^2+2^2+\ldots+n^2=\frac{n(n+1)(2n+1)}{6}.$$

Solution. Let $P(n) : 1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$.
$P(1)$ says that $1^2 = \frac{1\cdot 2\cdot 3}{6}$ which is clearly true.

For the inductive step, we assume that $P(n)$ holds and we derive from this that $P(n+1)$ holds.

$$\begin{aligned}
1^2+2^2+\ldots+n^2+(n+1)^2 &= \frac{n(n+1)(2n+1)}{6}+(n+1)^2\\
&=(n+1)\left[\frac{n(2n+1)}{6}+(n+1)\right]\\
&=(n+1)\frac{(n+2)(2n+3)}{6}\\
&=\frac{(n+1)(n+2)(2n+3)}{6},
\end{aligned}$$

which shows that $P(n+1)$ is true. Hence, by the Weak Principle of Induction, $P(n)$ is true for all $n \geq 1$.

Remark. On a similar note, one can prove several other standard identities, like

$$\begin{aligned} 1+2+\ldots+n &= \frac{n(n+1)}{2}, \\ 1+3+\ldots+(2n-1) &= n^2, \\ 1^3+2^3+\ldots+n^3 &= \left(\frac{n(n+1)}{2}\right)^2, \end{aligned}$$

to mention just a few. For such identities, induction turns out to be a very powerful tool whenever we know what the result should be, as for example proving

$$1^2+2^2+\ldots+n^2 = \frac{n(n+1)(2n+1)}{6}$$

without using induction becomes significantly more involved. One downside for proving such statements by induction is that it does not give the reader a proper intuition of the phenomenon behind the problem (in this case, Faulhaber's formula using the theory of Bernoulli numbers).

Example 2.4. (**Binomial Theorem**) Prove that for any positive integer n, the following formula holds:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k.$$

Solution. We prove the formula by induction on n. For $n = 1$,

$$\begin{aligned} (x+y)^1 &= x+y \\ &= \binom{1}{0} x^{1-0} y^0 + \binom{1}{1} x^{1-1} y^1 \\ &= \sum_{k=0}^{1} \binom{1}{k} x^{1-k} y^k. \end{aligned}$$

Assume now that the result holds for some $n \geq 1$ and consider $(x+y)^{n+1}$:

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n \\
&= (x+y)\left(\sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k\right) \\
&= \sum_{k=0}^{n}\binom{n}{k}x^{n+1-k}y^k + \sum_{j=0}^{n}\binom{n}{j}x^{n-j}y^{j+1} \\
&= \sum_{k=0}^{n}\binom{n}{k}x^{n+1-k}y^k + \sum_{j=0}^{n}\binom{n}{(j+1)-1}x^{(n+1)-(j+1)}y^{j+1} \\
&= \sum_{k=0}^{n}\binom{n}{k}x^{n+1-k}y^k + \sum_{k=1}^{n+1}\binom{n}{k-1}x^{n+1-k}y^k \\
&= \sum_{k=0}^{n+1}\left(\binom{n}{k}x^{n+1-k}y^k\right) - \binom{n}{n+1}x^0y^{n+1} + \\
&+ \sum_{k=0}^{n+1}\left(\binom{n}{k-1}x^{n+1-k}y^k\right) - \binom{n}{-1}x^{n+1}y^0 \\
&= \sum_{k=0}^{n+1}\left(\binom{n}{k}+\binom{n}{k-1}\right)x^{n+1-k}y^k \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}x^{n+1-k}y^k,
\end{aligned}$$

where the last equality follows from Pascal's identity. This establishes the result for $n+1$ and completes our proof.

Example 2.5. Prove that

$$\sum_{k=0}^{n}\binom{n+k}{k}\frac{1}{2^k} = 2^n.$$

Solution. We prove the result by induction on n. For $n=1$ we have to show that $1+1=2$, which is clear. Assume now that the identity holds for some positive integer $n \geq 1$ and let

$$f(n) = \sum_{k=0}^{n} \binom{n+k}{k} \frac{1}{2^k}.$$

Using Pascal's identity, we have

$$\begin{aligned} f(n+1) &= \sum_{k=0}^{n+1} \binom{n+1+k}{k} 2^{-k} = 1 + \sum_{k=1}^{n+1} \binom{n+k}{k-1} 2^{-k} + \sum_{k=1}^{n+1} \binom{n+k}{k} 2^{-k} \\ &= \frac{1}{2} \sum_{i=0}^{n} \binom{n+i+1}{i} 2^{-i} + \binom{2n+1}{n+1} 2^{-n-1} + f(n) \\ &= \frac{1}{2} f(n+1) + f(n), \end{aligned}$$

hence $f(n+1) = 2f(n)$, which is what we wanted.

Example 2.6. Let n be a positive integer. Prove that

$$\binom{n}{0}^{-1} + \binom{n}{1}^{-1} + \cdots + \binom{n}{n}^{-1} = \frac{n+1}{2^{n+1}} \left(\frac{2}{1} + \frac{2^2}{2} + \cdots + \frac{2^{n+1}}{n+1} \right).$$

Solution. The proof is by induction on n. Denote the left hand side by a_n, and the right hand side by b_n. The base case is clear and we only need to show that a_n and b_n satisfy the same recurrence.

A recurrence for b_n is simple:

$$b_n = \frac{n+1}{2n} b_{n-1} + 1.$$

So we only need to prove that

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

We have

$$\frac{n+1}{2n} \binom{n-1}{i}^{-1} = \frac{(n+1)i!(n-i-1)!}{2(n!)}.$$

To express this in terms of binomial coefficients of base n, we write $n+1$ as $(i+1)+(n-i)$ and conclude that

$$\frac{n+1}{2n}\binom{n-1}{i}^{-1} = \frac{((i+1)+(n-i))i!(n-i-1)!}{2(n!)} = \frac{1}{2}\binom{n}{i+1}^{-1} + \frac{1}{2}\binom{n}{i}^{-1}.$$

By summing these relations and using the fact that $\binom{n}{0} = \binom{n}{n} = 1$, we conclude $a_n = \frac{n+1}{2n}a_{n-1} + 1$.

Example 2.7. Prove that for any $n \geq 1$ we have

$$\left[\frac{1}{2}\right] + \left[\frac{2}{2}\right] + \ldots + \left[\frac{n}{2}\right] = \left[\frac{n}{2}\right] \cdot \left[\frac{n+1}{2}\right],$$

Solution. We prove the result by induction on n. For $n = 1$, the statement holds since $\left[\frac{1}{2}\right] = \left[\frac{1}{2}\right] \cdot \left[\frac{2}{2}\right] = 0$. Also, for $n = 2$ we have that

$$\left[\frac{1}{2}\right] + \left[\frac{2}{2}\right] = 1 \quad \text{and} \left[\frac{2}{2}\right] \cdot \left[\frac{3}{2}\right] = 1,$$

so the identity also holds in this case.

Now assume that the identity holds for some $n \geq 1$, i.e.

$$\left[\frac{1}{2}\right] + \left[\frac{2}{2}\right] + \ldots + \left[\frac{n}{2}\right] = \left[\frac{n}{2}\right] \cdot \left[\frac{n+1}{2}\right].$$

Adding $\left[\frac{n+1}{2}\right] + \left[\frac{n+2}{2}\right]$ to both sides we get

$$\begin{aligned}
\left[\frac{1}{2}\right] + \left[\frac{2}{2}\right] + \ldots + \left[\frac{n+2}{2}\right] &= \left[\frac{n}{2}\right] \cdot \left[\frac{n+1}{2}\right] + \left[\frac{n+1}{2}\right] + \left[\frac{n+2}{2}\right] \\
\left[\frac{n}{2}\right] \cdot \left[\frac{n+1}{2}\right] + \left[\frac{n+1}{2}\right] + \left[\frac{n+2}{2}\right] &= \left[\frac{n}{2}\right] \cdot \left[\frac{n+1}{2}\right] + \left[\frac{n}{2}\right] + \left[\frac{n+1}{2}\right] + 1 \\
&= \left(\left[\frac{n}{2}\right] + 1\right)\left(\left[\frac{n+1}{2}\right] + 1\right) \\
&= \left[\frac{n+2}{2}\right] \cdot \left[\frac{n+3}{2}\right].
\end{aligned}$$

Hence the conclusion also holds for $n+1$, so we are done.

Example 2.8. Prove that for any $n \geq 1$ we have

$$\sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \binom{n}{k} = 1 + \frac{1}{2} + \ldots + \frac{1}{n}.$$

Solution. We prove the identity by induction on n. For $n = 1$, we get $1 = 1$, so the base case is verified.

Assume now the result for some $n \geq 1$. To prove it for $n + 1$, it suffices to show that

$$\frac{1}{n+1} = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k} \binom{n+1}{k} - \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \binom{n+1}{k}.$$

But

$$\sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k} \binom{n+1}{k} - \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \binom{n+1}{k}$$

$$= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \left(\binom{n+1}{k} - \binom{n}{k} \right) + \frac{(-1)^{n+1}}{n+1}.$$

We now use Pascal's identity:

$$\binom{n+1}{k} - \binom{n}{k} = \binom{n}{k-1}.$$

Also,

$$\binom{n}{k-1} \frac{1}{k} = \frac{1}{n+1} \binom{n+1}{k},$$

so the equality to prove becomes

$$\frac{1}{n+1} = \frac{1}{n+1} \left(\sum_{k=1}^{n} \binom{n+1}{k} (-1)^{k-1} + (-1)^{n+1} \right).$$

But now $(-1)^{k-1} = (-1)^{k+1}$, so using the Binomial Theorem we have

$$
\begin{aligned}
&\sum_{k=1}^{n}\binom{n+1}{k}(-1)^{k-1} + (-1)^n \\
&= -\left(-\binom{n+1}{1} + \binom{n+1}{2} + \ldots + (-1)^n\binom{n+1}{n} + (-1)^{n-1}\right) \\
&= -\left(\sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k} - 1 - (-1)^{n+1} + (-1)^{n-1}\right) \\
&= -\left((1-1)^{n+1} - 1\right) = 1,
\end{aligned}
$$

as we wanted.

We conclude this chapter by establishing some identities regarding the Fibonacci numbers. Let $(F_n)_{n\geq 0}$ be the Fibonacci sequence defined by $F_0 = 0, F_1 = 1$ and $F_{n+1} = F_n + F_{n-1}$ for $n \geq 1$. We shall also require the following facts about 2×2 matrices: If

$$
A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},
$$

we have

$$
A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}
$$

and

$$
A \cdot B = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.
$$

For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we also define its determinant by $\det(A) = ad - bc$. An easy check shows that for two 2×2 matrices A and B one has $\det(AB) = \det(A)\det(B)$. For more details, the interested reader can consult any standard textbook on Linear Algebra.

Example 2.9. Prove that for any $n \geq 1$ one has $\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n$.

Solution. The base case $n = 1$ is clear.

Assume now that

$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n,$$

for some $n \geq 1$.

Then

$$\begin{aligned} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n+1} &= \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} F_{n+1} + F_n & F_{n+1} \\ F_n + F_{n-1} & F_n \end{pmatrix} \\ &= \begin{pmatrix} F_{n+2} & F_{n+1} \\ F_{n+1} & F_n \end{pmatrix}, \end{aligned}$$

as required.

Remark. The above identity allows us to prove several other useful results about the Fibonacci numbers. Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, so that the above identity reads

$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = A^n.$$

Then:

1. By taking determinants of both sides in the equality

   $$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = A^n$$

   we obtain $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$.

2. Using the fact that $A^{m+n-1} = A^m \cdot A^{n-1}$ (which holds for any square matrix) we deduce that $F_{m+n} = F_{m+1}F_n + F_mF_{n-1}$. This is also known as **Cassini's identity**. A similar identity for F_{m+n+p} is obtained by exploiting the condition $A^{m+n+p-1} = A^m \cdot A^n \cdot A^{p-1}$.

3. From $\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = A^n$, we obtain that $A^n = F_n \cdot A + F_{n-1} \cdot I_2$. For $n = 2$, this reads $A^2 = A + I_2$, which further implies $A^{n+2} = A^{n+1} + A^n$. Now

$$A^{2n} = (A^2)^n = (A + I_2)^n = \sum_{k=0}^{n} \binom{n}{k} A^k. \tag{*}$$

By identifying the corresponding entries in the above relation, we derive the identities

$$F_{2n} = \sum_{k=0}^{n} \binom{n}{k} F_k, \quad F_{2n+1} = \sum_{k=0}^{n} \binom{n}{k} F_{k+1}.$$

By multiplying both sides of (*) by A^m (m a positive integer) and identifying the corresponding elements on the $(1,2)$ entry, we find

$$F_{2n+m} = \sum_{k=0}^{n} \binom{n}{k} F_{k+m}.$$

4. From $A^2 = A + I_2$ we obtain $A^2 - A = I_2$. Thus

$$(A^2 - A)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k A^{2(n-k)} A^k = \sum_{k=0}^{n} \binom{n}{k} (-1)^k A^{2n-k}.$$

At the same time, $\left(A^2 - A\right)^n = I_2^n = I_2$, so combining this with the above relation and looking at the $(1,2)$ entry, we get

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k F_{2n-k} = F_0 = 0.$$

Generalizing this in the same manner as above, we obtain the more general result

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k F_{2n-k+m} = F_m.$$

5. From $A^n = F_n \cdot A + F_{n-1} \cdot I$, one gets

$$A^{nm} = (F_n A + F_{n-1} I)^m = \sum_{k=0}^{m} \binom{m}{k} F_n^{m-k} F_{n-1}^k \cdot A^{m-k},$$

which by componentwise identification gives further

$$F_{nm} = \sum_{k=0}^{m} \binom{m}{k} F_n^{m-k} F_{n-1}^k F_{m-k}.$$

Example 2.10. Let $(F_n)_{n\geq 0}$ be the Fibonacci sequence defined as before. Prove that for any $N \geq 1$ we have the identity:

$$\sum_{n=0}^{N} \frac{1}{F_{2^n}} = 3 - \frac{F_{2^N-1}}{F_{2^N}}.$$

Solution. We do induction on N. The base case is $N = 1$ which gives

$$\frac{1}{F_1} + \frac{1}{F_2} = 3 - \frac{F_1}{F_2} \Leftrightarrow 1 + 1 = 3 - 1,$$

so $P(1)$ is true.

Assume now that the result holds for some $N \geq 1$. Then

$$\begin{aligned}\sum_{n=0}^{N+1} \frac{1}{F_{2^n}} &= \left(\sum_{n=0}^{N} \frac{1}{F_{2^n}}\right) + \frac{1}{F_{2^{N+1}}} \\ &= 3 - \frac{F_{2^N-1}}{F_{2^N}} + \frac{1}{F_{2^{N+1}}} \\ &= 3 - \frac{F_{2^N-1}F_{2^{N+1}} - F_{2^N}}{F_{2^N}F_{2^{N+1}}}.\end{aligned}$$

Now

$$\frac{F_{2^N-1}F_{2^{N+1}} - F_{2^N}}{F_{2^N}F_{2^{N+1}}} = \frac{F_{2^N-1}F_{2^N}(F_{2^N+1} + F_{2^N-1}) - F_{2^N}}{F_{2^N}F_{2^{N+1}}},$$

by applying the Cassini's identity to $F_{2^{N+1}}$.

Further,

$$\frac{F_{2^N-1}F_{2^N}(F_{2^N+1}+F_{2^N-1})-F_{2^N}}{F_{2^N}F_{2^{N+1}}} = \frac{F_{2^N-1}F_{2^N+1}+F_{2^N-1}^2-1}{F_{2^{N+1}}}$$
$$= \frac{F_{2^N}^2+F_{2^N-1}^2}{F_{2^{N+1}}},$$

by the determinant identity above.

Finally,

$$\frac{F_{2^N}^2+F_{2^N-1}^2}{F_{2^{N+1}}} = \frac{F_{2^{N+1}-1}}{F_{2^{N+1}}},$$

by applying Cassini's identity to $F_{2^{N+1}-1}$. This establishes the induction step and hence the result.

2.2 Proposed Problems

Problem 2.1. (**GMB 1997**) Determine the sequence $a_1, a_2, \ldots,$ of positive reals, such that for all positive integers k, one has

$$a_1 + 2^2 \cdot a_2 + 3^2 \cdot a_3 + \ldots + k^2 \cdot a_k = \frac{k(k+1)}{2}\left(a_1 + a_2 + \ldots + a_k\right).$$

Problem 2.2. Prove that for all positive integers n,

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{2n-1} - \frac{1}{2n} = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n}.$$

Problem 2.3. Prove that for any positive integer n one has

$$(n+1)! = 1 + \frac{1!^2}{0!} + \frac{2!^2}{1!} + \ldots + \frac{n!^2}{(n-1)!}.$$

Problem 2.4. Prove that:

$$\sum_{k=0}^{n} \binom{n-k+1}{k} = F_{n+2},$$

for any non-negative integer n.

Problem 2.5. Show that for any positive integer n, we have

$$\sum_{r=1}^{n} \frac{1}{r}\binom{n}{r} = \sum_{r=1}^{n} \frac{2^r - 1}{r}.$$

Problem 2.6. The Bernoulli numbers $(B_n)_{n \geq 0}$ are given by the following recurrence:

$$B_0 = 1 \quad \text{and} \quad \sum_{i=0}^{m} \binom{m+1}{i} B_i = 0, \quad \text{for} \quad m > 0.$$

Prove that

$$1^k + 2^k + \ldots + (n-1)^k = \frac{1}{k+1} \sum_{i=0}^{k} \binom{k+1}{i} B_i \cdot n^{k+1-i},$$

for all non-negative integers n and k.

Problem 2.7. Let n be a positive integer. Prove that

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}i^k = 0,$$

for all $0 \leq k \leq n-1$.

Chapter 3

Functions and Functional Equations

3.1 Theory and Examples

Let A and B be two (not necessarily finite) sets. A **function** $f : A \to B$ is a correspondence between A and B which associates to each element of A precisely one element of B. The set A is called the **domain** of the function and B is called the **range**. The **image** of the function f, denoted by $\mathrm{Im}(f)$ is the set $\mathrm{Im}(f) = \{f(a) : a \in A\}$. For example, $f : \mathbb{N} \to \mathbb{Z}$, $f(x) = -x$ has domain $\mathbb{N}$, range $\mathbb{Z}$ and image $\mathbb{Z}_{\leq 0}$.

If $f : A \to B$ is a function and $C \subseteq B$ a subset, we define the **preimage** of C, denoted by $f^{-1}(C)$ by

$$f^{-1}(C) = \{a \in A : f(a) \in C\}.$$

For example, when $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, one has $f^{-1}(\{1,4\}) = \{-2,-1,1,2\}$, while $f^{-1}(-2) = \emptyset$.

A function $f : A \to B$ is called **injective** if whenever a and b are distinct elements of A, $f(a)$ and $f(b)$ are distinct elements of B. A function $f : A \to B$ is called **surjective** if for any element $y \in B$ there is an element $x \in A$ such that $f(x) = y$. A function $f : A \to B$ is **bijective** if it is both injective and surjective.

Let us look at a few examples:

1. Let $A = \{1,2,3\}$, $B = \{2,3\}$. The function $f : A \to B$ given by $f(1) = 2, f(2) = 3, f(3) = 2$ is surjective, but not injective, because both $f(1)$ and $f(3)$ are assigned the same value in B, namely 2.

2. Let $A = B = \mathbb{N}$, where $\mathbb{N}$ stands for the non-negative integers. Then the function $f : A \to B$ given by $f(n) = n+1$, for any $n \in \mathbb{N}$ is injective, but not surjective, since there is no element x in A for which $f(x) = 0$. However, if we considered $A = B = \mathbb{Z}$ with $f(n) = n+1$, then one can check that f is both injective and surjective, hence bijective.

3. For $A = B = \mathbb{R}$ and $f : A \to B$ given by $f(x) = x^2$, f is not injective, since for example $f(-1) = f(1) = 1$, nor surjective, because we cannot have $f(x) = -1$ for any $x \in A$.

$f : A \to B$ is **invertible** if and only if there is a function $g : B \to A$ such that $g(f(a)) = a, \forall a \in A$ and $f(g(b)) = b, \forall b \in B$. When f is invertible, we denote its inverse by f^{-1}. It can be proved that a function $f : A \to B$ is invertible if and only if it is bijective.

A function $f : \mathbb{R} \to \mathbb{R}$ is **even** if $f(x) = f(-x)$ for all $x \in \mathbb{R}$. $f : \mathbb{R} \to \mathbb{R}$ is an **odd** function if $f(-x) = -f(x)$ for any $x \in \mathbb{R}$.

A function f is an **increasing** function if whenever $x < y$ we have $f(x) \leq f(y)$. It is strictly increasing if $x < y$ implies $f(x) < f(y)$. Similarly, a function f is a **decreasing** function if $x < y$ implies $f(x) \geq f(y)$. It is called strictly decreasing if $x < y$ implies $f(x) > f(y)$.

There are several strategies one has to keep in mind when it comes to solving functional equations. We shall focus on this chapter on those types whose solution relies heavily on induction. The simplest examples are when we are asked to prove (or we can easily spot) what the function should be. Let us consider a few such questions to start with:

Example 3.1. Let $f : \mathbb{N}^* \to \mathbb{Z}$ be a function with the following properties:
1) $f(2) = 2$;
2) $f(mn) = f(m)f(n)$ for all m and n;
3) $f(m) > f(n)$ whenever $m > n$.
Prove that $f(n) = n$, for all $n \in \mathbb{N}^*$.

Solution. We will prove the result by induction on $n \geq 1$. Setting $m = 1$, $n = 2$ in condition 2) gives $f(2) = f(1) \cdot f(2)$, which combined with condition 1) shows that $f(1) = 1$. Setting $m = n = 2$ we find that $f(4) = 4$ and since $f(2) < f(3) < f(4)$, from condition 3) we get $2 < f(3) < 4$, which implies $f(3) = 3$. This shows that the result holds for the few first values of n.

Assume now that the result is true for all $k \in \mathbb{N}^*$ with $k \leq 2n$, some $n \geq 2$. By 2), we have

$$f(2n+2) = f(2(n+1)) = f(2)f(n+1) = 2n+2.$$

Also, $f(2n) < f(2n+1) < f(2n+2)$, therefore $2n < f(2n+1) < 2n+2$. Hence, $f(2n+1) = 2n+1$. This completes the induction step and our proof.

Things can get more complicated if we do not work over the positive integers and we may need to get more creative about the way we choose to proceed with our induction. A good example which illustrates some ideas to bear in mind is the following:

Example 3.2. Let $f : \mathbb{Q}^+ \to \mathbb{Q}^+$ be a function which satisfies

$$\begin{cases} f(x) + f\left(\frac{1}{x}\right) = 1. \\ f(2x) = 2f(f(x)) \end{cases}.$$

Show that $f\left(\frac{2012}{2013}\right) = \frac{2012}{4025}$.

Solution. The only such function is $f(x) = \frac{x}{1+x}$. It is easy to check that this function satisfies the conditions above. We want to prove by induction that it is unique. Any positive rational x can be written uniquely as $x = p/q$ for coprime positive integers p, q. Define the complexity of x to be $p + q$. We will prove uniqueness by induction on the complexity of x.

For the base case $p + q = 2$, i.e. $x = 1$, notice that plugging $x = 1$ into the first condition gives $2f(1) = 1$ or $f(1) = 1/2$, as desired.

Now suppose we have proved the uniqueness for all rational numbers with complexity less than n. We will need to analyze two cases.

First, suppose $n \geq 4$ is even and $x = p/q$ with p, q coprime and $p + q = n$. Since $n \geq 4$, we cannot have $p = q$. If $p < q$, then notice that $p/(q-p)$ and $2p/(q-p)$ have complexity less than n. The first of these clearly has complexity at most $p + (q - p) = q < n$. For the second the numerator and denominator are both even, hence the complexity is at most $p + \frac{q-p}{2} = \frac{p+q}{2} < n$. (We say "at most" here because one might imagine further cancellation. This cannot actually occur since any further cancellation would mean p and q have a common factor, but "at most" is good enough for us.) Thus applying the second condition to $p/(q - p)$ and using the induction hypothesis we get

$$\frac{2p}{q+p} = f(2p/(q-p)) = 2f(f(p/(q-p))) = 2f(p/q),$$

or $f(x) = f(p/q) = \frac{p}{q+p} = \frac{x}{1+x}$. If $p > q$, then we apply the first condition and the result just proven to get

$$f(x) = f(p/q) = 1 - f(q/p) = 1 - \frac{q}{p+q} = \frac{p}{p+q} = \frac{x}{1+x}.$$

This completes this case.

Now suppose $n \geq 3$ is odd and $x = p/q$ with p, q coprime and $p + q = n$. In this case $p/(q - p)$ still has complexity less than n, but $q - p$ is odd so $2p/(q-p)$ has complexity n. Thus from the second condition and the induction hypothesis, we only get

$$f(2p/(q-p)) = 2f(p/q).$$

The first condition gives $f(p/q) + f(q/p) = 1$ and clearly q/p also has complexity n. Thus instead of being able to immediately solve for $f(x)$ we get a collection of linear equations relating $f(r)$ for various rational numbers r all of the same complexity. We need to show that this collection has a unique solution. We give one method for this.

Pick a rational number $x_0 = \frac{2k}{n-2k}$ with complexity n and even numerator. Hence we get

$$f(x_0) = f\left(\frac{2k}{n-2k}\right) = 2f\left(\frac{k}{n-k}\right).$$

If k is even, define $x_1 = \frac{k}{n-k}$ so that $f(x_0) = 2f(x_1)$. Otherwise $n - k$ is even and we define $x_1 = \frac{n-k}{k}$. Hence using the first condition we find that $f(x_0) = 2(1 - f(x_1))$. Thus we can relate $f(x_0)$ to the value of f at another rational x_1 with complexity n and even numerator. Iterating this construction we can build a chain $x_0, x_1, x_2, \dots$ of such rationals with $f(x_k) = c_k \pm 2f(x_{k+1})$ for some integer c_k. There are only finitely many possibilities for x_i, hence this chain must eventually cycle. Further, we can compute x_i from x_{i+1} (if $x_{i+1} < 1$, then $x_i = \frac{2x_{i+1}}{1-x_{i+1}}$, and if $x_{i+1} > 1$, then $x_i = \frac{2}{x_{i+1}-1}$). Thus the cycle must begin with x_0. Hence there is some least integer $r > 0$ such that $x_r = x_0$ and iterating our formulas we find

$$f(x_0) = C \pm 2^r f(x_r) = C \pm 2^r f(x_0),$$

for some integer C and some choice of sign. Thus we can uniquely solve for $f(x_0)$. Since we already saw above that $f(x) = \frac{x}{1+x}$ is a solution, we must have $f(x_0) = \frac{x_0}{1+x_0}$. This also holds for x with even denominator by using the first condition $f(x) + f(1/x) = 1$.

This completes our induction. In particular, we have

$$f\left(\frac{2012}{2013}\right) = \frac{2012}{2012+2013} = \frac{2012}{4025}.$$

Unlike the above two questions, for most examples we would be asked to determine the value for the function ourselves. For such questions, it is good to be familiar with a series of standard functions (e.g. $f(x) = x$, $f(x) = x^2$, $f(x) = \sin(x)$, $f(x) = e^x$, etc.) and their properties, so that we are able to spot if the given function resembles any of those. This can make our detective work easier. Let us look at a rather easier example first:

Example 3.3. Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ such that for every $n \in \mathbb{N}^*$ we have

$$f(f(n)) + f(n) = 2n.$$

Solution. We naturally expect to have $f(n) = n$ as the only solution. Let $E(n)$ denote the relation $f(f(n)) + f(n) = 2n$, valid for all $n \in \mathbb{N}^*$.

We shall prove by strong induction that $f(n) = n$ for all $n \in \mathbb{N}^*$.

For $n = 1$, $E(1)$ reads $f(f(1)) + f(1) = 2$. Since $f(f(1)) \geq 1$, $f(1) \geq 1$, and we achieve the minimum, we must have $f(1) = 1$, $f(f(1)) = 1$, establishing the base case.

Now assume that $f(n) = n$ for all $n < k$, some $k \geq 2$. For $n = k$, $E(k)$ reads $f(f(k)) + f(k) = 2k$. If $f(k) < k$, then $f(f(k)) = f(k)$ by the induction hypothesis, and so $f(f(k)) + f(k) = 2f(k) < 2k$, which is a contradiction. If $f(k) > k$, then $f(f(k)) = 2k - f(k) < k$, and so $E(f(k))$ gives us $f(f(f(k))) + f(f(k)) = 2f(k)$. Since $f(f(k)) < k$, we have $f(f(f(k))) = f(f(k)) < k$ from the induction hypothesis, so the left hand side is less than $2k$. But $f(k) > k$, so the right hand side is greater than $2k$, which is a contradiction. Thus $f(k) = k$. This completes the induction and therefore the solution to our question.

Example 3.4. (**Canada MO 2015**) Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ such that

$$(n-1)^2 < f(n)f(f(n)) < n^2 + n, \quad \text{for all} \quad n \in \mathbb{N}^*.$$

Solution. We prove by induction on n that $f(n) = n$. The base case is $n = 1$, for which we get by substituting in the condition from the hypothesis that $0 < f(1)f(f(1)) < 2$. This necessarily implies $f(1) = 1$, so the base case is proved.

Assume now that $f(k) = k$ for all $k < n$ (some $n \geq 2$) and assume by contradiction that $f(n) \neq n$. We distinguish two cases:

Case 1. If $f(n) \leq n - 1$, then $f(f(n)) = f(n)$ from the induction hypothesis and $f(n)f(f(n)) = f(n)^2 \leq (n-1)^2$, contradicting the condition from the hypothesis.

Case 2. If $f(n) = M \geq n + 1$, then

$$(n+1)f(M) \leq f(n)f(f(n)) < n^2 + n.$$

Thus $f(M) < n$ and hence $f(f(M)) = f(M)$ and

$$f(M)f(f(M)) = f(M)^2 < n^2 \leq (M-1)^2,$$

which is also a contradiction. This completes the induction and our proof.

If things are a bit more complicated and we cannot tell what the function should be straight away, a good idea is to start by computing the first few values (whenever possible) and see if we can spot a pattern from that.

Example 3.5. (**AoPS**) Find all functions $f : \mathbb{Z} \to \mathbb{R}$ such that $f(1) = \frac{5}{2}$ and for any $m, n \in \mathbb{Z}$ we have

$$f(m)f(n) = f(m+n) + f(m-n).$$

Solution. We can compute the first few values for f to find $f(0) = 2$, $f(1) = 2.5$, $f(2) = 4.25$ and $f(3) = 8.125$. This suggests that our function is $f(n) = 2^n + 2^{-n}$. Notice that it will suffice to show this for $n \geq 0$, since setting $m = 0$ gives $2f(n) = f(n) + f(-n)$, so our function is even. Therefore, we proceed to proving by induction that $f(n) = 2^n + 2^{-n}$ for all $n \geq 0$.

The first few base cases were proved above.

Assume now that the result holds for all positive integers up to some $n \geq 1$. To show it for n, plug $m = n - 1$, $n = 1$ to get

$$f(n-1)f(1) = f(n) + f(n-2).$$

All values are known from the induction hypothesis, except for $f(n)$. Substituting $f(n-1) = 2^{n-1} + 2^{-(n-1)}$, $f(1) = \frac{5}{2}$ and $f(n-2) = 2^{n-2} + 2^{-(n-2)}$ gives us $f(n) = 2^n + 2^{-n}$, as required. This completes our proof.

For some problems, even if we are able to spot what the function should be, we need to exploit several properties of the function which are given in the hypothesis before we are able to prove that our guess is correct by induction. A rather famous example is the following:

Example 3.6. (**IMO 2009**) Determine all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ which satisfy the property that for all positive integers a and b, there exists a non-degenerate triangle with sides of lengths

$$a, f(b) \text{ and } f(b + f(a) - 1).$$

(A triangle is non-degenerate if its vertices are not collinear.)

Solution. First notice that if a triangle has sides of lengths 1, a, b with a, b positive integers, then by the triangle inequality we must have that $a = b$.

For $a = 1$, the above remark tells us that $f(b) = f(b + f(1) - 1)$.

Notice that $f(1) = 1$, as otherwise $f(1) - 1 > 0$, which from the above relation means that f repeats itself every $f(1) - 1$ numbers. This means that

f can take only finitely many values, so if we take a sufficiently large, a, $f(b)$, $f(b+f(a)-1)$ cannot be a triangle, because $a-f(b) > f(b+f(a)-1)$.

Now, setting $b=1$, we obtain that $a, 1, f(f(a))$ must be the sides of a triangle. This implies that $f(f(a)) = a$ by one of the previous remarks.

Claim. $f(n) = (n-1)f(2) - (n-2)$, for $n \geq 3$.

Proof. From $f(f(a)) = a$, f is bijective, so we now know that a, b, $f(f(a) + f(b) - 1)$ can be the side lengths of a triangle. This implies that

$$f(f(a) + f(b) - 1) < a + b.$$

If we take $a=b=2$ we then obtain that $f(2f(2)-1) < 4$, i.e. $f(2f(2)-1) \in \{1, 2, 3\}$.
The value 1 is not possible, as we would have $2f(2) - 1 = 1$, i.e. $f(2) = 1$, contradicting the bijectivity of f. The value 2 is also not possible, since we would get $2f(2) - 1 = f(2)$, i.e. $f(2) = 1$, which is again a contradiction. Therefore, we must have $2f(2) - 1 = f(3)$, proving the base case $n = 3$.

For the induction step we use a similar argument, by taking $a = 2$, $b = n$, and arguing that $f(f(2) + f(n) - 1) = n + 1$.

We have thus the result $f(n) = (n-1)f(2) - (n-2)$. In particular, this tells us that f is strictly increasing. Since we already saw that f is bijective and $f(1) = 1$, this means we must have $f(2) = 2$. Hence the claim reads $f(n) = 2(-1) - (n-2) = n$ for all $n \geq 3$. This completes our proof, showing that $f(n) = n$ is the only solution.

Example 3.7. (**IMO 2008 shortlist**) For an integer m, denote by $t(m)$ the unique number in $\{1, 2, 3\}$ such that $m + t(m)$ is a multiple of 3. A function $f : \mathbb{Z} \to \mathbb{Z}$ satisfies $f(-1) = 0$, $f(0) = 1$, $f(1) = -1$ and

$$f(2^n + m) = f(2^n - t(m)) - f(m) \quad \text{for all integers} \quad m, n \geq 0 \quad \text{with} \quad 2^n > m.$$

Prove that $f(3p) \geq 0$ holds for all integers $p \geq 0$.

Solution. The given conditions determine f uniquely on the positive integers The signs of $f(1), f(2), \ldots$ seem to change quite erratically. However, values of the form $f(2^n - t(m))$ are sufficient to compute directly any functional value. Indeed, let $n > 0$ have base 2 representation

$$n = 2^{a_0} + 2^{a_1} + \ldots + 2^{a_k},\ a_0 > a_1 > \ldots > a_l \geq 0,$$

and let $n_j = 2^{a_j} + 2^{a_{j+1}} + \ldots + 2^{a_k}, j = 0, \ldots, k$. Repeated applications of the recurrence from the hypothesis show that $f(n)$ is an alternating sum of the quantities $f(2^{a_j} - t(n_{j+1}))$ plus $(-1)^{k+1}$ (we do not need to know the exact formula for our proof).

Hence we shall focus on looking at the values $f(2^n - 1)$, $f(2^n - 2)$ and $f(2^n - 3)$. There are six cases, namely

$$\begin{aligned} t(2^{2k} - 3) &= 2, & t(2^{2k} - 2) &= 1,\\ t(2^{2k} - 1) &= 3, & t(2^{2k+1} - 3) &= 1,\\ t(2^{2k+1} - 2) &= 3, & t(2^{2k+1} - 1) &= 2. \end{aligned}$$

Claim. For all integers $k \geq 0$ the following equalities hold:

$$\begin{aligned} f(2^{2k+1} - 3) &= 0, & f(2^{2k+1} - 2) &= 3^k,\\ f(2^{2k+1} - 1) &= -3^k, & f(2^{2k+2} - 3) &= -3^k,\\ f(2^{2k+2} - 2) &= -3^k, & f(2^{2k+2} - 1) &= 2 \cdot 3^k. \end{aligned}$$

Proof. We proceed by induction on k. The base case $k = 0$ comes down to checking that $f(2) = -1$ and $f(3) = 2$; the given values $f(-1) = 0$, $f(0) = 1$, $f(1) = -1$ are also needed. Suppose that the claim holds for $k - 1$ for some $k \geq 1$. For $f(2^{k+1} - t(m))$, the recurrence formula and the induction hypothesis yield

$$\begin{aligned} f(2^{2k+1} - 3) &= f(2^{2k} + (2^{2k} - 3))\\ &= f(2^{2k} - 2) - f(2^{2k} - 3) = -3^{k-1} + 3^{k-1} = 0;\\ f(2^{2k+1} - 2) &= f(2^{2k} + (2^{2k} - 2))\\ &= f(2^{2k} - 1) - f(2^{2k} - 2) = 2 \cdot 3^{k-1} + 3^{k-1} = 3^k;\\ f(2^{2k+1} - 1) &= f(2^{2k} + (2^{2k} - 1))\\ &= f(2^{2k} - 3) - f(2^{2k} - 1) = -3^{k-1} - 2 \cdot 3^{k-1} = -3^k. \end{aligned}$$

For $f(2^{2k+2}-t(m))$, we use the three equalities we have just established:

$$\begin{aligned}
f(2^{2k+2}-3) &= f(2^{2k+1}+(2^{2k+1}-3)) \\
&= f(2^{2k+1}-1)-f(2^{2k+1}-3) = -3^k - 0 = -3^k; \\
f(2^{2k+2}-2) &= f(2^{2k+1}+(2^{2k+1}-2)) \\
&= f(2^{2k+1}-3)-f(2^{2k}-2) = 0 - 3^k = -3^k; \\
f(2^{2k+2}-1) &= f(2^{2k+1}+(2^{2k+1}-1)) \\
&= f(2^{2k+1}-2)-f(2^{2k+1}-1) = 3^k + 3^k = 2\cdot 3^k.
\end{aligned}$$

So we have established the claim.

A closer look at the six cases shows that $f(2^n-t(m)) \geq 3^{(n-1)/2}$ if $2^n-t(m)$ is divisible by 3, and $f(2^n-t(m)) \leq 0$ otherwise. On the other hand, note that $2^n-t(m)$ is divisible by 3 if and only if 2^n+m is. Therefore, for all non-negative integers m and n,

i) $f(2^n-t(m)) \geq 3^{(n-1)/2}$ if 2^n+m is divisible by 3;

ii) $f(2^n-t(m)) \leq 0$ if 2^n+m is not divisible by 3.

One more (direct) consequence of the claim is that

$$|f(2^n-t(m))| \leq \frac{2}{3}\cdot 3^{n/2} \text{ for all } m,n\geq 0.$$

The last inequality enables us to find an upper bound for $|f(m)|$ for m less than a given power of 2. We prove by induction on n that $|f(m)|\leq 3^{n/2}$ holds true for all integers $m,n\geq 0$ with $2^n>m$.

The base case $n=0$ is clear as $f(0)=1$. For the inductive step from n to $n+1$, let m and n satisfy $2^{n+1}>m$. If $m<2^n$, we are done by the inductive hypothesis. If $m\geq 2^n$, then $m=2^n+k$, where $2^n>k\geq 0$. Now, by $|f(2^n-t(k))|\leq \frac{2}{3}\cdot 3^{n/2}$ and the inductive assumption,

$$\begin{aligned}
|f(m)| &= |f(2^n-t(k))-f(k)| \leq |f(2^n-t(k))|+|f(k)| \\
&\leq \frac{2}{3}\cdot 3^{n/2}+3^{n/2} < 3^{(n+1)/2}.
\end{aligned}$$

The induction is complete.

We proceed to prove that $f(3p) \geq 0$ for all integers $p \geq 0$. Since $3p$ is not a power of 2, its binary expansion contains at least two summands. Hence, one can write $3p = 2^a + 2^b + c$, where $a > b$ and $2^b > c \geq 0$. Applying the recurrence formula twice yields,

$$f(3p) = f(2^a + 2^b + c) = f(2^a - t(2^b + c)) - f(2^b - t(c)) + f(c).$$

Since $2^a + 2^b + c$ is divisible by 3, we have $f(2^a - t(2^b + c)) \geq 3^{(a-1)/2}$ by i). Since $2^b + c$ is not divisible by 3, we have $f(2^b - t(c)) \leq 0$ by ii). Finally, $|f(c)| \leq 3^{b/2}$, as $2^b > c \geq 0$, so that $f(c) \geq -3^{b/2}$. Therefore, $f(3p) \geq 3^{(a-1)/2} - 3^{b/2}$, which is non-negative, because $a > b$.

So far we have mostly looked at examples where we ended up finding the function explicitly. However, the universe of problems which involve functions and their properties goes way beyond that. We conclude this section by presenting a few beautiful miscellaneous examples and the nice tools their solutions employ.

Example 3.8. (**Canada MO 1990**) Let $f : \mathbb{N}^* \to \mathbb{R}$ be a function which satisfies the following properties:

1) $f(1) = 1$, $f(2) = 2$;

2) $f(n+2) = f(n+2-f(n+1)) + f(n+1-f(n))$.

Show that $0 \leq f(n+1) - f(n) \leq 1$.

Solution. The first important observation is that since the domain of f is $\mathbb{N}^*$, from 2) we have that $f(n+1-f(n))$ is defined for all $n \in \mathbb{N}$, so $n+1-f(n) \geq 1$ and thus $f(n) \leq n$, for all $n \in \mathbb{N}^*$. Moreover, $n+1-f(n) \in \mathbb{N}^*$ implies $f(n) \in \mathbb{Z}$, for all $n \in \mathbb{N}^*$.

We now show that $f(n) > 0$ for all natural n by strong induction.

For the base case we have $f(1) = 1$, $f(2) = 2$ and from 2) we immediately deduce that $f(3) = 2 > 0$.

Now assume that $f(n) > 0$ for all $n \leq k$ (where $k \geq 3$). Then

$$f(k+1) = f(k+1-f(k)) + f(k-f(k-1)),$$

but $1 \leq k+1-f(k) \leq k$ and $1 \leq k-f(k-1) \leq k-1$, so

$$f(k+1-f(k)) > 0, \quad f(k-f(k-1)) > 0 \implies f(k+1) > 0,$$

which completes the induction.

Hence, $f(n) > 0$ for all n, and since $f(n)$ is an integer, we conclude that $f(n) \in \mathbb{N}^*$, for all natural $n \in \mathbb{N}^*$.

Notice that since $f(n) \in \mathbb{N}^*$ for all n, the statement $0 \leq f(n+1) - f(n) \leq 1$ (for all n) is equivalent to the statement $f(n+1) \in \{f(n), f(n)+1\}$ (for all n). We prove the latter by induction:

By above we have that $f(2) = f(1) + 1$ and $f(3) = f(2)$, so the base cases are verified.

Now suppose that the statement is true for all $n \leq k$. We have two cases:

Case 1. If $f(k) = f(k-1)$ then we also have that $f(k-2) \in \{f(k), f(k)-1\}$. If we let $a = k - f(k) < k$, then

$$\begin{aligned} f(k+1) &= f(k+1-f(k)) + f(k-f(k-1)) \\ &= f(k-f(k)+1) + f(k-f(k)) \\ &= f(a+1) + f(a). \end{aligned}$$

Also,

$$\begin{aligned} f(k) &= f(k-f(k-1)) + f(k-1-f(k-2)) \\ &= f(a) + f(k-1-f(k-2)) \in \{f(a)+f(a-1), 2f(a)\}. \end{aligned}$$

If $f(k) = 2f(a)$, then $f(k+1) - f(k) = f(a+1) - f(a) \leq 1$, and we are done. If $f(k) = f(a) + f(a-1)$ (which happens when $f(k-2) = f(k)$), then $f(k+1) - f(k) = f(a+1) - f(a-1) \leq 2$, but if $f(a+1) - f(a-1) = 2$, then $f(a-1) = f(a) - 1$ and $f(a+1) = f(a) + 1$, so

$$f(k) = f(k-1) = f(k-2) = 2f(a) - 1.$$

But in this case we have

$$f(k-1) = f(k-1-f(k-2)) + f(k-2-f(k-3)),$$

so $2f(a) - 1 = f(a-1) + f(k-2-f(k-3))$, which implies

$$f(k-2-f(k-3)) = f(a),$$

but $f(k-2-f(k-3))$ is either $f(a-1)=f(a)-1$ or $f(a-2)\le f(a-1)$. This is a contradiction.

So $0\le f(a+1)-f(a-1)\le 1$, and this case is proved.

Case 2. If $f(k)=f(k-1)+1$, then we if we let $a=k-f(k)<k$, we obtain

$$f(k+1)=f(k+1-f(k))+f(k-f(k-1))=2f(a+1),$$

and

$$\begin{aligned}f(k)&=f(k-f(k-1))+f(k-1-f(k-2))\\&=f(a+1)+f(k-1-f(k-2))\in\{f(a+1)+f(a),2f(a+1)\},\end{aligned}$$

so $f(k+1)-f(k)\in\{f(a+1)-f(a),0\}$, and $0\le f(a+1)-f(a)\le 1$. This completes the proof of the second case as well.

This completes the induction, so $0\le f(n+1)-f(n)\le 1$, for all natural n.

Example 3.9. (**AMM**) Given a positive integer number k, define the function $f:\mathbb{N}^*\to\mathbb{N}^*$ by

$$f(n)=\begin{cases}1, & \text{if } n\le k+1;\\ f(f(n-1))+f(n-f(n-1)), & \text{if } n>k+1.\end{cases}$$

Show that for every $n\in\mathbb{N}^*$, the preimage $f^{-1}(n)$ is a finite non-empty set of consecutive positive integers.

Solution. We first prove by induction on n that $f(n)-f(n-1)\in\{0,1\}$, for all $n\in\mathbb{N}^*$. The base cases $n=1,\ldots,k+1$ are given in the hypothesis.

For the induction step, suppose that the assumption holds for all values less than $n+1$, for some $n\ge k+1$. Note that this implies $f(m)<m$ for $2\le m\le n$.

If $f(n)-f(n-1)=0$, then

$$\begin{aligned}f(n+1)-f(n)&=f(f(n))+f(n+1-f(n))-f(f(n-1))-f(n-f(n-1))\\&=f(n-f(n)+1)-f(n-f(n))\in\{0,1\}.\end{aligned}$$

If $f(n) - f(n-1) = 1$, we have

$$\begin{aligned} f(n+1) - f(n) &= f(f(n)) + f(n+1-f(n)) - f(f(n-1)) - f(n-f(n-1)) \\ &= f(f(n-1)+1) - f(f(n-1)) \in \{0,1\}. \end{aligned}$$

This completes our induction. We are left to prove that the preimage is a non-empty set, which is equivalent to showing that f is unbounded. Assume the contrary, so there exist $a, b \in \mathbb{N}^*$ such that all of $f(a), f(a+1), f(a+2), \ldots$ are equal to b. Setting $n = a+b$ into the relation in the hypothesis gives us

$$f(a+b) = f(f(a+b-1)) + f(a+b-f(a+b-1)) = f(b) + f(a+b-b) = f(b) + b > b,$$

which is a contradiction. This completes our proof.

Example 3.10. (**USAMO 2000**) Call a real-valued function f *very convex* if

$$\frac{f(x)+f(y)}{2} \geq f\left(\frac{x+y}{2}\right) + |x-y|$$

holds for all real numbers x and y. Prove that no very convex function exists.

Solution. We prove by induction on $n \geq 0$ that the given inequality implies

$$\frac{f(x)+f(y)}{2} - f\left(\frac{x+y}{2}\right) \geq 2^n|x-y|, \quad \text{for all} \quad n \geq 0.$$

This will yield a contradiction, since for fixed x and y, the right hand side gets arbitrarily large, while the left hand side remains fixed. The base case $n = 0$ holds from the hypothesis.

Assume now that the result holds for some $n \geq 0$. For $a, b \in \mathbb{R}$ we have

$$\frac{f(a)+f(a+2b)}{2} \geq f(a+b) + 2^{n+1}|b|,$$

$$f(a+b) + f(a+3b) \geq 2(f(a+2b) + 2^{n+1}|b|),$$

and

$$\frac{f(a+2b)+f(a+4b)}{2} \geq f(a+3b) + 2^{n+1}|b|.$$

By adding these three inequalities and cancelling common terms we obtain

$$\frac{f(a)+f(a+4b)}{2} \geq f(a+2b) + 2^{n+3}|b|.$$

Now setting $x = a$, $y = a + 4b$ gives us

$$\frac{f(x)+f(y)}{2} \geq f\left(\frac{x+y}{2}\right) + 2^{n+1}|x-y|,$$

which completes our induction.

3.2 Proposed Problems

Problem 3.1. Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ which satisfy simultaneously the following properties:

1) $f(m) > f(n)$ for all $m > n$;
2) $f(f(n)) = 4n + 9$;
3) $f(f(n) - n) = 2n + 9$.

Problem 3.2. (**IMO 2007 shortlist**) Consider those functions $f : \mathbb{N}^* \mapsto \mathbb{N}^*$ which satisfy the condition

$$f(m+n) \geq f(m) + f(f(n)) - 1,$$

for all $m, n \in \mathbb{N}$. Find all possible values of $f(2007)$.

Problem 3.3. Let $f : \mathbb{N} \to \mathbb{N}$ be a function such that $f(0) = 1$ and

$$f(n) = f\left(\left[\frac{n}{2}\right]\right) + f\left(\left[\frac{n}{3}\right]\right),$$

for all $n \geq 1$. Prove that

$$f(n-1) < f(n) \Leftrightarrow n = 2^k 3^h, \quad \text{for some} \quad k, h \in \mathbb{N}.$$

Problem 3.4. (**AMM 10728**) Find all functions $f : \mathbb{Z} \to \mathbb{R}$ satisfying

$$f(a^3 + b^3 + c^3) = f(a)^3 + f(b)^3 + f(c)^3,$$

whenever $a, b, c \in \mathbb{Z}$.

Problem 3.5. (**APMO 2008**) Consider the function $f : \mathbb{N} \to \mathbb{N}$ defined by the following conditions :

i) $f(0) = 0$;
ii) $f(2n) = 2f(n)$, for all $n \in \mathbb{N}$;
iii) $f(2n+1) = n + 2f(n)$, for all $n \in \mathbb{N}$.

a) Determine the three sets

$$L = \{n | f(n) < f(n+1)\},\ E = \{n | f(n) = f(n+1)\}, \text{ and}$$
$$G = \{n | f(n) > f(n+1)\};$$

b) For each $k \geq 0$, find a formula for $a_k = \max\{f(n) : 0 \leq n \leq 2^k\}$ in terms of k.

Problem 3.6. (**India 2000**) Suppose $f : \mathbb{Q} \to \{0, 1\}$ is a function with the property that for $x, y \in \mathbb{Q}$, if $f(x) = f(y)$, then $f(x) = f((x + y)/2) = f(y)$. If $f(0) = 0$ and $f(1) = 1$, show that $f(q) = 1$ for all rational number q greater than or equal to 1.

Problem 3.7. Find all functions $f : \mathbb{Q}^+ \to \mathbb{Q}^+$, such that

$$f(x) + f\left(\frac{1}{x}\right) = 1, \quad \text{and} \quad f(1 + 2x) = \frac{1}{2}f(x),$$

for any $x \in \mathbb{Q}^+$.

Problem 3.8. Find all functions $f : [0, +\infty) \to [0, 1]$, such that for any $x \geq 0$, $y \geq 0$

$$f(x)f(y) = \frac{1}{2}f(yf(x)).$$

Problem 3.9. (**China 2013**) Prove that there exists only one function $f : \mathbb{N}^* \to \mathbb{N}^*$ satisfying the following two conditions:

i) $f(1) = f(2) = 1$;

ii) $f(n) = f(f(n-1)) + f(n - f(n-1))$ for $n \geq 3$.

For each integer $m \geq 2$, find the value of $f(2^m)$.

Problem 3.10. (**Silk Road MC**) Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ which satisfy

$$2f(mn) \geq f(m^2 + n^2) - f(m)^2 - f(n)^2 \geq 2f(m)f(n).$$

Problem 3.11. (**Turkey**) Find all functions $f : \mathbb{Q}^+ \to \mathbb{Q}^+$ such that

$$f\left(\frac{x}{x+1}\right) = \frac{f(x)}{x+1}, \quad f(x) = x^3 f\left(\frac{1}{x}\right), \quad \text{for all} \quad x \in \mathbb{Q}^+.$$

Problem 3.12. Find all functions $f : \mathbb{Z} \to \mathbb{Z}$ that satisfy

$$f(m+n) + f(mn-1) = f(m)f(n) + 2,$$

for all integers m, n.

Problem 3.13. (**Estonia 2000**) Find all functions $f : \mathbb{N} \to \mathbb{N}$ such that

$$f(f(f(n))) + f(f(n)) + f(n) = 3n \quad \text{for all} \quad n \in \mathbb{N}.$$

Problem 3.14. Prove that there exists a unique function $f : \mathbb{Q}^+ \to \mathbb{Q}^+$ satisfying all the following conditions:

a) If $0 < q < \frac{1}{2}$, then $f(q) = 1 + f\left(\frac{q}{1-2q}\right)$;

b) If $1 < q \leq 2$, then $f(q) = 1 + f(q-1)$;

c) $f(q)f\left(\frac{1}{q}\right) = 1$ for all $q \in \mathbb{Q}^+$.

Chapter 4

Inequalities

4.1 Theory and Examples

Before we illustrate the different approaches one can use to prove an inequality by induction, we need to introduce some basic concepts regarding inequalities. To ease some of the following definitions, we regard an inequality as a function f in one or more variables which we need to prove that has a certain relative position with respect to zero (greater than zero, less than zero, etc), when the variables are subject to some constraints.

For example, for the inequality $\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b}\geq\frac{3}{2}$, valid for any positive reals a,b,c, we define $f(a,b,c)=\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b}-\frac{3}{2}$ and we have to prove that $f(a,b,c)\geq 0$ whenever a,b,c are positive reals.

A multivariable function $f(x_1,x_2,\ldots x_n)$ is called **symmetric** if for any $1\leq i,j\leq n$ we have

$$f(x_1,x_2,\ldots x_i,\ldots x_j,\ldots x_n)=f(x_1,x_2,\ldots x_j,\ldots x_i,\ldots x_n),$$

i.e. if whenever we swap any two variables we do not change the function.

Example. $f(a,b,c)=a^2+b^2+c^2$ is symmetric, whereas

$$f(a,b,c)=a^3+b^2+c^3$$

is not symmetric, as $f(a,c,b)=a^3+c^2+b^3\neq f(a,b,c)$.

When we deal with a symmetric inequality, we can choose any particular ordering among the variables, i.e. we can assume without loss of generality that $a \leq b \leq c$ (or any other permutation). This follows from the fact that any permutation can be written as product of transpositions.

A multivariable function $f(x_1, x_2, \dots x_n)$ is called **cyclic** if

$$f(x_1, x_2, \dots x_{n-1}, x_n) = f(x_n, x_1, x_2, \dots x_{n-1}).$$

Example. $f(a,b,c) = \frac{a}{b} + \frac{b}{c} + \frac{c}{a}$ is cyclic, while $f(a,b,c) = \frac{a}{b} + \frac{a}{c} + \frac{b}{a}$ is not, since $f(c,a,b) = \frac{c}{a} + \frac{c}{b} + \frac{a}{c} \neq f(a,b,c)$.

If we are given a cyclic inequality, we can assume that one particular variable is the largest or the smallest. However, we cannot choose a certain ordering among all the variables, as it was the case when we had a symmetric inequality.

A multivariable function $f(x_1, x_2, \dots x_n)$ is called **homogeneous** if whenever t is a real number, we have $f(tx_1, tx_2, \dots, tx_n) = t^k f(x_1, x_2, \dots x_n)$, where k is the degree of f.

Example. $f(a,b,c) = a^2 + b^2 + c^2 - ab - bc - ac$ is homogeneous, while $f(a,b) = a^2 - b$ is not, since $f(ta, tb) \neq t^2 f(a,b)$.

When an inequality $f(x_1, x_2, \dots, x_n) \geq 0$ is homogeneous of order k, i.e.

$$f(tx_1, tx_2, \dots, tx_n) = t^k f(x_1, x_2, \dots, x_n),$$

without loss of generality, we can assume that the sum of the variables is constant, in order to bring the inequality in a simpler form, easier to prove it. In general, we can consider more complicated expressions, like $g(x_1, x_2, \dots, x_n) = constant$, where $g(x_1, x_2, \dots, x_n)$ is a homogeneous expression of any order, as long as these fit into the framework of the problem.

Now that we have established these basic notions, let us look at some examples where we can solve inequalities using induction:

Example 4.1. (**Bernoulli's Inequality**) For any real $x \geq -1$, show that if n is a positive integer, then $(1+x)^n \geq 1 + nx$.

Solution. We prove the result by induction on n:

Let $P(n)$: $(1+x)^n \geq 1 + nx$, for all $x \in \mathbb{R}$, $x \geq -1$.

$(1+x)^1 = 1 + 1 \cdot x$, so $P(1)$ is true.

Assume that $P(n)$ is true for some $n \geq 1$. We must show that $P(n+1)$ is true.

$$(1+x)^{n+1} = (1+x)^n \cdot (1+x) \geq (1+nx)(1+x),$$

since $P(n)$ was assumed to be true, and $1+x \geq 0$. Now

$$(1+nx)(1+x) = 1 + (n+1)x + nx^2 \geq 1 + (n+1)x,$$

which is what we wanted. So $P(n+1)$ is true, thus, by induction, $P(n)$ is true for all $n \geq 0$.

Remark. The inequality actually follows from the following, more general:

Example 4.2. (**Bernoulli's Inequality**) Let x_i, $i = 1, 2, \ldots, n$ be real numbers with the same sign (i.e. all positive or all negative), greater than -1. Then we have

$$(1+x_1)(1+x_2)\cdots(1+x_n) \geq 1 + x_1 + x_2 + \ldots + x_n.$$

Solution. We prove the inequality by induction on n. For $n = 1$, we have $1 + x_1 \geq 1 + x_1$.

Suppose now that the inequality holds for n arbitrary real numbers $x_i \geq -1$ with the same sign. Now consider $n+1$ arbitrary real numbers $x_i \geq -1$ with $i = 1, 2, \ldots, n+1$ with the same sign.

Since $x_1, x_2, \ldots, x_{n+1}$ have the same sign, we have that

$$(x_1 + x_2 + \ldots + x_n)x_{n+1} \geq 0. \tag{*}$$

Hence

$$(1+x_1)(1+x_2)\cdots(1+x_{n+1}) \geq (1 + x_1 + x_2 + \ldots + x_n)(1 + x_{n+1}),$$

using the assumption that $P(n)$ is true. But now

$$\begin{aligned} &(1 + x_1 + x_2 + \ldots + x_n)(1 + x_{n+1}) \\ &= 1 + x_1 + \ldots + x_n + x_{n+1} + (x_1 + \ldots + x_n)x_{n+1} \\ &\geq 1 + x_1 + \ldots + x_{n+1}, \end{aligned}$$

using (*).

Example 4.3. Let $n \geq 3$ be a positive integer and let $x_1, x_2, \ldots, x_n$ be positive reals. Prove that

$$\frac{x_1}{x_1+x_2} + \frac{x_2}{x_2+x_3} + \ldots + \frac{x_n}{x_n+x_1} < n-1.$$

Solution. The base case is $n = 3$ and we have to show that

$$\frac{x_1}{x_1+x_2} + \frac{x_2}{x_2+x_3} + \frac{x_3}{x_1+x_2} < 2,$$

which, after multiplying both sides by -1 and adding 3, is equivalent to

$$\frac{x_2}{x_1+x_2} + \frac{x_3}{x_2+x_3} + \frac{x_1}{x_1+x_3} > 1.$$

The last inequality is true since one has $\frac{x_2}{x_1+x_2} > \frac{x_2}{x_1+x_2+x_3}$ and the analogous inequalities, which summed give exactly the desired relation.

Assume now that $P(n)$ holds for some $n \geq 3$ and we want to show that $P(n+1)$ holds. To make use of the induction hypothesis, one thing we could try is to delete one of $x_1, x_2, \ldots, x_{n+1}$ and prove that the obtained expression in n variables decreases by at most 1. The induction hypothesis would then do the job.

If we eliminate x_i, we are left with

$$\frac{x_1}{x_1+x_2} + \ldots + \frac{x_{i-1}}{x_{i-1}+x_{i+1}} + \frac{x_{i+1}}{x_{i+1}+x_{i+2}} + \ldots.$$

Thus the expression would decrease by

$$\frac{x_{i-1}}{x_{i-1}+x_i} + \frac{x_i}{x_{i+1}+x_i} - \frac{x_{i-1}}{x_{i-1}+x_{i+1}}.$$

So it suffices to prove that

$$\frac{x_{i-1}}{x_{i-1}+x_i} + \frac{x_i}{x_{i+1}+x_i} - \frac{x_{i-1}}{x_{i-1}+x_{i+1}} \leq 1.$$

Notice that if $x_{i+1} \leq x_i$ (here $x_{n+1} = x_1$), then $\frac{x_{i-1}}{x_{i-1}+x_i} \leq \frac{x_{i-1}}{x_{i-1}+x_{i+1}}$ and $\frac{x_i}{x_i+x_{i+1}} < 1$. So by taking x_i to be the largest of $x_1, x_2, \ldots, x_{n+1}$ we ensure the condition we want for our induction to work. This completes our proof.

Example 4.4. (**Titu's Lemma**) Let $a_1, a_2, \ldots, a_n \in \mathbb{R}$ and $b_1, b_2, \ldots, b_n$ be positive reals. Prove that

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} + \ldots + \frac{a_n^2}{b_n} \geq \frac{(a_1 + a_2 + \ldots + a_n)^2}{b_1 + b_2 + \ldots + b_n}.$$

Solution. For the proof by induction, the base case $n = 1$ is clear. Notice that for the induction step, it suffices to prove the inequality for $n = 2$, since we then have

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} + \ldots + \frac{a_n^2}{b_n} \geq \frac{(a_1 + a_2 + \ldots + a_{n-1})^2}{b_1 + b_2 + \ldots + b_{n-1}} + \frac{a_n^2}{b_n} \geq \frac{(a_1 + a_2 + \ldots + a_n)^2}{b_1 + b_2 + \ldots + b_n}.$$

We are left to prove the case $n = 2$, that is

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} \geq \frac{(a_1 + a_2)^2}{b_1 + b_2}.$$

Multiplying by $b_1 b_2 (b_1 + b_2)$ we have to show

$$(a_1^2 b_2 + a_2^2 b_1)(b_1 + b_2) \geq (a_1 + a_2)^2 b_1 b_2.$$

Further, after expanding the parentheses,

$$(a_1^2 + a_2^2) b_1 b_2 + a_1^2 b_2^2 + a_2^2 b_1^2 \geq (a_1^2 + a_2^2) b_1 b_2 + 2 a_1 b_1 a_2 b_2,$$

which is true since $a_1^2 b_2^2 + a_2^2 b_1^2 \geq 2 a_1 b_1 a_2 b_2$ by completing the square.

Example 4.5. Let $a_1, a_2, \ldots, a_n \in \mathbb{R}_+$ such that $a_1 a_2 \ldots a_n = 1$. Prove that

$$a_1 + a_2 + \ldots + a_n \geq n.$$

Solution. We proceed by induction on n. For $n = 1$, things are clear. Assume now that the result holds for n variables and we want to prove it for $n + 1$.

Consider the variables $a_1, a_2, \ldots, a_{n+1}$. Without loss of generality, we can assume that a_n is the smallest of the variables and a_{n+1} is the largest.

As $a_1 a_2 \ldots a_n a_{n+1} = 1$, this implies that $a_n \leq 1$ and $a_{n+1} \geq 1$, so that $(1 - a_n)(1 - a_{n+1}) \leq 0 \Leftrightarrow a_n + a_{n+1} - 1 \geq a_n a_{n+1}$. We now make use of the case of n variables by considering the variables a_1, a_2, ..., a_{n-1} and $a_n a_{n+1}$. These are n variables that have product equal to 1 so we know that

$$a_1 + a_2 \dots a_{n-1} + a_n a_{n+1} \geq n.$$

But from above, $a_n a_{n+1} \leq a_n + a_{n+1} - 1$, which implies

$$a_1 + a_2 + \dots + a_n + a_{n+1} \geq n + 1,$$

establishing the inequality we wanted to prove.

Remark. If in the above inequality we set

$$a_1 = \frac{x_1}{\sqrt[n]{x_1 x_2 \dots x_n}}, \dots, a_n = \frac{x_n}{\sqrt[n]{x_1 x_2 \dots x_n}}, \text{ for some positive reals } x_1, \dots, x_n,$$

we obtain the well-known *AM-GM inequality:*

$$\frac{x_1 + x_2 + \dots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \dots x_n}.$$

We would now like to prove a very strong result, known as Suranyi's Inequality. Before we do this, we need to introduce a few more concepts:

Definition. Let $a_1 \leq a_2 \leq \dots \leq a_n$ and $b_1 \leq b_2 \leq b_n$ be any real numbers. We call $S(n) = a_1 b_1 + a_2 b_2 + \dots a_n b_n$ the **Sorted sum** of $a_1, \dots a_n, b_1, \dots b_n$ and $R(n) = a_1 b_n + a_2 b_{n-1} + \dots a_n b_1$ the **Reversed sum.**
For $c_1, c_2, \dots, c_n$ being any permutation of the numbers $b_1, b_2, \dots, b_n$, we call $P(n) = a_1 c_1 + a_2 c_2 + \dots a_n c_n$ a **Permuted sum** of $a_1, \dots a_n, b_1, \dots b_n$.

Example 4.6. (**Rearrangement Inequality**) With the above notations, we have

$$S(n) \geq P(n) \geq R(n).$$

Solution. We prove the result by induction on n.

For $n = 1$, we get $S(1) = P(1) = R(1)$.

We now assume $S(n) \geq P(n)$ for arbitrary reals $a_1 \leq a_2 \leq \dots \leq a_n$ and $b_1 \leq b_2 \leq b_n$ and we prove that $S(n+1) \geq P(n+1)$ for arbitrary reals $a_1 \leq a_2 \leq \dots \leq a_{n+1}$ and $b_1 \leq b_2 \leq b_{n+1}$, where $n \geq 1$.

Since the $c_1, \dots c_{n+1}$ is just a permutation of $b_1, \dots b_{n+1}$, there is some i such that $b_{n+1} = c_i$ and $c_{n+1} = b_j$. From the ordering $a_1 \leq a_2 \leq \dots \leq a_{n+1}$ and $b_1 \leq b_2 \dots \leq b_{n+1}$ we have

$$(a_{n+1} - a_i)(b_{n+1} - b_j) \geq 0.$$

This gives $a_i b_j + a_{n+1} b_{n+1} \geq a_i b_{n+1} + a_{n+1} b_j$, hence

$$a_i b_j + a_{n+1} b_{n+1} \geq a_i c_i + a_{n+1} c_{n+1}.$$

This implies that in the sum $P(n+1)$, if we swap c_i and c_{n+1} we obtain a sum as least as big as the original one. But once we swapped c_i and c_{n+1}, we reduce the problem to showing $S(n) \geq P(n)$, which is our inductive assumption. So we have established $S(n+1) \geq P(n+1)$, thus $S(n) \geq P(n)$ for all n.

The inequality $P(n) \geq R(n)$ now follows easily from $S(n) \geq P(n)$ by replacing $b_1 \leq b_2 \leq \dots \leq b_n$ by $-b_n \leq -b_{n-1} \leq \dots \leq -b_1$.

We shall also need the following consequence of the Rearrangement inequality:

Example 4.7. (**Chebyshev's inequality**) If $a_1 \leq a_2 \leq \dots \leq a_n$ and $b_1 \leq b_2 \leq \dots \leq b_n$ are two increasing sequences of real numbers then

$$\frac{a_1 b_1 + a_2 b_2 + \dots + a_n b_n}{n} \geq \frac{a_1 + a_2 + \dots + a_n}{n} \cdot \frac{b_1 + b_2 + \dots + b_n}{n}.$$

Solution. By the Rearrangement inequality

$$\begin{aligned}
a_1 b_1 + a_2 b_2 + \dots + a_n b_n &= a_1 b_1 + a_2 b_2 + \dots + a_n b_n \\
a_1 b_1 + a_2 b_2 + \dots + a_n b_n &\geq a_1 b_2 + a_2 b_3 + \dots + a_n b_1 \\
a_1 b_1 + a_2 b_2 + \dots + a_n b_n &\geq a_1 b_3 + a_2 b_4 + \dots + a_n b_2 \\
&\dots \\
a_1 b_1 + a_2 b_2 + \dots + a_n b_n &\geq a_1 b_n + a_2 b_1 + \dots + a_n b_{n-1}
\end{aligned}$$

And the result follows by summing up all those and dividing by n^2.

We are now ready to prove the promised result:

Example 4.8. (**Suranyi's Inequality**) Prove that for any positive numbers $a_1, a_2, \dots a_n$, the following inequality holds:

$$(n-1)(a_1^n + a_2^n + \dots a_n^n) + na_1 a_2 \dots a_n$$
$$\geq (a_1 + a_2 + \dots + a_n)(a_1^{n-1} + a_2^{n-1} + \dots + a_n^{n-1}).$$

Solution. For $n = 1$ we have nothing to prove. Assume now that the inequality is true for n numbers and let us prove it for $n+1$ numbers.

Due to the symmetry and homogeneity of the inequality, it is enough to prove it under the conditions $a_1 \geq a_2 \geq \dots \geq a_{n+1}$ and $a_1 + a_2 + \dots + a_n = 1$. We need to prove that:

$$n\sum_{i=1}^{n} a_i^{n+1} + na_{n+1}^{n+1} + na_{n+1}\prod_{i=1}^{n} a_i + a_{n+1}\prod_{i=1}^{n} a_i - (1+a_{n+1})\left(\sum_{i=1}^{n} a_i^n + a_{n+1}^n\right) \geq 0.$$

From the inductive hypothesis we have

$$(n-1)(a_1^n + a_2^n + \dots + a_n^n) + na_1 a_2 \dots a_n \geq a_1^{n-1} + a_2^{n-1} + \dots + a_n^{n-1}.$$

Therefore

$$na_{n+1}\prod_{i=1}^{n} a_i \geq a_{n+1}\sum_{i=1}^{n} a_i^{n-1} - (n-1)a_{n+1}\sum_{i=1}^{n} a_i^n.$$

Using this last inequality, it remains to prove that:

$$\left(n\sum_{i=1}^{n} a_i^{n+1} - \sum_{i=1}^{n} a_i^n\right) - a_{n+1}\left(n\sum_{i=1}^{n} a_i^n - \sum_{i=1}^{n} a_i^{n-1}\right)$$
$$+a_{n+1}\left(\prod_{i=1}^{n} a_i + (n-1)a_{n+1}^n - a_{n+1}^{n-1}\right) \geq 0.$$

Now we will break this inequality into

$$a_{n+1}\left(\prod_{i=1}^{n} a_i + (n-1)a_{n+1}^n - a_{n+1}^{n-1}\right) \geq 0$$

and

$$\left(n\sum_{i=1}^{n} a_i^{n+1} - \sum_{i=1}^{n} a_i^n\right) - a_{n+1}\left(n\sum_{i=1}^{n} a_i^n - \sum_{i=1}^{n} a_i^{n-1}\right) \geq 0.$$

The first of these two inequalities is easy, since by Bernoulli's inequality

$$\begin{aligned}\prod_{i=1}^{n} a_i &= \prod_{i=1}^{n}(a_i - a_{n+1} + a_{n+1}) \\ &\geq a_{n+1}^n + a_{n+1}^{n-1} \cdot \sum_{i=1}^{n}(a_i - a_{n+1}) = a_{n+1}^{n-1} - (n-1)a_{n+1}^n.\end{aligned}$$

For the second inequality, we can rewrite it as

$$n\sum_{i=1}^{n} a_i^{n+1} - \sum_{i=1}^{n} a_i^n \geq a_{n+1}\left(n\sum_{i=1}^{n} a_i^n - \sum_{i=1}^{n} a_i^{n-1}\right).$$

Since

$$n\sum_{i=1}^{n} a_i^n - \sum_{i=1}^{n} a_i^{n-1} \geq 0$$

(using Chebyshev's inequality) and $a_{n+1} \leq \frac{1}{n}$ from $a_{n+1} \leq \ldots \leq a_1$ and $a_1 + \ldots + a_n = 1$, it is enough to prove that

$$n\sum_{i=1}^{n} a_i^{n+1} - \sum_{i=1}^{n} a_i^n \geq \frac{1}{n}\left(n\sum_{i=1}^{n} a_i^n - \sum_{i=1}^{n} a_i^{n-1}\right).$$

This inequality now follows from the fact that $na_i^{n+1} + \frac{1}{n}a_i^{n-1} \geq 2a_i^n$, for all i. This establishes our inductive step and completes the proof.

We are also going to prove the well-known AM-GM and Cauchy-Schwarz inequalities by induction. To do this, we need to introduce a very useful variation of the induction principle, called **Cauchy induction**. The way this works is that instead of proving that $P(n) \Rightarrow P(n+1)$, we prove that $P(n) \Rightarrow P(2n)$ and that $P(n) \Rightarrow P(n-1)$. It is clear that in this way, starting from any base case, we can get to any non-negative integer n, hence proving $P(n)$ for all possible values of n.

Example 4.9. (**Cauchy-Schwarz**) Let $a_1, \ldots, a_n, b_1, \ldots, b_n$ be any real numbers. Then

$$(a_1^2 + \ldots + a_n^2)(b_1^2 + \ldots + b_n^2) \geq (a_1 b_1 + \ldots + a_n b_n)^2.$$

Solution. We use Cauchy induction. For $n = 1$ the inequality is clear. Also for $n = 2$ we have to prove that $(a_1^2 + a_2^2)(b_1^2 + b_2^2) \geq (a_1 b_1 + a_2 b_2)^2$, which after expanding the parentheses on both sides and simplification is equivalent to $a_1^2 b_2^2 + a_2^2 b_1^2 \geq 2a_1 a_2 b_1 b_2 \Leftrightarrow (a_1 b_2 - a_2 b_1)^2 \geq 0$.

Assume now that the inequality holds for some $n \geq 2$ and let us prove that it holds for $2n$. We need to show that

$$(a_1^2 + \ldots + a_{2n}^2)(b_1^2 + \ldots + b_{2n}^2) \geq (a_1 b_1 + \ldots + a_{2n} b_{2n})^2.$$

Let $x_1^2 = a_1^2 + \ldots + a_n^2$ (note that this substitution is possible as $a_1^2 + \ldots + a_n^2 \geq 0$),

$$x_2^2 = a_{n+1}^2 + \ldots + a_{2n}^2, \ y_1^2 = b_1^2 + \ldots + b_n^2, \ y_2^2 = b_{n+1}^2 + \ldots + b_{2n}^2.$$

Then from the base case with two variables, we know that

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) \geq (x_1 y_1 + x_2 y_2)^2.$$

$P(n)$ implies further that

$$x_1 y_1 = \sqrt{(a_1^2 + \ldots + a_n^2)(b_1^2 + \ldots + b_n^2)} \geq (a_1 b_1 + \ldots + a_n b_n),$$

and similarly

$$x_2 y_2 \geq (a_{n+1} b_{n+1} + \ldots + a_{2n} b_{2n}).$$

Therefore

$$(x_1 y_1 + x_2 y_2)^2 \geq (a_1 b_1 + \ldots + a_{2n} b_{2n})^2,$$

which is what we wanted.

Now we are left to prove that $P(n) \Rightarrow P(n-1)$. In other words, assuming that for any n variables we have

$$(a_1^2 + \ldots + a_n^2)(b_1^2 + \ldots + b_n^2) \geq (a_1 b_1 + \ldots + a_n b_n)^2,$$

we want to prove that for any $n - 1$ variables

$$(a_1^2 + \ldots + a_{n-1}^2)(b_1^2 + \ldots + b_{n-1}^2) \geq (a_1 b_1 + \ldots + a_{n-1} b_{n-1})^2.$$

This follows from $P(n)$ by simply setting $a_n = b_n = 0$.

Example 4.10. (**AM-GM inequality**) Let $x_1, x_2, \ldots x_n$ be non-negative reals. Then

$$\frac{x_1 + \ldots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \ldots x_n}.$$

Equality holds if and only if the numbers are equal.

Solution 1. Our base case will be $n = 2$, for which we need to show that

$$\frac{x_1 + x_2}{2} \geq \sqrt{x_1 x_2}.$$

This follows from writing the difference as a perfect square:

$$\frac{x_1 + x_2}{2} - \sqrt{x_1 x_2} = \frac{(\sqrt{x_1} - \sqrt{x_2})^2}{2}$$

Equality holds if and only if $x_1 = x_2$.

Now assume that the inequality is true for some $n \geq 2$. Then we have

$$\begin{aligned}\frac{x_1 + \ldots + x_{2n}}{2n} &= \frac{\frac{x_1 + \ldots + x_n}{n} + \frac{x_{n+1} + \ldots + x_{2n}}{n}}{2} \\ &\geq \frac{\sqrt[n]{x_1 \ldots x_n} + \sqrt[n]{x_{n+1} \ldots x_{2n}}}{2} \\ &\geq \sqrt[2n]{x_1 \ldots x_{2n}}.\end{aligned}$$

Here we used the inequality for n twice and then the inequality for $n = 2$ and the numbers $\sqrt[n]{x_1 \ldots x_n}$ and $\sqrt[n]{x_{n+1} \ldots x_{2n}}$. Equality holds if and only if $x_1 = \ldots = x_n$, $x_{n+1} = \ldots = x_{2n}$ and their geometric means are also equal, which implies $x_1 = x_2 = \ldots = x_{2n}$.

Finally, let us assume that the inequality holds for n variables and prove that this implies that it holds for $n-1$ variables. The idea is to plug in x_n as the arithmetic mean of the other terms and cancel it out. So set

$$x_n = \frac{x_1 + \ldots + x_{n-1}}{n-1}.$$

From AM-GM for n numbers which we assumed to be true we have that

$$\frac{x_1 + \ldots + x_{n-1} + \frac{x_1 + \ldots + x_{n-1}}{n-1}}{n} \geq \sqrt[n]{x_1 x_2 \ldots x_{n-1} \cdot \frac{x_1 + \ldots + x_{n-1}}{n-1}}.$$

This implies that

$$\frac{x_1 + \ldots + x_{n-1}}{n-1} \geq \sqrt[n]{x_1 x_2 \ldots x_{n-1} \cdot \frac{x_1 + \ldots + x_{n-1}}{n-1}}.$$

Raising this expression to the n-th power yields

$$\left(\frac{x_1 + \ldots + x_{n-1}}{n-1}\right)^n \geq x_1 x_2 \ldots x_{n-1} \cdot \frac{x_1 + \ldots + x_{n-1}}{n-1},$$

and cancellation yields

$$\left(\frac{x_1 + \ldots + x_{n-1}}{n-1}\right)^{n-1} \geq x_1 x_2 \ldots x_{n-1},$$

which implies the desired inequality.

Remark. A variant of the Cauchy induction is the following: assume we have to prove an inequality involving variables $x_1, \ldots, x_n$. We start by showing as before that the base case holds and that $P(n) \Rightarrow P(2n)$. If the base case was $n = 1$ or $n = 2$, this would prove the result for the situation when n is a power of 2. For a general n, we choose a k sufficiently large such that $2^k \geq n$. We then use the fact that $P(2^k)$ holds and set some convenient particular values for the variables $x_{n+1}, \ldots, x_{2^k}$ which enable us to derive the result for the remaining variables $x_1, \ldots, x_n$. To illustrate how this variant works, we use to give another proof of the $AM - GM$ inequality:

Second solution to AM-GM. The first part of the proof which establishes the base case $n = 2$ and the implication $P(n) \Rightarrow P(2n)$ is the same as in the first solution.

Now given n a positive integer, choose k such that $2^k > n$ and set

$$a_{n+1} = \ldots = a_{2^k} = \frac{a_1 + \ldots a_n}{n}.$$

Then using the fact that $P(2^k)$ is true, we have

$$a_1 + \ldots + a_n + (2^k - n)\tfrac{a_1 + \ldots + a_n}{n} \geq 2^k \sqrt[2^k]{a_1 a_2 \ldots a_n (\tfrac{a_1 + \ldots + a_n}{n})^{2^k - n}}$$

which further simplifies to $1 \geq a_1 \dots a_n \cdot (\frac{a_1+\dots+a_n}{n})^{-n}$, which is exactly the sought inequality.

Another example for which we use the second form of Cauchy induction is the following:

Example 4.11. (**USSR 1990**) All coefficients of a quadratic polynomial $f(x) = ax^2 + bx + c$ are positive and $a + b + c = 1$. Prove that the inequality

$$f(x_1) \cdot \dots \cdot f(x_n) \geq 1$$

holds for all positive numbers $x_1, \dots, x_n$, satisfying $x_1 \cdot \dots \cdot x_n = 1$.

Solution. First observe that if $x_1 = 1$, we have$f(x_1) = a + b + c = 1$. We now prove that for any positive reals x and y we have

$$f(x) \cdot f(y) \geq \left(f\left(\sqrt{xy}\right)\right)^2. \tag{1}$$

Let $z := \sqrt{xy}$. Then one has

$$f(x) \cdot f(y) - (f(z))^2 = a^2\left(x^2y^2 - z^4\right) + b^2\left(xy - z^2\right) + c^2(1-1)$$

$$+ab\left(x^2y + xy^2 - 2z^3\right) + ac\left(x^2 + y^2 - 2z^2\right) + bc\,(x + y - 2z)$$

$$= ab\left(\sqrt{x^2y} - \sqrt{xy^2}\right)^2 + ac(x-y)^2 + bc\left(\sqrt{x} - \sqrt{y}\right)^2 \geq 0.$$

We now prove by induction that whenever n is a power of 2, for all positive reals $x_1, \dots, x_n$ the following holds:

$$f(x_1) \cdot \dots \cdot f(x_n) \geq \left(f\left(\sqrt[n]{x_1 \cdot \dots \cdot x_n}\right)\right)^n.$$

Assume that this is true for $n = 2^k$. Then using the inductive hypothesis and (1) we obtain

$$\begin{aligned} f(x_1) \cdot \dots \cdot f\left(x_{2^{k+1}}\right) &= \left(f(x_1) \cdot \dots \cdot f(x_{2^k})\right) \cdot \left(f\left(x_{2^k+1}\right) \cdot \dots \cdot f\left(x_{2^{k+1}}\right)\right) \\ &\geq \left(f\left(\sqrt[2^k]{x_1 \cdot \dots \cdot x_{2^k}}\right) \cdot f\left(\sqrt[2^k]{x_{2^k+1} \cdot \dots \cdot x_{2^{k+1}}}\right)\right)^{2^k} \\ &\geq \left(f\left(\sqrt[2^{k+1}]{x_1 \cdot \dots \cdot x_{2^{k+1}}}\right)\right)^{2^{k+1}}, \end{aligned}$$

and so the statement is also true for $n = 2^{k+1}$.

Suppose now that n is arbitrary and $x_1 \cdot \ldots \cdot x_n = 1$. Let k be the positive integer such that $2^{k-1} < n \le 2^k$. Let us add, if necessary, $x_{n+1} = x_{n+2} = \ldots = x_{2^k} = 1$. Since $f(x_{n+1}) = f(x_{n+2}) = \ldots = f(x_{2^k}) = 1$, we may write

$$f(x_1) \cdot \ldots \cdot f(x_n) = f(x_1) \cdot \ldots \cdot f(x_{2^k}) \ge \left(f\left(\sqrt[2^k]{x_1 \cdot \ldots \cdot x_{2^k}}\right)\right)^{2^k} = 1.$$

This completes our proof.

4.2 Proposed Problems

Problem 4.1. There are $n \geq 1$ real numbers with non-negative sum written on a circle. Prove that one can enumerate them $a_1, a_2, \dots, a_n$ such that they are consecutive on the circle and

$$a_1 \geq 0,\ a_1 + a_2 \geq 0, \dots,\ a_1 + a_2 + \dots + a_{n-1} \geq 0,\ a_1 + a_2 + \dots + a_n \geq 0.$$

Problem 4.2. Let $a_1, a_2, \dots, a_n$ be positive real numbers. Prove that

$$\frac{a_1}{(1+a_1)^2} + \frac{a_2}{(1+a_1+a_2)^2} + \cdots + \frac{a_n}{(1+a_1+\cdots+a_n)^2} < \frac{a_1+\cdots+a_n}{1+a_1+\cdots+a_n}.$$

Problem 4.3. Show that $2(\sqrt{n+1}-1) < 1 + \dfrac{1}{\sqrt{2}} + \dots + \dfrac{1}{\sqrt{n}} < 2\sqrt{n}.$

Problem 4.4. Let x be a real number. Prove that for all positive integers n,

$$|\sin nx| \leq n|\sin x|.$$

Problem 4.5. Let $n > 2$. Find the least constant k such that for any $a_1, \dots, a_n > 0$ with product 1 we have

$$\frac{a_1a_2}{(a_1^2+a_2)(a_2^2+a_1)} + \frac{a_2a_3}{(a_2^2+a_3)(a_3^2+a_2)} + \dots + \frac{a_na_1}{(a_n^2+a_1)(a_1^2+a_n)} \leq k.$$

Problem 4.6. Let $a_1, a_2, \dots, a_n$ be integers, not all zero, such that $a_1 + a_2 + \cdots + a_n = 0$. Prove that

$$|a_1 + 2a_2 + \cdots + 2^{k-1}a_k| > \frac{2^k}{3},$$

for some $k \in \{1, 2, \dots n\}$.

Problem 4.7. Let $n \geq 2$ and $a_1, a_2, \dots, a_n \in (0,1)$ with $a_1a_2\dots a_n = A^n$. Show that

$$\frac{1}{1+a_1} + \dots + \frac{1}{1+a_n} \leq \frac{n}{1+A}.$$

Problem 4.8. (**APMO 1999**) Let $a_1, a_2, \ldots$ be a sequence of real numbers satisfying $a_{i+j} \leq a_i + a_j$ for all $i, j = 1, 2, \ldots$. Prove that

$$a_1 + \frac{a_2}{2} + \frac{a_3}{3} + \cdots + \frac{a_n}{n} \geq a_n$$

for each positive integer n.

Problem 4.9. (**Tuymaada 2000**) Let $n \geq 2$ be a positive integer and $x_1, \ldots, x_n$ be real numbers such that $0 < x_k \leq \frac{1}{2}$, for all $k = 1, 2, \ldots, n$. Prove that

$$\left(\frac{n}{x_1 + x_2 + \ldots + x_n} - 1\right)^n \leq \left(\frac{1}{x_1} - 1\right) \cdot \ldots \cdot \left(\frac{1}{x_n} - 1\right).$$

Problem 4.10. (**China**) Let $a_1, \ldots, a_n$ be real numbers. Prove that the following two statements are equivalent:

a) $a_i + a_j \geq 0$ for all $i \neq j$;

b) If $x_1, \ldots, x_n$ are non-negative real numbers whose sum is 1, then

$$a_1 x_1 + \ldots + a_n x_n \geq a_1 x_1^2 + \ldots + a_n x_n^2.$$

Problem 4.11. (**USAMO 2000**) Let $a_1, b_1, a_2, b_2, \ldots, a_n, b_n$ be non-negative real numbers. Prove that

$$\sum_{i,j=1}^{n} \min\{a_i a_j, b_i b_j\} \leq \sum_{i,j=1}^{n} \min\{a_i b_j, a_j b_i\}.$$

Problem 4.12. (**Romania TST 1981**) Let $n \geq 1$ be a positive integer and let $x_1, x_2, \ldots, x_n$ be real numbers such that $0 \leq x_n \leq x_{n-1} \leq \ldots \leq x_3 \leq x_2 \leq x_1$. We consider the sums

$$s_n = x_1 - x_2 + \ldots + (-1)^n x_{n-1} + (-1)^{n+1} x_n;$$
$$S_n = x_1^2 - x_2^2 + \ldots + (-1)^n x_{n-1}^2 + (-1)^{n+1} x_n^2.$$

Show that $s_n^2 \leq S_n$.

Problem 4.13. Let $1 = x_1 \leq x_2 \leq \ldots \leq x_{n+1}$ be non-negative integers. Prove that

$$\frac{\sqrt{x_2 - x_1}}{x_2} + \ldots + \frac{\sqrt{x_{n+1} - x_n}}{x_{n+1}} < 1 + \frac{1}{2} + \ldots + \frac{1}{n^2}.$$

Problem 4.14. Prove that

$$\sum_{i=0}^{n} |\sin(2^i x)| \leq 1 + \frac{\sqrt{3}}{2} n,$$

where n is a non-negative integer.

Problem 4.15. Prove the following inequality

$$2(a^{2012} + 1)(b^{2012} + 1)(c^{2012} + 1) \geq (1 + abc)(a^{2011} + 1)(b^{2011} + 1)(c^{2011} + 1),$$

where $a > 0$, $b > 0$, $c > 0$.

Problem 4.16. Prove that for any $n \in \mathbb{Z}$, $n \geq 14$ and any $x \in \left(0, \frac{\pi}{2n}\right)$ the following inequality holds:

$$\frac{\sin 2x}{\sin x} + \frac{\sin 3x}{\sin 2x} + \ldots + \frac{\sin(n+1)x}{\sin nx} < 2 \cot x.$$

Problem 4.17. Let $A_n = \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$, $n \geq 2$. Prove that

$$e^{A_n} > \sqrt[n]{n!} \geq 2^{A_n}.$$

Problem 4.18. Show that if $x_1, x_2, \ldots, x_n \in (0, 1/2)$, then

$$\frac{x_1 x_2 \ldots x_n}{(x_1 + x_2 + \ldots + x_n)^n} \leq \frac{(1 - x_1) \ldots (1 - x_n)}{((1 - x_1) + \ldots + (1 - x_n))^n}.$$

Problem 4.19. (**ELMO 2013 shortlist**) Let n be a fixed positive integer. Initially, n 1's are written on a blackboard. Every minute, David picks two numbers x and y written on the blackboard, erases them, and writes the number $(x + y)^4$ on the blackboard. Show that after $n - 1$ minutes, the number written on the blackboard is at least $2^{\frac{4n^2-4}{3}}$.

Chapter 5

Sequences and Recurrences

5.1 Theory and Examples

When we are given a sequence that satisfies certain properties (such as a recurrence relation) and we are asked to find the general term in closed form, we can try guessing $a_n = f(n)$ in several ways and then prove that our guess is correct by induction. We list a few key strategies that are helpful in this process:

a) Starting from a_1, we may try to calculate the first few terms $a_1, a_2, \ldots$ until we spot a possible formula.

b) It might be easier to actually calculate the ratios $\frac{a_{n+1}}{a_n}$ for $n = 1, 2, \ldots$ and then guess a rule which can be proved by induction.

c) The guess may become easier if we know that our sequence converges to some value a. In this case, we can substitute a_{n+1} and a_n with a in the recurrence relation and find the value of a. Then studying the value of $a_n - a$, we might spot a pattern.

d) By calculating the first few terms, we may notice some correlation between the given sequence and some other famous sequences that we know, for example the Fibonacci sequence.

e) Guess a functional form for a_n with a small number of free parameters and try to solve for these parameters. For example, one might guess $a_n = P(n)$ is a polynomial of low degree in n or an exponential $a_n = Cr^n$.

We present below some examples that use some of these ideas.

Example 5.1. The sequence $(a_n)_{n\geq 0}$ satisfies

$$a_{m+n} + a_{m-n} = \frac{1}{2}(a_{2m} + a_{2n}) \quad \text{for all integers} \quad m, n \geq 0 \quad \text{with} \quad m \geq n.$$

If $a_1 = 1$, find a_{1995}.

Solution. If we set $m = n$, we get that

$$a_{2m} + a_0 = \frac{1}{2}(a_{2m} + a_{2m}) \implies a_0 = 0.$$

Also

$$a_{1+0} + a_{1-0} = \frac{1}{2}(a_2 + a_0) \implies a_2 = 4.$$

We now use strong induction to prove that $a_k = k^2$ for all $k \geq 0$. The base cases $k = 0, 2$ are verified above and $a_1 = 1$ from hypothesis.
Setting $n = 0$, $m = k$ in the recurrence relation, we get

$$a_k + a_k = \frac{1}{2}(a_{2k} + 0) \implies a_{2k} = 4a_k$$

Setting $m = 2k$, $n = 1$ and using the above formula, we obtain

$$a_{2k+1} + a_{2k-1} = \frac{1}{2}(a_{4k} + a_2) = \frac{1}{2}(4a_{2k} + a_2),$$

and hence

$$a_{2k+1} = 2a_{2k} - a_{2k-1} + 2.$$

Assume now that $a_j = j^2$ for all $2 \leq j < n$. If $n = 2k$ is even, we obtain

$$a_n = 4a_k = (2k)^2,$$

and if $n = 2k + 1$ is odd, we obtain

$$a_n = 2a_{2k} - a_{2k-1} + 2 = 2(2k)^2 - (2k-1)^2 + 2 = (2k+1)^2.$$

Hence $a_n = n^2$ for all n. In particular, $a_{1995} = 1995^2$.

Example 5.2. Define the sequence $(x_n)_{n\geq 0}$ by

1) $x_n = 0$ if and only if $n = 0$ and
2) $x_{n+1} = x^2_{[\frac{n+3}{2}]} + (-1)^n x^2_{[\frac{n}{2}]}$ for all $n \geq 0$.

Find x_n in closed form.

Solution. Setting $n = 0$ and $n = 1$ yields $x_1 = x_1^2$ and $x_2 = x_2^2$, hence $x_1 = x_2 = 1$. From the given condition we obtain

$$x_{2n+1} = x^2_{n+1} + x^2_n \text{ and } x_{2n} = x^2_{n+1} - x^2_{n-1}.$$

Subtracting these relations implies

$$x_{2n+1} - x_{2n} = x^2_n + x^2_{n-1} = x_{2n-1},$$

hence

$$x_{2n+1} = x_{2n} + x_{2n-1},\ n \geq 1. \tag{1}$$

We induct on n to prove that

$$x_{2n} = x_{2n-1} + x_{2n-2},\ n \geq 1. \tag{2}$$

Indeed, $x_2 = x_1 + x_0$ and assume that (2) is true up to n. Then

$$\begin{aligned} x_{2n+2} - x_{2n} &= x^2_{n+2} - x^2_n - x^2_{n+1} + x^2_{n-1} \\ &\overset{(*)}{=} (x_{n+1} + x_n)^2 - x^2_n - x^2_{n+1} + (x_{n+1} - x_n)^2 \\ &= x^2_{n+1} + x^2_n = x_{2n+1}, \end{aligned}$$

as claimed (the equality (*) holds because of (1) and the induction hypothesis).

From relations (1) and (2) it follows that $x_{n+2} = x_{n+1} + x_n$ for all $n \geq 0$. Because $x_0 = 0$ and $x_1 = 1$, the sequence $(x_n)_{n\geq 0}$ is the Fibonacci sequence, hence

$$x_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

Example 5.3. Define the sequence $(a_n)_{n\geq 0}$ by $a_0 = 0$, $a_1 = 1$, $a_2 = 2$, $a_3 = 6$ and

$$a_{n+4} = 2a_{n+3} + a_{n+2} - 2a_{n+1} - a_n,\ n \geq 0.$$

Prove that n divides a_n for all $n > 0$.

Solution. From the hypothesis it follows that $a_4 = 12$, $a_5 = 25$ and $a_6 = 48$. We have

$$\frac{a_1}{1} = \frac{a_2}{2} = 1,\ \frac{a_3}{3} = 2,\ \frac{a_4}{4} = 3,\ \frac{a_5}{5} = 5,\ \frac{a_6}{6} = 8.$$

This shows that $\frac{a_n}{n} = F_n$ for all $n = 1, \ldots, 6$, where $(F_n)_{n\geq 1}$ is the Fibonacci sequence.

We prove by induction that $a_n = nF_n$ for all n. Indeed, assuming that $a_k = kF_k$ for $k \leq n+3$, we have

$$\begin{aligned}
a_{n+4} &= 2(n+3)F_{n+3} + (n+2)F_{n+2} - 2(n+1)F_{n+1} - nF_n \\
&= 2(n+3)F_{n+3} + (n+2)F_{n+2} - 2(n+1)F_{n+1} - n(F_{n+2} - F_{n+1}) \\
&= 2(n+3)F_{n+3} + 2F_{2n+2} - (n+2)F_{n+1} \\
&= 2(n+3)F_{n+3} + 2F_{2n+2} - (n+2)(F_{n+3} - F_{n+2}) \\
&= (n+4)(F_{n+3} + F_{n+2}) \\
&= (n+4)F_{n+4},
\end{aligned}$$

as desired.

Example 5.4. Let $(a_n)_{n\geq 0}$ be the sequence defined by $a_0 = 0$, $a_1 = 1$ and

$$\frac{a_{n+1} - 3a_n + a_{n-1}}{2} = (-1)^n,$$

for all integers $n > 0$. Prove that a_n is a perfect square for all $n \geq 0$.

Solution. Recall that the Fibonacci sequence is given by $F_1 = 1$, $F_2 = 1$ and $F_{n+2} = F_{n+1} + F_n$, $(n \geq 1)$. We claim that $a_n = F_n^2$ and we prove this by induction on $n \geq 0$.

The base cases are an immediate check:

$$a_0 = f_0^2 = 0,\ a_1 = f_1^2 = 1,\ a_2 = f_2^2 = 1.$$

Assume that the statement is true for some $n \geq 2$, we will prove it for $n+1$. We have

$$\frac{a_{n+1} - 3a_n + a_{n-1}}{2} + \frac{a_n - 3a_{n-1} + a_{n-2}}{2} = (-1)^n + (-1)^{n-1} = 0.$$

Thus $a_{n+1} = 2a_n - 2a_{n-1} - a_{n-2}$ and it is enough to prove that

$$F_{n+1}^2 = 2F_n^2 + 2F_{n-1}^2 - F_{n-2}^2, \text{ or}$$

$$(F_n + F_{n-1})^2 = 2F_n^2 + 2F_{n-1}^2 - (F_n - F_{n-1})^2,$$

which is clearly true after expanding the parentheses on both sides.

Example 5.5. Consider $a_n = \sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\cdot^{\cdot^{\cdot}}}}}$, a tower of n terms equal to $\sqrt{2}$. Prove that a_n is increasing and bounded above by 2.

Solution. We will prove by induction on n that $a_n < 2$ and $a_n < a_{n+1}$.
For $n = 1$, clearly $\sqrt{2} < 2$ and $\sqrt{2}^{\sqrt{2}} > \sqrt{2}$.
Assuming the result for some $n \geq 1$, by the induction hypothesis we have

$$a_{n+1} = \sqrt{2}^{a_n} < \sqrt{2}^2 = 2.$$

Also,

$$a_{n+1} = \sqrt{2}^{a_n} > \sqrt{2}^{a_{n-1}} = a_n.$$

Example 5.6. (**USAMO 1997 shortlist**) Let $a_1 = a_2 = 97$ and

$$a_{n+1} = a_n a_{n-1} + \sqrt{(a_n^2 - 1)(a_{n-1}^2 - 1)},\ n > 1.$$

Prove that

a) $2 + 2a_n$ is a perfect square;

b) $2 + \sqrt{2 + 2a_n}$ is a perfect square.

Solution. The expressions $a^2 - 1$ and $2 + 2a$ ask for the substitution

$$a = \frac{1}{2}\left(b^2 + \frac{1}{b^2}\right).$$

The equality $\frac{1}{2}\left(b^2 + \frac{1}{b^2}\right) = 97$ implies $2 + 2a = \left(b + \frac{1}{b}\right)^2 = 196$, hence

$$b + \frac{1}{b} = 14.$$

Setting $b = c^2$ yields $2 + \sqrt{2+2a} = \left(c + \frac{1}{c}\right)^2 = 16$, thus $c + \frac{1}{c} = 4$. Let $c = 2 + \sqrt{3}$. We will prove by induction that

$$a_n = \frac{1}{2}\left(c^{4F_n} + \frac{1}{c^{4F_n}}\right),\ n \geq 1,$$

where F_n is the n^{th} Fibonacci number.
Indeed, this is true for $n = 1$, $n = 2$ and, assuming that

$$a_k = \frac{1}{2}\left(c^{4F_k} + \frac{1}{c^{4F_k}}\right),\ k \leq n$$

implies

$$\begin{aligned} a_{n+1} &= \frac{1}{4}\left(c^{4F_n} + \frac{1}{c^{4F_n}}\right)\left(c^{4F_{n-1}} + \frac{1}{c^{4F_{n-1}}}\right) \\ &+ \frac{1}{4}\left(c^{4F_n} - \frac{1}{c^{4F_n}}\right)\left(c^{4F_{n-1}} - \frac{1}{c^{4F_{n-1}}}\right) \\ &= \frac{1}{2}\left(c^{4F_{n+1}} + \frac{1}{c^{4F_{n+1}}}\right). \end{aligned}$$

Then

$$2 + 2a_n = 2 + c^{4F_n} + \frac{1}{c^{4F_n}} = \left(c^{2F_n} + \frac{1}{c^{2F_n}}\right)^2$$

and

$$2 + \sqrt{2 + 2a_n} = 2 + c^{2F_n} + \frac{1}{c^{2F_n}} = \left(c^{F_n} + \frac{1}{c^{F_n}}\right)^2.$$

To finish the question, notice that $x_m = c^m + \frac{1}{c^m}$ is an integer for all positive integers m since it satisfies $x_0 = 2$, $x_1 = 4$, and $x_m = 4x_{m-1} - x_{m-2}$ for $m \geq 2$.

Example 5.7. Determine the general term of the sequence $(a_n)_{n\geq 1}$ given by

$$a_1 = 1, \quad a_{n+1} = \frac{1}{16}\left(1 + 4a_n + \sqrt{1 + 24a_n}\right).$$

Solution. For this question, we will use the approach where we first prove that the sequence converges to some value a, and after we substitute a_{n+1} and a_n with a in the recurrence relation, we find the value of a. Clearly the sequence has only positive terms. We prove that the sequence is decreasing. To do this, it suffices to show that $\frac{1}{3} \leq a_n$ for all n, since we then have

$$\begin{aligned} a_n - a_{n+1} &= \frac{1}{16}\left(16a_n - 1 - 4a_n - \sqrt{1+24a_n}\right) \\ &= \frac{1}{16}\left(12a_n - 1 - \sqrt{1+24a_n}\right). \end{aligned}$$

Now $12a_n - 1 - \sqrt{1+24a_n} \geq 0$ is equivalent to $12a_n - 1 \geq \sqrt{1+24a_n}$. Since $a_n \geq \frac{1}{3}$, we can square both sides and we get

$$144a_n^2 - 48a_n \geq 0, \quad \text{which is true, as} \quad a_n \geq \tfrac{1}{3}.$$

We show that $a_n \geq \frac{1}{3}$ by induction. The base case is verified from the hypothesis. For the inductive step, we have

$$a_{n+1} = \frac{1}{16}\left(1 + 4a_n + \sqrt{1+24a_n}\right) \geq \frac{1}{16}\left(1 + 4\frac{1}{3} + \sqrt{1 + 24\frac{1}{3}}\right) = \frac{1}{3}.$$

This completes our induction. So a_n is decreasing and bounded below, hence convergent to some value $a \geq 0$. Substituting a for a_{n+1} and a_n in the recurrence relation we find that $a = 0$ or $a = \frac{1}{3}$. As $a_n \geq \frac{1}{3}$ for all n, we must have that $a = \frac{1}{3}$. Then

$$a_1 - \frac{1}{3} = \frac{1}{2} + \frac{1}{6}, \quad a_2 - \frac{1}{3} = \frac{1}{2^2} + \frac{1}{3 \cdot 2^3},$$

$$a_3 - \frac{1}{3} = \frac{1}{2^3} + \frac{1}{3 \cdot 2^5}, \quad a_4 - \frac{1}{3} = \frac{1}{2^4} + \frac{1}{3 \cdot 2^7}.$$

Now we claim that

$$a_n = \frac{1}{3} + \frac{1}{2^n} + \frac{2}{3 \cdot 4^n}.$$

This claim can be proved explicitly by induction. We have already established the base cases. Assuming it for n, we have

$$a_{n+1} = \frac{1}{16}\left(1 + 4\left(\frac{1}{3} + \frac{1}{2^n} + \frac{2}{3 \cdot 4^n}\right) + \sqrt{1 + 24\left(\frac{1}{3} + \frac{1}{2^n} + \frac{2}{3 \cdot 4^n}\right)}\right).$$

After noting that

$$1 + 24\left(\frac{1}{3} + \frac{1}{2^n} + \frac{2}{3 \cdot 4^n}\right) = \left(3 + \frac{4}{2^n}\right)^2,$$

and expanding and simplifying the terms accordingly, the above expression gives the desired formula for a_{n+1}, completing our induction.

Example 5.8. Consider the sequences $(u_n)_{n\geq 1}$, $(v_n)_{n\geq 1}$ defined by $u_1 = 3$, $v_1 = 2$ and $u_{n+1} = 3u_n + 4v_n$, $v_{n+1} = 2u_n + 3v_n$, $n \geq 1$. Define $x_n = u_n + v_n$, $y_n = u_n + 2v_n$, $n \geq 1$. Prove that $y_n = \left[x_n\sqrt{2}\right]$, for all $n \geq 1$.

Solution. We prove by induction that

$$u_n^2 - 2v_n^2 = 1, \; n \geq 1. \tag{1}$$

For $n = 1$, the claim is true from hypothesis. Assuming that the equality is true for some n, we have

$$u_{n+1}^2 - 2v_{n+1}^2 = (3u_n + 4v_n)^2 - 2(2u_n + 3v_n)^2 = u_n^2 - 2v_n^2 = 1,$$

hence (1) is true for all $n \geq 1$. We now prove that

$$2x_n^2 - y_n^2 = 1, \, n \geq 1. \tag{2}$$

Indeed,

$$2x_n^2 - y_n^2 = 2(u_n + v_n)^2 - (u_n + 2v_n)^2 = u_n^2 - 2v_n^2 = 1,$$

as claimed. It follows that

$$\left(x_n\sqrt{2} - y_n\right)\left(x_n\sqrt{2} + y_n\right) = 1, \, n \geq 1.$$

Notice that $x_n\sqrt{2} + y_n > 1$, so

$$0 < x_n\sqrt{2} - y_n < 1, \, n \geq 1.$$

Hence $y_n = \left[x_n\sqrt{2}\right]$, as claimed.

Example 5.9. (**IMO 2013 shortlist**) Let n be a positive integer, and consider a sequence $a_1, a_2, \ldots, a_n$ of positive integers. Extend it periodically to an infinite sequence $a_1, a_2, \ldots$ by defining $a_{n+i} = a_i$ for all $i \geq 1$. If

$$a_1 \leq a_2 \leq \ldots \leq a_n \leq a_1 + n \tag{1}$$

and

$$a_{a_i} \leq n + i - 1 \quad \text{for} \quad i = 1, 2, \ldots, n, \tag{2}$$

prove that

$$a_1 + \ldots + a_n \leq n^2.$$

Solution. We provide a solution to this problem by induction on n:

We begin by verifying the initial case $n = 1$. For this we get $a_{a_1} \leq 1$, and since for $n = 1$, the sequence is constant, this implies $a_1 = a_{a_1} \leq 1$, so the result follows.

For the general case we first rule out the simple cases when $a_1 = 1$ (which implies $a_2, \ldots, a_n \leq n + 1$), $a_1 = n$ (which implies $a_n \leq n$), or $a_n \leq n$, as for these the inequality holds trivially. We also see that if $a_>n$, then we get $n < a_1 \leq a_{a_1} \leq n$, a contradiction. So we may assume $a_1 < n < a_n$.

Now consider the number t with $1 \leq t \leq n - 1$ such that

$$a_1 \leq a_2 \leq \ldots \leq a_t \leq n < a_{t+1} \leq \ldots \leq a_n.$$

Since $1 \leq a_1 \leq n$ and $a_{a_1} \leq n$ by (2), we have $a_1 \leq t$.

Define the sequence $d_1, \ldots, d_{n-1}$ by

$$d_i = \begin{cases} a_{i+1} - 1 & \text{if } i \leq t - 1; \\ a_{i+1} - 2 & \text{if } i \geq t, \end{cases}$$

and extend it periodically modulo $n - 1$. One may verify that this sequence also satisfies the hypotheses of the problem. The induction hypothesis then gives

$$d_1 + \ldots + d_{n-1} \leq (n-1)^2,$$

which implies that

$$\sum_{i=1}^{n} a_i = a_1 + \sum_{i=2}^{t}(d_{i-1}+1) + \sum_{i=t+1}^{n}(d_{i-1}+2) \leq t + (t-1) + 2(n-t) + (n-1)^2 = n^2.$$

This completes our induction and hence also the proof of the problem.

Example 5.10. (**Russia 2000**) The sequence $a_1, a_2, \ldots, a_{2000}$ of real numbers satisfies the condition

$$a_1^3 + a_2^3 + \ldots + a_n^3 = (a_1 + a_2 + \ldots + a_n)^2, \quad \text{for all} \quad 1 \leq n \leq 2000.$$

Prove that every element of the sequence is an integer.

Solution. We prove by induction on n both that a_n is an integer and that

$$a_1 + a_2 + \ldots + a_n = \frac{N_n(N_n + 1)}{2}$$

for a non-negative integer N_n. We extend this sum to the case $n = 0$, for which we use $N_0 = 0$, and this will be the base case of our induction.

Assume now that the claim holds for some $n \geq 0$ and we prove it for $n+1$. We are given that

$$\left(\sum_{i=1}^{n+1} a_i\right)^2 = \sum_{i=1}^{n+1} a_i^3,$$

which is equivalent to

$$\left(\frac{N_n(N_n + 1)}{2} + a_{n+1}\right)^2 = \left(\frac{N_n(N_n + 1)}{2}\right)^2 + a_{n+1}^3.$$

By expanding and factoring, the above equality becomes

$$a_{n+1}(a_{n+1} - (N_n + 1))(a_{n+1} + N_n) = 0.$$

Thus $a_{n+1} \in \{0, N_n + 1, -N_n\}$, so that a_{n+1} is an integer.
To finish the induction, we need to determine N_{n+1}. If $a_{n+1} = 0$, then we may set $N_{n+1} = N_n$. If $a_{n+1} = N_n + 1$, then

$$\sum_{i=1}^{n} a_i + a_{n+1} = \frac{N_n(N_n + 1)}{2} + (N_n + 1) = \frac{(N_n + 2)(N_n + 1)}{2},$$

so we may set $N_{n+1} = N_n + 1$. Finally, if $a_{n+1} = -N_n$, then

$$\sum_{i=1}^{n} a_i + a_{n+1} = \frac{N_n(N_n + 1)}{2} - N_n = \frac{N_n(N_n - 1)}{2},$$

so we set $N_{n+1} = N_n - 1$. This completes the induction step and the proof to our problem.

5.2 Proposed Problems

Problem 5.1. The sequence $(a_n)_{n\geq 1}$ is defined by $a_1 = 2$ and

$$a_{n+2} = \frac{2+a_n}{1-2a_n}, \ n \geq 1.$$

Prove that all its terms are nonzero.

Problem 5.2. Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ be sequences defined by

$$a_{n+3} = a_{n+2} + 2a_{n+1} + a_n, \ n = 0, 1, \ldots, \ a_0 = 1, \ a_1 = 2, \ a_2 = 3$$

and

$$b_{n+3} = b_{n+2} + 2b_{n+1} + b_n, \ n = 0, 1, \ldots, \ b_0 = 3, \ b_1 = 2, \ b_2 = 1.$$

How many integers do the sequences have in common?

Problem 5.3. (**India 1996**) Define a sequence $(a_n)_{n\geq 1}$ by $a_1 = 1$ and $a_2 = 2$ and $a_{n+2} = 2a_{n+1} - a_n + 2$ for $n \geq 1$. Prove that for any $m \geq 1$, $a_m a_{m+1}$ is also a term in this sequence.

Problem 5.4. (**Russia 2000**) Let $a_1, a_2, \ldots, a_n$ be a sequence of non-negative real numbers, not all zero. For $1 \leq k \leq n$, let

$$m_k = \max_{1\leq i\leq k} \frac{a_{k-i+1} + a_{k-i+2} + \ldots + a_k}{i}.$$

Prove that for any $\alpha > 0$, the number of integers k which satisfy $m_k > \alpha$ is less than $\frac{a_1+a_2+\ldots+a_n}{\alpha}$.

Problem 5.5. (**USAMO 2003**) Let $n \neq 0$. For every sequence of integers

$$A = a_0, a_1, a_2, \ldots, a_n \quad \text{satisfying} \quad 0 \leq a_i \leq i, \quad \text{for} \quad i = 0, \ldots, n,$$

define another sequence

$$t(A) = t(a_0), t(a_1), t(a_2), \ldots, t(a_n)$$

by setting $t(a_i)$ to be the number of terms in the sequence A that precede the term a_i and are different from a_i. Show that, starting from any sequence A as above, fewer than n applications of the transformation t lead to a sequence B such that $t(B) = B$.

Problem 5.6. (**Russia 2000**) Let $a_1, a_2, \ldots$ be a sequence with $a_1 = 1$ satisfying the recursion

$$a_{n+1} = \begin{cases} a_n - 2 & \text{if } a_n - 2 \notin \{a_1, a_2, \ldots, a_n\} \quad \text{and} \quad a_n - 2 > 0; \\ a_n + 3 & \text{otherwise.} \end{cases}$$

Prove that for every positive integer $k > 1$, we have $a_n = k^2 = a_{n-1} + 3$ for some n.

Problem 5.7. Let $(x_n)_{n\geq 1}$ be defined by the relations $x_1 = 1$, and

$$x_{n+1} = \frac{x_n}{n} + \frac{n}{x_n}, \ n \geq 1.$$

Prove that $\lfloor x_n^2 \rfloor = n$, for all $n \geq 4$.

Problem 5.8. (**Russia 2000**) For any odd integer $a_0 > 5$, consider the sequence $a_0, a_1, a_2, \ldots$, where

$$a_{n+1} = \begin{cases} a_n^2 - 5 & \text{if } a_n \quad \text{is odd,} \\ \frac{a_n}{2} & \text{if } a_n \quad \text{is even,} \end{cases}$$

for all $n \geq 0$. Prove that this sequence is not bounded.

Problem 5.9. (**China 1997**) Let $(a_n)_{n\geq 1}$ be a sequence of non-negative real numbers satisfying $a_{n+m} \leq a_n + a_m$ for all positive integers m, n. Prove that if $n \geq m$, then

$$a_n \leq m a_1 + \left(\frac{n}{m} - 1\right) a_m.$$

Problem 5.10. Let $n \geq 2$ be an integer. Show that there exist $n+1$ numbers $x_1, x_2, \ldots, x_{n+1} \in \mathbb{Q} \setminus \mathbb{Z}$, so that $\{x_1^3\} + \{x_2^3\} + \ldots + \{x_n^3\} = \{x_{n+1}^3\}$, where $\{x\}$ is the fractional part of x.

Problem 5.11. Let $(a_n)_{n\geq 1}$ be a sequence of real numbers such that

$$a_1 = a_2 = 1 \quad \text{and} \quad a_{n+2} = a_{n+1} + \frac{a_n}{3^n}, \quad \text{for} \quad n \geq 1.$$

Prove that $a_n \leq 2$ for any $n \geq 1$.

Problem 5.12. (**IMO 1995**) The positive real numbers $x_0, x_1, \ldots, x_{1995}$ satisfy $x_0 = x_{1995}$ and

$$x_{i-1} + \frac{2}{x_{i-1}} = 2x_i + \frac{1}{x_i},$$

for $i = 1, 2, \ldots, 1995$. Find the maximum value that x_0 can have.

Problem 5.13. (**INMO 2010**) Define a sequence $(a_n)_{n\geq 0}$ by $a_0 = 0$, $a_1 = 1$ and

$$a_n = 2a_{n-1} + a_{n-2}, \quad \text{for} \quad n \geq 2.$$

a) For every $m > 0$ and $0 \leq j \leq m$, prove that $2a_m$ divides $a_{m+j} + (-1)^j a_{m-j}$.
b) Suppose that 2^k divides n for some natural numbers n and k. Prove that 2^k divides a_n.

Problem 5.14. (**IMO 2013 shortlist**) Let n be a positive integer and let $a_1, \ldots, a_{n-1}$ be arbitrary real numbers. Define the sequences $u_0, \ldots, u_n$ and $v_0, \ldots, v_n$ inductively by $u_0 = u_1 = v_0 = v_1 = 1$, and

$$u_{k+1} = u_k + a_k u_{k-1}, \; v_{k+1} = v_k + a_{n-k} v_{k-1} \text{ for } k = 1, \ldots, n-1.$$

Prove that $u_n = v_n$.

Problem 5.15. (**IMO 2006 shortlist**) The sequence of real numbers $a_0, a_1, a_2, \ldots$ is defined recursively by

$$a_0 = -1, \quad \sum_{k=0}^{n} \frac{a_{n-k}}{k+1} = 0, \quad \text{for} \quad n \geq 1.$$

Show that $a_n > 0$ for $n \geq 1$.

Problem 5.16. Let $(a_n)_{n\geq 0}$ be a sequence defined by

$$a_0 = a_1 = 47 \quad \text{and} \quad 2a_{n+1} = a_n a_{n-1} + \sqrt{(a_n^2 - 4)(a_{n-1}^2 - 4)}.$$

Prove that $a_n + 2$ is a perfect square for any $n \geq 0$.

Problem 5.17. Let $a_0, a_1, a_2, \ldots$ be an increasing sequence of non-negative integers such that every non-negative integer can be expressed uniquely in the form $a_i + 2a_j + 4a_k$, where i, j and k are not necessarily distinct. Determine a_{1998}.

Problem 5.18. A sequence of integers $a_1, a_2, a_3, \ldots$ is defined as follows: $a_1 = 1$ and for $n \geq 1$, a_{n+1} is the least integer greater than a_n such that $a_i + a_j \neq 3a_k$ for any i, j and k in $\{1, 2, 3, \ldots, n+1\}$, not necessarily distinct. Determine a_{1998}.

Problem 5.19. Let k be a positive integer. The sequence $a_1, a_2, a_3, \ldots$ is defined by $a_1 = k+1$, and $a_{n+1} = a_n^2 - ka_n + k$, $n \geq 1$. Prove that a_m and a_n are coprime (for $m \neq n$).

Problem 5.20. (**Bulgaria TST 2011**) Let $(x_n)_{n\geq 1}$ be a sequence defined by $x_1 = \frac{2}{3}$ and

$$x_{n+1} = \frac{3x_n + 2}{3 - 2x_n}, \quad \text{for all} \quad n \geq 1.$$

Is this sequence eventually periodic?

Problem 5.21. (**St. Petersburg**) Let $(x_n)_{n\geq 1}$ and $(y_n)_{n\geq 1}$ be two sequences given by $x_1 = \frac{1}{10}$, $y_1 = \frac{1}{8}$ and

$$x_{n+1} = x_n + x_n^2, \quad y_{n+1} = y_n + y_n^2 \quad \text{for} \quad n \geq 1.$$

Prove that for any positive integers m and n, we cannot have $x_n = y_m$.

Problem 5.22. (**Taiwan 2000**) Let $f : \mathbb{N}^* \to \mathbb{N}$ be defined recursively by $f(1) = 0$ and

$$f(n) = \max_{1\leq j\leq \lfloor \frac{n}{2} \rfloor} \{f(j) + f(n-j) + j\},$$

for all $n \geq 2$. Determine $f(2000)$.

Problem 5.23. (**Taiwan 1997**) Let $n > 2$ be an integer. Suppose that $a_1, a_2, \ldots, a_n$ are positive real numbers such that

$$k_i = \frac{a_{i-1} + a_{i+1}}{a_i}$$

is a positive integer for all i (here $a_0 = a_n$ and $a_{n+1} = a_1$). Prove that

$$2n \leq k_1 + k_2 + \ldots + k_n \leq 3n.$$

Problem 5.24. (**Zeckendorf**) Prove that any positive integer N can be represented uniquely as a sum of distinct and non-consecutive terms of the Fibonacci sequence:

$$N = \sum_{j=1}^{m} F_{i_j}, \quad i_j - i_{j-1} \geq 2.$$

Problem 5.25. Let $a > 1$ be a real number which is not an integer. Prove that the sequence $(a_n)_{n\geq 0}$ defined by

$$a_n = [a^{n+1}] - a[a^n]$$

is not periodic. Here, $[x]$ denotes the integer part of x.

Problem 5.26. (**China 2004**) For a given real number a and a positive integer n, prove that:

i) There exists exactly one sequence of real numbers $x_0, x_1, \ldots, x_n, x_{n+1}$ such that

$$\begin{cases} x_0 = x_{n+1} = 0, \\ \frac{1}{2}(x_i + x_{i+1}) = x_i + x_i^3 - a^3, \ i = 1, 2, \ldots, n. \end{cases}$$

ii) the sequence $x_0, x_1, \ldots, x_n, x_{n+1}$ in i) satisfies $|x_i| \leq |a|$ where $i = 0, 1, \ldots, n+1$.

Problem 5.27. (**IMO 2010 Shortlist**) A sequence $x_1, x_2, \ldots$ is defined by

$$x_1 = 1 \quad \text{and} \quad x_{2k} = -x_k, \quad x_{2k-1} = (-1)^{k+1} x_k, \quad \text{for all} \quad k \geq 1.$$

Prove that for all $n \geq 1$, we have

$$x_1 + x_2 + \ldots + x_n \geq 0.$$

Problem 5.28. Let a_n be the number of strings of length n which contain only the digits 0 and 1 and such that no two 1's can be distance two apart. Find a formula for a_n in closed form.

Problem 5.29. (**IMO 2009 shortlist**) Let n be a positive integer. Given a sequence $\varepsilon_1, \ldots, \varepsilon_{n-1}$ with $\varepsilon_i = 0$ or $\varepsilon_i = 1$ for each $i = 1, \ldots, n-1$, the sequences $a_0, \ldots, a_n$ and $b_0, \ldots, b_n$ are constructed by the following rules:

$$a_0 = b_0 = 1, \quad a_1 = b_1 = 7,$$

$$a_{i+1} = \begin{cases} 2a_{i-1} + 3a_i, & \text{if } \varepsilon_i = 0, \\ 3a_{i-1} + a_i, & \text{if } \varepsilon_i = 1, \end{cases} \qquad \text{for each } i = 1, \dots, n-1,$$

$$b_{i+1} = \begin{cases} 2b_{i-1} + 3b_i, & \text{if } \varepsilon_{n-i} = 0, \\ 3b_{i-1} + b_i, & \text{if } \varepsilon_{n-i} = 1, \end{cases} \qquad \text{for each } i = 1, \dots, n-1.$$

Prove that $a_n = b_n$.

Problem 5.30. (**IMO 2008 shortlist**) Let a_0, a_1, a_2, ... be a sequence of positive integers such that the greatest common divisor of any two consecutive terms is greater than the preceding term; in symbols, $\gcd(a_i, a_{i+1}) > a_{i-1}$. Prove that $a_n \geq 2^n$ for all $n \geq 0$.

Chapter 6

Number Theory

6.1 Theory and Examples

We start by giving a comprehensive review some of the basic notions and results which we shall employ in this chapter:

Given $a, b \in \mathbb{Z}$, $a \neq 0$, we say a **divides** b, a is a **factor** of b or $a \mid b$ if there exists $c \in \mathbb{Z}$ such that $b = ac$. For any b, ± 1 and $\pm b$ are always factors of b. The other factors are called **proper factors**.

The **Division Algorithm** is a theorem which asserts that given two integers a and b with $b \neq 0$, there exist unique integers q and r such that

$$a = qb + r, \quad \text{and} \quad 0 \leq r < |b|.$$

A **common factor** of two integers a and b is an integer $c \in \mathbb{Z}$ such that $c \mid a$ and $c \mid b$. The **highest common factor** or **greatest common divisor** (gcd for short) of two numbers $a, b \in \mathbb{N}$ is a number $d \in \mathbb{N}$ such that d is a common factor of a and b, and if c is also a common factor, $c \mid d$. We will denote the gcd of a and b by $\gcd(a, b)$ or simply (a, b). It exists and it is uniquely determined by a and b. If $d = \gcd(a, b)$, then $d \mid (ua + vb)$, for all $u, v \in \mathbb{Z}$ (in fact d is the smallest positive linear combination of a and b). **Bézout's theorem** says that if $a, b \in \mathbb{N}$ and $c \in \mathbb{Z}$, then there exist $u, v \in \mathbb{Z}$ with $c = ua + vb$ if and only if $(a, b) \mid c$. **Euclid's Algorithm** provides an explicit procedure for computing $\gcd(a,b)$, namely if we continuously break down a and b by the following procedure (stopping when we get a zero remainder):

$$\begin{aligned} a &= q_1 b + r_1 \\ b &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\vdots \\ r_{n-2} &= q_n r_{n-1}, \end{aligned}$$

then the highest common factor is r_{n-1}. We say a and b are **coprime** if $\gcd(a,b) = 1$.

A positive integer p is prime if $p > 1$ and the only factors of p are ± 1 and $\pm p$. The **Fundamental Theorem of Arithmetic** (which we will prove shortly) says that every positive integer is expressible as product of primes in exactly one way. Euclid proved that there are infinitely many primes and **Chebyshev's Theorem** says that for any $n > 3$, there is always a prime between n and $2n-2$. For a positive integer a and a prime p, we will frequently use the notation $\nu_p(a)$ to denote the power of p in the prime factorization of a (e.g. $\nu_3(18) = 2$, $\nu_2(11) = 0$). We can extend this notation to rational numbers by defining $\nu_p\left(\frac{m}{n}\right) = \nu_p(m) - \nu_p(n)$.

For two integers a and b and a positive integer n we write $a \equiv b \pmod n$ to mean $n \mid (a-b)$. Standard properties such as $a \equiv b \pmod n$ and $c \equiv d \pmod n$ imply $ac \equiv bd \pmod n$ can be proved from the definition.

Chinese Remainder Theorem. Let $(m,n) = 1$ and $a, b \in \mathbb{Z}$. Then there is a unique solution (modulo mn) to the simultaneous congruences

$$\begin{cases} x \equiv a \pmod m \\ x \equiv b \pmod n \end{cases},$$

i.e. there exists x satisfying both and every other solution is $\equiv x \pmod{mn}$.

Fermat's Little Theorem. If a is a positive integer and p is a prime, then $a^p \equiv a \pmod p$.

We will prove this theorem in the examples below, too.

For a given positive integer n, we let $\phi(n)$ denote the cardinality of the set $\{a \in \mathbb{N}^* : a \leq n \quad \text{and} \quad (a,n) = 1\}$ (i.e. the number of positive integers less than or equal to n and coprime to n).

Euler's Theorem. If a and n are positive integers and $(a,n) = 1$, then $a^{\phi(n)} \equiv 1 \pmod{n}$. This generalizes Fermat's Little Theorem from above.

For two positive integers a and n with $(a,n) = 1$, the **order** of a modulo n is the smallest positive integer k such that $a^k \equiv 1 \pmod{n}$. Notice that Euler's Theorem proves that the order always exists and that it divides $\phi(n)$.

Given a positive integer n, an integer a is called a **quadratic residue** modulo n, if there exists an integer x such that $x^2 \equiv a \pmod{n}$. Otherwise, a is called a **quadratic non-residue**. For a prime number p, and an integer a one defines the Legendre symbol by

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p; \\ -1 & \text{if } a \text{ is a quadratic non-residue modulo } p; \\ 0 & \text{if } (a,p) > 1. \end{cases}$$

Standard properties of the Legendre symbol follow from the properties of moduli. For more details on this, as well as the applications of the Legendre symbol in other areas of Number Theory, we invite the reader to consult any standard texts on Elementary Number Theory.

From the Division Algorithm, we know that for each positive integer n and any integer a, there exists a unique integer $0 \leq b \leq n-1$ such that $a \equiv b \pmod{n}$. For a given positive integer n, the relation $a \sim b$ given by $a \sim b$ if $a \equiv b \pmod{n}$ is an equivalence relation. By the observation we have made, we can choose the equivalence class representatives to be the numbers $0, 1, \ldots, n-1$. We let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} = \{0, 1, \ldots, n-1\}$ be the set of these classes. If for $a \in \mathbb{Z}$ we denote by $[a] \in \mathbb{Z}_n$ its corresponding equivalence class, we can define an additive operation on $\mathbb{Z}_n$ in the natural way as $c + d = [c + d]$, for any $c, d \in \mathbb{Z}_n$ (so for example $2 + 3 = 1$ in $\mathbb{Z}_4$). In a similar manner, if we look at the set $\mathbb{Z}_n^* = (\mathbb{Z}/n\mathbb{Z})^* = \{0 < a < n : \gcd(a,n) = 1\}$ instead, we can define a multiplicative operation on $\mathbb{Z}_n^*$ via $c \cdot d = [c \cdot d]$. In general, we will be concerned with the case when $n = p$ is a prime, so that $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$, $\mathbb{Z}_p^* = \{1, 2, 3, \ldots, p-1\}$, and the operation defined on $\mathbb{Z}_p$ is addition, while on $\mathbb{Z}_p^*$ it is the multiplication, as described before.

6.1.1 The p and $\left\lfloor \frac{n}{p} \right\rfloor$ Technique

A particularly nice variant of induction which we have not met yet and is specific to Number Theory is the following: rather than proving $P(0)$ and $P(n) \Rightarrow P(n+1)$, we prove that $P(n)$ is true by induction on the number of (prime) factors of n (not necessarily distinct). One of the most famous applications of this idea is:

Example 6.1. (**Fundamental Theorem of Arithmetic**) Prove that every natural number $n \geq 2$ is expressible as a product of primes in exactly one way. In particular, if

$$p_1 p_2 \dots p_k = q_1 q_2 \dots q_l,$$

where p_i, q_i are primes, but not necessarily distinct, then $k = l$ and $q_1, \dots, q_l$ are $p_1, \dots, p_k$ in some order.

Solution. For the proof, we will need the following result, which follows from Bézout's theorem and we leave it as an exercise for the reader: if p is a prime number and $p \mid ab$ then $p \mid a$ or $p \mid b$.

Let us now prove the original question. Assume the contrary. Then (by the well-ordering principle) there exists a minimal n that cannot be written as a product of primes.

If n is a prime, then n is a product of primes. Otherwise, write $n = ab$, where $1 < a, b < n$. By minimality of n, both a and b are products of primes. Hence so is n. Contradiction.

For the second part, we do induction on $k + l$, where k and l are the number of prime factors in two representations of n, as in the hypothesis. When $k + l = 2$, we must have $k = 1$, $l = 1$ (otherwise we get that product of two primes is equal to 1, which is a contradiction). But then we have $p_1 = q_1$, which is what we wanted.

Now let $p_1 \cdots p_k = q_1 \cdots q_l$. We know that $p_1 | q_1 \cdots q_l$. Then $p_1 | q_1(q_2 q_3 \cdots q_l)$. Thus $p_1 | q_i$ for some i. Without loss of generality, assume $i = 1$. Then $p_1 = q_1$ since both are primes. Thus $p_2 p_3 \cdots p_k = q_2 q_3 \cdots q_l$. From the induction hypothesis for $l + k - 2$, we are done.

Example 6.2. (**OMM 2002**) Let n be a positive integer. Does n^2 have more positive divisors of the form $4k + 1$ or of the form $4k - 1$?

Solution. Denote by $M(n)$ the number of positive divisors of n^2 of the form $4k-1$ and by $P(n)$ the number of positive divisors of n^2 of the form $4k+1$.

We prove by strong induction that $M(n) < P(n)$. For $n = 2^k$ we have $M(2^k) = 0 < 1 = P(2^k)$.

Let p be a prime such that $p \nmid n$, $p \equiv 1 \pmod 4$. Then every divisor of $(p^\alpha n)^2 = p^{2\alpha} n^2$ of the form $4k-1$ is a power of p times a divisor of n^2 of this form and similarly for divisors of the form $4k+1$. Hence $M(p^\alpha n) = (2\alpha+1)M(n)$ and $P(p^\alpha n) = (2\alpha+1)P(n)$, so $M(n) < P(n)$ forces $M(p^\alpha n) < P(p^\alpha n)$.

Now consider a prime p such that $p \nmid n$, $p \equiv -1 \pmod 4$. Then every divisor of $(p^\alpha n)^2 = p^{2\alpha} n^2$ of the form $4k-1$ is either p^{2r} times a divisor of n^2 of the form $4k-1$ (for some $0 \le r \le \alpha$) or p^{2r-1} times a divisor of n^2 of the form $4k+1$ (for some $1 \le r \le \alpha$), and the reverse for divisors of the form $4k+1$. Thus $M(p^\alpha n) = (\alpha+1)M(n) + \alpha P(n)$ and $P(p^\alpha n) = (\alpha+1)P(n) + \alpha M(n)$, so $M(n) < P(n)$ again forces $M(p^\alpha n) < P(p^\alpha n)$.

An example where the number of factors of N that we induct on do not have to be necessarily prime is the following:

Example 6.3. (**St. Petersburg**) The number N equals the product of 200 different positive integers. Prove that N has at least 19901 different divisors (including 1 and the number itself).

Solution. We show by induction that if N equals the product of k distinct positive integers, then N has at least $\frac{k(k-1)}{2} + 1$ distinct divisors. The base case $N = 1$ is trivial.

For the induction step, assume $M = a_1 \cdot \ldots \cdot a_{k+1} = N \cdot a_{k+1}$. We can assume without loss of generality that a_{k+1} is the largest among the a_i's (otherwise, we simply reorder them). By the induction hypothesis, the number $N = \frac{M}{a_{k+1}}$ has at least $\frac{k(k-1)}{2} + 1$ distinct divisors. Note that $\frac{N}{a_i} \cdot a_{k+1}$ are all divisors of M, which are bigger than N, since $a_{k+1} > a_i$ for $i = 1, \ldots, k$. So we have at least k new divisors for a total of at least $\frac{k(k-1)}{2} + 1 + k = \frac{(k+1)k}{2} + 1$ distinct divisors.

We will now prove a beautiful result, called the Erdös-Ginzburg-Ziv theorem. For this, we need the following auxiliary result, which also uses induction:

Example 6.4. (**Cauchy-Davenport**) Let $p \geq 3$ be a prime number. Then for any two non-empty subsets A and B of $\mathbb{Z}/p\mathbb{Z}$, we have

$$|A+B| \geq \min(|A|+|B|-1, p),$$

where $A+B=\{a+b: \ a \in A, b \in B\}$.

Solution. We will prove the statement using strong induction on $|A|$. For the base case, note that when $|A|=1$ we have $|A+B|=|B|$ (since $A+B$ is just a translation of B) and thus there is nothing to prove.

For the induction step, let A be a set with $|A|>1$. Because $|A|$ has at least two elements, we can assume without loss of generality that A contains 0, by translating the set if necessary (notice that this operation does not affect the statement of the problem). Let x be one of the other non-zero elements. We can also assume that $1 \leq |B| < p$, as otherwise there is not much to prove either. Thus, there is an element n such that $nx \in B$ but $(n+1)x \notin B$. Translating by $-nx$ in B, we find that $0 \in B$ and $x \notin B$.

Now note that $A \cup B + A \cap B \subseteq A+B$. The key observation here is that $A \cap B$ contains 0 and is not the whole of A because $x \notin B$. So we can use the induction hypothesis to obtain

$$|A+B| \geq |A \cup B + A \cap B| \geq \min(|A \cap B| + |A \cup B| - 1, p).$$

Using the well-known identity $|A \cap B| + |A \cup B| = |A| + |B|$, we obtain the result.

Example 6.5. (**Erdös-Ginzburg-Ziv**) Prove that from any $2n-1$ numbers one can choose n whose sum is divisible by n.

Solution. This problem uses induction on the number of prime divisors of n.

Firstly, let us prove the simpler part of the question, which is the induction step. We prove that if the problem holds for $n=a$ and $n=b$, then it also holds for $n=ab$:

Consider $2ab-1$ numbers. We select $2a-1$ of them and from the induction hypothesis, from these $2a-1$ numbers we can select a group of a numbers whose sum is divisible by a. Discarding this set of a elements, we have $(2b-1)a-1$ numbers left. From these, we choose $2a-1$ more numbers producing one more group of a numbers whose sum is divisible by a. In general, if we have selected $k < 2b-1$ groups of a numbers whose sum is divisible by a, we have $(2b-k)a > 2a-1$ numbers left, so we can select one more such group. Thus, at the end we will have $2b-1$ groups of a numbers whose sum is divisible by a. Let these groups be $G_1, G_2, \ldots, G_{2b-1}$ with the corresponding sums $aS_1, aS_2, \ldots, aS_{2b-1}$. From the induction hypothesis on b, we can find $1 \leq i_1 \leq \ldots \leq i_b \leq 2b-1$ such that $S_{i_1} + S_{i_2} + \ldots + S_{i_b}$ is divisible by b. Therefore the ab numbers from the groups $G_{i_1}, G_{i_2}, \ldots, G_{i_b}$ have the total sum divisible by ab, which is what we needed.

We are now left to prove the base case, namely that for a prime p, given any $2p-1$ numbers one can choose p out of them whose sum is divisible by p. Since for our sums we are only interested in the residues modulo p, we may assume without loss of generality that the $2p-1$ numbers are in the set $\{0, 1, \ldots, p-1\}$. Let us denote these $2p-1$ numbers by $a_1 \leq a_2 \leq \ldots \leq a_{2p-1}$. If $a_i = a_{i+p-1}$ for some $i \leq p-1$, then we must have

$$a_i + a_{i+1} + \ldots + a_{i+p-1} = pa_i \equiv 0 \pmod{p}.$$

Otherwise, let us define $A_i = \{a_i, a_{i+p-1}\}$, for $1 \leq i \leq p-1$. By repeated application of the Cauchy-Davenport Lemma, we conclude that

$$|A_1 + A_2 + \ldots + A_{p-1}| = p,$$

and hence every element of $\mathbb{Z}_p$ is a sum of precisely $p-1$ of the first $2p-2$ elements of our sequence. In particular, $-a_{2p-1}$ is such a sum, which gives us the required p-subset whose sum is 0 in $\mathbb{Z}_p$.

6.1.2 Divisibility

Other typical applications of Induction in Number Theory are the divisibility problems. We will look at a few ideas that arise in this type of questions. We begin with an easy example which involves only algebraic manipulations:

Example 6.6. Prove that for all $n \in \mathbb{N}$ we have that $17 \mid 2^{5n+3} + 5^n \cdot 3^{n+2}$.

Solution. Let $P(n): \quad 17 \mid 2^{5n+3} + 5^n \cdot 3^{n+2}$.

We have $2^3 + 5^0 \cdot 3^2 = 17$, so $P(0)$ is true.

Assume now that $P(n)$ is true for some $n \geq 0$.

We must show that $17 \mid (2^{5n+8} + 5^{n+1} \cdot 3^{n+3})$. We perform the following algebraic manipulations:

$$2^{5n+8} = 2^{5n+3} \cdot 2^5 = 2^{5n+3}(34 - 2) = 2^{5n+3} \cdot 34 - 2^{5n+3} \cdot 2,$$
$$5^{n+1} \cdot 3^{n+3} = 5^n \cdot 3^{n+2} \cdot 15 = 5^n \cdot 3^{n+2} \cdot 17 - 5^n \cdot 3^{n+2} \cdot 2.$$

Therefore, in order to prove $P(n+1)$, it suffices to prove that

$$17 \mid 2^{5n+3} \cdot 2 + 5^n \cdot 3^{n+2} \cdot 2,$$

which follows from the fact that $P(n)$ was assumed to be true. By the Weak Principle of Induction, we have the conclusion.

A rather nice application of Induction concerning divisibility problems is the following:

Example 6.7. (Fermat's Little Theorem). Let p be a prime number. Then p divides $n^p - n$ for all positive integers n.

Solution. We prove the result by induction on n. The base case $n = 1$ is true, as $p \mid (1^p - 1) = 0$.

Assume now that the result holds for some $n \geq 1$, i.e. $p \mid (n^p - n)$. Then

$$\begin{aligned}(n+1)^p - (n+1) &= n^p + \binom{p}{1}n^{p-1} + \ldots + \binom{p}{p-1}n + 1 - (n+1) \\ &= n^p - n + \sum_{i=1}^{p-1}\binom{p}{i}n^{p-i}.\end{aligned}$$

Now $\binom{p}{i} = \dfrac{p!}{(p-i)!i!}$ and for $1 \leq i \leq p-1$, p cannot divide $i!$ or $(p-i)!$. Hence we conclude that $\binom{p}{i}$ is divisible by p for $0 < i < p$. Now $n^p - n$ is divisible by p from the induction hypothesis, so $p \mid ((n+1)^p - (n+1))$, proving the result.

Let us look at some more examples:

Example 6.8. (**Kvant M2277**) Prove that the sum

$$S = 1^n + 2^n + \ldots + (2^k - 1)^n$$

is divisible by 2^{k-1} for all positive integers n and k.

Solution. If n is odd, then we can rewrite the sum as

$$S = (1^n+(2^k-1)^n)+(2^n+(2^k-2)^n)+\ldots+((2^{k-1}-1)^n+(2^{k-1}+1)^n)+2^{n(k-1)},$$

and we can clearly see that S is divisible by 2^{k-1} in this case.

If n is even, we prove the result by induction on k, with the base case $k = 1$ being clear. Assuming the result for some $k \geq 1$, we know that $a^n \equiv (2^{k+1}-a)^n \pmod{2^k}$, hence

$$1^n + \ldots + (2^{k+1} - 1)^n = (1^n + (2^{k+1} - 1)^n) + \ldots$$

$$+((2^k - 1)^n + (2^k + 1)^n) + 2^{nk} \equiv 2(1^n + 2^n + \ldots + (2^k - 1)^n) + 2^{kn} \pmod{2^k}.$$

From the induction hypothesis we know that $1^n+2^n+\ldots+(2^k-1)^n$ is divisible by 2^{k-1}, thus $1^n + 2^n + \ldots + (2^{k+1} - 1)^n$ must be divisible by 2^k, completing our proof.

The following more challenging problem was set by Estonia and got shortlisted for the 47-th IMO:

Example 6.9. (**IMO 2006 shortlist**) For all positive integers n, show that there exists a positive integer m such that n divides $2^m + m$.

Solution. We will prove by induction on $d \geq 1$ that, for every positive integer N, there exist positive integers $b_0, b_1, \ldots, b_{d-1}$ such that, for each $i = 0, 1, 2, \ldots, d-1$ we have $b_i > N$ and

$$2^{b_i} + b_i \equiv i \pmod{d}.$$

Notice that this yields the claim of our question by taking $m = b_0$.

The base case $d = 1$ is trivial. For the induction step, take an $a > 1$ and assume that the statement holds for all $d < a$. Note that as i ranges over $\mathbb{N}$, the remainders of 2^i modulo a repeat periodically starting with some exponent

M. Let k be the length of the period; this means that $2^{M+k'} \equiv 2^M \pmod{a}$ holds only for those k' which are multiples of k. Note further that the period cannot contain all the a remainders, since 0 is either missing or it is the only number in the period. Thus $k < a$.

Let $d = \gcd(a, k)$ and let $a' = a/d$, $k' = k/d$. Since $0 < k < a$, we also have $0 < d < a$. By the induction hypothesis, there exist positive integers $b_0, b_1, \ldots, b_{d-1}$ such that $b_i > \max\{2^M, N\}$ and

$$2^{b_i} + b_i \equiv i \pmod{d}, \; i = 0, 1, 2, \ldots, d-1. \tag{1}$$

For each $i = 0, 1, \ldots, d-1$ consider the sequence

$$2^{b_i} + b_i, 2^{b_i+k} + (b_i + k), \ldots, 2^{b_i+(a'-1)k} + (b_i + (a'-1)k). \tag{2}$$

Modulo a, these numbers are congruent to

$$2^{b_i} + b_i, 2^{b_i} + (b_i + k), \ldots, 2^{b_i} + (b_i + (a'-1)k),$$

respectively. The d sequences contain $a'd = a$ numbers altogether. We will now prove that no two of these numbers are congruent modulo a.

Suppose that

$$2^{b_i} + (b_i + mk) \equiv 2^{b_j} + (b_j + nk) \pmod{a}, \tag{3}$$

for some values of $i, j \in \{0, 1, \ldots, d-1\}$ and $m, n \in \{0, 1, \ldots, a'-1\}$. Since d is a divisor of a, we also have

$$2^{b_i} + (b_i + mk) \equiv 2^{b_j} + (b_j + nk) \pmod{d}.$$

Since d is a divisor of k, using (1), we obtain that $i \equiv j \pmod{d}$. As $i, j \in \{0, 1, \ldots, d-1\}$, this simply means that $i = j$. Substituting this into (3) yields $mk \equiv nk \pmod{a}$. Therefore $mk' \equiv nk' \pmod{a}$; and since a' and k' are coprime, we get $m \equiv n \pmod{a'}$. Hence also $m = n$.

It follows that the a numbers that make up the d sequences in (2) satisfy all the requirements. They are certainly all greater than N because we chose each $b_i > \max\{2^M, N\}$. So the statement holds for a, completing the induction.

Example 6.10. (**TOT 2009**) Denote by $[n]!$ the product

$$1 \cdot 11 \cdot \ldots \cdot \underbrace{11\ldots1}_{n \text{ ones}} \quad (n \text{ factors in total}).$$

Prove that $[n+m]!$ is divisible by $[n]! \cdot [m]!$.

Solution. Define $f(n) = \underbrace{11\ldots1}_{n \text{ ones}}$ and $f(0) = 1$, so that $[0]! = 1$. Also define

$$\begin{bmatrix} n \\ k \end{bmatrix} = \frac{[n]!}{[k]![n-k]!},$$

for $0 \le k \le n$.

We use induction on n to prove that $\begin{bmatrix} n \\ k \end{bmatrix}$ is always a positive integer for all $n \ge 1$. For $n = 0$ we have

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \frac{[0]!}{[0]![0]!} = 1.$$

Assume now that the result holds for some $n \ge 0$. Then, for $n+1$, we have

$$\begin{aligned}
\begin{bmatrix} n+1 \\ k \end{bmatrix} &= \frac{[n+1]!}{[k]![n+1-k]!} \\
&= \frac{[n]!f(n+1)}{[k]![n+1-k]!} \\
&= \frac{[n]!f(n+1-k) \cdot 10^k}{[k]![n-k]!f(n+1-k)} + \frac{[n]!f(k)}{[k-1]![n+1-k]!} \\
&= 10^k \begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} n \\ k-1 \end{bmatrix}.
\end{aligned}$$

Both of the above terms are integers, so our induction is complete. This shows in particular that for any non-negative integers m and n one has that

$$\begin{bmatrix} m+n \\ n \end{bmatrix} = \frac{[m+n]!}{[m]![n]!}$$

is a positive integer, hence $[m+n]!$ is divisible by $[m]![n]!$, as required.

Example 6.11. Let a and b be positive integers such that $a-b>1$. Prove that there exist infinitely many positive integers n for which

$$n^2 \mid (a^n - b^n).$$

Solution. We will prove the result by explicitly constructing a sequence $(n_k)_{k\geq 1}$ which consists of infinitely many positive integers satisfying the required property:
We let $n_1 = 1$, and for $k \geq 1$ we set

$$n_{k+1} = \frac{a^{n_k} - b^{n_k}}{n_k}.$$

We prove by induction on $k \geq 1$ that $n_k \in \mathbb{N}$ and $n_k^2 \mid (a^{n_k} - b^{n_k})$.
The result is clear when $k = 1$, as $n_1 = 1$ and $1 \mid (a - b)$.
For the induction step, assume that the result holds for some $k \geq 1$ and we prove that it also holds for $k+1$.
By the induction hypothesis $n_k^2 \mid (a^{n_k} - b^{n_k})$, so

$$n_{k+1} = \frac{a^{n_k} - b^{n_k}}{n_k} \in \mathbb{N} \quad \text{and} \quad n_k \mid n_{k+1}.$$

Let $n_{k+1} = n_k \cdot m_k$, where $m_k \in \mathbb{Z}_{>0}$. Then

$$a^{n_{k+1}} - b^{n_{k+1}} = a^{n_k \cdot m_k} - b^{n_k \cdot m_k}$$

$$\begin{aligned}
&= (a^{n_k} - b^{n_k})((a^{n_k})^{m_k-1} + (a^{n_k})^{m_k-2}b^{n_k} + \ldots + (b^{n_k})^{m_k-1}) \\
&= n_k n_{k+1}((a^{n_k})^{m_k-1} - (b^{n_k})^{m_k-1} + ((a^{n_k})^{m_k-2} \\
&\quad -(b^{n_k})^{m_k-2})b^{n_k} + \ldots + m_k(b^{n_k})^{m_k-1}) \\
&= n_k n_{k+1}((a^{n_k} - b^{n_k})c + m_k(b^{n_k})^{m_k-1}) \\
&= n_k n_{k+1}(n_k n_{k+1} c + m_k(b^{n_k})^{m_k-1}) = n_{k+1}^2 \cdot d,
\end{aligned}$$

where c and d are integers. Hence, it follows that

$$n_{k+1}^2 \mid a^{n_{k+1}} - b^{n_{k+1}}.$$

This completes our induction.

Finally, we show that the sequence (n_k) is increasing. Note that

$$\begin{aligned} n_{k+1} = \frac{a^{n_k} - b^{n_k}}{n_k} &= \frac{(a-b)(a^{n_k-1} + a^{n_k-2}b + \ldots + b^{n_k-1})}{n_k} \\ &\geq \frac{2(3^{n_k-1} + 3^{n_k-2} + \ldots + 1)}{n_k} \\ &\geq \frac{2(n_k + n_k - 1 + \ldots + 1)}{n_k} = n_k + 1 > n_k, \end{aligned}$$

where we have used that $3^0 < 3^1 < \ldots < 3^{n-1}$ and that $3^{n-1} \geq n$, for $n \geq 1$. Therefore, $n_{k+1} > n_k$, which completes the solution.

Example 6.12. (**AMM**) Let d, k, q be positive integers with k odd. Find the power of 2 in the factorization of $\sum\limits_{n=1}^{2^d \cdot k} n^q$.

Solution. We prove that if q is even or $q = 1$, then the answer is $d - 1$, and otherwise the answer is $2(d-1)$. We will do this by induction on d. When $q = 1$, the result follows immediately by direct computation. Also, if $d = 1$, then one can see that the total sum is odd, showing that the exponent must be 0. These are the base cases for our proof by induction.

Assume now that $d > 1$. We can rewrite the given sum as

$$(2^d k)^q - (2^{d-1}k)^q + \sum_{j=1}^{2^{d-1}k} (j^q + (2^d k - j)^q).$$

If q is even, we take the sum modulo 2^d and we find that

$$\sum_{n=1}^{2^d k} n^q \equiv 2 \sum_{n=1}^{2^{d-1}k} n^q \pmod{2^d}.$$

From the induction hypothesis, we know that $\sum\limits_{n=1}^{2^{d-1}k} n^q \equiv 2^{d-2} \pmod{2^{d-1}}$, so

$$\sum_{n=1}^{2^d k} n^q \equiv 2^{d-1} \pmod{2^d},$$

as we wanted.

If $q \geq 3$ is odd, we have by the binomial theorem that

$$n^q + (2^d k - n)^q \equiv 2^d qkn^{q-1} \pmod{2^{2d-1}},$$

so $\sum_{n=1}^{2^d k} n^q \equiv 2^d qk \sum_{n=1}^{2^{d-1}k} n^{q-1} \pmod{2^{2d-1}}$. Since $q-1 \geq 2$ is even, we can make use of the result we established above to conclude that

$$\sum_{n=1}^{2^{d-1}k} n^{q-1} \equiv 2^{d-2} \pmod{2^{d-1}},$$

so

$$\sum_{n=1}^{2^d k} n^q \equiv 2^{2d-2} \pmod{2^{2d-1}},$$

completing our proof.

6.1.3 Representations

Another very popular category of problems from Induction in Number Theory are the questions where we are asked to prove that a number or a sequence of numbers can be represented in a certain form. We illustrate this theme below, starting with the following, slightly easier example:

Example 6.13. Prove that any number between 1 and $n!$ can be written as a sum of at most n distinct divisors of $n!$.

Solution. We proceed by induction on n. For $n = 3$, the claim is true.

Assume that the hypothesis holds for $n-1$. Let $1 < k < n!$ and let k', q be the quotient and the remainder when k is divided by n. Hence $k = k'n + q$, $0 \leq q < n$ and $0 \leq k' < \frac{k}{n} < \frac{n!}{n} = (n-1)!$.

From the inductive hypothesis, there are integers $d_1' < d_2' < \ldots < d_s'$, $s \leq n-1$ such that $d_i' \mid (n-1)!$, $i = 1, 2, \ldots, s$ and $k' = d_1' + d_2' + \ldots + d_s'$. Hence $k = nd_1' + nd_2' + \ldots + nd_s' + q$.

Now, if $q = 0$, then $k = d_1 + d_2 + \ldots + d_s$, where $d_i = nd_i'$, $i = 1, 2, \ldots, s$ are distinct divisors of $n!$.

If $q \neq 0$, then $k = d_1 + d_2 + \ldots + d_{s+1}$, where $d_i = nd_i'$, $i = 1, 2, \ldots, s$, and $d_{s+1} = q$. It is clear that $d_i \mid n!$, for $i = 1, 2, \ldots, s$ and $d_{s+1} \mid n!$, since $q < n$. On the other hand, $d_{s+1} < d_1 < d_2 < \ldots < d_s$, because $d_{s+1} = q < n \leq nd_1' = d_1$. Therefore k can be written as a sum of at most n distinct divisors of $n!$, as claimed.

Example 6.14. A sequence $c_1, c_2, \ldots$ is called *perfect* if every natural number m with $1 \leq m \leq c_1 + c_2 + \ldots + c_n$ can be represented as

$$m = \frac{c_1}{a_1} + \frac{c_2}{a_2} + \ldots + \frac{c_n}{a_n},$$

where $a_1, a_2, \ldots, a_n$ are natural numbers. Find the maximum possible value of c_n for a perfect sequence.

Solution. Let $b_1, b_2, \ldots, b_n$ be a permutation in increasing order of $c_1, c_2, \ldots, c_n$. We know from hypothesis that $b_k + b_{k+1} + \ldots + b_n$ can be represented as

$$b_k + b_{k+1} + \ldots + b_n = \sum_{i=1}^{n} \frac{b_i}{a_i}.$$

We cannot have $a_k = a_{k+1} = \ldots = a_n = 1$. So let $j \geq k$ be the largest possible number such that $a_j > 2$. We deduce that

$$b_k + b_{k+1} + \ldots + b_n \leq b_1 + b_2 + \ldots + b_{j-1} + \frac{b_j}{2} + b_{j+1} + \ldots + b_n,$$

hence $b_j \leq 2(b_1 + b_2 + \ldots + b_{k-1})$. We get the relation

$$b_k \leq 2(b_1 + b_2 + \ldots + b_{k-1}).$$

However, $b_1 \leq 2$, as otherwise $b_1 + b_2 + \ldots + b_n - 1$ could not be represented as $\frac{b_1}{a_1} + \frac{b_2}{a_2} + \ldots + \frac{b_n}{a_n}$. Thus

$$b_2 \leq 2b_1 \leq 4, \quad b_3 \leq 2(b_2 + b_1) = 12, \ldots$$

We prove by induction that $b_k \leq 4 \cdot 3^{k-2}$, for $k \geq 2$. The base case is true from above.

The induction step follows from the fact that

$$b_k \leq 2(b_1 + b_2 + \ldots + b_{k-1}) = 2(2 + 4 + \ldots + 4 \cdot 3^{k-3}) = 4 \cdot 3^{k-2}.$$

Thus $c_k \leq b_k \leq 4 \cdot 3^{k-2}$. It remains to prove that the sequence defined by

$$c_1 = 2, \quad c_k = 4 \cdot 3^{k-2} \quad \text{for} \quad k > 1$$

is perfect.

We prove by induction on k that if $m \leq c_1 + c_2 + \ldots + c_k = 2 \cdot 3^{k-1}$, then we can write m as $\sum_{i=1}^{m} \frac{c_i}{a_i}$. The base is clear, so we want to prove the statement for $k+1$, assuming it for k. If $m = 1$, we put $a_i = 2 \cdot 3^{k-1}$. If $1 < m \leq 2 \cdot 3^{k-1} + 1$, then we can put $a_k = c_k$ and use the induction hypothesis with k numbers for $m - 1$. If $2 \cdot 3^{k-1} + 1 < m \leq 4 \cdot 3^{k-1}$ then we set $a_k = 2$ and use the induction hypothesis for $m - 2 \cdot 3^{k-1}$. If $4 \cdot 3^{k-1} < m$ then we set $a_k = 1$ and use the induction hypothesis for $m - 4 \cdot 3^{k-1}$.

Example 6.15. Prove that every positive rational number r can be written as

$$r = \frac{1}{q_1} + \frac{1}{q_2} + \ldots + \frac{1}{q_m},$$

where $q_1, q_2, \ldots, q_m$ are distinct integers. Moreover, if $r < 1$, then we may assume $q_1 \mid q_2$, $q_2 \mid q_3$, $\ldots$, $q_{m-1} \mid q_m$.

Solution. We begin by proving the second assertion (for $r < 1$). Let $r = \frac{p}{q}$ in lowest terms. We prove the statement by induction on p, with the base case $p = 1$ being clear.

For $p > 1$, we have by the way we defined r that $p \nmid q$. Let $k = 1 + \left[\frac{q}{p}\right]$, so $0 < kp - q < p < q$. Then by induction hypothesis, $k\left(\frac{p}{q} - \frac{1}{k}\right) = \frac{kp-q}{q}$ can be represented as $\frac{1}{q_1} + \frac{1}{q_2} + \ldots + \frac{1}{q_m}$ with $q_i | q_{i+1}$. Then

$$\frac{p}{q} = \frac{1}{k} + \frac{kp - q}{kq} = \frac{1}{k} + \frac{1}{kq_1} + \frac{1}{kq_2} + \ldots + \frac{1}{kq_m},$$

and this is the desired representation.

For the general part, as the sequence

$$S(n) = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$$

is divergent, we can find an n such that

$$1 + \frac{1}{2} + \ldots + \frac{1}{n} \leq r < 1 + \frac{1}{2} + \ldots + \frac{1}{n} + \frac{1}{n+1}.$$

Therefore, if we let

$$s = r - 1 - \frac{1}{2} - \ldots - \frac{1}{n},$$

we have $0 \leq s < \frac{1}{n+1}$. By applying the result we established in the first part for $(n+1)s$, we find distinct $q_1, q_2, \ldots, q_m$ such that

$$(n+1)s = \frac{1}{q_1} + \frac{1}{q_2} + \ldots + \frac{1}{q_m}.$$

Hence

$$r = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \frac{1}{(n+1)q_1} + \ldots + \frac{1}{(n+1)q_m}.$$

Clearly all fractions are different. This completes our proof.

Example 6.16. Show that every positive rational number can be written as a quotient of products of factorials of (not necessarily distinct) primes, and the representation is unique up to rearranging and cancelling common factors.

Solution. Given a positive rational number r, we can prove both the existence and uniqueness parts by induction on the largest prime p with the property $\nu_p(r) \neq 0$.

The base case is $r = 2^k$ for some integer $k \in \mathbb{Z}$, in which case $r = (2!)^k$.

For the induction step, if $p > 2$ and $k = \nu_p(r)$, then $\frac{r}{(p!)^k}$ is a rational number all whose prime divisors are less than p and hence by induction it can be written in the required format. This proves the existence.

For the uniqueness, suppose $r = \prod_{j=1}^{s} (p_j!)^{k_j}$ for some primes $p_1 < p_2 < \cdots < p_s$ and non-zero integers k_j. Then we see that $\nu_{p_s}(r) = k_s \neq 0$. Hence

$p_s = p$ and $k_s = \nu_p(r)$ and by induction $\prod_{j=1}^{s-1}(p_j!)^{k_j}$ is the unique representation for $\dfrac{r}{(p!)^k}$.

Example 6.17. (**Moscow 2014**) Prove that for all positive integers n, there exists a positive integer k_n such that the decimal expansion of k_n^2 starts with n digits of 1 and ends with a block of length n which consists only of the digits 1 and 2.

Solution. We first prove by induction on n that there exists an integer m_n whose decimal expansion ends with 1 and the decimal expansion of m_n^2 ends with a block of length n consisting only of 1's and 2's. For the base case $n = 1$, simply take $m_1 = 1$.

Assume now that the result holds for some $n \geq 1$ and m_n is the desired number. Define $p_a = m_n + a \cdot 10^n$, for some $1 \leq a \leq 9$ to be chosen later. Notice that the last digit of p_a is also 1. We have

$$p_a^2 = m_n^2 + 2am_n10^n + a^2 10^{2n}.$$

The number m_n^2 ends with a block of length n consisting of 1's and 2's, the number $2am_n10^n$ ends with n zeros and $a^2 10^{2n}$ ends with $2n$ zeros. Let the $(n+1)$-st digit (from right to left) of m_n^2 be b. Since m_n ends with 1, the $(n+1)$-st digit of $2a \cdot 10^n$ is $2a \pmod{10}$. Therefore, the $(n+1)$-st digit of p_a^2 is $b + 2a \pmod{10}$. If b is odd, we take a such that $b + 2a \equiv 1 \pmod{10}$. Otherwise, we take a such that $b + 2a \equiv 2 \pmod{10}$. We complete the induction step by taking $m_{n+1} = p_a$.

Now set

$$c_n = \underbrace{11\dots1}_{n} \cdot 10^{4n} \quad \text{and} \quad d_n = c_n + 10^{4n}.$$

We have

$$\sqrt{d_n} - \sqrt{c_n} = \frac{d_n - c_n}{\sqrt{d_n} + \sqrt{c_n}} = \frac{10^{4n}}{\sqrt{d_n} + \sqrt{c_n}} > \frac{10^{4n}}{2 \cdot 10^{3n}} > 1.$$

Therefore, in the interval $(\sqrt{c_n}, \sqrt{d_n})$ there exists an integer u_n and its square will start with n digits of 1. We now take l larger than the number of digits of $2u_nm_n$ and m_n^2 and we consider the number $k_n = u_n \cdot 10^l + m_n$. One can see that k_n^2 satisfies the required conditions of the problem.

Example 6.18. (**St. Petersburg**) Prove that for any $n \geq 1$, the number $n!$ has the following property: to any divisor of $n!$, except $n!$, we can add another divisor of $n!$ so that the sum is again a divisor of $n!$.

Solution. We proceed by induction on n. The base cases $n = 1$ and $n = 2$ are immediate.

Let us now perform the induction step from n to $n+1$. If $k \mid (n+1)!$, then we write $k = ab$, where $a \mid n!$ and $b \mid n+1$. If $a \neq n!$, then by the induction hypothesis, there is d with $d \mid n!$ and $(a+d) \mid n!$. Then bd and $k+bd = (a+d)b$ are divisors of $(n+1)!$.

If $a = n!$, then $b \neq n+1$. If $n+1$ is composite, then $n+1 = bc$ for some b, c and we can write k also as $k = (n+1)a'$ where $a' = \frac{n!}{c}$ and for such a representation, the above proof works.

Finally, if $n+1$ is prime, then we must have $b = 1$ and $k = n!$. But then

$$n! + (n-1)! = (n+1) \cdot (n-1)!.$$

This completes our proof.

Example 6.19. (**IMO 1998**) For any positive integer n, let $\tau(n)$ denote the number of its positive divisors (including 1 and itself). Determine all positive integers m for which there exists a positive integer n such that $\frac{\tau(n^2)}{\tau(n)} = m$.

Solution. If $n = p_1^{a_1}p_2^{a_2}\dots p_k^{a_k}$ then $\tau(n) = (a_1+1)(a_2+1)\dots(a_k+1)$. Thus, the problem asks what integers n can be represented as

$$\frac{(2a_1+1)(2a_2+2)\dots(2a_k+1)}{(a_1+1)(a_2+1)\dots(a_k+1)},$$

for positive integers $a_1, a_2, \dots, a_k$.

Firstly, as $(2a_1+1)(2a_2+1)\dots(2a_k+1)$ is odd, any such n has to be odd. Now let us prove by strong induction on n that every odd n can be represented in such a way. The induction step would be done if we found an m, $m < n$ such that

$$\frac{n}{m} = \frac{(2x_1+1)(2x_2+1)\dots(2x_k+1)}{(x_1+1)(x_2+1)\dots(x_k+1)},$$

because then we could apply the induction hypothesis to m. Now if we choose $x_{i+1} = 2x_i$, we get

$$\frac{(2x_1+1)(2x_2+1)\dots(2x_k+1)}{(x_1+1)(2x_1+1)(2x_2+1)\dots(2x_{k-1}+1)} = \frac{2x_k+1}{x_1+1} = \frac{2^k x_1+1}{x_1+1}.$$

Thus if we can find an odd $m < n$ such that $\frac{m}{n} = \frac{t+1}{2^k t+1}$ for some $k \in N$, the induction step is finished. Now as m should be odd we must have $t = 2s$, so it suffices to have $\frac{2s+1}{2^{k+1}s+1} n \in N$ for some $s, k \in N$. So n must have a divisor of form $\frac{2^{k+1}s+1}{(2^{k+1}s+1, 2s+1)}$. As

$$(2^{k+1}s+1, 2s+1) | 2^k(2s+1) - (2^{k+1}s+1) = 2^k - 1,$$

if we set $2s+1 = (2m+1)(2^k-1)$ or $s = 2^{k-1}(2m+1) - m - 1$ then

$$\frac{2^{k+1}s+1}{(2^{k+1}s+1, 2s+1)} = \frac{2^{k+1}s+1}{2^k-1} = 2s+2m+1 = 2^k(2m+1)-1.$$

This is the desired breakthrough, because if we write $n+1 = 2^k(2m+1)$ then $2^k(2m+1) - 1 = n$ and because of our reasoning above we can perform the induction step. Explicitly,

$$n = \frac{4s+1}{2s+1}\frac{8s+1}{4s+1}\dots\frac{2^{k+1}s+1}{2^k s+1}(2m+1),$$

where $s = 2^{k-1}(2m+1) - m - 1$ for m, k such that $n+1 = 2^k(2m+1)$. Thus applying the induction hypothesis for $2m+1$ we get the claim.

6.2 Proposed Problems

Problem 6.1. Show that for any positive integer n, 3^n divides $\underbrace{11\dots11}_{3^n\text{ones}}$.

Problem 6.2. Prove that for any two positive integers a and m , the sequence

$$a, a^a, a^{a^a}, \dots$$

is eventually constant modulo m.

Problem 6.3. (**Poland 1998**) Let x, y be real numbers such that the numbers $x+y$, x^2+y^2, x^3+y^3, and x^4+y^4 are all integers. Prove that for all positive integers n, the number x^n+y^n is an integer.

Problem 6.4. (**Ibero American 2012**) Let a, b, c, d be integers such that the number $a-b+c-d$ is odd and it divides the number $a^2-b^2+c^2-d^2$. Show that for every positive integer n, $a-b+c-d$ divides $a^n-b^n+c^n-d^n$.

Problem 6.5. (**AoPS**) Let m and n be positive integers with $\gcd(m, n) = 1$. Compute

$$\gcd(5^m+7^m, 5^n+7^n).$$

Problem 6.6. Let n be a non-negative integer. Prove that the numbers $0, 1, 2, \dots, n$ can be rearranged into a sequence $a_0, a_1, \dots, a_n$ such that $a_i + i$ is a perfect square for all $0 \le i \le n$.

Problem 6.7. We call a positive n triangular if it can be written as

$$n = \frac{a(a+1)}{2},$$

for some positive integer a. Show that $11\dots1_9$ is triangular, where the subscript denotes the fact that $11\dots1$ is written in base 9.

Problem 6.8. Let a, b, m be positive integers such that $\gcd(b, m) = 1$. Prove that the set $\{a^n + bn : n = 1, 2, \dots, m^2\}$ contains a complete set of residues modulo m.

Problem 6.9. (**Kömal**) Let a and n be two positive integers such that $a^n - 1$ is divisible by n. Prove that the numbers $a+1$, a^2+2, ..., a^n+n are distinct modulo n.

Problem 6.10. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = ax^2 + bx + c$, where a, b, c are positive integers. Let n be some given positive integer. Prove that for any positive integer m, there exist n consecutive integers $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that each of the numbers $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)$ has at least m distinct prime factors.

Problem 6.11. (**Iran 2005**) Find all $f : \mathbb{N}^* \to \mathbb{N}^*$ such that for every $m, n \in \mathbb{N}^*$,

$$f(m) + f(n) \mid m + n.$$

Problem 6.12. (**Kvant M2252**) Prove that

$$1 + 3^3 + 5^5 + \ldots + (2^n - 1)^{2^n - 1} \equiv 2^n \pmod{2^{n+1}}, \quad \text{for} \quad n \geq 2.$$

Problem 6.13. (**GMA 2013**) Prove that for any positive integer n and any prime p, the sum

$$S_n = \sum_{k=0}^{\lfloor \frac{n}{p} \rfloor} (-1)^k \binom{n}{kp}$$

is divisible by $p^{\lfloor \frac{n-1}{p-1} \rfloor}$.

Problem 6.14. (**Bulgaria 1996**) Let $k \geq 3$ be an integer. Show that there exist odd positive integers x and y with $2^k = 7x^2 + y^2$.

Problem 6.15. (**USAMO 1998**) Prove that for each $n \geq 2$, there is a set S of n integers such that $(a-b)^2$ divides ab for every distinct $a, b \in S$.

Problem 6.16. (**Brazil 2011**) Prove that there exist positive integers $a_1 < a_2 < \ldots < a_{2011}$ such that for all $1 \leq i < j \leq 2011$ we have $\gcd(a_i, a_j) = a_j - a_i$.

Problem 6.17. (**Bulgaria TST**) Let $a, m \geq 2$ and let $\text{ord}_m^a = k$ (i.e. $a^k \equiv 1 \pmod{m}$ and $a^s \not\equiv 1 \pmod{m}$ for any $0 < s < k$). Prove that if t is an odd number such that every prime that divides t also divides m and $\gcd\left(t, \frac{a^k - 1}{m}\right) = 1$, then $\text{ord}_{mt}^a = kt$.

Problem 6.18. (**China TST**) Prove that for all positive integers m, n there exists an integer k such that $2^k - m$ has at least n distinct prime divisors.

Problem 6.19. (**Serbia**) Prove that for all positive integers m there exists a positive integer $k \geq 2$ such that $3^k - 2^k - k$ is divisible by m.

Problem 6.20. Let k be a positive integer. Prove that for all non-negative integers m there exists a positive integer n with at least m prime factors (not necessarily distinct) such that $2^{kn^2} + 3^{kn^2}$ is divisible by n^3.

Problem 6.21. Prove that for all positive integers k there exists an integer n which has exactly k prime divisors and $n^3 \mid (2^{n^2} + 1)$.

Problem 6.22. (**Polish Training Camp**) Let k be a positive integer. The sequence $(a_n)_{n\geq 1}$ is given by

$$\sum_{d|n} da_d = k^n, \quad \text{for all} \quad n \geq 1.$$

Prove that every term of the sequence is an integer.

Problem 6.23. Prove that there is a permutation $(a_1, a_2, \ldots, a_n)$ of $(1, 2, 3, \ldots, n)$ such that none of the numbers $a_1, a_1+a_2, \ldots, a_1+a_2+\ldots+a_{n-1}$ is a perfect square.

Problem 6.24. Prove that any integer can be represented in infinitely many ways as $\pm 1^2 \pm 2^2 \pm 3^2 \pm \ldots \pm t^2$, for a convenient t and a suitable choice of the signs $+$ and $-$.

Problem 6.25. (**Romania TST 2013**) Find all positive integers n that can be written as

$$n = \frac{(a_1^2 + a_1 - 1)(a_2^2 + a_2 - 1) \cdots (a_k^2 + a_k - 1)}{(b_1^2 + b_1 - 1)(b_2^2 + b_2 - 1) \cdots (b_k^2 + b_k - 1)},$$

for some positive integers $a_i, b_i \in \mathbb{N}^*$ and some $k \in \mathbb{N}^*$.

Problem 6.26. (**USA TST 2006**) Let n be a positive integer. Find, with proof, the least positive integer d_n which cannot be expressed in the form

$$\sum_{i=1}^{n}(-1)^{a_i}2^{b_i},$$

where a_i and b_i are non-negative integers for each i.

Problem 6.27. **(USAMO 2003)** Prove that for every positive integer n there exists an n-digit number divisible by 5^n, all of whose digits are odd.

Problem 6.28. **(IMO 2004)** We call a positive integer *alternating* if every two consecutive digits in its decimal representation are of different parity. Find all positive integers n such that n has a multiple which is alternating.

Problem 6.29. **(IMO shortlist 2000)** Does there exist a positive integer n such that n has exactly 2000 prime divisors and n divides $2^n + 1$?

Problem 6.30. (**IMO shortlist 2002**) Let $p_1, p_2, \ldots, p_n$ be distinct primes greater than 3. Show that $2^{p_1 p_2 \cdots p_n} + 1$ has at least 4^n divisors.

Problem 6.31. (**IMO 1988**) Show that if a, b and $q = \dfrac{a^2+b^2}{ab+1}$ are non-negative integers, then $q = \gcd(a,b)^2$.

Problem 6.32. (**IMO 1999 shortlist**) Prove that there exist two strictly increasing sequences (a_n) and (b_n) such that $a_n(a_n+1)$ divides b_n^2+1 for every natural n.

Problem 6.33. Prove that for any two positive integers n and m, we have

$$\gcd(F_n, F_m) = F_{\gcd(n,m)}.$$

Problem 6.34. Let n be a positive integer which is not divisible by 3. Show that $x^3+y^3 = z^n$ has at least one solution (x, y, z) with x, y, z positive integers.

Problem 6.35. Let n be a positive integer. What is the largest number of elements that one can choose from the set $A = \{1, 2, \ldots, 2n\}$ such that the sum any two chosen numbers is composite?

Problem 6.36. (**Bulgaria 1999**) Find the number of positive integers n, $4 \leq n \leq 2^k - 1$, whose binary representations do not contain three equal consecutive digits.

Problem 6.37. Prove that every positive integer is a sum of one or more numbers of the form $2^r 3^s$, where r and s are non-negative integers and no summand divides another (for example, $23 = 9 + 8 + 6$).

Problem 6.38. Let $p \geq 3$ be a prime and let $a_1, a_2, \ldots, a_{p-2}$ be a sequence of positive integers such that p does not divide either a_k or $a_k^k - 1$ for all $k = 1, 2, \ldots, p-2$. Prove that the product of some elements of the sequence is congruent to 2 modulo p.

Problem 6.39. Let n be a positive integer. Prove that the number of ordered pairs (a, b) of relatively prime positive divisors of n is equal to the number of divisors of n^2.

Problem 6.40. Prove that for each integer $n \geq 3$, there are n pairwise distinct positive integers such that each of them divides the sum of the remaining $n-1$.

Problem 6.41. (**USAMO 2008**) Show that for all positive integers n we can find distinct positive integers $a_1, a_2, \ldots a_n$ such that $a_1 \cdot a_2 \cdots a_n - 1$ is the product of two consecutive integers.

Problem 6.42. We begin with a triple (a, b, c) of positive reals a, b, c, such that $a \leq b \leq c$. Starting with the given triple, at each step, we perform the transformation

$$(x, y, z) \to (|x - y|, |y - z|, |z - x|).$$

Prove that we can eventually reach a 0 in one of the triples if and only if there exist positive integers $n \geq k \geq 0$ such that and

$$nb = ka + (n - k)c.$$

Problem 6.43. (**Poland 2000**) A sequence $p_1, p_2, \ldots$ of prime numbers satisfies the following condition: for $n \geq 3$, p_n is the greatest prime divisor of $p_{n-1} + p_{n-2} + 2000$. Prove that the sequence is bounded.

Problem 6.44. Prove that for all positive integers m, there exists an integer n such that $\phi(n) = m!$.

Problem 6.45. (**Bulgaria 2012**) Let p be an odd prime and let $a_1, \ldots, a_{2p-1}$ be distinct integers lying in the interval $[1, p^2]$ such that their sum is divisible by p. Prove that there exist positive integers $b_1, \ldots, b_{2p-1}$ none of which is divisible by p, such that their representation in base p contains only 1 and 0 and such that $\sum\limits_{j=1}^{2p-1} a_j b_j$ is divisible by p^{2012}.

Problem 6.46. (**Poland 2010**) Prove that there exists a set A consisting of 2010 positive integers such that for any nonempty subset $B \subseteq A$, the sum of the elements in B is a perfect power greater than 1.

Problem 6.47. (**AMM**) For a positive integer m, we let $\sigma(m)$ be the sum

$$\sigma(m) = \sum_{\substack{1 \le d < m \\ d \mid m}} d.$$

Prove that for every integer $t \geq 1$ there exists an m such that

$$m < \sigma(m) < \sigma(\sigma(m)) < \ldots < \sigma^t(m),$$

where $\sigma^k = \underbrace{\sigma \circ \sigma \circ \ldots \circ \sigma}_{k \text{ terms}}$.

Problem 6.48. (**Moscow 2013**) For a positive integer m, we denote by $S(m)$ the sum of digits of m. Prove that for any positive integer n, there exists an integer k such that

$$S(k) = n, \quad S(k^2) = n^2, \quad S(k^3) = n^3.$$

Problem 6.49. Establish whether there exist positive integers $a_1 < a_2 < \ldots < a_n < \ldots$ such that every number in the sequence

$$a_1^2,\ a_1^2 + a_2^2, \ldots, a_1^2 + a_2^2 + \ldots + a_n^2, \ldots$$

is the square of a positive integer.

Problem 6.50. For which pairs (a, b) of positive integers do there exist only finitely many positive integers n such that $n^2 \mid a^n + b^n$?

Chapter 7

Combinatorics

7.1 Theory and Examples

Combinatorics is among the most popular topics when it comes to mathematical contests. Hence, it is no surprise that lots of the olympiad problems that are solved using induction are combinatorics questions. To illustrate better some of the major themes that occur in combinatorics and use induction, we divided this chapter into three separate sections: Partitions and configurations, Graph Theory and Combinatorial Geometry.

7.1.1 Partitions and Configurations

Let us recall some of the basic definitions that occur when we look at partitions of objects:

A **permutation** of order n is a bijective function from the set $\{1, 2, \ldots, n\}$ onto itself. The number of permutations of order n is $n!$.

A **combination** of order n over k is a k-element subset of $\{1, 2, \ldots, n\}$. For $0 \leq k \leq n$, the number of such combinations is given by the formula

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}.$$

Let us prove by induction the following well-known theorem, known as the Pigeonhole Principle:

Example 7.1. (**Pigeonhole Principle**) Let n and k be positive integers. If a set of $nk+1$ different elements is partitioned into n mutually disjoint subsets, then there is at least one set which contains at least $k+1$ elements.

Solution. We will prove this by induction on $n \geq 1$. For the base case $n=1$, everything is clear, as we only have one set containing $k+1$ elements.

Assume now that the result holds for some $n \geq 1$, and consider $(n+1)k+1$ elements which are distributed among $n+1$ mutually disjoint subsets. Let us call these subsets $A_1, A_2, \ldots, A_{n+1}$. If A_{n+1} contains at least $k+1$ elements, then we are done. Otherwise, there are at most k elements in A_{n+1}, so the subsets $A_1, A_2, \ldots, A_n$ contain at least $(n+1)k+1-k = nk+1$ elements. By the induction hypothesis, at least one of $A_1, A_2, \ldots, A_n$ must contain $k+1$ elements, so we are done.

Remark. There is also an alternative direct proof of the above theorem, which uses only proof by contradiction.

Another theorem which is among first results one learns while studying partitions of sets is the following:

Example 7.2. (**The Inclusion-Exclusion Principle**) If $A_1, A_2, \ldots, A_n$ are some sets, then

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = \sum_{i=1}^{n} |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \ldots$$

$$+(-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} |A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}| + \ldots + (-1)^{n-1} |A_1 \cap \ldots \cap A_n|.$$

Solution. Let $P(n)$ be the statement: "For any sets $A_1 \ldots A_n$, we have

$$|A_1 \cup \cdots \cup A_n| = \sum_{i} |A_i| - \sum_{i<j} |A_i \cap A_j| + \cdots \pm |A_i \cap A_2 \cap \cdots \cap A_n|".$$

$P(1)$ is trivially true. $P(2)$ is also true, since

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|,$$

by basic properties of sets.

Assume now that the result holds for all positive integers less than or equal to some $n \geq 2$. Now given $A_1 \ldots A_{n+1}$, let $B_i = A_i \cap A_{n+1}$ for $1 \leq i \leq n$. Observe that $B_i \cap B_j = A_i \cap A_j \cap A_{n+1}$. Likewise, $B_i \cap B_j \cap B_k = A_i \cap A_j \cap A_k \cap A_{n+1}$. By the induction hypothesis applied for 2 and n sets, respectively, we have

$$\begin{aligned}|A_1 \cup A_2 \cup \ldots \cup A_{n+1}| &= |A_1 \cup \ldots \cup A_n| + |A_{n+1}| - |(A_1 \cup \ldots \cup A_n) \cap A_{n+1}| \\ &= |A_1 \cup \ldots \cup A_n| + |A_{n+1}| - |B_1 \cup \ldots \cup B_n| \\ &= \sum_{i \leq n} |A_i| - \sum_{i<j \leq n} |A_i \cap A_j| + \ldots + |A_{n+1}| \\ &\quad - \sum_{i \leq n} |B_i| + \sum_{i<j \leq n} |B_i \cap B_j| - \ldots \end{aligned} \tag{1}$$

Notice that $\sum_{i \leq n} |B_i| = \sum_{i \leq n} |A_i \cap A_{n+1}|$. Hence

$$\sum_{i<j \leq n} |A_i \cap A_j| + \sum_{i \leq n} |B_i| = \sum_{i<j \leq n+1} |A_i \cap A_j|,$$

and similarly for higher order intersections.
Substituting this back into (1), we obtain that

$$|A_1 \cup A_2 \cup \cdots \cup A_{n+1}| = \sum_{i \leq n+1} |A_i| - \sum_{i<j \leq n+1} |A_i \cap A_j| + \ldots,$$

which shows that $P(n+1)$ holds. This completes the induction step and the proof of the result.

A common theme in Combinatorics is that each problem comes with its own universe. Therefore, the best method through which one can master the subject is to simply see lots of different problems with various ideas behind them. In the following, we will present a series of such examples, of different levels of difficulty:

Example 7.3. (**St. Petersburg**) We have a 100×100 grid of points in the plane. We can draw lines that cover all the points in the grid, except the bottom left point. What is the minimum number of such lines that are needed?

Solution. The answer is 198. We can take 99 vertical lines and 99 horizontal ones, covering all rows and columns not containing the bottom left point.

We prove more generally that for an $m \times n$ grid we need $m + n - 2$ lines. We do this by induction as follows:

The base case is when m or n is 1, and in this case we have $n - 1$ collinear points to cover, but we are not allowed to draw a line containing any two of them, since it will contain the forbidden point. So the minimum is $n - 1$.

Now if $m > 1$ and $n > 1$, assume that the result is true for all grids smaller than $m \times n$. Consider the border of non-forbidden points. There are $2m + 2n - 5$ points on it. If some line contains one full side of this border, then the result is easy by induction, removing that line (side) (i.e. $m \to m - 1$ or $n \to n - 1$). Otherwise, any line intersects at most 2 points on this border, so we must have at least $\lfloor \frac{2m+2n-5}{2} \rfloor + 1 = m + n - 2$ lines. This completes our proof.

Example 7.4. Let n be an integer greater than 1. Find the least number of rooks such that no matter how we place them on an $n \times n$ chessboard, there will be two rooks that do not attack each other, but at the same time they will be under attack by a third rook.

Solution. We show that the least number of rooks is $2n - 1$. We label the squares of the $n \times n$ chess board so that the bottom left one is $a_{1,1}$ and the top right one is $a_{n,n}$. First notice that by placing the rooks on $a_{1,2}$, $a_{1,3}, \dots,$ $a_{1,n}$, $a_{2,1}$, $a_{3,1}, \dots,$ $a_{n,1}$, we see that $2n - 2$ rooks are not sufficient.

We will prove by induction on n that $2n-1$ rooks are sufficient. For $n = 2$, the result is clear.

We now suppose that the result is true for $n = k \geq 1$. By placing the $2k+1$ rooks on the $(k + 1) \times (k + 1)$ chessboard, there is at least one row containing one rook or no rooks. Otherwise, the total number of rooks is greater than or equal to $2k + 2$, which is not true. Similarly there is at least one column containing one rook or no rooks. Select any such row and any such column and delete them from the $(k + 1) \times (k + 1)$ chessboard.

We combine the undeleted parts of the $(k + 1) \times (k + 1)$ chessboard to obtain a $k \times k$ chessboard which contains at least $2k - 1$ rooks. Select any $2k - 1$ rooks. By the induction assumption, they are sufficient. It follows that

$2k+1$ rooks are sufficient for the $(k+1)\times(k+1)$ chessboard. This completes the solution.

Example 7.5. On a circular route, there are n identical cars. Together they have enough gas for one car to make a complete tour. Prove that there is a car that can make a complete tour by taking gas from all the cars that it encounters.

Solution. We will prove this by induction on n. It is clearly true for $n = 1$.

Now assume for n and prove for $n+1$. Consider $n+1$ cars. Clearly, one of them, say A, has enough gas to get to the next one, say B.

Now consider a new configuration in which all cars are in the same position and have the same amount of gas, except B, which is removed, and A which is left in the same position, but it is also given the gas of B.

By the induction hypothesis, there exists a car which can do the complete tour. We claim that it can do the complete tour also in the configuration with $n+1$ cars.

Indeed, it will have the same amount of gas in both configurations in all places except between A and B. But we have chosen A such that it has enough gas to reach B, so the car will not stop between A and B. This concludes the proof.

Example 7.6. (**TOT 2002**) There is a large pile of cards. On each card, one of the numbers $1, 2, \ldots, n$ is written. It is known that the sum of all numbers of all the cards is equal to $k \cdot n!$ for some integer k. Prove that it is possible to arrange cards into k stacks so that the sum of numbers written on the cards in each stack is equal to $n!$.

Solution. We begin by proving a little lemma:

Lemma. From any set $\{a_1, \ldots, a_n\}$ of n integers, one can choose a number or several numbers with their sum divisible by n.

Proof. Let us assume that none of the numbers is divisible by n. Now consider the numbers $b_1 = a_1$, $b_2 = a_1 + a_2, \ldots$, $b_n = a_1 + a_2 + \ldots + a_n$. If none of them is divisible by n, then at least two numbers b_j and b_l $(k < l)$ have the same remainders. Then their difference $a_{j+1} + \ldots + a_l$ is divisible by n.

We now prove our questions using induction on n. If $n = 1$, then the only number is 1 and it is written on each card, hence every card forms the required group with sum $1!$ on its own.

Assume now that we established the result for some $n \geq 1$, i.e. if the sum of the numbers on all cards is $k \cdot n!$, then the cards could be arranged into k stacks with the sum of the numbers in each stack equal to $n!$.

We call a *supercard* any group of cards with sum $l \cdot (n+1)$, $l = 1, \ldots, n$. We also call l a *supercard value.* Any card with the number $n+1$ on it is a supercard of value 1. From the rest of the cards with numbers $1, \ldots, n$ we form supercards as follows: we choose any $n+1$ cards; then using the above lemma, we can choose several of them with the sum divisible by $n+1$; these will form a supercard by definition. The algorithm stops when we are left with less that $n+1$ cards. But now, the sum of the leftovers must also be divisible by $n+1$, since the total sum and the sum on each supercard are both divisible by $n+1$. This means that they form themselves a supercard, whose sum does not exceed $n(n+1)$.

So we have a pile of supercards with values $1, \ldots, n$ and the total sum of the values is equal to $\frac{k \cdot (n+1)!}{n+1} = k \cdot n!$. Now using the induction hypothesis, we can split the supercards into k stacks with the sum of the values in each being equal to $n!$. Hence the sum of the cards in each stack is $n! \cdot (n+1) = (n+1)!$, as we wanted.

Example 7.7. (**St. Petersburg**) Santa Claus has $n \geq 2$ different presents and several identical bags. Each bag contains exactly two objects: either two bags, or a bag and a present, or two presents. In particular, the big bag of presents Santa Claus has contains two objects. How many ways are there to split the presents into bags?

Solution. We show by induction on n that the number of ways is

$$(2n-3)!! := 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-3)$$

and also that the number of needed bags is $n - 1$.

For the induction step, assume that we have $n + 1$ presents.

The $(n+1)$-st present lies in some bag, together with either another bag or another present. If we remove the $(n+1)$-st present and the bag it lies in (only the bag, not the content), then the remaining "setup"is a valid distribution of presents. Conversely, given a valid distribution of n presents, to obtain a valid distribution of $n+1$ presents, it suffices to "add"(concatenate) the $(n+1)$-st present to any of the existing objects (the object can be a bag or a present), and encompass both in a new bag.

By induction, for n presents we use $2n-1$ objects: n presents and $n-1$ bags, so to add a new object we have $2n-1$ choices for the existing object. It is easy to see that different configurations for n presents (which can be done in $(2n-3)!!$ ways by the induction hypothesis), and different choices of the $2n-1$ objects which we will concatenate the new object to - correspond to different configurations for $n+1$ objects, so for $n+1$ presents the answer is $(2n-1)!!$.

Example 7.8. Consider an $m \times n$ grid with all its entries being integers, where $m > 1$ and $n > 1$ are positive integers. We know that if the difference of two integers of the same row is greater than 1, then the same row contains a number which is bigger than one of the two numbers and smaller than the other one. Prove that there is an integer which appears in every row or in every column of the grid.

Solution. We begin by proving the following result:

Lemma. Consider a finite sequence of integers such that, if the difference of two of its elements is greater than 1, then there is some integer greater than one of those numbers and smaller than the other one which is also an element of the sequence. Then any integer greater than one of those numbers and smaller than the other one is an element of the sequence.

Proof. Let $a > b$ be two integers of the sequence such that $a - b > 1$ and let c be such that $b < c < a$. We prove by induction on $a - b$ that c is also a member of the sequence.

The base case is $a - b = 2$, which follows immediately from the hypothesis.

Assume now that the result holds when $a - b = 2, 3, \ldots, k$, $k \in \mathbb{Z}$, $k \geq 2$. We show that it also holds for $a - b = k + 1$. From hypothesis, there is a number c such that $b < c < a$. Then $a - c \leq k$ and $c - b \leq k$, so from the induction hypothesis, all the numbers between b and c and all numbers between c and a are taken, establishing the result for $k + 1$.

Let us now proceed to the proof of the initial problem. First consider the set A consisting of the smallest number in each column. Now let k be the largest element of A.

If k occurs in all columns, then the statement is proven.

If k does not occur in some column, then we consider that column together with the column where k is written. Pick now any row. If k is not in that row, then according to the lemma and the choice of k, in one of the two common squares of that row and the considered two columns there is a number greater k and in the other there is a number smaller than k. Again, according to the lemma we obtain that k is written in the chosen row. It follows that the number k is written in every row, establishing the result.

Example 7.9. Let S be a set with 2002 elements, and let N be an integer with $0 \leq N \leq 2^{2002}$. Prove that it is possible to color every subset of S either black or white so that the following conditions hold:

a) the union of any two white subsets is white;

b) the union of any two black subsets is black;

c) there are exactly N white subsets.

Solution. We replace 2002 by n and prove the statement by induction on n. The base case is clear.

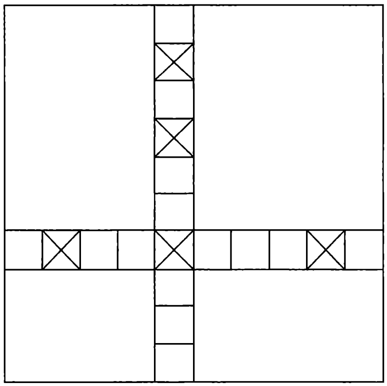
Assume now that the result holds for all integers which are less than some $n \geq 0$ and we prove it for n. In order to do this, we add in the induction hypothesis the extra condition that the empty set is white (it cannot be done if $N = 0$, but this case is simple and we exclude it below).

If $\left\lceil \frac{N+1}{2} \right\rceil = k$, then $1 \leq k \leq 2^n$. If $S = \{1, 2, \ldots, n\}$ then we may partition the power set of $\{1, 2, \ldots, n-1\}$ in black and white subsets as in the condition such that there are k white subsets $S_1, S_2, \ldots, S_k$. For $n = 2k$, we can take the subsets $S_1, S_1 \cup \{n\}, S_2, S_2 \cup \{n\}, \ldots, S_k, S_k \cup \{n\}$ to be white and the rest black. This coloring satisfies the condition of the problem: If $S_i \cup S_j = S_k$ then $(S_i \cup \{n\}) \cup (S_j \cup \{n\}) = S_k \cup \{n\}$ and analogously for the second cases. The union of two black subsets is black by the same principle. For $n = 2k - 1$, as the empty set is among $S_1, S_2, \ldots, S_k$ by the induction step, $\{n\}$ is among $S_1, S_1 \cup \{n\}, S_2, S_2 \cup \{n\}, \ldots, S_k, S_k \cup \{n\}$. Recoloring $\{n\}$ black, the conditions of a) and b) still hold, as it is easy to see using the very same principle above. It it clear that in both cases the empty set is white.

Example 7.10. A set M consisting of unit squares from an $n \times n$ table is called *convenient* if each row and each column of the table contains at least two squares belonging to the set. For each $n \geq 5$ determine the maximum $m \in \mathbb{Z}$ for which there exists a convenient set M of m squares such that when we remove any square from M, the resulting set is not convenient.

Solution. We begin by proving that $m \leq 4n - 8$.

We call a *cross* the union of some row and some column of the $n \times n$ table such that both the row and the column contain at least 3 unit squares from the given convenient set M, and moreover, their intersection is a square in M (see the figure below).



The upper bound on m now follows from the following:

Claim. If some set M contains $m \geq 4n - 7$ squares of the $n \times n$ table, then there exists at least one cross in the table.

We prove the claim by induction with respect to n. We first prove the claim when $n = 5, 6, 7$:

Consider one of the rows or the columns of the table which contains the maximum number k of squares from M among all the rows and columns. Note that $k \geq 3$. Let us assume without loss of generality that these k squares are at the very top of the leftmost column.

Assume that any of the top k rows contains at most two squares from M, as otherwise the claim would be proved. Hence, the total number of squares from M in the top k rows is at most $2k$. The remaining $n-k$ rows contain in total at most $k(n-k)$ squares. It follows that

$$4n-7 \leq m \leq 2k + k(n-k), \quad \text{or} \quad k^2-(n+2)k+4n-7 \leq 0.$$

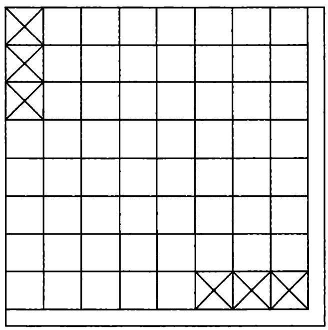
However, the discriminant of the left hand side is

$$D = (n+2)^2 - 4(4n-7) = (n-4)(n-8) < 0, \quad \text{for} \quad n = 5, 6, 7.$$

Thus, $k^2-(n+2)k+4n-7 > 0$. This gives a contradiction.

For the induction step, we prove that if the statement holds when $n \leq k$, $k \geq 7$, then it also holds when $n = k+3$.

Firstly, notice that there is a row containing at least 3 squares from the given set M (otherwise $m \leq 2n$). Similarly, there is a column containing at least 3 squares from M. Consider such a row and such a column. Without loss of generality, we can assume that this column is the leftmost one in the table and the squares from M belonging to it are all at the very top of the column. We make the corresponding assumption for the row (see the figure below).

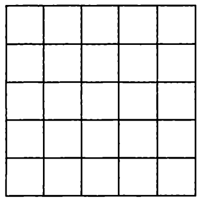


Consider one of the top 3 rows. If it contains at least 3 squares from M, then this row together with the leftmost column give us the needed cross. Hence we may assume that any of the three top rows, as well as any of the three rightmost columns contain at most 2 squares from M. Thus we have in total at most 12 squares from M in the shaded area. Now, since

$$m \geq 4n-7 = 4(k+3)-7 = 4k-7+12,$$

we conclude that the remaining $k \times k$ table contains at least $4k - 7$ squares from M. Therefore, by the induction hypothesis, this $k \times k$ table contains a cross which is also a cross for the initial $n \times n$ table. This completes the proof of our claim.

To complete the question we illustrate below the example when $m = 4n - 8$.



7.1.2 Graph Theory

A **graph** is an ordered pair $G = (V, E)$, where V is the vertex set, whose elements are the vertices (or the nodes) of the graph and E is the edge set, whose elements are the connections between vertices (called edges) of the graph. The **order** of a graph is the number $|V|$ of vertices in it. The **size** of a graph is the number of edges, $|E|$. The trivial graph is the one with $|V| = 1$ and $|E| = 0$. For two graphs $G = (V, E)$ and $G' = (V', E')$, we say that G' is a **subgraph** of G if $V' \subseteq V$ and $E' \subseteq E$.

An **undirected graph** is a graph where all the edges are bidirectional. In contrast, a graph where the edges point in a direction is called a **directed graph**. For $u, v \in V$, we write (u, v) to represent that the arc is oriented from x to y. When the graph is undirected, we usually write $\{u, v\}$ for the edge between u and v instead.

The notion of a graph can be adapted to allow additional features. For our purposes, it will suffice to consider only edges joining two distinct vertices and at most one edge joining each pair of vertices. Such graphs are called **simple graphs**.

The **degree** of a node, denoted by $\deg(v)$ is the number of edges incident to it.

For $u, v \in V$, a $u - v$ **walk** is a sequence of vertices which starts at u, ends at v and two consecutive vertices in the sequence are connected by an edge from E (i.e. they are **adjacent**). The **length** of the walk is defined as the number of edges used in that walk. A **trail** is a $u - v$ walk where no edge is repeated. A **circuit** is a trail for which $u = v$ (the first and last vertex in the sequence are the same). A **path** is a $u - v$ walk where no vertex is repeated and a **cycle** is a walk for which $u = v$ but there is no further repetition of vertices. A **Hamiltonian cycle** is a cycle which uses each node of the graph exactly once and an **Euler circuit** is a circuit which uses every edge exactly once. We say that $u, v \in V$ are **connected** if there is a $u - v$ path. A graph is **connected** if any pair $u, v \in V$ is connected.

An undirected graph is called a **tree** if any two vertices are connected by exactly one path. Equivalently, a tree is an undirected graph which is connected and has no cycles. A **forest** is a disjoint union of trees.

A **complete graph** is a graph where all pairs of vertices are connected by an edge. A complete graph with $|V| = n$ is denoted by K_n.

A **bipartite graph** is a graph $G = (V, E)$ whose vertex set V may be partitioned into two disjoint sets A and B such that every edge $e \in E$ has one endpoint in A and the other endpoint in B. The set A is often called the input set and B is called the output set. A **perfect matching** in a bipartite graph G is an injective map $f : A \to B$ such that for every $x \in A$, there is an edge $e \in E$ with endpoints x and $f(x)$. For any subset $X \subseteq A$, define $N(X)$ to be the set of all vertices $y \in B$ that are endpoints of edges with one endpoint in X.

Example 7.11. (**Hall's Matching Theorem**). Let $G = (V, E)$ be a bipartite graph with input set A and output set B. There exists a perfect matching $f : A \to B$ if and only if for every subset $X \subseteq A$ we have

$$|N(X)| \geq |X|. \tag{1}$$

Solution. We prove the result by induction on $|A|$. The base case is $|A| = 1$, for which the result is clear.

Assume now that the result holds if $|A| \leq n$, some $n \geq 1$, and consider a bipartite graph G whose input set A has $n + 1$ elements. We distinguish two cases:

Case 1. For every proper subset $X \subset A$, we have $|N(X)| \geq |X| + 1$. Then we choose any $x \in A$ and $y \in N(\{x\})$ (notice that $N(\{x\})$ has at least one element from the induction hypothesis). Let G^* be the bipartite graph with input set $A^* = A \backslash \{x\}$, output set $B^* = B \backslash \{y\}$, and whose edges are the same as those of G, except for those incident to either x or y, which are deleted.

Then the bipartite graph G^* satisfies the condition (1) and by the induction hypothesis, there is a perfect matching in G^*. This perfect matching extends to a perfect matching in the original graph G by setting $f(x) = y$.

Case 2. There exists a proper subset $X \subset A$ such that $N(X) = |X|$. Then for such a proper subset X, we construct bipartite graphs G^* and G^{**} with input sets $A^* = X$ and $A^{**} = A \setminus X$, output sets $B^* = N(X)$, $B^{**} = B \setminus N(X)$ and edges inherited from the original graph G. We shall use the induction hypothesis to show that there is a perfect matching in each of G^* and G^{**}. Combining the two matchings, this would give a perfect matching for G.

Notice that neither A^* nor A^{**} has cardinality greater than n, since $X = A^*$ is a proper subset of A. Therefore it suffices to show that the condition (1) holds for both G^* and G^{**}. For the graph G^* this is easy, since for any subset $Y \subseteq X = A^*$, the corresponding set $N(Y)$ is the same in G^* and G.

For the graph G^{**}, if there was a subset $Y \subseteq A^{**} = A \setminus X$ whose corresponding set $N^{**}(Y)$ in G^{**} had fewer than $|Y|$ elements, then in the graph G, the set $N(Y \cup X)$ would have at most $|Y \cup X| - 1$ elements, because $N(Y \cup X) = N^{**}(Y) \cup N(X)$. This is impossible, because the graph G satisfies (1). Hence G^{**} also satisfies (1), completing our proof.

A **planar graph** is a graph G drawn in the plane in such a way that the edges do not meet except at common endpoints. The edges of a planar graph divide the plane up into regions which are called **faces**. One of the faces extends out to infinity and is called unbounded. The rest are bounded.

Example 7.12. (**Euler's Theorem.**) In a connected planar graph with V vertices, F faces, and E edges

$$V + F = E + 2.$$

Solution. We prove the result by inducting on the number E of edges. If the connected planar graph G has no edges, it is an isolated vertex and $V+F-E = 1+1-0 = 2$. This is the base case.

For the induction step, assume that the result holds for all graphs with fewer than E edges. We will prove it holds for graphs with E edges. Choose any edge e. Either side of the edge e borders a face. If they are different faces, then following the remaining sides of either face gives a path between the endpoints of e. Thus we can remove e without disconnecting the graph. This will cause the two faces bordering e to merge. Thus removing e reduces E and F by one. Thus the result follows by the induction hypothesis.

Suppose the two sides of e border the same face. Then we could draw a short segment crossing e and join its endpoints with a curve inside this face. This gives a simple closed curve that crosses e exactly once and crosses no other edge of G. Thus deleting e will split G into two components, one inside this curve and one outside. (There must be two components since one endpoint of e is inside the curve and the other is outside and there can be at most two since each component must contain an endpoint of e.) Suppose these components have V_i vertices, E_i edges, and F_i faces for $i = 1, 2$. Since each has fewer than E edges, by the induction hypothesis $V_i - E_i + F_i = 2$ for $i = 1, 2$. Since no vertices were added or lost $V = V_1 + V_2$. Since only e was deleted, $E = E_1 + E_2 + 1$. Since the components have $F_i - 1$ internal faces and one (shared) external face, $F = F_1 + F_2 - 1$. Thus adding the Euler's formulas for the two components gives $V - (E - 1) + (F + 1) = 4$ or $V - E + F = 2$. This completes the induction step and the proof.

Example 7.13. (**Ramsey's Theorem.**) Let $k \geq 2$ be a positive integer and let $n_1, n_2, \ldots, n_k$ be any positive integers. Then there exists a number $R(n_1, \ldots, n_k)$ such that if the edges of a complete graph of order $R(n_1, \ldots, n_k)$ are colored with k different colors, then for some $i = 1, \ldots, k$, the graph contains a complete subgraph of order n_i whose edges are all of color i. (The minimum possible $R(n_1, \ldots, n_k)$ is called the *Ramsey number* for $n_1, \ldots, n_k$).

Solution. We first prove the result when $k = 2$, using induction on $n_1 + n_2$. Notice that for any $n_1 \geq 1$, we have $R(n_1, 1) = R(1, n_2) = 1$, since a complete graph with only one node has no edges. This establishes the base case.

For the induction step, we assume that both $R(n_1 - 1, n_2)$ and $R(n_1, n_2 - 1)$ exist for some positive integers n_1 and n_2, and we prove that $R(n_1, n_2)$ exists. We shall prove the following:

$$R(n_1, n_2) \leq R(n_1 - 1, n_2) + R(n_1, n_2 - 1).$$

To show the above inequality, consider a complete graph G on $R(n_1 - 1, n_2) + R(n_1, n_2 - 1)$ vertices. Let v be a vertex of this graph and partition $G \setminus \{v\}$ into two sets A and B such that for all $w \in G \setminus \{v\}$, we have $w \in A$ if $\{v, w\}$ has color 1 and $w \in B$ if $\{v, w\}$ has color 2.

Notice that G has $|A| + |B| + 1$ vertices, so either $|A| \geq R(n_1 - 1, n_2)$, or $|B| \geq R(n_1, n_2 - 1)$. Without loss of generality, assume that $|A| \geq R(n_1 - 1, n_2)$. If A has a K_{n_2} which is all of color 2, then so does G and we are done. Otherwise, A has a $K_{n_1 - 1}$ which is of color 1, so $A \cup \{v\}$ has a subgraph K_{n_1} of color 1, by the way we defined A. This completes our induction for the case $k = 2$.

For the general case, we induct on the number k of colors. The base case $k = 2$ was proved above (notice also that the case $k = 1$ can be considered, but this is obvious). Assume now that $k > 2$ and that the result holds for all integers less than k. We prove that

$$R(n_1, \ldots, n_k) \leq R(n_1, \ldots, n_{k-2}, R(n_{k-1}, n_k)).$$

Let G be a complete graph on $R(n_1, \ldots, n_{k-2}, R(n_{k-1}, n_k))$ vertices whose edges are colored with k colors. The key step is the following trick:

Assume for the moment that the colors $k-1$ and k are the same, so G is now colored with only $k-1$ colors. From the induction hypothesis, we know that either G has a subgraph K_{n_i} all colored in i for some $1 \leq i \leq k-2$, or it has a subgraph $K_{R(n_{k-1}, n_k)}$ all colored in the color $k-1$ (which is the same as k for now). In the former situation, our assumption on the colors $k-1$ and k was never used, so we are done. In the latter situation, we remove our assumption on the colors $k-1$ and k and by the way we defined $R(n_{k-1}, n_k)$, we have that $K_{R(n_{k-1}, n_k)}$ contains either a $K_{n_{k-1}}$ of color $k-1$, or a K_{n_k} of color k. This completes our proof.

A **k-clique** in a graph G is a set of k vertices such that any pair of them is connected by an edge. Equivalently, a k-clique is a subgraph which is a copy of the complete graph K_k with k vertices.

Example 7.14. (Turan's Theorem.) If $G(V, E)$ is a graph on n vertices without a k-clique, then

$$|E| \leq \frac{(k-2)n^2}{2(k-1)}.$$

Solution. We prove the result by induction on n. The base cases given by the small values for n are an immediate check.

Let G be a graph on $V = \{1, \ldots, n\}$ vertices without k-cliques and which has the maximal number of edges. G definitely contains $(k-1)$-cliques, since otherwise we could simply add an edge. So let A be a $(k-1)$-clique, $B = V \setminus A$, so that $|B| = n - k + 1$. We know that A contains $\binom{k-1}{2}$ edges. We estimate the number e_B of edges in B and the number e_{AB} of edges between A and B. From the induction hypothesis, we have that

$$e_B \leq \frac{(k-2)(n-k+1)^2}{2(k-1)}.$$

As G has no k-clique, every vertex in B must be adjacent to at most $k-2$ vertices in A, from which we obtain that

$$e_{AB} \leq (k-2)(n-k+1).$$

Adding the obtained inequalities, we obtain that

$$|E| \leq \binom{k-1}{2} + \frac{(k-2)(n-k+1)^2}{2(k-1)} + (k-2)(n-k+1) = \frac{(k-2)n^2}{2(k-1)}.$$

Remark. The upper bound given in the theorem can be achieved for a graph G with $n = q(k-1)$ vertices where the vertices are divided into $k-1$ groups of size q and two vertices are joined by an edge if and only if they are in different groups. If n is not a multiple of $k-1$, then the bound above is not quite sharp. With a little more work on the base cases, one can show that for general n the graph with no k-clique and the maximum number of edges can be built as follows: divide the vertices into $k-1$ groups of $\lfloor \frac{n}{k-1} \rfloor$ or $\lfloor \frac{n}{k-1} \rfloor + 1$ elements, then join two vertices by an edge if and only if they belong to different groups.

Example 7.15. (**St. Petersburg**) Is it possible to choose several points in space and connect some of them by line segments, such that each vertex is the endpoint of exactly 3 segments, and any closed path has length at least 30?

Solution. Yes. In graph theoretic language, we show by induction that there exists a graph in which every vertex has degree precisely 3, and every cycle has length at least n.

The base case $n = 3$ is straightforward -consider a tetrahedron.

Assume now that the result holds for some $n \geq 3$ and consider such a graph for n. Assume it has m edges. Label the edges from 1 to m and for every edge k, write the number 2^k near one of its endpoints. Let M be a large number, say $M > n \cdot 2^m$. Replace each vertex by a group of M new vertices, labelled from 0 to $M-1$.

For every edge in the original graph, draw M edges in the new graph by the following rule: assume the edge is k and it joins vertices a and b, and that 2^k is written near a. Let A and B be the groups of M new vertices corresponding to a and b. For every $i = 0, \ldots, M-1$ join the vertex i from A to the vertex

$i + 2^k$ in B (indices are mod M). Note that the vertex i of group B will be connected to the vertex $i - 2^k$ of group A.

It is clear that every vertex in the new graph is connected to exactly 3 others (one in each group corresponding to the 3 edges in the initial graph). It remains to show that any cycle in the new graph has length at least $n + 1$. Assume the contrary, and let $a_1 a_2 \ldots a_\ell$ be a cycle of length $\ell \leq n$. Consider the groups corresponding to these vertices. They give us a cycle $b_1 \ldots b_\ell$ in the original graph (possibly self-intersecting). Note that this cycle cannot use the same edge twice in a row, as in the new graph the neighbours of any vertex belong to different groups. So we can extract a simple cycle from $b_1 \ldots b_\ell$. Since in the original graph every cycle has length $\geq n$, it follows that we must have $\ell = n$ and $b_1 \ldots b_\ell$ is itself a simple cycle.

Now note that the number of a_{i+1} differs from the number of a_i by $\pm 2^{k_i}$ (mod M), where k_i is the number of the edge $a_i' a_{i+1}'$ in the original graph. Thus

$$\pm 2^{k_1} \pm \ldots \pm 2^{k_\ell} \equiv 0 \pmod{M},$$

for some $\pm$'s. Since $M > n \cdot 2^m$, it follows that the above expression on the left side must equal 0. Since all the k_i's are distinct, this is impossible. This completes our proof.

Example 7.16. Prove that if a graph G with n vertices is a tree (connected and containing no cycles), then it has exactly $n - 1$ edges.

Solution. The key observation is the following. Suppose we select a vertex a vertex of G and start to "walk" on the edges from that vertex, forming a trail $v_1, v_2, \ldots$. The trail cannot self-intersect, as then we would have a cycle, so the graph would not be a tree. Thus the walk will eventually stop at some vertex of degree 1.

We now prove the statement by induction on n. Things are clear for the base case $n = 1$.

Now assume we have proved the result for trees with $n \geq 1$ vertices. Consider a tree with $n + 1$ vertices and using the remark above find a vertex of degree 1. We remove that vertex together with its adjacent edge. The remaining graph is a tree (it has no cycles and it is still connected as no path can go through the removed vertex) and has n vertices, so we can apply induction:

we have $n-1$ edges in the graph with n vertices which, together with the removed edge, make n edges in the graph with $n+1$ vertices.

Example 7.17. Let G be a connected graph with an even number of vertices. Prove that you can select a subset of edges of G such that each vertex is incident to an odd number of selected edges.

Solution. All connected graphs have a tree as a spanning subgraph, so it suffices to prove the problem for trees. We do this by induction on the number of vertices. The base case is clear.

Suppose now that the statement is true for even integers less than n, with n even. In a tree T with n vertices, choose an arbitrary vertex v with degree $d > 1$. Then deleting v from T results in a forest of d trees: $T_1, T_2, \dots T_d$. Let

$$U_i = \begin{cases} T_i & \text{if } |V(T_i)| \quad \text{is even;} \\ T_i \cup \{v\} & \text{otherwise.} \end{cases}$$

Note that since T has an even number of vertices, $T - \{v\}$ has an odd number of vertices, so an odd number of T_i have an odd number of vertices and an odd number of U_i contain v. Each U_i is a tree with an even number of vertices, so the inductive hypothesis can be applied to each of them. The edges corresponding to two different U_i will never overlap, and if we combine all the selected edges, the only vertex that might not work is v. But since v is contained in an odd number of U_i and an odd number of odd numbers sum to an odd number, it will have an odd number of edges adjacent to it. So this selection works and the induction is done.

Example 7.18. Let G be a simple graph with $2n+1$ vertices and at least $3n+1$ edges. Prove that there exists a cycle having an even number of edges.

Solution. We will proceed by induction on the number of vertices of the graph to show that whenever we have $f(n) = \left\lceil \frac{3(n-1)}{2} \right\rceil + 1$ edges there is an even cycle. The base case $n = 1$ is clear.

Suppose we proved the result for up to $n-1$ vertices, where $n \geq 2$. If our graph G has n vertices and $f(n)$ edges, note that if we can find a vertex v so that $\deg(v) \in \{0, 1\}$ then $G \setminus \{v\}$ has an even cycle by the induction hypothesis.

Suppose we can find adjacent vertices u and v of degree 2. Say u is adjacent to u_1 and v and v is adjacent to u and v_1. If there is already an edge from u_1 to v_1, then we already have a cycle of length 4. Otherwise, we can consider $G \setminus \{u, v\}$ with an extra edge from u_1 to v_1. This reduces the number of vertices by 2 and the number of edges by 2. Thus by the induction hypothesis this graph has an even cycle. If this cycle uses the new edge from u_1 to v_1, then it becomes a cycle with 2 more vertices in G, but still of even length.

Now we are ready to prove the induction step. Let $P = v_1v_2\dots v_m$ be a path of longest length in G. Then there are edges v_1v_1' and v_nv_n' with $v_1', v_n' \in P$ because $\deg(v_1), \deg(v_n) \geq 2$ and P is maximal. Now there are two cases:

Case 1. If $\deg(v_1) \geq 3$. Then there is another edge v_1x with $x \in P$ and $x \neq v_1'$. Consider the union of the edges v_1x, v_1v_1' and the paths from v_1 to v_1' and from v_1 to x. They form a cycle with a chord and therefore we have an even cycle. (The chord splits the cycle into two other cycles. The sum of the lengths of these three cycles is even).

Case 2. If $\deg(v_1) = 2$, by the observation above we have $\deg(v_2) \geq 3$ and so there is an edge v_2y. If $y \in P$, then we can again find a cycle with a chord as above. If y is not in P, then its degree cannot be 1, but if y was connected to another $z \notin P$ then this contradicts the maximality of P. So $z \in P$. Now the union of v_1v_1', v_2y, yz and the path from v_1 to whichever is the farthest between z and v_1' is a cycle with a chord and so the induction step is complete.

Example 7.19. (**Russia 1995**) In the city of Dujinsk every intersection has exactly three streets meeting at that intersection (each street joins two intersections). The streets are colored in three colors: red, black and white in such a way that from every intersection has a street of each color. We call an intersection positive in the colors of the streets can be read as red, black and white in clockwise direction and negative if this can be done counterclockwise. Prove that the number of positive intersections differs from the number of negative intersections by a multiple of 4.

Solution. To tackle the question more easily, we consider a planar graph with nodes and edges instead of the city with intersections and streets. Now, in this new graph, one of the first ideas that comes in mind is induction on the number of nodes. So how can we use it? We cannot generally remove an edge or a node while still fulfilling the hypotheses. However, we can remove one edge if the following conditions hold: Suppose that on one face of the graph we have three consecutive edges with the colors red, black, red. Let these edges be AB, BC, CD in that order. Then the two white edges BE, CF must lie outside the face. Now we can remove B, C by joining A, D with a red edge and E, F with a white one. It is easy to see that if we draw these edges to follow approximately the contours of the paths $A - B - C - D$ and $E - B - C - F$, then we get a planar graph and we ensure that the signs of A, D, E, F do not change. We also see that B and C had opposite signs so the problem would follow by induction in this case.

What happens if cannot find such a configuration? This means that on every face the colors cycle so in particular every face has a length divisible by 3. However, we can easily handle this case if we remember one of the crucial facts in planar graph theory: in a planar graph, there is a vertex with degree less than 6, and by duality, a face with length at most 5. Since the faces must have length divisible by 3, we deduce that we have a triangle, say ABC with AB red, BC black and CA white. Now let AD, BE, CF be the other edges emerging from the vertices of the triangle. We can remove A, B, C by creating a new vertex T inside ABC and by joining it to D, E, F by edges of black, white and red color, respectively. We see that all of A, B and C had the same sign, which is opposite to the sign of T and so our difference has changed by ± 4, which is what we need.

The proof of the base case for this induction (when we have just four vertices) is left to the reader.

7.1.3 Combinatorial Geometry

As its name suggests, Combinatorial geometry is a field which blends elements from combinatorics and geometry. It deals mainly with arrangements of geometric objects and their discrete properties. We shall consider mostly problems that use topics such as covering, coloring, tiling or decomposition. One of the theorems that lies at the heart of Combinatorial Geometry is:

Example 7.20. (**Helly's Theorem**) Given $n \geq 4$ convex figures in the plane such that any three of them have a common point, prove that all n figures have a common point.

Solution. We will prove the statement by induction on n. We first prove the case $n = 4$.

Denote the given figures by M_1, M_2, M_3 and M_4. Let A_i be the intersection point of all the figures except M_i. Two arrangements for the points A_i are possible:

Case 1. One of the points, say A_4 lies inside the triangle formed by the remaining points. Since A_1, A_2, A_3 belong to the convex figure M_4, all points of $A_1A_2A_3$ also belong to M_4. Therefore A_4 belongs to M_4 and by construction it also belongs to the other figures.

Case 2. If $A_1A_2A_3A_4$ is a convex quadrilateral, let O be the intersection point of diagonals A_1A_3 and A_2A_4. Let us prove that O belongs to all the given figures. Both points A_1 and A_3 belong to figures M_2 and M_4, therefore the segment A_1A_3 belongs to these figures. Similarly, the segment A_2A_4 belongs to figures M_1 and M_3. It follows that the intersection point of segments A_1A_3 and A_2A_4 belongs to all the given figures.

Assume now that the statement holds for $n \geq 4$ figures and let us prove it for $n+1$ figures. Given convex figures $M_1, \ldots, M_n$, M_{n+1}, every three of which have a common point, let us consider the figures $M_1, \ldots, M_{n-1}, M_n'$, where M_n' is the intersection of M_n and M_{n+1}. Clearly M_n' is also a convex figure. Let us now prove that any three of the new figures have a common point. We only need to look at the triples of figures that contain M_n', since for the others, the statement is true by assumption. But now from the case $n = 4$ we know that for any i and j, the figures M_i, M_j, M_n and M_{n+1} always have a common point. So M_i, M_j and M_n' will have a common point. By the induction hypothesis, we are done.

Let us now study some further examples.

Example 7.21. The sides and the diagonals of a regular n-gon X are colored in k colors such that:

i) For each color a and any two vertices A, B of X either the segment AB is colored in a or there is a vertex V such that AC and BC are colored in a.

ii) The sides of every triangle with vertices among the vertices of X are colored in at most two colors.

Prove that $k \leq 2$.

Solution. Assume we have at least three colors, say $1, 2, 3$. Let A, B, C be points such that edge AB has color 1 and edges AC, BC have color 2. There exists a vertex D such that DB and DC have color 3. As the triangles DAB, DAC are colored in at most two colors, we conclude that DA is also of color 3. Then we add a vertex E such that EC, ED have color 1. As ECB, EDB are colored in at most two colors, EB also has color 1. So as EDA and ECA must be colored in two colors, EA has also color 1. This suggests us to try to prove the following statement:

"There exist vertices $X_1, X_2, \ldots, X_n$ such that the edge X_iX_j has color $1 + (j+1) \pmod 3$, for $i < j$."

The proof is done by induction on n, the base case being the vertices $A, \ldots, E$ described above.

For the induction step, notice that the side $X_{n-1}X_n$ has color $1 + (n+1) \pmod 3$, therefore there exists a vertex X such that both $X_{n-1}X$ and X_nX have color $1 + (n+2) \pmod 3$. Clearly X is not among $X_1, X_2, \ldots, X_{n-2}$ because otherwise XX_n would be of color $1 + (n+1) \pmod 3$ which is different from $1 + (n+2) \pmod 3$. Then, as the triangles $XX_{n-1}X_{n-2}$ and XX_nX_{n-2} are colored in at most two colors each, we conclude that XX_{n-2} has also color $1 + (n+2) \pmod 3$. Similarly, for $1 \leq i \leq n-3$, the triangles $XX_nX_i, XX_{n-1}X_i, XX_{n-2}X_i$ are all colored in at most two colors, so we conclude that XX_i has color $1 + (n+2) \pmod 3$. So we may set $X_{n+1} = X$. The induction is complete.

From the above result, we conclude that the polygon has infinitely many vertices, which is a contradiction.

For the next example, we will require the following famous result, known as the Sylvester-Gallai Theorem, which asserts that given a configuration of $n \geq 3$ points in the plane, either all of the points are collinear, or there is a line which passes through exactly two points of the configuration.

Example 7.22. We have $n \geq 3$ points, not all of them collinear. Prove that these points determine at least n lines.

Solution. We will proceed by induction on n. For $n = 3$, the result is clear.

Assume now the statement is true for some $n \geq 3$. For the points $\{P_1, \ldots, P_{n+1}\}$, by the Sylvester-Gallai theorem, there is a line passing through only two of them, without loss of generality say that it is the line connecting P_{n+1} and P_n. Note that if all of $P_1, \ldots, P_n$ are collinear, we can easily get $n+1$ lines from connecting P_{n+1} to each of them and are done. Otherwise, removing P_{n+1} and the line connecting it to P_n we get a set of n points that are not all collinear. Thus we can apply the induction hypothesis and get at least n lines. Adding the extra line determined by P_n and P_{n+1}, we get at least $n+1$ lines.

Example 7.23. The set M consists of the four vertices of a square with side length 20 and 1999 more points in the interior of this square. Prove that there is a triangle with area at most $\frac{1}{10}$, formed by the vertices of M.

Solution. It is enough to prove that the square can be divided into at least 4000 triangles with vertices in M, so one of them will have area at most $\frac{20^2}{4000} = \frac{1}{10}$. Clearly the number 1999 should not be special, so let us prove by induction on n that if M contains the vertices of the square and n points in the interior, then we can partition the square into at least $2n+2$ triangles formed by the vertices of M. The base case $n = 1$ is true, because a square with a point in the interior can be broken into 4 triangles.

Now, assume we have proven the assertion for $n = k$ and let us prove it for $n = k+1$. Take a point P from M, different from the vertices if the square. If we remove P, we can divide the square into $2k+2$ triangles with vertices in $M \setminus \{P\}$ by the induction hypothesis. Therefore, P either falls inside a triangle or it falls on the side of one of the triangles. If P falls inside the triangle XYZ, we can divide XYZ into three triangles PXY, PYZ, PZX, thus increasing the number of triangles by 2. If P falls on a segment AB, then AB belongs to at least two different triangles ABC, ABD. We can divide them into the triangles APC, BPC, APD, BPD, again increasing the number of triangles by 2. So we get at least $2k+4$ triangles and the induction step is proved.

Example 7.24. **(IMO 2013 shortlist)** In the plane, 2013 red points and 2014 blue points are marked so that no three of the marked points are collinear.

One needs to draw k lines not passing through the marked points and dividing the plane into several regions. The goal is to do it in such a way that no region contains points of both colors.

Find the minimal value of k such that the goal is attainable for every configuration of 4027 points.

Solution. The answer is $k = 2013$. We will actually prove the following, more general statement:

If n points in the plane, no three of which are collinear, are colored in red and blue arbitrarily, then it suffices to draw $\lfloor \frac{n}{2} \rfloor$ lines to reach the goal.

Proof. We proceed by induction on n. If $n \leq 2$, then the statement is obvious. Now assume that $n \geq 3$, and consider a line l containing two marked points A and B such that all the other marked points are on one side of l; thus l is a line containing a side of the convex hull of the marked points. In particular, A and B are on the convex hull of the marked points.

Remove for a moment the points A and B. By the induction hypothesis, for the remaining configuration it suffices to draw $\lfloor \frac{n}{2} \rfloor - 1$ lines to reach the goal. Now return the points A and B back. Three cases are possible:

Case 1. If A and B have the same color, then one may draw a line parallel to l and separating A and B from the other points. Obviously, the obtained configuration of $\lfloor \frac{n}{2} \rfloor$ points works.

Case 2. If A and B have different colors, but they are separated by some drawn line, then again the same line parallel to l works.

Case 3. Finally, assume that A and B have different colors and lie in one of the regions defined by the drawn lines. By the induction hypothesis, this region contains no other points of one of the colors. Without loss of generality, assume the only blue point it contains is A. Then it suffices to draw a line separating A from all the other points, which can be done since A is on the convex hull of the marked points.

Thus the step of the induction is proved and we are done.

Example 7.25. (**JBMO 2004**) Consider a convex polygon having n vertices, $n \geq 4$. We arbitrarily decompose the polygon into triangles having all the vertices among the vertices of the polygon, such that no two of the triangles have interior points in common. We paint in black the triangles that have two sides that are also sides of the polygon, in red if only one side of the triangle is also a side of the polygon and in white those triangles that have no sides that are sides of the polygon.

Prove that there are two more black triangles that white ones.

Solution. We prove by induction that there are at least two black triangles, and that $b = w+2$ where b (resp. w) denote the number of black (resp. white) triangles.

First, for $n = 4$ there are two black triangles and no red and white. This establishes the base case.

Now, let $n > 4$ be given, and assume that the result holds for any k such that $4 \leq k < n$. Let us consider an n-gon colored as in the statement of the problem. Let d be any diagonal (not side) which has been drawn for the triangulation. Then d separates the n-gon into two disjoint polygons. If any of these polygons is a triangle, then it is a black one in the original coloring. If none of these polygons is a triangle, then use the induction hypothesis for each of the two sub-polygons. This gives us at least 2 black triangles on each side of d. Furthermore, for each side at most one of these black triangles have d as a side. Thus, in any case, there is at least one black triangle on each side of d, which ensures that there are at least two black triangles for the n-gon.

Now let $T = ABC$ be a black triangle (which does exist from above), with vertices in that order on the n-gon. Consider the $n-1$-gon obtained by deleting T and replacing it by the edge AC. The coloring induced on this $(n-1)$-gon is the same as the initial one, except for the triangle T' having AC as side. Note that since AC is a diagonal of the n-gon, then T' was red or white in the original coloring, and since it is a side of the $(n-1)$-gon, it is now black or red, respectively. From the induction hypothesis, we have $b' = w' + 2$ for the $(n-1)$-gon.

If T' is black for the induced coloring, then it was red for the initial one. Adding T, it leads to $b = b' - 1 + 1 = b'$ and $w = w'$.

If T' is red for the induced coloring, then it is white for the initial coloring. Adding T leads to $b = b' + 1$ and $w = w' + 1$.

Thus, in either case, we have $b = w+2$ as desired, which ends the induction step and the proof.

Example 7.26. (**Moscow 1996**) We are given a set A consisting of some lines in the plane. It is known that for any subset B of A consisting of k^2+1, $(k \geq 3)$ lines, there exist k points such that any line which is in B passes through at least one of these points. Prove that we can also choose k points such that each line in A passes through at least one of them.

Solution. Let us first introduce some terminology. For a set X, we say that the lines in X pass through k nodes if there exist k distinct points in the plane such that each line in X passes through at least one of these points. We also say that the statement $S(n,k)$ is true if from the fact that any n lines of X pass through k nodes we can deduce that all lines in X pass through k nodes. With this setup, we are asked to prove that $S(n,k)$ holds for $n = k^2+1$, $k \geq 3$.

We shall prove the statement by induction on k:

$$S(3,1) \Rightarrow S(6,2) \Rightarrow S(10,3) \Rightarrow \ldots \Rightarrow S(k^2+1,k).$$

The base case for our induction is clear, since $S(3,1)$ says that if any three lines of a set pass through the same point, then all lines of that set pass through the same point.

First, assume that $S((k-1)^2+1, k-1)$ $(k \geq 4)$ holds and A is a set of lines such that any k^2+1 lines pass through k nodes. Let us consider k^2+1 lines in A and the k points they pass through. Then one of the k points, call it P, is part of at least $k+1$ lines. Let M denote the set of all lines in A passing through P and let A' be the set of lines in A which are not in M. We shall prove that for A' we have that any $(k-1)^2+1$ lines pass through $k-1$ nodes:

Indeed, to any set of $(k-1)^2+1$ lines in A' we add up to $2k-1$ of the lines in M and if necessary, some extra lines from A' to get a total of k^2+1 lines. Then, from the hypothesis of the problem, these k^2+1 lines pass through k nodes. Let N be the set of these nodes. We show that $P \in N$. Indeed, there are at least $k+1$ lines in M passing through P, so if $P \notin N$, then the lines pass through at least $k+1$ points, while we assumed that k points are enough. So it must be that $P \in N$. So excluding P, the remaining $k-1$ points are contained by all lines in A' not passing through P, so in particular by the $(k-1)^2+1$ lines from A' that we initially considered. From the induction hypothesis, $S((k-1)^2+1, k-1)$ is true, so all lines in A' pass through $k-1$ nodes. But then it follows that all lines in A pass through these $k-1$ nodes together with the point P, so k points in total. This finishes the proof of the induction step.

We still need to prove the first two steps of our chain of implications, $S(3,1) \Rightarrow S(6,2) \Rightarrow S(10,3)$. We leave it to the reader to check that these follow by just making slightly changes to the argument for the inductive step.

7.2 Proposed Problems

Problem 7.1. (**IMO 2002**) Let n be a positive integer and S the set of points $(x,\, y)$ in the plane, where x and y are non-negative integers such that $x + y < n$. The points of S are colored in red and blue so that if (x, y) is red, then (x', y') is red as long as $x' \leq x$ and $y' \leq y$. Let A be the number of ways to choose n blue points such that all their x-coordinates are different and let B be the number of ways to choose n blue points such that all their y-coordinates are different. Prove that $A = B$.

Problem 7.2. (**IMO 2013 shortlist**) Let n be a positive integer. Find the smallest integer k with the following property: Given any real numbers $a_1, \ldots, a_d$ such that $a_1 + a_2 + \ldots + a_d = n$ and $0 \leq a_i \leq 1$ for $i = 1, 2, \ldots, d$, it is possible to partition these numbers into k groups (some of which may be empty) such that the sum of the numbers in each group is at most 1.

Problem 7.3. (**TOT 2002**) The spectators are seated in a row with no empty places. Each is in a seat which does not match the spectator's ticket. An usher can order two spectators in adjacent seats to trade places unless one of them is already seated correctly. Is it true that from any initial arrangement, the usher can place all the spectators in their correct seats?

Problem 7.4. (**USSR 1991**) Several (more than two) consecutive positive integers $1, 2, \ldots, n$ are written on a blackboard. In one move, it is permitted to erase any pair of numbers, say p and q and replace them by the numbers $p + q$ and $|p - q|$ instead. In several moves, a student was able to make all numbers on the blackboard equal to k. Find all possible values of k.

Problem 7.5. (**USSR 1990**) We are given $4m$ coins, among which exactly half of the coins are counterfeit. All genuine coins have equal weights, all counterfeit coins also have equal weights, but a counterfeit coin is lighter than a genuine one. How can one determine all counterfeit coins in no more than $3m$ weighings, using a balance without weights?

Problem 7.6. (**IMO 2006 shortlist**) An (n, k)-tournament is a contest with n players held in k rounds such that:

i) Each player plays in each round, and every two players meet at most once.

ii) If player A meets player B in round i, player C meets player D on round i, and player A meets player C in round j, then player B meets player D in round j.

Determine all pairs (n,k) for which there exists an (n,k)-tournament.

Problem 7.7. In an $m \times n$ rectangular board of mn unit squares, *adjacent* squares are ones with a common edge, and a *path* is a sequence of squares in which any two consecutive squares are adjacent. Each square of the board can be colored black or white. Let N denote the number of colorings of the board such that there exists at least one black path from the left edge of the board to its right edge, and let M denote the number of colorings in which there exist at least two non-intersecting black paths from the left edge to the right edge. Prove that $N^2 \geq 2^{mn}M$.

Problem 7.8. (**IMO shortlist 1998**) Let $U = \{1, 2, \ldots, n\}$, where $n \geq 3$. A subset S of U is said to be *split* by an arrangement of the elements of U if an element not in S occurs in the arrangement somewhere between two elements of S. For example, 13542 splits $\{1,2,3\}$ but not $\{3,4,5\}$. Prove that for any $n-2$ subsets of U, each containing at least 2 and at most $n-1$ elements, there is an arrangement of the elements of U which splits all of them.

Problem 7.9. Let $n \neq 4$ be a positive integer. Consider a set $S \subseteq \{1, 2, \ldots, n\}$ with $|S| > \left\lceil \frac{n}{2} \right\rceil$. Prove that there exist $x, y, z \in S$ with $x + y = 3z$.

Problem 7.10. A word consists of n letters from the alphabet $\{a, b, c, d\}$. A word is called *convoluted* if it has two consecutive identical blocks of letters. For example, *caab* and *cababdc* are convoluted, but *abcab* is not. Prove that the number of non-convoluted words with n letters is greater than 2^n.

Problem 7.11. (**IMO 2006 shortlist**) We have $n \geq 2$ lamps $L_1, \ldots, L_n$ in a row, each of them being either on or off. Every second we simultaneously modify the state of each lamp as follows: if the lamp L_i and its neighbors (only one neighbor for $i = 1$ or $i = n$, two neighbors for other i) are in the same state, then L_i is switched off; otherwise, L_i is switched on.

Initially, all the lamps are off except the leftmost one which is on.

a) Prove that there are infinitely many integers n for which all the lamps will eventually be off.

b) Prove that there are infinitely many integers n for which the lamps will never be all off.

Problem 7.12. (**IMO 2005 shortlist**) Let k be a fixed positive integer. A company has a special method to sell sombreros. Each customer can convince up to two persons to buy a sombrero after he/she buys one; convincing someone already convinced does not count. Each of these new customers can convince up to two others and so on. If a customer convinces two people to but sombreros and each of these two people in turn makes at least k persons buy sombreros (directly or indirectly), then that customer wins a free instructional video. Prove that if n persons bought sombreros, then at most $\frac{n}{k+2}$ of them got videos.

Problem 7.13. Let r be a positive integer. Consider an infinite collection of sets, each having r elements such that each two of these sets are not disjoint. Prove that there is a set with $r - 1$ elements that is not disjoint from any member of the collection.

Problem 7.14. We have n finite sets $A_1, A_2, \ldots, A_n$ such that the intersection of any collection of them has an even number of elements, except for the intersection of all subsets, which has an odd number of elements. Find the least possible number of elements that $A_1 \cup A_2 \cup \ldots \cup A_n$ can have.

Problem 7.15. (**USSR 1991**) A $k \times l$ minor of an $n \times n$ table consists of all cells which lie on the intersection of any k rows with any l columns. The number $k+l$ is called the semiperimeter of this minor. It is known that several minors, each of semiperimeter not less than n, jointly cover the main diagonal of the table. Prove that these minors jointly cover at least half of all cells.

Problem 7.16. (**IMO 2013 shortlist**) Let $n \geq 2$ be an integer. Consider all circular arrangements of the numbers $0, 1, \ldots, n$; the $n+1$ rotations of an arrangement are considered to be equal. A circular arrangement is called *beautiful* if, for any four distinct numbers $0 \leq a, b, c, d \leq n$ with $a+c = b+d$, the chord joining the numbers a and c does not intersect the chord joining numbers b and d. Let M be the number of beautiful arrangements of $0, 1, \ldots, n$. Let N be the number of pairs (x, y) of positive integers such that $x + y \leq n$ and $gcd(x, y) = 1$. Prove that

$$M = N + 1.$$

Problem 7.17. Prove that among any $2k+1$ integers, each having absolute value at most $2k-1$, one can always choose three that add up to 0.

Problem 7.18. Prove that a cube C can be divided into n cubes for all $n \geq 55$.

Problem 7.19. For a permutation $a_1, a_2, \ldots, a_n$ of the numbers $1, 2, \ldots, n$ one is allowed to change the places of any consecutive blocks, that is from

$$a_1, \ldots, a_i, a_{i+1}, \ldots, a_{i+p}, a_{i+p+1}, \ldots, a_{i+q}, a_{i+q+1}, \ldots, a_n$$

one can obtain

$$a_1, \ldots a_i, a_{i+p+1}, \ldots, a_{i+q}, a_{i+1}, \ldots, a_{i+p}, a_{i+q+1}, \ldots, a_n.$$

Find the least number of such changes after which one can obtain n, $n-1, \ldots,$ 2, 1 from $1, 2, \ldots, n$.

Problem 7.20. Let m circles intersect at points A and B. We write numbers using the following algorithm: We write 1 at points A and B, at every midpoint of an open arc AB we write 2, then at the midpoint of the arc between every two points with numbers written we write the sum of the two numbers and so on repeating n times. Let $r(n, m)$ be the number of appearances of the number n after writing all of them on our m circles.

a) Determine $r(n, m)$;

b) For $n = 2006$, find the least m for which $r(n, m)$ is a perfect square.

Example. For a single arc AB the numbers written on successive rounds are: $1-1$; $1-2-1$; $1-3-2-3-1$; $1-4-3-5-2-5-3-4-1$; $1-5-4-7-3-8-5-7-2-7-5-8-3-7-4-5-1$.

Problem 7.21. (**Hungary 2000**) Given a positive integer k and more than 2^k integers, prove that a set S of $k+2$ of these numbers can be selected such that for any positive integer $m \le k+2$, all the m-element subsets of S have different sums of elements.

Problem 7.22. (**Russia 2000**) There is a finite set of congruent square cards, placed on a rectangular table with their sides parallel to the sides of the table. Each card is colored in one of k colors. For any k cards of different colors, it is possible to pierce two of them with a single pin. Prove that all the cards of some color can be pierced by $2k-2$ pins.

Problem 7.23. (**Russia 2000**) Each of the numbers $1, 2, \ldots, N$ is colored black or white. We are allowed to simultaneously change the colors of any three numbers in arithmetic progression. For which numbers N can we always make all the numbers white?

Problem 7.24. For $n \ge 1$, let O_n be the number of $2n$-tuples $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ with all entries being either 0 or 1 and for which the sum $x_1y_1 + \ldots + x_ny_n$ is odd, and E_n the number of $2n$-tuples of same type for which the sum is even. Prove that

$$\frac{O_n}{E_n} = \frac{2^n - 1}{2^n + 1}.$$

Problem 7.25. (**TOT 2001**) In a row there are 23 boxes such that for $1 \le k \le 23$, there is a box containing exactly k balls. In one move, we can double the number of balls in any box by taking balls from another box which has more. Is it always possible to end up with k balls in the k-th box for $1 \le k \le 23$?

Problem 7.26. Consider a $2^n \times 2^n$ square. Prove that after removing a 1×1 square from one of its corners, the remaining region can be tiled by "corners" (a corner is a 2×2 square with one of the four corner unit squares removed).

Problem 7.27. Consider 10×10 grid square, such that in some of its squares there are written ten ones, ten twos, ..., ten nines and one ten (such that in each square there is at most one number written). Prove that one can choose ten squares from different rows, such that in the chosen squares we have the numbers $1, 2, \ldots, 10$.

Problem 7.28. (**IMO 2009**) Let $a_1, a_2, \ldots a_n$ be different positive integers and M a set of $n-1$ positive integers not containing the number $s = a_1 + a_2 + \ldots + a_n$. A grasshopper is going to jump along the real axis. It starts at the point 0 and makes n jumps to the right of lengths $a_1, a_2, \ldots, a_n$ in some order. Prove that the grasshopper can organize its jumps in such a way that it never falls in any point of M.

Problem 7.29. (**IMO shortlist 2004**) For a finite graph G, let $f(G)$ be the number of triangles and $g(G)$ the number of tetrahedra formed by edges of G. Find the least constant c such that

$$g(G)^3 \leq c \cdot f(G)^4,$$

for every graph G.

Problem 7.30. Prove that a graph with $\binom{n+k-2}{k-1}$ vertices contains either a K_n or a $\overline{K_k}$ i.e. either n mutually connected vertices or k mutually not connected vertices.

Problem 7.31. In a simple graph with a finite number of vertices each vertex has degree at least three. Prove that the graph contains an even cycle.

Problem 7.32. For a positive integer n, let S be a set of $2^n + 1$ elements. Let f be a function from the set of two-element subsets of S to $\{0, \ldots, 2^{n-1} - 1\}$. Assume that for any elements x, y, z of S, one of $f(\{x, y\}), f(\{y, z\}), f(\{z, x\})$ is equal to the sum of the other two. Prove that there exist a, b, c in S such that $f(\{a, b\}), f(\{b, c\}), f(\{c, a\})$ are all equal to 0.

Problem 7.33. Let G be a graph on n vertices such that there are no K_4 subgraphs in it. Prove that G contains at most $\left(\frac{n}{3}\right)^3$ triangles.

Problem 7.34. (**Moscow 2000**) In a country, there is at least one road going out of each city (each road connects exactly two cities). We call a city *marginalized* if there is only one road going out of it. It is known that it is not possible to get out of one city and then to get back into it using a closed circuit. The cities were split into two sets so that no two cities belonging to the same set are connected by a road. Assuming that there are at least as many cities in the first set as in the second, show that the first set must contain a marginalized city.

Problem 7.35. (**Moscow 2001**) There are 20 teams each belonging to a different city that play some football games among themselves, such that each team plays a home game and at most two away games. Prove that we can schedule the games in such a way that each team does not play more than one game per day and all games are played in three days.

Problem 7.36. (**Five color theorem**) Prove that the vertices of every planar graph can be colored with 5 colors such that each edge has its corresponding vertices of different colors.

Problem 7.37. In a simple graph with a finite number of vertices, each vertex has degree at least three. Prove that the graph contains a cycle whose length is not divisible by 3.

Problem 7.38. (**China 2000**) A table tennis club wishes to organize a doubles tournament, a series of matches where in each match one pair of players competes against a pair of two different players. Let a player's *match number* for a tournament be the number of matches he or she participates in. We are given a set $A = \{a_1, a_2, \ldots, a_k\}$ of distinct positive integers all divisible by 6. Find with proof the minimal number of players among whom we can schedule a doubles tournament such that

i) each participant belongs to at most 2 pairs;

ii) any two different pairs have at most 1 match against each other;

iii) if two participants belong to the same pair, they never compete against each other;

iv) the set of the participants' match numbers is exactly A.

Problem 7.39. (**Poland 2000**) Given a natural number $n \geq 2$, find the smallest integer k with the following property: Every set consisting of k cells of an $n \times n$ table contains a non-empty subset S such that in every row and in every column of the table there are an even number of cells belonging to S.

Problem 7.40. (**Austrian-Polish MO 2000**) We are given a set of 27 distinct points in the plane, no three collinear. Four points from this set are vertices of a unit square; the other 23 points lie inside this square. Prove that there exist three distinct points X, Y, Z in this set such that $[XYZ] \leq \frac{1}{48}$.

Problem 7.41. (**Moscow 1999**) We consider in the plane a convex polygon such that each of its sides is a segment which is colored towards the outside of the polygon (i.e. we regard a part of the segment as colored, and the other one not). Inside the polygon we draw some diagonals which again, have a side that is colored and one which is not. Prove that one of the polygons that was formed while partitioning the initial polygon must also have its sides colored towards the outside.

Problem 7.42. (**USSR 1989**) A fly and a spider are on a 1×1 meter square ceiling. In one second, the spider can jump from its position to the middle of any of the four segments which join it to the vertices of the ceiling. The fly does not move. Prove that in eight moves the spider can be within 1 centimeter of the fly.

Problem 7.43. Prove that if the plane is divided into parts ("countries") by lines and circles, then the obtained map can be painted in two colors so that the parts separated by an arc or a segment are of distinct colors.

Problem 7.44. (**IMO 2006 shortlist**) Let S be a finite set of points in the plane such that no three of them are on a line. For each convex polygon P whose vertices are in S, let $a(P)$ be the number of vertices of P, and let $b(P)$ be the number of points of S which are outside P. A line segment, a point and the empty set are considered as convex polygons of 2, 1 and 0 vertices, respectively. Prove that for every real number x we have

$$\sum_P x^{a(P)}(1-x)^{b(P)} = 1,$$

where the sum is taken over all convex polygons with vertices in S.

Problem 7.45. (**IMO 2006 shortlist**) A holey triangle is an upward equilateral triangle of side length n with n upward unit triangular holes cut out. A diamond is a $60^\circ - 120^\circ$ unit rhombus. Prove that a holey triangle T can be tiled with diamonds if and only if the following condition holds: Every upward equilateral triangle of side length k in T contains at most k holes, for $1 \le k \le n$.

Problem 7.46. (**IMO 2005 shortlist**) Let $n \ge 3$ be a given positive integer. We wish to label each side and each diagonal of a regular n-gon $P_1 \dots P_n$ with a positive integer less than or equal to r so that:

i) every integer between 1 and r occurs as a label;

ii) in each triangle $P_iP_jP_k$, two of the labels are equal and greater than the third.

Given these conditions:

a) Determine the largest positive integer r for which this can be done.

b) For that value of r, how many such labellings are there?

Problem 7.47. In a convex n-gon ($n \ge 403$), there are $200n$ diagonals drawn. Prove that one of them intersects at least 10000 others.

Problem 7.48. Consider a convex n-gon such that no three of its diagonals are concurrent. In how many parts do the diagonals divide the n-gon?

Problem 7.49. We are given some unit squares which are translations of each other in the plane, such that from any $n+1$, at least two intersect. Prove that we can place at most $2n-1$ needles in the plane, such that every square is stabbed by at least one needle.

Chapter 8

Games

8.1 Theory and Examples

The two key concepts that we will deal with throughout this chapter are the notion of a *game* and the one of a *strategy*.

While a game is simply defined as a structured form of play, a **mathematical game** is a game whose rules, strategies and outcomes are clearly defined by mathematical objects. Examples include chess, tic-tac-toe and others.

In a game, **strategy** for a player is a choice for each position he might find himself in of one of the options that the player has in that position. Note that since his choices are set, this may limit the set of positions in which he can find himself. However the choices of any other participants to the game are not determined, so his strategy must reflect this. Once a player has chosen a strategy, this fully determines the actions he will take at any stage of the game.

Throughout the chapter, when we say that a player has a strategy, we shall understand that a strategy can be found, even though for some examples this may take too long in real life. As the following problem shows, for many games we can prove that a strategy exists, so we can theoretically establish who is going to win the game given perfect play, even before the game starts. Nevertheless, finding the actual strategy can take too much computational time in practice, which fortunately keeps the fun of these games still alive.

Example 8.1. Two players play a game where they take turns to make a move and which always finishes with the victory of one of the players. The game is also designed so that it ends in at most n steps, for some fixed positive integer n. Prove that either the first or the second player has a winning strategy.

Solution. We prove the statement by induction on the number n of steps after which the game certainly finishes. Let us call the two players A and B and we assume that A makes the first move.

If $n = 1$, then certainly one of the players has a winning strategy. Indeed, only player A will get to move. If any of his moves leads to him winning, then a winning strategy for player A is to make this move. If every one of his moves leads to a win for B, then a winning strategy for B is to let A move arbitrarily.

Assume now that the result holds for some $n \geq 1$. We shall prove that it also holds for $n + 1$.

After any move by player A we are left with a game that certainly ends in at most n steps. Hence by the inductive hypothesis one of the two players has a winning strategy. If any move by player A leaves a position in which player A has a winning strategy, then making this move and then following that winning strategy is a winning strategy for player A. Otherwise, it must be that whatever first move player A makes, we are left in a situation where player B has a winning strategy. Then a winning strategy for player B is to let A make his first move arbitrarily and then follow the resulting winning strategy. This completes our induction step.

Remark. The same argument which we used in the induction step applies to prove the more general result where we only require that the game finishes in a finite number of steps, not necessarily bounded by a fixed n.

Example 8.2. (**Putnam training**) For the following game, we assume that there is an unlimited supply of tokens. Two players arrange several piles of tokens in a row. By turns, each of them takes one token from one of the piles and adds at will as many tokens as he or she wishes to piles placed to the left of the pile from which the token was taken (so if the token was taken from the leftmost pile, then no new tokens can be added). Assuming that the game ever finishes, the player that takes the last token wins. Prove that, no matter how they play, the game will eventually end after finitely many steps.

Solution. We will prove the result by induction on the number n of piles:

For $n = 1$, we have only one pile and because each player must take at least one token from that pile, the number of tokens in the pile will decrease at each move until it is empty.

Now assume that the game with n piles must end eventually, for some $n \geq 1$. We shall prove that this also holds for $n+1$ piles. First note that the players cannot keep taking tokens only from the first n piles, since by the induction hypothesis, the game with n piles must eventually end. So at some point, one player must take a token from the $(n+1)$-st pile. Now, no matter how many tokens were added to the other n piles after that, it is still true that the players cannot keep taking tokens only from the first n piles forever, so eventually someone will take another token from the $(n+1)$-st pile. Consequently, the number of tokens in that pile will continue decreasing until it is empty. Once this happens, we only have n piles left, so by the induction hypothesis the game will end in finitely many steps.

Example 8.3. Consider the following two-player game. There are 56 candies on the table. Players play by turns (one after another). In each turn it is allowed to take either one, three, or five candies. The winner is the person who takes the last set of candies. Given perfect play on both sides, who will win?

Solution. Let us prove by mathematical induction that if the number of the candies is of the form $6n - 4$, where $n \in \mathbb{N}^*$, then the second player wins.

If $n = 1$, then the first player takes one candy and the second player takes the other one.

Assume now that the result holds for some $n \geq 1$. We show that it also holds for $n+1$. We have $6(n+1) - 4 = 6n + 2$ candies. If the first player takes a candies, then the second player takes $6 - a$ candies. Hence, the game will be left with $6k - 4$ candies. By the induction hypothesis, the second player wins. This completes our induction. As $56 = 6 \cdot 10 - 4$, we know from our proof that the second player wins.

Example 8.4. On an infinite chessboard consisting of unit squares (x, y) with $x, y \geq 0$ two players play the following game: initially a king is positioned somewhere on the board, but not on $(0, 0)$, and they alternatively move it either down or left or down-left. The player who *wins* is the one who moves the king into the $(0,0)$ square. Find the initial positions of the king for which the first player wins.

Solution. We will prove by induction on $x+y \geq 1$ that the first player wins if and only if the original square (x,y) doesn't have $2 \mid x$ and $2 \mid y$.

Indeed, for $x+y=1$ or $x=y=1$, the first player can move directly on $(0,0)$ and win.

For $x+y>1$, if $2 \mid x$ and $2 \mid y$, the first player can only only move on $(x-1,y)$, $(x-1,y-1)$ or $(x,y-1)$, and neither of these has both coordinates even. So, by induction, the second player has a winning strategy.

Else, one of $(x-1,y)$, $(x-1,y-1)$ or $(x,y-1)$ has both coordinates even, so the first player can move to it, from which he has a winning strategy (by induction).

Example 8.5. (**Baltic Way, 2013**) A positive integer is written on a blackboard. Players A and B play the following game: in each move one has to choose a divisor m of the number n written on the blackboard with $1<m<n$ and replace n with $n-m$. Player A makes the first move, then players move alternately. The player who can't make a move loses the game. For which starting numbers is there a winning strategy for player B?

Solution. Call a number n *protected* if the second player has a winning strategy when the game starts with the number n. The game ends when there are no moves, hence either $n=1$ or n is prime. In these cases, the first player has no move and thus loses, so these are protected. Suppose we are at a protected value. Since the second player has a winning strategy, regardless of how the current player moves, he cannot leave a protected number (from which he would win). Thus from a protected value every move must leave an unprotected value. Suppose we are at an unprotected value. Then the first player must have a winning strategy, hence some move must leave a protected value (from which he wins).

We will show by induction on n that the protected values are either $n = 2^{2k-1}$ for some positive integer k or n odd.

The base cases are when $n=1$ or n is prime. These are protected and have either $n=2=2^{2-1}$ or n odd. For the inductive step there are four cases.

Suppose $n = 2^{2k-1}$. Then the possible moves are to choose $m = 2^j$ with $1 \leq j < 2k-1$. These moves will leave us at $n - m = 2^j(2^{2k-j-1} - 1)$. This is always even and will only be a power of 2 if $j = 2k-2$, when it is an even power of 2, namely $n - m = 2^{2k-2}$. By the induction hypothesis none of these outcomes is protected, hence n is protected.

Suppose n is odd. Then the possible moves are write $n = km$ for $1 < k, m < n$ and to move to $n - m = (k-1)m$. These are all even since $k - 1$ is even and are not powers of 2 since they have a non-trivial odd factor m. By the induction hypothesis none of these outcomes is protected, hence n is protected.

Suppose $n = 2^{2k}$. Then choosing $m = 2^{2k-1}$ will leave a protected outcome $n - m = 2^{2k-1}$. Hence n is unprotected.

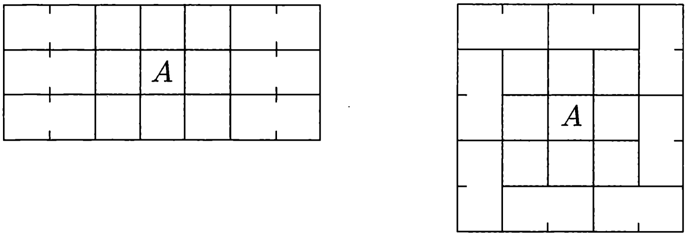
Suppose $n = 2^k l$, where $l > 1$ is odd. Then choosing $m = l$ will leave a protected (odd) outcome $n - m = (2^k - 1)l$. Hence n is unprotected.

Example 8.6. Let m, n be odd numbers greater than 1, such that $4 \mid (m - n)$. Consider the following two-player game: players play by turns (one after another) and put asterisks in the squares of an $m \times n$ grid rectangle (in each square it is allowed to put only one asterisk). The first player is allowed to put an asterisk in any square, except the central square. In the following turns, a player can put an asterisk in a square having a common vertex with the square corresponding to the previous asterisk. A player wins if he puts an asterisk in the central square or if his opponent cannot move. Given perfect play on both sides, who will win?

Solution. Let us denote by A the 3×3 grid square consisting of the central square in the $m \times n$ grid rectangle and the eight squares having a common vertex with it. Denote by B the figure consisting of the squares that do not belong to A. In particular, if $m = n = 3$, then B is empty.

For $m + n > 6$ we prove by mathematical induction with respect to $m + n$ that B can be divided into dominoes.

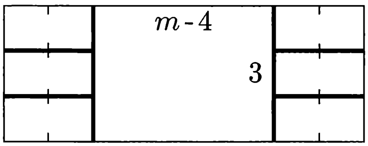
Notice that we cannot have $m + n = 8$, since any pair of odd numbers that sum to 8 does not satisfy the condition $4 \mid (m - n)$. So the base case is $m + n = 10$. Without loss of generality, let $m \geq n$. Then either $m = 7$, $n = 3$ or $m = 5$, $n = 5$. The proofs for both cases are illustrated in the figure below.



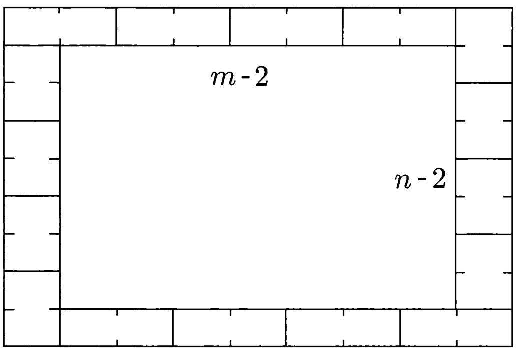
a) *b)*

Assume now that the statement holds for $m+n \leq k$, where $k > 9$, $k \in \mathbb{N}$, and we prove that it also holds for $m+n = k+1$. Again, we can assume by symmetry that $m \geq n$. We distinguish two cases:

Case 1. If $n = 3$, then the statement holds true for an $(m-4) \times 3$ grid rectangle. From the figure below, it follows that the statement holds for an $m \times 3$ grid rectangle, too.



Case 2. If $n \geq 5$, then the statement holds true for an $(m-2) \times (n-2)$ grid rectangle. From the figure below, it follows that the statement also holds for an $m \times n$ grid rectangle.



This completes our induction. Back to our game, if the first player puts an asterisk in any square of B, then the second player can put an asterisk in the other square of the domino which contains that square. Obviously, sooner or later, the first player will put an asterisk (for the first time) in one of the squares of A (not the central one). Then the second player can put an asterisk in the central square of A and win the game.

Example 8.7. (**Russia 2000**) Two pirates divide their loot, consisting of two sacks of coins and one diamond. They decide to use the following rules. On each turn, one pirate chooses a sack and takes $2m$ coins from it, keeping m for himself and putting the rest into the other sack. The pirates alternate taking turns until no more moves are possible; the first pirate unable to make a move loses the diamond, and the other pirate takes it. For what initial numbers of coins can the first pirate guarantee that he will obtain the diamond?

Solution. We claim that if there are x and y coins left in the two sacks, respectively, then the next player P_1 to move has a winning strategy if and only if $|x-y|>1$. Otherwise, the other player P_2 has a winning strategy.

We prove the above claim by induction on the total number of coins, which is $x+y$. If $x+y=0$, then no moves are possible and the next player to move does not have a winning strategy.

Assume now that the claim is true when $x+y\leq n$ for some non-negative n and we prove that it is also true when $x+y=n+1$.

First consider the case $|x-y|\leq 1$. Assume that a move is possible; otherwise, the next player P_1 automatically loses, which proves our claim. The next player must take $2m$ coins from one sack, say the one containing x coins, and put m coins into the sack containing y coins. Hence the new difference between the numbers of coins in the sacks is

$$|(x-2m)-(y+m)|\geq|-3m|-|y-x|\geq 3-1=2.$$

At this point, there are now a total of $x+y-m$ coins in the sacks, and the difference between the numbers of coins in the two sacks is at least 2. Thus, by the induction hypothesis, P_2 has a winning strategy. This proves the claim when $|x-y|\leq 1$.

We now treat the case $|x-y| \geq 2$. We can assume without loss of generality that $x > y$. The pirate P_1 would like to find an m such that $2m \leq x$, $m \geq 1$ and

$$|(x-2m)-(y+m)| \leq 1.$$

The number $m = \lceil \frac{x-y-1}{3} \rceil$ satisfies the last two inequalities from above and we also claim that $2m \leq x$. Indeed, $x - 2m$ is non-negative, because it differs by at most 1 from the positive number $y + m$. After taking $2m$ coins from the sack with x coins, P_1 leaves a total of $x + y - m$ coins, where the difference between the numbers of coins in the sacks is at most 1. Hence, by the induction hypothesis, the other player P_2 has no winning strategy, so P_1 must have. This completes the proof of our claim.

It follows that the first pirate can guarantee that he will obtain the diamond if and only if the number of coins initially in the sacks differ by at least 2.

8.2 Proposed Problems

Problem 8.1. Three players play the following game. There are 54 candies on the table. Players play by turns (one after another). In each turn it is allowed to take either one, three, or five candies, such that the same number of candies cannot be taken in two consecutive turns. The winner is the person who takes the last set of candies. Given perfect play of all the players, who will win?

Problem 8.2. Let a_0, a_1, a_2,..., a_{2016} be distinct positive numbers. Consider the following two-player game: players take turns to write the numbers a_0, a_0, a_1, a_2,..., a_{2016} instead of $*$ in the polynomial $*x^{2016} + *x^{2015} + ... + *$ (one number in each turn). If the polynomial obtained in the end has an integer root, then the second player wins; otherwise, the first player wins. Given perfect play by both sides, who will win?

Problem 8.3. A regular pack of 52 cards (with 26 red cards and 26 black cards) is shuffled and dealt out to you one card at a time. At any moment, based on what you have seen so far, you can say "I predict that the next card will be red". You can only make this prediction once. Which strategy gives you the greatest chance of being right?

Problem 8.4. On an infinite chessboard consisting of unit squares (x, y) with $x, y \geq 0$ two players play the following game: initially a king is positioned somewhere on the board, but not on $(0, 0)$, and they alternatively move it either down or left or down-left. The player who *loses* is the one who moves the king into the $(0, 0)$ square. Find the initial position of the king for which the first player wins.

Problem 8.5. (**TOT 2003**) In a game, Boris has 1000 cards numbered $2, 4, \ldots, 2000$, while Anna has 1001 cards numbered $1, 3, \ldots, 2001$. The game lasts 1000 rounds. In an odd-numbered round, Boris plays any card of his. Anna sees it and plays a card of hers. The player whose card has the larger number wins the round, and both cards are discarded. An even-numbered round is played in the same manner, except that Anna plays first. At the end of the game, Anna discards her unused card. What is the maximal number of rounds each player can guarantee to win, regardless of how the opponent plays?

Problem 8.6. There are $n > 1$ balls in a box. Now two players A and B are going to play a game. At first, A can take out $1 \leq k < n$ ball(s). When one player takes out m ball(s), then the next player can take out $1 \leq \ell \leq 2m$ ball(s). The person who takes out the last ball wins. Find all positive integers n such that B has a winning strategy.

Problem 8.7. (**Russia 2002**) We are given one red and $k > 1$ blue cells, and a pack of $2n$ cards, numbered from 1 to $2n$. Initially, the pack is situated on the red cell and arranged in an arbitrary order. In each move, we are allowed to take the top card from one of the cells and place it either onto the top of another cell on which the number on the top card is greater by 1, or onto an empty cell. Given k, what is the maximal n for which it is always possible to move all the cards onto a single blue cell?

Problem 8.8. (**Mathematical Reflections**) A and B play the following game on a $2n + 1 \times 2m + 1$ board: A has a pawn in the bottom left corner (square $(1, 1)$) and wants to get to the top right corner (square $(2n+1, 2m+1)$). At each turn, A moves the pawn in an adjacent square (having a common edge) and B either does nothing or blocks a square for the rest of the game, but in such a way that A can still get to the top right corner. Prove that B can force A to do at least $(2n+1)(2m+1) - 1$ moves before reaching the top right corner.

Problem 8.9. (**44th tournament of Ural, 4th tour**). Consider the following two-player game: there are given two piles of stones, one consisting of 1914 stones and the other of 2014. In his turn, a player is allowed to remove either two stones from the bigger pile or one stone from the smaller pile; if at some stage the piles are equal, then he is allowed to remove from one of the piles either one or two stones. A player loses when cannot play his turn anymore. Given perfect play by both sides, who will win?

Chapter 9

Miscellaneous Topics

Even though each of the three sections that are included in this chapter (Geometry, Calculus and Algebra) could have been a chapter in its own right, we decided to group them into a single one for the following key reasons: unlike the other chapters where we started by introducing all the necessary theory before discussing any examples, for each of these three sections this aim would have been too ambitious, as the background required to understand all the applications is too large to be covered in 2 or 3 pages of theory. Moreover, for each of the previous chapters, we aimed to give a decent overview of the main ideas or topics that can arise in each field we discussed; this is also out of reach when it comes to these three topics, as an exhaustive overview would require a separate book for each of them. Consequently, we are content to discuss just a few from those numerous ideas, and refer the interested reader to other sources for more (see for example [8], [12]).

9.1 Geometry

9.1.1 Examples

The applications of Induction in Geometry are surprisingly vast: one can use Induction to prove questions that ask for the existence of a specific configuration (certain arrangement of points, a circumcircle of a figure, a line/figure subject to some constraints, etc), construction problems (e.g. dividing a certain segment into n equal parts using just a ruler), problems of locus, geometric inequalities and many more. We give a small sample of these in the examples below, leaving some other very beautiful questions to be tackled by the passionate readers in the Proposed Problems section.

Example 9.1.1. If $A_1, \dots, A_n$ are any points in the plane, with any three not collinear, then there is a convex polygon P such that some of the A_i are vertices of P and the rest of the points are inside P. (The polygon P is called the *convex hull* of the points $A_1, \dots, A_n$.)

Solution. We proceed by induction on n. For $n = 3$, the conclusion is immediate because the points form a triangle.

Assume now that the statement is true for some $n \geq 3$. Suppose we are given a set of $n+1$ points. By the induction hypothesis we can find a convex polygon $P = B_1 B_2 \dots B_m$ such that the vertices B_i are among $A_1, \dots, A_n$ and the remaining vertices of $A_1, \dots, A_n$ are inside P.

Now consider the vertex A_{n+1}. There will be some pair B_i, B_j with $i < j$ such that the angle $\angle B_i A_{n+1} B_j$ is maximal. Suppose there is some $k \neq i, j$ for which B_k lies outside this angle. The three rays $A_{n+1}B_i$, $A_{n+1}B_j$ and $A_{n+1}B_k$ cannot lie in a half-plane, since this would make one of the angles $\angle B_i A_{n+1} B_k$ or $\angle B_k A_{n+1} B_j$ larger than $\angle B_i A_{n+1} B_j$, contradicting its maximality. Thus A_{n+1} lies in the interior of triangle $B_i B_j B_k$ and the polygon P is still the one we need.

Otherwise, every B_k with $k \neq i, j$ lies inside the angle $\angle B_i A_{n+1} B_j$. Without loss of generality, assume the union of P and the triangle $B_i A_{n+1} B_j$ is the polygon $Q = A_{n+1} B_i B_{i+1} \dots B_j$. Since Q lies inside the angle $\angle B_i A_{n+1} B_j$, the interior angles of Q at B_i and B_j are less than $180°$. The interior angles at the other B's are interior angles of P hence less than $180°$ and the interior angle at A_{n+1} is $\angle B_i A_{n+1} B_j < 180°$. Thus Q is convex, its vertices are some subset of $A_1, \dots, A_{n+1}$, and since it contains P, all the remaining A_i are inside Q. Thus Q is the required polygon.

Example 9.1.2. A point O is inside a convex n-gon $A_1 \dots A_n$ (this includes the case when O is on the boundary). Prove that among the angles $A_i O A_j$, there are not fewer than $n-1$ non-acute ones.

Solution. We prove the result by induction on n. For $n = 3$, things are clear.

Now let us consider the n-gon $A_1 \dots A_n$, where $n \geq 4$. Choose p, q, r such that O lies inside the triangle $A_pA_qA_r$. Let A_k be a vertex of the given n-gon distinct from the points A_p, A_q and A_r. By removing A_k from the n-gon $A_1 \dots A_n$, we get a $(n-1)$-gon and we can apply the induction hypothesis to it. This gives at least $n-2$ non-acute angles that do not involve A_k. The point O must lie inside one of the triangles with vertices A_k and two of A_p, A_q, A_r. Suppose without loss it is $A_kA_qA_r$. Then $\angle A_kOA_q + \angle A_kOA_p = 360° - \angle A_qOA_r \geq 180°$. Hence at least one of these is a non-acute angle involving A_k. This completes our proof.

Example 9.1.3. Let $A_1A_2 \dots A_{2n}$ be a polygon inscribed in a circle. We know that all the pairs of its opposite sides except one are parallel. Prove that for any odd n, the remaining pair of sides is also parallel and for any even n, the lengths of the exceptional sides are equal.

Solution. We prove the result by induction on $n \geq 4$. For the $n = 4$, the result is immediate. Now consider the hexagon $ABCDEF$ with $AB \parallel DE$ and $BC \parallel EF$. We prove that $CD \parallel AF$:

Notice that $AB \parallel DE$ implies that $\angle ACE = \angle BFD$ and since $BC \parallel EF$, it follows that $\angle CAE = \angle BDF$. Triangles ACE and BDF have two pairs of equal angles, and therefore, their third angles are also equal. The equality of these angles implies the equality of the arcs $\overset{\frown}{AC}$ and $\overset{\frown}{DF}$, thus chords CD and AF are parallel.

Now assume that the statement is proved for the $2(n-1)$-gon and let us prove it for the $2n$-gon. Let $A_1 \dots A_{2n}$ be a $2n$-gon in which $A_1A_2 \parallel A_{n+1}A_{n+2}$, $\dots$, $A_{n-1}A_n \parallel A_{2n-1}A_{2n}$. Let us consider the $2(n-1)$-gon $A_1A_2 \dots A_{n-1}A_{n+1} \dots A_{2n-1}$. By the induction hypothesis, for n odd we have $A_{n-1}A_{n+1} = A_{2n-1}A_1$, and for n even we have $A_{n-1}A_{n+1} \parallel A_{2n-1}A_1$.

Let us now study the triangles $A_{n-1}A_nA_{n+1}$ and $A_{2n-1}A_{2n}A_1$. We first treat the case when n is even. Then vectors $\overrightarrow{A_{n-1}A_n}$ and $\overrightarrow{A_{2n-1}A_{2n}}$, as well as $\overrightarrow{A_{n-1}A_{n+1}}$ and $\overrightarrow{A_{2n-1}A_{2n}}$ are parallel and oppositely directed. Hence $\angle A_nA_{n-1}A_{n+1} = \angle A_1A_{2n-1}A_{2n}$ and $A_nA_{n+1} = A_{2n}A_1$ since they are chords that cut equal arcs.

Now let n be odd. Then $A_{n-1}A_{n+1} = A_{2n-1}A_1$, i.e. $A_1A_{n-1} \parallel A_{n+1}A_{2n-1}$.

In hexagon $A_{n-1}A_nA_{n+1}A_{2n-1}A_{2n}A_1$ we have $A_1A_{n-1} \parallel A_{n+1}A_{2n-1}$ and $A_{n-1}A_n \parallel A_{2n-1}A_{2n}$. Hence, from the base case of the hexagon that we discussed above, we have $A_nA_{n+1} \parallel A_{2n}A_1$, as required. This completes our proof

Example 9.1.4. (**Moscow Math Circles**) Prove that one can cut any two polyhedrons of equal volume into several tetrahedrons of pairwise equal volumes.

Solution. First, we divide both polyhedrons into arbitrary tetrahedrons. To do this, one can first divide a polyhedron into pyramids by connecting its inner point with the vertices.

Once we have divided our polyhedron into tetrahedrons, we select the smallest of the tetrahedrons we obtained. Then we cut a tetrahedron of the same size from one of the remaining tetrahedrons in the other collection. Now we finish the question by applying induction on the number of tetrahedrons.

We conclude this section by presenting two more challenging problems:

Example 9.1.5. Let $A_1A_2\dots A_n$ $(n \geq 4)$ be a convex polygon. Denote by R_i the radius of the circumcircle of the triangle $A_{i-1}A_iA_{i+1}$ (here $i = 2, 3, \dots, n$ and $A_{n+1} = A_1$). Prove that if $R_2 = R_3 = \dots = R_n$ then one can draw a circumcircle for the polygon $A_1A_2\dots A_n$.

Solution. We begin by proving a few lemmas:

Lemma 1. If the convex quadrilateral $MNPK$ cannot be inscribed in a circle, but the radii of the circumcircles of the triangles MNP and MNK are equal, then

$$\angle MPN + \angle MKN = 180^\circ.$$

Proof. From the hypotheses, the circumcircles of the triangles MNP and MNK are symmetric about the line MN. Hence,

$$\angle MKN + \angle MPN = \angle MK'N + \angle MPN = 180^\circ,$$

where K' is the symmetric of K with respect to the line MN.

Lemma 2. Let $ABCDE$ be a convex pentagon such that $R_B = R_C = R_D$, where R_B, R_C and R_D denote the radii of the circumcircles of the triangles ABC, BCD and CDE, respectively. Then at least one of the quadrilaterals $ABCD$ or $BCDE$ can be inscribed in a circle.

Proof. Assume by contradiction that neither of the quadrilaterals $ABCD$ or $BCDE$ can be inscribed in a circle. Then using the previous lemma, it follows that

$$\angle CBD + \angle CED = \angle BAC + \angle BDC = 180^\circ,$$

hence

$$360^\circ = (\angle BAC + \angle CBD) + (\angle BDC + \angle CED) < 180^\circ + 180^\circ.$$

This gives a contradiction and establishes the result of the lemma.

Note that the proof of Lemma 2 still works just fine if the points A and E coincide. Hence in this case we conclude that $ABCD$ can be inscribed in a circle. This is the $n = 4$ case of the problem.

Lemma 3. If $ABCDEF$ is a convex hexagon such that $R_B = R_C = R_D = R_E$ (with the notations as in the previous lemma), then one of the quadrilaterals $ABCD$ or $CDEF$ can be inscribed in a circle.

Proof. Assume by contradiction that one cannot draw a circumcircle for either of the specified quadrilaterals. Then using the first lemma that we proved above, we have that

$$\angle BAC + \angle BDC = \angle DCE + \angle DFE = 180^\circ,$$

while according to the second lemma, it is possible to draw a circumcircle for the quadrilateral $BCDE$ (consider the pentagon $ABCDE$).

Hence $\angle DBE = \angle DCE$ and $\angle CEB = \angle BDC$. Thus

$$360^\circ = (\angle BAC + \angle DBE) + (\angle DFE + \angle CEB) < 180^\circ + 180^\circ.$$

This gives a contradiction, so the lemma holds.

Again, the proof of Lemma 3 is fine if A and F coincide.

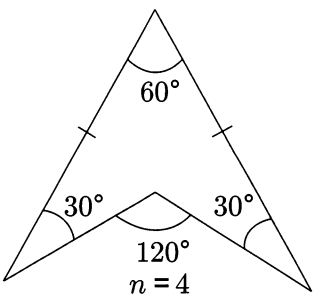
Having proved these lemmas, we proceed to the proof of the problem by mathematical induction. The $n = 4$ case was already handled above. For

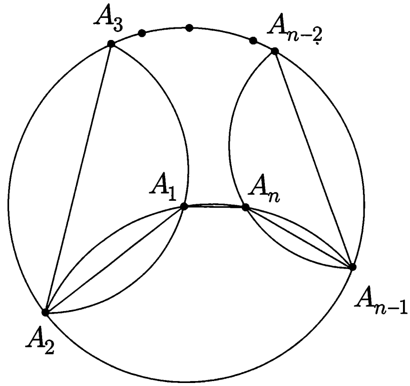
$n = 5$, applying Lemma 2 and Lemma 3 (with $A = F$) we see that at least two of the quadrilaterals $A_1A_2A_3A_4$, $A_2A_3A_4A_5$, $A_3A_4A_5A_1$ can be inscribed in a circle. From this it follows that it is possible to draw a circumcircle for the pentagon $A_1A_2A_3A_4A_5$.

For the induction step we show that if the statement holds true for $n \geq 5$, then it also holds for $n + 1$.

Consider the quadrilaterals $A_1A_2A_3A_4$, $A_2A_3A_4A_5$, $A_3A_4A_5A_6$ and $A_4A_5A_6A_7$ (when $n = 6$, as before, we set $A_7 = A_1$). Using **Lemma 2** and **Lemma 3** we have that a pair of consecutive quadrilaterals in the given list can be inscribed in a circle. Assume that these are $A_1A_2A_3A_4$ and $A_2A_3A_4A_5$ (the argument is essentially unchanged for the other possibilities). Then it is not difficult to check that the convex n-gon $A_1A_2A_4A_5 \ldots A_nA_{n+1}$ satisfies the conditions of the problem, so by the induction hypothesis, it can be inscribed in a circle. Moreover, the vertex A_3 is also on that circumference, because it is possible to draw a circumcircle for quadrilateral $A_1A_2A_3A_4$. This completes our proof.

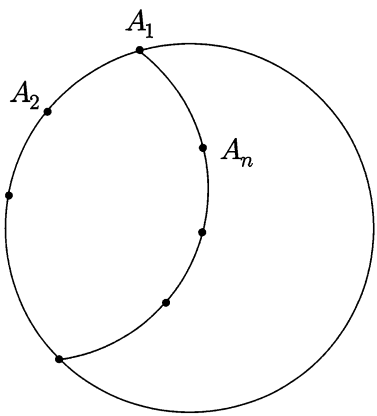
Remark. Non convex polygons for which $R_2 = R_3 = \ldots = R_n$ exist:



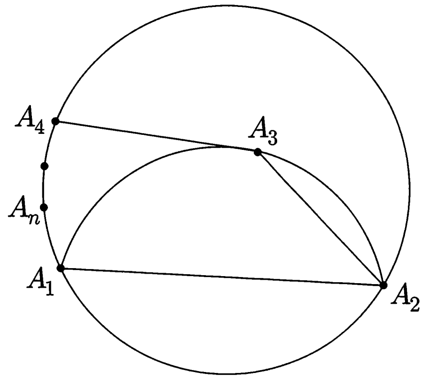


$n \geq 5$

Remark. If in a convex n-gon only $n - 2$ of the radii from the question are equal, then in general, the conclusion does not hold. See for example the figures below:



or



Example 9.1.6. (**IMO 2006**) Let $P = A_1A_2 \dots A_n$ be a convex polygon. For each side A_iA_{i+1} let T_i be the triangle of largest area having A_iA_{i+1} as a side and another vertex of the polygon as its third vertex. Let S_i be the area of T_i, and S the area of the polygon. Prove that $\sum S_i \geq 2S$.

Solution. To prove the problem by induction we need to reduce somehow the number of vertices of the polygon. We can either remove a vertex, a side or move continuously a vertex until the polygon degenerates into one with less vertices or edges.

Consider

$$f(P) = \sum[T_i] - 2[P].$$

Let $l_i = A_iA_{i+1}$ and V_i be the third vertex of T_i. Note that V_i is uniquely determined unless there is a side l_j parallel to l_i, in this case V_i can be either of its endpoints.

For $n = 3$ the assertion is clear and for $n = 4$ we have

$$S_1 + S_2 + S_3 + S_4 \geq [A_1A_2A_3] + [A_2A_3A_4] + [A_3A_4A_1] + [A_4A_1A_2] = 2S.$$

Next, we shall use induction of step 2. We are now going to do the following operation, as much as we can:

(*) Pick up a side A_iA_{i+1} which is not parallel to any of the other sides of the polygon . Now let us try to move $X = A_i$ on the line $A_{i-1}A_i$, ensuring that while we move it, the polygon remains convex and also that the line A_iA_{i+1}

never becomes parallel to any of the other sides. We claim that f is linear in XA_{i-1}. Indeed, for any side l_k, V_k remains unchanged (actually V_k could change if in the process of moving A_i one of the sides l_j would become parallel to l_k, then V_k would "jump" from one endpoint of l_j to another. However this is not the case because we ensure A_iA_{i+1} is not parallel to other sides of the polygon, and they do not change directions). Therefore, T_k is either constant, or has one vertex $X = A_i$ and the other two vertices fixed. In any case, S_k is clearly linear in A_iX, and obviously so is S. Thus f, as a linear function, takes its minimal values at the extremities. Now what could be the extremities? We could encounter one of the following cases:

a) $A_i = A_{i-1}$, in which the polygon degenerates into $A_1 \dots A_{i-1}A_{i+1} \dots A_n$, and we use induction.

b) A_i goes to infinity, in which case the inequality is easy to prove.

c) A_i becomes collinear with $A_{i+1}A_{i+2}$, in which the polygon degenerates into $A_1 \dots A_iA_{i+2} \dots A_n$, and we use induction again.

d) A_iA_{i+1} becomes parallel to one of the sides of the polygon. In this case the number of pairs of parallel sides in P increases.

We are done in the cases a),b),c). If we encounter case d), repeat (*) again and so on. Eventually we reach a polygon in which all sides are divided into pairs of parallel ones. In this case, we can deduce that n is even and $l_i || l_{i+\frac{n}{2}}$ (we work modulo n). Assume now that $n \geq 6$ because for $n = 4$ we actually have equality. Let $m = \frac{n}{2}$. We can see that $S_i + S_{i+m} = [A_iA_{i+1}A_{i+m}A_{i+m+1}]$, so

$$f(P) = \sum [A_iA_{i+1}A_{m+i}A_{m+i+1}] - 2S.$$

We can suppose that $\angle A_1A_2A_3 + \angle A_2A_3A_4 > 180$, that is the rays $(A_1A_2$ and $(A_4A_3$ intersect at a point X. Let rays $(A_{m+1}A_{m+2}$ and $(A_{m+4}A_{m+3}$ intersect at a point Y. We claim that $f(P) \geq f(A_1XA_4 \dots A_{m+1}YA_{m+4} \dots A_n)$. This would provide the final step in the problem, since this new polygon has $n-2$ sides. Now

$$f(P) - f(A_1XA_4 \dots A_{m+1}YA_{m+4} \dots A_n)$$

$$= 2[A_2XA_3] + 2[A_{m+2}YA_{m+3}] + [A_2A_3A_{m+2}A_{m+3}] + [A_1A_2A_{m+1}A_{m+2}]$$

$$+[A_3A_4A_{m+3}A_{m+4}] - [A_1XA_{m+1}Y] - [A_4XA_{m+4}Y].$$

But

$$[A_1XA_{m+1}Y] - [A_1A_2A_{m+1}A_{m+2}] = [A_2XA_{m+2}Y] = [A_2YX] + [XYA_{m+2}],$$

and similarly for $[A_4XA_{m+4}Y]$. The problem then reduces to

$$2[A_2XA_3] + 2[A_{m+2}YA_{m+3}] + [A_2A_3A_{m+2}A_{m+3}]$$

$$-[A_2YX] - [XYA_{m+2}] - [A_3YX] - [XYA_{m+3}] \geq 0.$$

However,

$$[A_2YX] + [XYA_{m+2}] + [A_3YX] + [XYA_{m+3}] = [A_2YX] + [XYA_{m+2}]$$

$$+([XA_3A_{m+2}] + \frac{1}{2}(\overline{XA_3}, \overline{A_{m+2}Y})) + ([A_2YA_{m+3}] + \frac{1}{2}(\overline{YA_{m+3}}, \overline{A_2X}))$$

$$= [A_2A_3A_{m+2}A_{m+3}] + [XA_2A_3] + [YA_{m+2}A_{m+3}]+$$

$$+\frac{1}{2}(\overline{XA_3}, \overline{A_{m+2}Y}) + \frac{1}{2}(\overline{YA_{m+3}}, \overline{A_2X})$$

(where $(\overline{u}, \overline{v})$ denotes the area of the parallelogram determined by vectors $\overline{u}$ and $\overline{v}$).

Therefore, we are left to prove that

$$[XA_2A_3] + [YA_{m+2}A_{m+3}] \geq \frac{1}{2}(\overline{XA_3}, \overline{A_{m+2}Y}) + \frac{1}{2}(\overline{YA_{m+2}}, \overline{A_2X}).$$

However, since the triangles $XA_2A_3, YA_{m+2}A_{m+3}$ are similar, we deduce that both $\frac{1}{2}(\overline{XA_3}, \overline{A_{m+2}Y})$ and $\frac{1}{2}(\overline{YA_{m+3}}, \overline{A_2X})$ are equal to $\sqrt{[XA_2A_3][XA_{m+2}A_{m+3}]}$, and we finish the problem by the AM-GM inequality.

9.1.2 Proposed Problems

Problem 9.1.1. a) Prove that any n-gon can be cut into triangles by non-intersecting diagonals.

b) Prove that the sum of the inner angles of any n-gon is equal to $(n-2)180°$. Hence prove that the number of triangles into which an n-gon is cut by non-intersecting diagonals is equal to $n-2$.

Problem 9.1.2. For any positive integer $n > 1$, prove that there exist 2^n points in the plane, no three collinear, such that no $2n$ of them form a convex polygon.

Problem 9.1.3. Let $A_1 \dots A_n$ be a convex polygon inscribed in a circle such that among its vertices, there are no two which form a diameter. Prove that if among the triangles $A_pA_qA_r$, as p, q, r range over $1, \dots, n$, there is at least one acute triangle, then there are at least $n-2$ such acute triangles.

Problem 9.1.4. a) Prove that the projection of a point P of the circumscribed circle of a cyclic quadrilateral $ABCD$ onto the Simson lines of this point with respect to triangles BCD, CDA, DAB and BAC lie on one line. (*This line is called the Simson line of P with respect to the inscribed quadrilateral*).

b) Prove by induction that we can similarly define the Simson line of a point P with respect to an inscribed n-gon as the line that contains the projections of point P on the Simson lines of all $(n-1)$-gons obtained by deleting one of the vertices of the n-gon.

Problem 9.1.5. On the circle of radius 1 with center O there are given $2n+1$ points $P_1, \dots P_{2n+1}$ which lie on one side of a diameter. Prove that

$$|\overrightarrow{OP_1} + \ldots + \overrightarrow{OP_{2n+1}}| \geq 1.$$

Problem 9.1.6. Let l_1 and l_2 be two parallel lines and $A, B \in l_1$, $A \neq B$. Using only a ruler, divide the segment AB into n equal parts, where $n \geq 2$.

Problem 9.1.7. Given $2n+1$ points in the plane such that no three of them are collinear, construct a $(2n+1)$-gon (self-intersecting polygons are allowed) for which the given points serve as the midpoints of its sides.

Problem 9.1.8. We are given a convex polygon in the plane. Show that we can pick a triangle formed by three of its vertices such that the sum of squares of its side lengths is at least the sum of squares of the side lengths of the polygon.

Problem 9.1.9. (**TOT 2001**) Prove that there exist 2001 convex polyhedra such that any three of them do not have any common points, but any two of them touch each other (i.e. they have at least one common boundary point but no common inner points).

Problem 9.1.10. (**IMO 1992**) Let S be a finite set of points in three-dimensional space. Let S_x, S_y, S_z be the sets consisting of the orthogonal projections of the points of S onto the yz-plane, zx-plane and xy-plane, respectively. Prove that

$$|S|^2 \le |S_x| \cdot |S_y| \cdot |S_z|,$$

where $|A|$ denotes the number of elements in the finite set A.
(Note: the orthogonal projection of a point onto a plane is the foot of the perpendicular from that point to the plane).

Problem 9.1.11. Prove that any convex n-gon which is not a parallelogram can be enclosed by a triangle whose sides lie along three sides of the given n-gon.

Problem 9.1.12. Given a circle and n points in the plane, construct an n-gon (self-intersecting polygons are allowed) which is inscribed in the given circle and such that the lines determined by its sides pass through the given points.

9.2 Induction in Calculus

9.2.1 Examples

In this chapter we are going to show how Induction can be used to solve Calculus problems such as proving identities about functions and their derivatives, how to calculate limits or several integrals. Before we look at such examples, let us prove the following identity, known as Leibniz's formula:

Example 9.2.1. Let I be an interval and let $f, g : I \to \mathbb{R}$ be two n times differentiable functions. Prove that:

$$(fg)^{(n)} = \binom{n}{0} f^{(n)} g + \binom{n}{1} f^{(n-1)} g^{(1)} + \binom{n}{2} f^{(n-2)} g^{(2)} + \ldots + \binom{n}{n} f g^{(n)},$$

where $f^{(k)}$ denotes the k-th derivative of f.

Solution. We prove the result by induction on n. For $n = 1$ we obtain

$$(fg)' = f'g + fg',$$

which is the well-known product formula for the derivative.

Assume now that the result holds for some $n \geq 1$, so that we have

$$(fg)^{(n)} = \binom{n}{0} f^{(n)} g + \binom{n}{1} f^{(n-1)} g^{(1)} + \binom{n}{2} f^{(n-2)} g^{(2)} + \ldots + \binom{n}{n} f g^{(n)}.$$

By differentiating this relation, we obtain

$$\begin{aligned}
(fg)^{(n+1)} &= \binom{n}{0} \left(f^{(n+1)} g + f^{(n)} g^{(1)} \right) + \binom{n}{1} \left(f^{(n)} g^{(1)} + f^{(n-1)} g^{(2)} \right) + \ldots \\
&\quad + \binom{n}{n} \left(f^{(1)} g^{(n)} + f g^{(n+1)} \right) \\
&= \binom{n+1}{0} f^{(n+1)} g + \left(\binom{n}{0} + \binom{n}{1} \right) f^{(n)} g^{(1)} + \ldots \\
&\quad + \left(\binom{n}{n-1} + \binom{n}{n} \right) f^{(1)} g^{(n)} + \binom{n+1}{n+1} f g^{(n+1)}.
\end{aligned}$$

Using Pascal's identity

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k},$$

we obtain

$$(fg)^{(n+1)} = \binom{n+1}{0} f^{(n+1)} g + \binom{n+1}{1} f^{(n)} g^{(1)} + \ldots + \binom{n+1}{n+1} f g^{(n+1)},$$

so $P(n+1)$ holds, completing our induction.

As an example where we can prove an identity relating higher order derivatives of a function, we have the following:

Example 9.2.2. Let $f : (0, \infty) \to \mathbb{R}$ be an infinitely differentiable function. Show that

$$\left(x^{n-1} f\left(\frac{1}{x}\right)\right)^{(n)} = \frac{(-1)^n}{x^{n+1}} f^{(n)}\left(\frac{1}{x}\right), \quad \text{for any positive integer} \quad n,$$

where $f^{(n)}$ denotes the n-th derivative of f.

Solution. We prove the result by induction on n. For $n = 1$, we have to show that

$$\left(f\left(\frac{1}{x}\right)\right)' = \frac{-1}{x^2} f'\left(\frac{1}{x}\right),$$

which follows from the well-known formula $(f \circ g(x))' = f'(g(x)) \cdot g'(x)$.

Assume now that the identity holds for some $n \geq 1$. Then

$$\left(x^n f\left(\frac{1}{x}\right)\right)^{(n+1)} = \left(\left(x^{n-1} f\left(\frac{1}{x}\right)\right) \cdot x\right)^{(n+1)}$$

$$= \left(x^{n-1} f\left(\frac{1}{x}\right)\right)^{(n+1)} \cdot x + \binom{n+1}{1} \left(x^{n-1} f\left(\frac{1}{x}\right)\right)^{(n)} \quad \text{(by Leibniz's formula)}$$

$$= \left(\frac{(-1)^n}{x^{n+1}} f^{(n)}\left(\frac{1}{x}\right)\right)' \cdot x + (n+1)\frac{(-1)^n}{x^{n+1}} f^{(n)}\left(\frac{1}{x}\right) \quad \text{(induction hypothesis)}$$

$$= (-1)^n \left(\frac{-(n+1)}{x^{n+2}} f^{(n)}\left(\frac{1}{x}\right) + \frac{1}{x^{n+1}} \left(\frac{-1}{x^2}\right) f^{(n+1)}\left(\frac{1}{x}\right)\right) \cdot x$$

$$+ (n+1) \cdot \frac{(-1)^n}{x^{n+1}} f^{(n)}\left(\frac{1}{x}\right) = \frac{(-1)^{n+1}}{x^{n+2}} f^{(n+1)}\left(\frac{1}{x}\right).$$

This establishes the result for $n+1$ and completes our proof.

Example 9.2.3. Let $I_n = \int\limits_1^e (\ln(x))^n \, dx$. Prove that

$$I_n = e\sum_{k=0}^{n-2}(-1)^k \frac{n!}{(n-k)!} + (-1)^{n-1}n!, \quad \text{for} \quad n \geq 2.$$

Solution. We prove the result by induction on n. The base case is $n = 2$. We have

$$\begin{aligned}
I_2 = \int_1^e \ln^2(x) \, dx &= \int_1^e (x)' \ln^2(x) \, dx \\
&= x\ln^2(x)\Big|_1^e - \int_1^e x \cdot 2\ln(x) \cdot \frac{1}{x} \, dx \\
&= e - 2\int_1^e \ln(x) \, dx \\
&= e - 2\int_1^e (x)' \ln(x) \, dx \\
&= e - 2\left(x\ln(x)|_1^e - \int_1^e x \cdot \frac{1}{x} \, dx \right) \\
&= e - 2,
\end{aligned}$$

which proves the base case.

Assume now that the result holds for some $n \geq 2$. Let $f(x) = (\ln(x))^n$ and $g(x) = x$. We have

$$\begin{aligned}
I_n &= \int_1^e g'(x) \cdot f(x) \, dx \\
&= f(x) \cdot g(x)|_1^e - \int_1^e g(x) \cdot f'(x) \, dx \\
&= x\,(\ln(x))^n|_1^e - n\int_1^e (\ln(x))^{n-1} \, dx \\
&= e - nI_{n-1}.
\end{aligned}$$

Hence

$$\begin{aligned}I_{n+1} &= e-(n+1)I_n\\ &= e-e\sum_{k=0}^{n-2}(-1)^k\cdot(n+1)\cdot\frac{n!}{(n-k)!}+(-1)^n\cdot(n+1)\cdot n!\\ &= e\sum_{k=0}^{n-1}(-1)^k\cdot\frac{(n+1)!}{(n+1-k)!}+(-1)^n\cdot(n+1)!.\end{aligned}$$

This completes the proof for the induction step.

Remark. Notice that for $1\le x\le e$ we have that $0\le\ln(x)\le 1$, so $0\le I_n\le e-1$, which implies that $\lim_{n\to\infty}\frac{I_n}{n!}=0$. Using the formula that we derived for I_n in the above question, this implies that

$$\lim_{n\to\infty}\left(\frac{1}{2!}-\frac{1}{3!}+\frac{1}{4!}+\ldots+(-1)^n\frac{1}{n!}\right)=\frac{1}{e}.$$

The following problems give a sample of how induction can be used to compute limits of sequences, functions or inequalities that are solved using Calculus techniques:

Example 9.2.4. Consider the sequence $(x_n)_{n\ge 1}$ defined by

$$x_1=1\quad\text{and}\quad x_{n+1}=\frac{n}{x_n}+\frac{x_n}{n}\quad\text{for all}\quad n\ge 1.$$

Show that

$$\lim_{n\to\infty}\frac{x_n}{\sqrt{n}}=1.$$

Solution. From the recursion relation we immediately obtain that $x_2=x_3=2$. To show that

$$\lim_{n\to\infty}\frac{x_n}{\sqrt{n}}=1,$$

it suffices to find two functions $f,g:\mathbb{Z}_+\to\mathbb{R}$ for which we know that

$$\lim_{n\to\infty}\frac{f(n)}{\sqrt{n}}=\lim_{n\to\infty}\frac{g(n)}{\sqrt{n}}=1\quad\text{and}\quad g(n)\ge x_n\ge f(n),$$

for all n sufficiently large.

A natural choice for f is $f(n) = \sqrt{n}$. To figure out the choice for g, we study the conditions that g has to satisfy so that the statement

$$P(n): \quad g(n) \geq x_n \geq \sqrt{n} \quad \text{holds for all} \quad n \geq 3$$

can be proved by induction.

The base case reads $g(3) \geq 2 \geq \sqrt{3}$. Assume now that $P(n)$ is true for some $n \geq 3$. From $x_{n+1} = \frac{n}{x_n} + \frac{x_n}{n}$ and $g(n) \geq x_n \geq \sqrt{n}$ we obtain

$$x_{n+1}^2 = \frac{n^2}{x_n^2} + \frac{x_n^2}{n^2} + 2 \geq \frac{n^2}{g(n)^2} + \frac{n}{n^2} + 2 = \frac{n^2}{g(n)^2} + \frac{1}{n} + 2.$$

To deduce $x_{n+1} \geq \sqrt{n+1}$, we would like to have $\frac{n^2}{g(n)^2} + \frac{1}{n} + 2 \geq n+1$, for all $n \geq 3$. It would thus suffice to set $\frac{n^2}{g(n)^2} = n-1$, i.e. $g(n) = \frac{n}{\sqrt{n-1}}$. Notice that this choice for g satisfies $g(3) \geq 2$ and we can also prove by induction that $x_n \geq \sqrt{n}$ for all $n \geq 3$, as the above analysis shows.

It remains to check whether $x_n \leq \frac{n}{\sqrt{n-1}}$ for all $n \geq 3$. The base case is verified. Assume now that the statement holds for some $n \geq 3$. To determine whether we can perform the induction step, let us study the function

$$h(x) = \frac{n}{x} + \frac{x}{n}.$$

We have that $h'(x) = \frac{1}{n} - \frac{n}{x^2}$, so as long as $x < n$, we have that $h'(x) < 0$. In particular, h is decreasing on the interval $\left[\sqrt{n}, \frac{n}{\sqrt{n-1}}\right]$. We proved above that $x_n > \sqrt{n}$, so we have

$$x_{n+1} = h(x_n) \leq h(\sqrt{n}) = \sqrt{n} + \frac{1}{\sqrt{n}} = \frac{n+1}{\sqrt{n}}.$$

This establishes the induction step and thus

$$\frac{n}{\sqrt{n-1}} \geq x_n \geq \sqrt{n} \quad \text{for all} \quad n \geq 3.$$

Since

$$\lim_{n\to\infty} \frac{\frac{n}{\sqrt{n-1}}}{\sqrt{n}} = \lim_{n\to\infty} \frac{\sqrt{n}}{\sqrt{n}} = 1,$$

we must have that $\lim\limits_{n\to\infty} \frac{x_n}{\sqrt{n}} = 1$ as well.

Example 9.2.5. Evaluate the limit

$$\lim_{x\to 0} \frac{1-\cos(x)\cos(2x)\dots\cos(nx)}{x^2}, \quad \text{where} \quad n \in \mathbb{Z}, n \geq 1.$$

Solution. We will show that the limit exists for all $n \geq 1$ simultaneously with finding a suitable recurrence relation for it. So let us denote the required limit by L_n. For $n = 1$ we have

$$L_1 = \lim_{x\to 0} \frac{1-\cos(x)}{x^2} = \lim_{x\to 0} \frac{2\sin^2\left(\frac{x}{2}\right)}{x^2} = \frac{1}{2} \lim_{x\to 0} \left(\frac{\sin\left(\frac{x}{2}\right)}{\frac{x}{2}} \right)^2 = \frac{1}{2}.$$

We also have that

$$\begin{aligned} L_n - L_{n-1} &= \lim_{x\to 0} \frac{(1-\cos(nx))\cos(x)\cos(2x)\dots\cos((n-1)x)}{x^2} \\ &= \lim_{x\to 0} \frac{1-\cos(nx)}{x^2} = \lim_{x\to 0} \frac{2\sin^2\left(\frac{nx}{2}\right)}{x^2} \\ &= \frac{n^2}{2} \lim_{x\to 0} \left(\frac{\sin\left(\frac{nx}{2}\right)}{\frac{nx}{2}} \right)^2 = \frac{n^2}{2}. \end{aligned}$$

Hence by the induction hypothesis the limit L_n exists.
Using this to compute a few values for L_n, we are lead to guess

$$L_n = \frac{n(n+1)(2n+1)}{12}.$$

We prove this by induction on n. For $n = 1$, the result was proved above. Assume now that the result holds for $n-1$, $n \geq 2$, that is

$$L_{n-1} = \frac{(n-1)n(2n-1)}{12}.$$

From the recurrence relation $L_n - L_{n-1} = \frac{n^2}{2}$, we obtain that

$$\begin{aligned} L_n &= L_{n-1} + \frac{n^2}{2} \\ &= \frac{(n-1)n(2n-1)}{12} + \frac{n^2}{2} \\ &= \frac{n(2n^2+3n+1)}{12} \\ &= \frac{n(n+1)(2n+1)}{12}. \end{aligned}$$

This proves the result for n, completing our induction.

Example 9.2.6. Prove that for any positive integer n and any $x \in (0, \pi)$, we have

$$\frac{\sin(x)}{1} + \frac{\sin(2x)}{2} + \ldots + \frac{\sin(nx)}{n} > 0.$$

Solution. Let

$$S_n(x) = \frac{\sin(x)}{1} + \frac{\sin(2x)}{2} + \ldots + \frac{\sin(nx)}{n}.$$

We prove by induction that the statement

$$P(n): \quad S_n(x) > 0 \quad \text{if} \quad x \in (0, \pi)$$

is true for all positive integers n.

For $n = 1$ we have $S_1(x) = \sin x > 0$, since $x \in (0, \pi)$, so $P(1)$ is true. Assume now that $P(n)$ is true for some $n \geq 1$.

We have $\lim_{x \to 0} S_{n+1}(x) = 0$ and $\lim_{x \to \pi} S_{n+1}(x) = 0$. Moreover, we can analyze the extreme values for $S_{n+1}(x)$ by looking at the roots of $S'_{n+1}(x) = 0$.

$$S'_{n+1}(x) = \cos(x) + \cos(2x) + \ldots + \cos\left((n+1)x\right).$$

Now using the formula

$$\cos(x) + \cos(2x) + \ldots + \cos\left((n+1)x\right) = \frac{\cos\left(\frac{n+2}{2}x\right)\sin\left(\frac{n+1}{2}x\right)}{\sin\left(\frac{x}{2}\right)},$$

(which can be proved easily by induction) we have that $S'_{n+1}(x) = 0$ has the solutions

$$x = \frac{2k\pi}{n+1} \quad \text{and} \quad x = \frac{(2k+1)\pi}{n+2},$$

where k is an integer subject to $0 < \frac{2k\pi}{n+1} < \pi$ and $0 < \frac{(2k+1)\pi}{n+2} < \pi$, for the second solution. Since

$$S_{n+1}(x) = S_n(x) + \frac{\sin((n+1)\pi)}{n+1},$$

for the first solution we have

$$S_{n+1}\left(\frac{2k\pi}{n+1}\right) = S_n\left(\frac{2k\pi}{n+1}\right) + 0 > 0,$$

using the induction hypothesis, while for the second solution we obtain

$$\begin{aligned}
S_{n+1}\left(\frac{(2k+1)\pi}{n+2}\right) &= S_n\left(\frac{(2k+1)\pi}{n+2}\right) + \frac{\sin\left(\frac{(2k+1)(n+1)\pi}{n+2}\right)}{n+1} \\
&> \frac{\sin\left(\frac{(2k+1)(n+1)\pi}{n+2}\right)}{n+1} \\
&= \frac{\sin\left(2k\pi + \frac{(n+1-2k)\pi}{n+2}\right)}{n+1} \\
&= \frac{\sin\left(\frac{(n+1-2k)\pi}{n+2}\right)}{n+1},
\end{aligned}$$

where again, for the first inequality we used the induction hypothesis.
Now, since $0 < \frac{2k+1}{n+2} < 1$, we have that $0 < \frac{n+1-2k}{n+2} < 1$, so

$$\sin\left(\frac{(n+1-2k)\pi}{n+2}\right) > 0.$$

So we obtain that $S_{n+1}\left(\frac{(2k+1)\pi}{n+2}\right) > 0$, for this case as well.
As for the extreme values we have that $S_{n+1}(x) > 0$ and

$$\lim_{x\to 0} S_{n+1}(x) = \lim_{x\to\pi} S_{n+1}(x) = 0,$$

it must be that $S_{n+1}(x) > 0$ for all $x \in (0, \pi)$. Thus $P(n)$ holds for all positive integers n, establishing our result.

9.2.2 Proposed Problems

Problem 9.2.1. Let $f : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$, $f(x) = \arctan(x)$. Show that if n is a positive integer, we have that

$$f^{(n)}(0) = \begin{cases} 0 & \text{if } n = 2k; \\ (-1)^k (2k)! & \text{if } n = 2k+1, \end{cases}$$

where $f^{(n)}$ denotes the n-th derivative of f.

Problem 9.2.2. Let $f : [-1, 1] \to (-\frac{\pi}{2}, \frac{\pi}{2})$, $f(x) = \arcsin(x)$. Prove that

$$f^{(n)}(0) = \begin{cases} 0 & \text{if } n = 2k; \\ (1 \cdot 3 \cdot 5 \ldots \cdot (2k-1))^2 & \text{if } n = 2k+1, \end{cases}$$

where $f^{(n)}$ denotes the n-th derivative of f.

Problem 9.2.3. Let $f : [0, \infty) \to \mathbb{R}$ be given by

$$f(x) = \int_0^1 e^{-t} t^{x-1} \, dt.$$

Prove that for any non-negative integer n we have

$$f(n+1) = n! - \frac{1}{e} \sum_{k=0}^{n} \frac{n!}{(n-k)!}.$$

Problem 9.2.4. Prove that for any $n \in \mathbb{Z}$, $n \geq 1$, one has

$$\lim_{x \to 0} \frac{n! x^n - \sin(x)\sin(2x)\ldots\sin(nx)}{x^{n+2}} = \frac{n(2n+1)}{36} \cdot n!.$$

Problem 9.2.5. Prove that if m, n are non-negative integers with $m > n$, then

$$\int_0^{\pi} \cos^n(x)\cos(mx) \, dx = 0.$$

Using this, deduce the value of

$$J_n = \int_0^{\pi} \cos^n(x)\cos(nx) \, dx, \quad \text{for all non-negative integers} \quad n.$$

Problem 9.2.6. Let $a_1, a_2, \ldots, a_{2001}$ be non-zero real numbers. Prove that there exists a real number x, such that

$$\sin(a_1 x) + \sin(a_2 x) + \ldots + \sin(a_{2001} x) < 0.$$

Problem 9.2.7. Let $f : \mathbb{R} \to \mathbb{R}$ be a function defined by

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{if } x \neq 0; \\ 0 & \text{if } x = 0. \end{cases}$$

Prove that f is infinitely differentiable at 0 and $f^{(n)}(0) = 0$, where $f^{(n)}$ represents the n-th derivative of f.

Problem 9.2.8. Let $n > 1$ be a positive integer, $0 < a_1 < \ldots < a_n$ and let $c_1, \ldots, c_n$ be non-zero real numbers. Prove that the number of roots of the equation

$$c_1 \cdot a_1^x + \ldots + c_n \cdot a_n^x = 0$$

is not larger than the number of negative elements of the sequence $c_1 c_2$, $c_2 c_3$, $\ldots$, $c_{n-1} c_n$.

Problem 9.2.9. Prove that for any $|x| < 1$ and any positive integer k one has

$$\sum_{n=0}^{\infty} x^n \binom{n+1}{k} = \frac{x^{k-1}}{(1-x)^{k+1}}.$$

9.3 Induction in Algebra

9.3.1 Examples

Most of the problems that we will look at in this chapter involve polynomials. It is worth mentioning that the applications of Induction in Linear Algebra and Abstract Algebra go far beyond this and it would take us an entirely separate book to give a comprehensive exposition of them. We selected below some beautiful examples that resemble the common techniques one has to keep in mind for questions that involve Induction and polynomials. Further elegant applications are given in the Proposed problems section.

Example 9.3.1. (**St. Petersburg**) Let $P(X)$ be a polynomial of degree $n \geq 1$ with real coefficients such that $|P(x)| \leq 1$ for all $0 \leq x \leq 1$. Prove that

$$\left|P\left(-\frac{1}{n}\right)\right| \leq 2^{n+1} - 1.$$

Solution. If we let $f(x) = P\left(\frac{x}{n}\right)$, the problem is equivalent to showing that if $|f(x)| \leq 1$ for $x \in [0, n]$, then $|f(-1)| \leq 2^{n+1} - 1$. We do this by induction on n. The base case $n = 1$ is immediate.

For the induction step from $n-1$ to n, let f be of degree n and consider $g(x) := f(x) - f(x+1)$. Then $\frac{1}{2}g$ satisfies the condition that we want and it is of degree at most $n-1$. So

$$|f(-1)| = |f(0) + g(-1)| \leq 1 + 2(2^n - 1) = 2^{n+1} - 1,$$

which completes our proof.

Example 9.3.2. (**TOT 2005**) For any function $f(x)$, define $f^1(x) = f(x)$ and $f^n(x) = f(f^{n-1}(x))$ for any integer $n \geq 2$. Does there exist a quadratic polynomial $f(x)$ such that the equation $f^n(x) = 0$ has exactly 2^n real roots for every positive integer n?

Solution. The answer is yes. An example of such a function is $f(x) = x^2 - 2$. For $f(x) = 0$, we have that $x^2 = 2$ and the roots are $\pm\sqrt{2}$.

We claim that every root of $f^{n+1}(x) = 0$ has the form $r_{n+1} = \pm\sqrt{2 \pm r_n}$, for some root r_n of $f^n(x) = 0$. Indeed,

$$f^{n+1}(r_{n+1}) = f^n\left(\left(\pm\sqrt{2 \pm r_n}\right)^2 - 2\right) = f^n(\pm r_n) = 0.$$

Since $f^{n+1}(x)$ has degree equal to twice the degree of $f^n(x)$, these are all the roots.

We are now going to prove that $\pm r_n$ are real and $|r_n| < 2$ for all n. We do this by induction on n. For $n = 1$, this is certainly true from above.

Assume now that the result holds for some $n \geq 1$. Since $|r_n| < 2$, we have that $2 \pm r_n > 0$, so $r_{n+1} = \pm\sqrt{2 \pm r_n}$ are real. Moreover, $|2 \pm r_n| \leq 2 + |r_n| < 4$, hence $|r_{n+1}| < 2$.

Finally, observe that $\sqrt{2}$ and $-\sqrt{2}$ are distinct roots and that distinct roots of $f^n(x) = 0$ lead to distinct roots of $f^{n+1} = 0$.

Example 9.3.3. (**Iran TST 1998**) Let $p(X)$ be a polynomial with integer coefficients such that $p(n) > n$ for every positive integer n. Consider the sequence $(x_n)_{n \in \mathbb{N}^*}$ defined by

$$x_1 = 1 \quad \text{and} \quad x_{n+1} = p(x_n), \quad \text{for all} \quad n \geq 1.$$

We know that for any positive integer N, there exists a term of the above sequence which is divisible by N. Prove that $p(X) = X + 1$.

Solution. Notice that since $p(X)$ is a polynomial, it suffices to prove that $p(n) = n + 1$ for an infinite number of integers n. We will prove this by showing that $x_n = n$ by induction on n. For $n = 1$, the statement holds from the hypothesis.

Assume now that we have proved the statement for some $n \geq 1$. We know that $p(x_n) > x_n = n$. If $x_{n+1} = p(x_n) \neq n + 1$ then $x_{n+1} \geq n + 2$, as $p(X)$ has integer coefficients. Let $k = x_{n+1} - x_n$. The key observation is the following fact: for integers u and v we have that $(u - v) \mid (f(u) - f(v))$, for any polynomial f with integer coefficients. From the way we defined our sequence x_n, we have that

$$k \mid (x_{n+2} - x_{n+1}) \mid \ldots \mid (x_{M+1} - x_M) \quad \forall M \geq n + 1.$$

From the hypothesis, for $N = kx_{n+1}$, there is a $t \geq n+1$ such that $kx_{n+1} \mid x_t$. By setting $M = t-1$ in the above chain of divisibilities, we have $x_t \equiv x_n \pmod{k}$. So $(x_{n+1} - x_n) \mid x_n$. Thus $x_{n+1} \leq 2x_n = 2n$. Note that if $n = 1$, this implies $x_2 = 2$ so we may assume $n \geq 2$. On the other hand,

$$n - 1 = (x_n - x_1) \mid (x_{n+1} - x_2) = x_{n+1} - 2,$$

so we conclude that $x_{n+1} \geq 2n$ and $x_{n+1} = 2n$. If $n = 2$, then from $x_1 = 1$, $x_2 = 2$, and $x_3 = 4$, we conclude that $p(1) = 2$ and $p(2) = 4$. But then it is easy to see that x_n alternates between being 1 and 2 modulo 3, and never a multiple of 3. Thus we must have $x_3 = 3$ and we may assume $n \geq 3$. But then

$$n - 2 = (x_n - x_2) \mid (x_{n+1} - x_3) = 2n - 3,$$

which is a contradiction. Therefore, $x_{n+1} = n+1$, as required. This completes our proof.

Example 9.3.4. (**IMC 2014**) Prove that there exist positive real numbers $a_0, a_1, \ldots, a_n$ such that the polynomial $\pm a_n x^n \pm a_{n-1}x^{n-1} \pm \ldots \pm a_0$ has n distinct real roots for all possible choices of signs.

Solution. We prove the result by induction on $n \geq 1$. The base case $n = 1$ is immediate.

Assume now that the result holds for some $n \geq 1$ and let

$$P(x) = \pm a_n x^n \pm a_{n-1}x^{n-1} \pm \ldots \pm a_0$$

be such a polynomial. Then the polynomial

$$\overline{P(x)} = \pm a_n x^{n+1} \pm a_{n-1}x^n \pm \ldots \pm a_0 x,$$

where $a_0 \neq 0$ has $n+1$ distinct real roots. Since there are no double roots, the local extreme points of $\overline{P(x)}$ are not among the roots (recall that the local extreme points are the solutions to $\overline{P(x)}' = 0$). Let $s_1 < \ldots < s_n$ be the solutions to $\overline{P(x)}' = 0$. By the previous observation, there must be some $\varepsilon > 0$ such that $\left|\overline{P(s_i)}\right| > \varepsilon$, for all $i = 1, \ldots, n$.

We claim that the polynomial

$$Q(x) = \pm a_n x^{n+1} \pm a_{n-1}x^n \pm \ldots \pm a_0 x \pm \varepsilon$$

has $n+1$ distinct roots for all possible choices of signs. Notice that for $i = 1, \ldots, n-1$, the numbers $Q(s_i)$ and $Q(s_{i+1})$ have different signs, so there must be a root of Q in each of the intervals (s_i, s_{i+1}). Considering also the behaviour on $(-\infty, s_1)$ and $(s_n, +\infty)$, we deduce that $Q(x)$ has indeed $n+1$ distinct real roots. This completes our proof.

Example 9.3.5. Let $P(X) = a_d X^d + \ldots + a_0$ be a polynomial. Prove that $P(X)$ is divisible by $(X-1)^m$ if and only if for all $s = 0, 1, \ldots, m-1$ we have

$$a_n(n+1)^s + a_{n-1}n^s + \ldots + a_1 2^s + a_0 = 0.$$

Solution. Assume first that $P(X)$ is divisible by $(X-1)^m$ and write

$$P(X) = (X-1)^m Q(X),$$

where

$$Q(X) = \sum_{i=0}^{n-m} q_i X^{n-m-i}.$$

Since

$$(X-1)^m = \sum_{j=0}^{m} \binom{m}{j} (-1)^j X^{m-j},$$

we have

$$\begin{aligned} P(X) &= \left(\sum_{i=0}^{n-m} q_i X^{n-m-i}\right)\left(\sum_{j=0}^{m} \binom{m}{j} (-1)^j X^{m-j}\right) \\ &= \sum_{i=0}^{n-m} q_i \sum_{j=0}^{m} (-1)^j \binom{m}{j} X^{n-i-j}. \end{aligned}$$

The statement of the problem reduces to showing that

$$\sum_{i=0}^{n-m} q_i \left(\sum_{j=0}^{m} (-1)^j \binom{m}{j} (n-i-j+1)^s\right) = 0.$$

The term $\sum_{j=0}^{m}(-1)^j\binom{m}{j}(n-i-j+1)^s$ is equal to the m-th difference polynomial of X^s (evaluated at $n-i-1$) which is identically 0. This establishes the first implication.

For use below, note that the next term $\sum_{j=0}^{m}(-1)^j\binom{m}{j}(n-i-j+1)^m$ is the m-th difference of X^m. Hence it is the constant polynomial with value $m!$. Thus in particular, $\sum_{j=0}^{m}\binom{m}{j}(m+1-j)^m = m!$.

We prove the converse statement by induction on m. The base case $m=1$ is clear. Assume now that the statement is true for $m=k$, some $k \geq 1$. By the Division Algorithm, we have

$$P(X) = (X-1)^{k+1}Q(X) + R(X), \quad \text{where} \quad \deg(R(X)) \leq k.$$

Using the induction hypothesis, $P(X)$ is divisible by $(X-1)^k$, so we must have $R(X) = a(X-1)^k$. Notice that the polynomial $(X-1)^{k+1}Q(X)$ satisfies the conditions for $s = 0, 1, \ldots, k$, so we only need to check the condition for $R(X)$ with $s = k$. We have

$$R(X) = a\sum_{j=0}^{k}(-1)^j\binom{k}{j}X^{k-j},$$

thus

$$a((k+1)^k - k \cdot k^k + \ldots) = 0.$$

We saw above that this sum is $k!$ and hence nonzero, therefore have $a = 0$, which completes the proof.

9.3.2 Proposed Problems

Problem 9.3.1. (**Iran 1985**) Let α be an angle such that $\cos(\alpha) = \frac{p}{q}$, where p and q are two integers. Prove that the number $q^n \cos(n\alpha)$ is an integer, for any $n \in \mathbb{N}^*$.

Problem 9.3.2. Prove that there is a monic polynomial $P \in \mathbb{Z}[X]$ of degree n, such that $P(2\cos x) = 2\cos nx$ and $P(x + \frac{1}{x}) = x^n + \frac{1}{x^n}$.

Problem 9.3.3. Prove that there is no polynomial $P \in \mathbb{R}[X]$ of degree $n \geq 1$ such that $P(x) \in \mathbb{Q}$ for all $x \in \mathbb{R}\backslash\mathbb{Q}$.

Problem 9.3.4. Let $F(X)$ and $G(X)$ be two polynomials with real coefficients such that the points

$$(F(1), G(1)), (F(2), G(2)), \ldots, (F(2011), G(2011))$$

are vertices of a regular 2011-gon.
Prove that $\deg(F) \geq 2010$ or $\deg(G) \geq 2010$.

Problem 9.3.5. Is $\cos 1^\circ$ rational?

Problem 9.3.6. (**Poland 2000**) Let P be a polynomial of odd degree satisfying the identity

$$P(x^2 - 1) = P(x)^2 - 1.$$

Prove that $P(x) = x$ for all real x.

Problem 9.3.7. (**Bulgaria**) Prove that there exists a quadratic polynomial $f(X)$ such that $f(f(X))$ has 4 non-positive real roots and $f^n(X)$ has 2^n real roots, where f^n denotes the composite of f with itself n times.

Problem 9.3.8. Prove that if a rational function which is not a polynomial takes rational values at all positive integers, then it is the quotient of two coprime polynomials, both having integer coefficients.

Problem 9.3.9. (**IMO 2007 shortlist**) Let $n > 1$ be an integer. In the space, consider the set

$$S = \{(x, y, z) \mid x, y, z \in \{0, 1, \ldots, n\}, x + y + z > 0\}.$$

Find the smallest number of planes that jointly contain all $(n+1)^3 - 1$ points of S but none of them passes through the origin.

Problem 9.3.10. Prove that for every positive integer n there exists a polynomial $p(x)$ with integer coefficients such that $p(1), p(2), \ldots, p(n)$ are distinct powers of 2.

Problem 9.3.11. (**Moscow 2013**) Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $f(x) \in \mathbb{Z}$ for all $x \in \mathbb{Z}$. For every prime p, there exists a polynomial $Q_p(X)$ of degree less than 2013, with integer coefficients such that $f(n) - Q_p(n)$ is divisible by p for all positive integers n. Prove that there exists a polynomial $g(x) \in \mathbb{R}[X]$ such that for all positive integers we have $f(n) = g(n)$.

Chapter 10

Solutions

1 A Brief Overview of Induction

Problem 1.1. Prove that for all $n \geq 1$:

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} < \frac{1}{\sqrt{3n}}.$$

Solution. Let us look at what happens if we try to prove the statement by induction as given. For $n = 1$ we have $\frac{1}{2} < \frac{1}{\sqrt{3}}$, which is true as $2 > \sqrt{3}$.

For the inductive step, assume that

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} < \frac{1}{\sqrt{3n}},$$

for some $n \geq 1$. For the inductive step we would have to prove that

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} \cdot \frac{2n+1}{2n+2} < \frac{1}{\sqrt{3n+3}}$$

So let us make use of the inductive hypothesis. We have that

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} \cdot \frac{2n+1}{2n+2} < \frac{1}{\sqrt{3n}} \frac{2n+1}{2n+2}.$$

So if we managed to prove that

$$\frac{1}{\sqrt{3n}} \cdot \frac{2n+1}{2n+2} < \frac{1}{\sqrt{3n+3}},$$

we would be done. However, this inequality is false, since

$$\frac{1}{\sqrt{3n}} \cdot \frac{2n+1}{2n+2} < \frac{1}{\sqrt{3n+3}} \Leftrightarrow \left(\frac{2n+1}{2n+2}\right)^2 < \frac{n}{n+1},$$

which is further equivalent to

$$(2n+1)^2 < 4n(n+1) \Leftrightarrow 4 < 0.$$

However, if instead of $\frac{1}{\sqrt{3n}}$ we had something of the form $\frac{1}{\sqrt{3n+a}}$ for some constant a, at the inductive step we would have to prove that

$$\frac{1}{\sqrt{3n+a}} \cdot \frac{2n+1}{2n+2} < \frac{1}{\sqrt{3n+3+a}}.$$

This is equivalent to showing that

$$\left(\frac{2n+1}{2n+2}\right)^2 < \frac{3n+a}{3n+3+a}.$$

We have seen above that $a = 0$ is not quite enough to make this work. Also notice that the value of $\frac{3n+a}{3n+3+a}$ increases with a. So hopefully there is some positive real a for which $\frac{3n+a}{3n+3+a}$ is large enough to have

$$\left(\frac{2n+1}{2n+2}\right)^2 < \frac{3n+a}{3n+3+a}.$$

By cross multiplying and simplifying the terms, it turns out that $a = 1$ suffices. So we try to prove the stronger statement:

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} \leq \frac{1}{\sqrt{3n+1}} < \frac{1}{\sqrt{3n}}$$

The base case for $n = 1$ is clear. Now assume that the inequality holds for some n, so that

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n-1}{2n} \leq \frac{1}{\sqrt{3n+1}}$$

Now we multiply both sides by $\frac{2n+1}{2n+2}$. Then we have:

$$\frac{1}{2} \cdot \frac{3}{4} \cdots \frac{2n+1}{2n+2} \leq \frac{1}{\sqrt{3n+1}} \cdot \frac{2n+1}{2n+2}.$$

So it suffices to show

$$\frac{1}{\sqrt{3n+1}} \cdot \frac{2n+1}{2n+2} \leq \frac{1}{\sqrt{3n+4}}.$$

We have

$$\begin{aligned} \frac{1}{\sqrt{3n+1}} \cdot \frac{2n+1}{2n+2} &\leq \frac{1}{\sqrt{3n+4}} \\ (2n+1)\sqrt{3n+4} &\leq (2n+2)\sqrt{3n+1} \\ (4n^2+4n+1)(3n+4) &\leq (4n^2+8n+4)(3n+1) \\ 12n^3+28n^2+19n+4 &\leq 12n^3+28n^2+20n+4 \\ 0 &\leq n \end{aligned}$$

Problem 1.2. Show that for every positive integer $n \geq 2$ one has

$$\frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{n^2} < 1.$$

Solution. Notice that the sum on the left hand side is increasing so we must sharpen our inequality. Using similar ideas to those presented in the examples we can obtain the following stronger statement:

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

The above inequality can now be proved by induction. The base case is $n = 2$, for which we have to check $\frac{1}{4} < 1 - \frac{1}{2}$, which is clear.
Assume now that the result holds for some $n \geq 2$. We need to prove that

$$\frac{1}{2^2} + \ldots + \frac{1}{n^2} + \frac{1}{(n+1)^2} < 1 - \frac{1}{n+1}.$$

Using the induction hypothesis we have

$$\begin{aligned}\frac{1}{2^2}+\ldots+\frac{1}{n^2}+\frac{1}{(n+1)^2} &< 1-\frac{1}{n}+\frac{1}{(n+1)^2}\\ &< 1-\frac{1}{n+1},\end{aligned}$$

where the last inequality follows from $\frac{1}{(n+1)^2} < \frac{1}{n(n+1)}$. This completes the induction step and our proof.

Problem 1.3. (**China 2004**) Prove that every positive integer n, except a finite number of them, can be represented as a sum of 2004 positive integers: $n = a_1 + a_2 + \cdots + a_{2004}$, where $1 \le a_1 < a_2 < \cdots < a_{2004}$, and $a_i \mid a_{i+1}$ for all $1 \le i \le 2003$.

Solution. The first thing we should bear in mind with this kind of problems is that most likely the number 2004 plays no particular role in the actual setup. Hence we will try to prove the result when 2004 is replaced by a general number n. Also, when it comes to positive integers, the statement "all n, except a finite number of them" is implied by the statement "all $n \ge N$, for some positive integer N". So we begin by trying to prove the following stronger statement:

"For any positive integer k, there exists a positive integer N_k such that any integer $n \ge N_k$ can be represented as $n = \sum_{i=1}^{k} a_i$, with $1 \le a_1 < a_2 < \cdots < a_k$, and $a_i \mid a_{i+1}$ for all $1 \le i \le k-1$"

For the base case $k = 1$, we take $N_1 = 1$ and there is nothing to prove. Let us try to perform the induction step.

Let us assume the statement true for all values from 1 to some $k \ge 1$, and consider the case $k+1$.

Ideally, we would want to say $n = 1+(n-1)$, so by taking $N_{k+1} = N_k+1$, since $n \ge N_{k+1}$, we have $n-1 \ge N_k$ and considering now the representation for $n-1$ by $a_1 < a_2 < \ldots < a_k$, we obtain the representation $1 < a_1 < \ldots < a_k$ for n. However, this would imply that in the representation of $n-1$ we had $a_1 > 1$, which we cannot guarantee with our setup. For example, when we want to perform the inductive step with this idea from $k = 2$ to $k = 3$, we have $N_2 = 2$ and $N_3 = 3$, but to represent 4 as $4 = 1 + a_1 + a_2$, we would need to have that 3 can be represented as $3 = a_1 + a_2$, with $1 < a_1 < a_2$, which clearly cannot happen.

So to make this idea work, we need to strengthen our statement further. To figure out what the statement should be, let us consider another useful idea which is pretty common to this type of problems proved by induction. If we could prove the statement for prime powers (this includes the primes themselves), then for any positive integer n we can write $n = d \cdot m$, where d is a prime power and m is a positive integer. So as long as $d \geq N_{k+1}$, we obtain $n = (ma_1) + (ma_2) + \ldots + (ma_{k+1})$, where $d = a_1 + a_2 + \ldots + a_{k+1}$.

Now let us see what happens when we try to first prove the statement for prime powers. The case we first need to consider is when n is itself a prime. Since a prime cannot be factored further, it seems that we need to resort again to the idea of writing $n = 1 + (n - 1)$ and we run into the same problem as before. However, the extra piece of information that we now have is that except when $n = 3$, for any prime n, $n - 1$ will be composite. So we may be able to show that whenever n is not prime, we can write $n = \sum_{i=1}^{k} a_i$, with $1 \leq a_1 < a_2 < \cdots < a_k$, and $a_i \mid a_{i+1}$ for all $1 \leq i \leq k - 1$, and $1 < a_1$. Let us take this as our new statement. Namely we want to prove by induction the following:

"For any positive integer k, there exists a positive integer N_k such that any integer $n \geq N_k$ can be represented as $n = \sum_{i=1}^{k} a_i$, with $1 \leq a_1 < a_2 < \cdots < a_k$, and $a_i \mid a_{i+1}$ for all $1 \leq i \leq k - 1$, and $1 < a_1$ when n is not a prime."

As before, the base case $k = 1$ is clear.

Let us now assume the statement true for all values from 1 to some $k \geq 1$, and consider the case $k + 1$. With the prime power decomposition idea that we discussed above at the back of our mind, we consider the following cases:

Case 1. If $n > 3$ is a prime, write $n = 1 + (n - 1)$. If $n - 1 \geq N_k$, then $n - 1$ is even, thus not a prime, so it can be represented as $n - 1 = \sum_{i=1}^{k} b_i$, with $1 < b_1$.

Then we can write $n = \sum_{i=1}^{k+1} a_i$ by taking $a_1 = 1$ and $a_i = b_{i-1}$ for $2 \le i \le k+1$.

Case 2. If $n = 2^\ell \ge 2^5$, we write $n = 2^j + 2^j(2^{\ell-j} - 1)$ with $j = 1$ for odd ℓ, respectively $j = 2$ for even ℓ. This assures us that $2^{\ell-j} - 1$ is not a prime, and if $2^{\ell-j} - 1 \ge N_k$, it can be represented as $2^{\ell-j} - 1 = \sum_{i=1}^{k} b_i$, with $1 < b_1$, then we can write $n = \sum_{i=1}^{k+1} a_i$ by taking $a_1 = 2^j > 1$ and $a_i = 2^j b_{i-1}$, for $2 \le i \le k+1$.

Case 3. If $n = p^\ell \ge p^3$, with $3 \le p \le N_k$ an odd prime, we write $n = p + p(p^{\ell-1} - 1)$. $p^{\ell-1} - 1$ is not a prime, and if $p^{\ell-1} - 1 \ge N_k$, then it can be represented as $p^{\ell-1} - 1 = \sum_{i=1}^{k} b_i$, with $1 < b_1$. Therefore we can write $n = \sum_{i=1}^{k+1} a_i$ by taking $a_1 = p > 1$ and $a_i = pb_{i-1}$, for $2 \le i \le k+1$.

Case 4. Finally, if n is a large enough composite number, then it will have a factor d, either a prime larger than N_k, or a power of a prime not larger than N_k, but with exponent large enough to fall into one of the cases described above. Then $n = dm$, and we can represent $d = \sum_{i=1}^{k+1} a_i$, hence $n = \sum_{i=1}^{k+1} (ma_i)$, with $1 < ma_1$. Thus there exists a N_{k+1} with the required property.

This completes the induction proof of our statement and it implies in particular the statement of the original question.

Problem 1.4. The sequence $(a_n)_{n\ge1}$ is defined by

$$a_1 = \frac{1}{2}, \; a_{n+1} = \frac{2n-1}{2n+2} a_n.$$

Prove that $a_1 + a_2 + \ldots + a_n < 1$ for all $n \ge 1$.

Solution. Let us sharpen the condition to something of the form

$$a_1 + a_2 + \ldots + a_{n-1} + f(n) \cdot a_n \le 1,$$

where $f(n) \geq 1$ is a number depending on n to be properly defined later. We want to prove the statement of the problem of the induction, so we have to check the base case and the induction step. The base case gives $f(1)a_1 \leq 1$ so $f(1) \leq 2$. For the induction step we need

$$a_1 + a_2 + \cdots + a_{n-2} + f(n-1)a_{n-1} \geq a_1 + a_2 + \cdots + a_{n-1} + f(n)a_n,$$

i.e.

$$a_n \leq \frac{f(n-1)-1}{f(n)} a_{n-1}.$$

As $a_n = \dfrac{2n-3}{2n} a_{n-1} \Rightarrow f(n-1) - 1 \geq \dfrac{2n-3}{2n} f(n)$.
We see that $f(n-1) = 2(n-1)$, $f(n) = 2n$ satisfy the conditions with equality so if we take $f(x) = 2x$ then we get the identity

$$a_1 + a_2 + \ldots + 2na_n = 1,$$

which in particular implies the inequality we are interested in.

Problem 1.5. Consider S the set of all binary sequences of length n. It is partitioned into two sets A and B, each having 2^{n-1} elements. Two sequences, one from A and one from B, form a *tentacle* if they differ in only one position. Prove that there are at least 2^{n-1} tentacles.

Solution. An attempt of a direct induction fails, because we are restricted to the condition $|A| = |B| = 2^{n-1}$. So a first idea of employing the induction would be to generalize the statement to all partitions $S = A \cup B$. If $|A| = 2^n$, $|B| = 0$ we have no tentacles, and similarly if $|A| = 0$, $|B| = 2^n$.

Therefore, a reasonable claim would be that the number of tentacles is at least $\min\{|A|, |B|\}$. This claim is indeed provable by induction on n. The base case $n = 1$ is clear.

For the induction step, let S' be the subset of S formed of all binary sequences whose last digit is 0, and S'' the other sequences in S. Denote $A' = A \cap S'$, $A'' = A \cap S''$, $B' = B \cap S'$, $B'' = B \cap S''$. Let also $f(X), X \in S$ be the sequence of length $n-1$ obtained from X by dropping its last digit. (These constructions are natural as we want to use the induction hypothesis). Without loss of generality, $|A| \leq |B|$. Set $a = |A|$, $u = |A'|$, $v = |A''|$ so

$u+v=a$, $|B'|=2^{n-1}-u$, $|B''|=2^{n-1}-v$. Now it's clear by applying the induction hypothesis to the sets $f(S')=f(A')\cup f(B')$, $f(S'')=f(A'')\cup f(B'')$ that we have at least $\min\{u,2^{n-1}-u\}+\min\{v,2^{n-1}-v\}$ tentacles formed by pairs of sequences from A' and B' or A'' and B''. If $u\le 2^{n-2}, v\le 2^{n-2}$ we get at least $u+v$ sequences and we are done. Otherwise, suppose that one of u,v, say u is greater than 2^{n-2}. Observe that if $X\in A', Y\in B''$ and $f(X)=f(Y)$ then X,Y is a tentacle. But $|A'|+|B''|=u+2^{n-1}-v>2^{n-1}$ as $u>v$ from the conditions $u>2^{n-2}, u+v\le 2^{n-1}$. As f can take at most 2^{n-1} values, it is clear that we form at least $|A'|+|B''|-2^{n-1}$ tentacles composed of members of A' and B''. Thus we have a total of at least $2^{n-1}-u+v+u+2^{n-1}-v-2^{n-1}=2^{n-1}\ge|A|$ tentacles, so this case is also solved.

Problem 1.6. Let S be a set of points in the plane (not necessarily finite) and consider a complete graph G which has the points in S as its nodes. We color the edges of this graph with two colors. It is known (for example from Ramsey Theory) that if S has the same cardinality as $\mathbb{N}$, then we will be able to find a monochromatic complete subgraph whose vertex set also has the cardinality of $\mathbb{N}$. Is it true that if S has same cardinality as $\mathbb{R}$ we will find a monochromatic subgraph whose vertex set also has the cardinality of $\mathbb{R}$?

Solution. The answer is no. Let the two colors that we work with be red and blue. For simplicity, we identify the vertex set S with $\mathbb{R}$. Let $\le$ stand for the usual ordering on $\mathbb{R}$ and let $\preceq$ be a well-order on $\mathbb{R}$.

Now for two distinct vertices x and y in S, we color the edge between x and y with red if the two orders $\le$ and $\preceq$ agree on x,y (that is, if we choose the labels so that $x\le y$, then $x\preceq y$) and with blue otherwise. Assume now on the contrary that there is either a complete red subgraph whose vertex set A has the cardinality of $\mathbb{R}$ or a complete blue subgraph with this cardinality.

In the first case, we would have an uncountable set A on which the well-order $\preceq$ agrees with the usual order $\le$. To see that this cannot happen note that if we have any sequence a_i of elements of A, then the sets $B_i=\{x\in A: x\preceq a_i\}$ are countable. Hence the union $\bigcup B_i$ is countable. But it is possible to choose a sequence a_i of elements of A such that the every $x\in A$ has $x\le a_i$ for some i. (If A is not bounded above, we choose A_i to be any element of A with $A_i>i$. If A is bounded above, we let α be the supremum of the elements of A and choose $a_i > \alpha - 1/i$.) If the two orders agreed on A, this would make $\cup B_i = A$, hence uncountable.

In the second case the argument is the same, we just replace $\geq$ with $\leq$ throughout.

2 Sums, Products, and Identities

Problem 2.1. (**GMB 1997**) Determine the sequence $a_1, a_2, \ldots,$ of positive reals, such that for all positive integers k, one has

$$a_1 + 2^2 \cdot a_2 + 3^2 \cdot a_3 + \ldots + k^2 \cdot a_k = \frac{k(k+1)}{2}\left(a_1 + a_2 + \ldots + a_k\right).$$

Solution. When $k = 1$, the identity becomes $a_1 = a_1$. Let $a_1 = a > 0$.
When $k = 2$ we have $a_1 + 2^2 a_2 = 3(a_1 + a_2)$, hence $a_2 = 2a_1 = 2a$.
For $k = 3$ we get $a_1 + 2^2 a_2 + 3^2 a_3 = 6(a_1 + a_2 + a_3)$, therefore

$$9a + 9a_3 = 18a + 6a_3 \Rightarrow a_3 = 3a.$$

We prove by induction on k that $a_k = k \cdot a$, for all $k \geq 1$. We established the base cases above.
Assuming the result for some $k \geq 1$, using the hypothesis for $k+1$ numbers we have

$$\begin{aligned} & a + 2^3 a + 3^3 a + \ldots + k^3 a + (k+1)^2 a_{k+1} \\ &= \frac{(k+1)(k+2)}{2}\left(a + 2a + \ldots + ka + a_{k+1}\right), \end{aligned}$$

from which we get that

$$\begin{aligned} & \left((k+1)^2 - \frac{(k+1)(k+2)}{2}\right) a_{k+1} \\ &= \frac{(k+1)(k+2)}{2} \cdot \frac{k(k+1)}{2} \cdot a - a\left(1^3 + \ldots + k^3\right). \end{aligned}$$

Now using the fact that

$$1^3 + 2^3 + \ldots + k^3 = \left(\frac{k(k+1)}{2}\right)^2,$$

we obtain from the above relation that $a_{k+1} = (k+1)a$, as we wanted.

Problem 2.2. Prove that for all positive integers n,

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{2n-1} - \frac{1}{2n} = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n}.$$

Solution. We use mathematical induction to prove this statement. The base case $n = 1$ is true, $1 - \frac{1}{2} = \frac{1}{1+1}$. Assume that the statement is true for n:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{2n-1} - \frac{1}{2n} = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n}.$$

Let us prove it for $n+1$:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{2n-1} - \frac{1}{2n} + \frac{1}{2n+1} - \frac{1}{2n+2} = \frac{1}{n+2} + \frac{1}{n+2} + \ldots + \frac{1}{2n+2}.$$

Taking the difference it suffices to prove

$$\frac{1}{2n+1} - \frac{1}{2n+2} = \frac{1}{2n+1} + \frac{1}{2n+2} - \frac{1}{n+1},$$

which is clearly true.

Problem 2.3. Prove that for any positive integer n one has

$$(n+1)! = 1 + \frac{1!^2}{0!} + \frac{2!^2}{1!} + \ldots + \frac{n!^2}{(n-1)!}.$$

Solution. The identity clearly holds for $n = 1$. For the inductive step, assume that we have

$$(n+1)! = 1 + \frac{1!^2}{0!} + \frac{2!^2}{1!} + \ldots + \frac{n!^2}{(n-1)!}.$$

We will show that

$$(n+2)! = 1 + \frac{1!^2}{0!} + \frac{2!^2}{1!} + \ldots + \frac{n!^2}{(n-1)!} + \frac{(n+1)!^2}{n!}.$$

Using $P(n)$, this is equivalent to proving

$$(n+2)! = (n+1)! + \frac{(n+1)!^2}{n!},$$

which is equivalent to $n+2 = 1 + \dfrac{(n+1)!}{n!}$ or $n+2 = 1+n+1$, which is clear.

Problem 2.4. Prove that:

$$\sum_{k=0}^{n} \binom{n-k+1}{k} = F_{n+2},$$

for any non-negative integer n.

Solution. We prove the result by induction of step 2:

For $n = 0$, we have $\binom{1}{0} = F_2$, so $P(0)$ holds.

For $n = 1$, we need $\binom{2}{0} + \binom{1}{1} = F_3$, which is true.

Now let $n \geq 2$. Using Pascal's identity and applying induction for $n-1$ and $n-2$, we have:

$$\begin{aligned}
\sum_{k=0}^{n} \binom{n-k+1}{k} &= \sum_{k=0}^{n} \left[\binom{(n-1)-k+1}{k} + \binom{(n-1)-k+1}{k-1} \right] \\
&= \sum_{k=0}^{n-1} \binom{(n-1)-k+1}{k} + \sum_{k=0}^{n-2} \binom{(n-2)-k+1}{k} \\
&= F_{n+1} + F_n \\
&= F_{n+2}
\end{aligned}$$

Problem 2.5. Show that for any positive integer n, we have

$$\sum_{r=1}^{n} \frac{1}{r} \binom{n}{r} = \sum_{r=1}^{n} \frac{2^r - 1}{r}.$$

Solution. We prove the identity by induction on n. For $n = 1$, both sides become 1, so the result holds.

Assume now that the statement holds for some $n \geq 1$, that is

$$\sum_{r=1}^{n} \frac{1}{r} \binom{n}{r} = \sum_{r=1}^{n} \frac{2^r - 1}{r}. \tag{1}$$

To prove it for $n+1$, we have to show that

$$\sum_{r=1}^{n+1} \frac{1}{r} \binom{n+1}{r} = \sum_{r=1}^{n+1} \frac{2^r - 1}{r}.$$

So using (1), it suffices to show that

$$\sum_{r=1}^{n} \frac{1}{r}\left(\binom{n+1}{r} - \binom{n}{r}\right) + \frac{1}{n+1}\binom{n+1}{n+1} = \frac{2^{n+1}-1}{n+1},$$

or equivalently, using Pascal's identity, that

$$\sum_{r=1}^{n} \frac{1}{r}\binom{n}{r-1} + \frac{1}{n+1} = \frac{2^{n+1}-1}{n+1}.$$

But now

$$\frac{1}{r}\binom{n}{r-1} = \frac{1}{n+1}\binom{n+1}{r},$$

so

$$\begin{aligned}
\sum_{r=1}^{n} \frac{1}{r}\binom{n}{r-1} + \frac{1}{n+1} &= \frac{1}{n+1}\left(\sum_{r=1}^{n}\binom{n+1}{r} + 1\right) \\
&= \frac{1}{n+1}\left(\sum_{r=0}^{n+1}\binom{n+1}{r} - 2 + 1\right) \\
&= \frac{1}{n+1}\left((1+1)^{n+1} - 2 + 1\right) \\
&= \frac{2^{n+1}-1}{n+1},
\end{aligned}$$

as required.

Problem 2.6. The Bernoulli numbers $(B_n)_{n\geq 0}$ are given by the following recurrence:

$$B_0 = 1 \quad \text{and} \quad \sum_{i=0}^{m}\binom{m+1}{i}B_i = 0, \quad \text{for} \quad m > 0.$$

Prove that

$$1^k + 2^k + \ldots + (n-1)^k = \frac{1}{k+1}\sum_{i=0}^{k}\binom{k+1}{i}B_i \cdot n^{k+1-i},$$

for all non-negative integers n and k.

Solution. Set

$$S(n)=\sum_{i=0}^{k}\binom{k+1}{i}B_i n^{k+1-i}.$$

If we can prove that $S(1)=0$ and $S(n+1)-S(n)=(k+1)n^k$, then the conclusion of the problem follows by induction. Now

$$\begin{aligned}S(n+1)-S(n)&=\sum_{i=0}^{k}\binom{k+1}{i}B_i\left(\sum_{j=0}^{k-i}\binom{k+1-i}{j}n^j\right)\\&=\sum_{j=0}^{n+1}\left(\sum_{i=0}^{k-j}\binom{k+1-i}{j}\binom{k+1}{i}B_i n^j\right).\end{aligned}$$

But

$$\binom{k+1-i}{j}\binom{k+1}{i}=\binom{k+1}{j}\binom{k+1-j}{i}.$$

Thus we can write the expression as

$$\sum_{j=0}^{n+1}\binom{k+1}{j}\left(\sum_{i=0}^{k-j}\binom{k+1-j}{i}B_i\right)n^j.$$

But from the definition we deduce

$$\sum_{i=0}^{k-j}\binom{k+1-j}{i}B_i=0\quad\text{for}\quad j<k.$$

Therefore all powers of n of exponent less than k vanish and for k the coefficient is clearly $k+1$. So indeed $S(n+1)-S(n)=(k+1)n^k$, as desired.

Problem 2.7. Let n be a positive integer. Prove that

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}i^k=0,$$

for all $0\leq k\leq n-1$.

Solution. We prove the result by induction on n. The base case $n = 1$ is clear.

Assume that we have proved the identity for some $n \geq 1$. Notice that from the induction hypothesis we obtain that for any polynomial $p(X)$ of degree less than n we have

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}p(i) = 0.$$

In particular, for $p(i) = (i+1)^k, k < n$ we have

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}(i+1)^k = 0,$$

or

$$\sum_{i=1}^{n+1}(-1)^i\binom{n}{i-1}i^k = 0.$$

Adding this relation to $\sum_{i=0}^{n}(-1)^i\binom{n}{i}i^k = 0$ and using Pascal's Identity

$$\binom{n+1}{i} = \binom{n}{i} + \binom{n}{i-1},$$

we deduce that

$$\sum_{i=0}^{n+1}(-1)^i\binom{n+1}{k}i^k = 0.$$

Thus the condition holds for $k < n$ and we only need to show it for $k = n$. Now notice that

$$\sum_{i=0}^{n+1}(-1)^i\binom{n+1}{k}p(i) = 0,$$

for any polynomial $p(X)$ of degree less than n. Since there is a polynomial q of degree $n-1$ such that $x(x-1)\ldots(x-n+1) = x^n - q(x)$, we deduce that

$$\sum_{i=0}^{n+1}(-1)^i\binom{n+1}{k}i^k = \sum_{i=0}^{n+1}(-1)^i\binom{n+1}{k}i(i-1)\ldots(i-n+1).$$

But $i(i-1)\dots(i-n+1)$ is zero for all $i = 0, 1, \dots, n-1$ and it is $n!$, $(n+1)!$ for $i = n$, $i = n+1$, respectively. So

$$S = \sum_{i=0}^{n+1} (-1)^i \binom{n+1}{k} i(i-1)\dots(i-n+1)$$

becomes

$$S = (-1)^n \binom{n+1}{n} n! + (-1)^{n+1} \binom{n+1}{n+1} (n+1)! = 0.$$

3 Functions and Functional Equations

Problem 3.1. Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ which satisfy simultaneously the following properties:

1) $f(m) > f(n)$ for all $m > n$;
2) $f(f(n)) = 4n + 9$;
3) $f(f(n) - n) = 2n + 9$.

Solution. A natural guess suggested by 2) is that $f(n) = 2n + 3$. We will prove that this is the case in two steps. Firstly, we prove using strong induction that

$$f(n) = f(1) + 2(n-1), \quad \text{for all} \quad n \in \mathbb{N}^*.$$

Notice that the above identity clearly holds for $n = 1$ and for $n = 2$, we have that $f(f(2)-2) = f(f(1)) = 13$. Condition 1) implies that f is injective, so $f(2) = f(1) + 2$. For $n = 3$ we get that $f(f(3) - 3) = 15$, $f(f(2)) = 17$, $f(f(2) - 2) = 13$, so we have

$$f(2) > f(3) - 3 > f(2) - 2 \implies f(3) = f(2) + 2 = f(1) + 4.$$

Assume now that $f(k) = f(1) + 2(k-1)$, for all $k < n$ and $n \geq 4$.

We distinguish two cases:

Case 1. If n is odd, then from the induction hypothesis we have

$$f\left(f\left(\frac{n+1}{2}\right)\right) = 2n + 11 \text{ and } f\left(f\left(\frac{n-1}{2}\right)\right) = 2n + 7.$$

This shows that

$$f\left(\frac{n+1}{2}\right) > f(n) - n > f\left(\frac{n-1}{2}\right),$$

so

$$f(1) + n - 1 > f(n) - n > f(1) + n - 3.$$

Hence $f(n) = f(1) + 2n - 2 = f(1) + 2(n-1)$.

Case 2. If n is even, by the same idea as above we have

$$f\left(f\left(\frac{n}{2}\right)\right) = 2n + 9 = f(f(n) - n)) \implies f(n) - n = f\left(\frac{n}{2}\right) = f(1) + n - 2,$$

hence $f(n) = f(1) + 2(n-1)$, as we wanted.

This completes the proof of the fact that $f(n) = f(1) + 2(n-1)$. To find the value of $f(1)$, notice that for $n = 2$ we get

$$f(f(1)+2) = 17 \implies f(1) + 2(f(1)+1) = 17 \implies f(1) = 5.$$

Therefore, $f(n) = 2n + 3$ is indeed the only solution.

Problem 3.2. (**IMO 2007 shortlist**) Consider those functions $f : \mathbb{N}^* \mapsto \mathbb{N}^*$ which satisfy the condition

$$f(m+n) \geq f(m) + f(f(n)) - 1,$$

for all $m, n \in \mathbb{N}$. Find all possible values of $f(2007)$.

Solution. Let $P(m,n)$ be the assertion that

$$f(m+n) \geq f(m) + f(f(n)) - 1$$

We claim that $f(2007) \in \{1, 2, \ldots, 2008\}$.

For all $k \in \{1, 2, \ldots, 2007\}$ a function such that $f(2007) = k$ can be given by $f(n) = 1$ for $1 \leq n \leq 2006$, $f(2007) = k$ and $f(n) = n$ for $2008 \leq n$. A function such that $f(2007) = 2008$ can be given by $f(n) = n$ for $2007 \nmid n$ and $f(n) = n + 1$ for $2007 \mid n$.

Now we will prove that $f(n) \leq n + 1$. Assume for contradiction that there exists k such that $f(k) = k + c$ where $c > 1$. By $P(n, m-n)$ where $m > n$, it follows that $f(m) \geq f(n) + f(f(m-n)) - 1 \geq f(n)$. Hence f is non-decreasing.

Now we will prove by induction that $f(ik) \geq ik + i(c-1) + 1$ for all $i \in \mathbb{N}$. Note that we have shown the case where $i = 1$. Now assume that the claim holds for $i = 1, 2, \ldots, n$. By $P(nk, k)$, it follows that

$$f((n+1)k) \geq f(nk) + f(f(k)) - 1 = f(nk) + f(k+c) - 1,$$

and since f is non-decreasing we further have that

$$f(nk) + f(k+c) - 1 \geq f(nk) + k + c - 1 = (n+1)k + (n+1)(c-1) + 1.$$

This implies that for any m, there exists $p \in \mathbb{N}$ such that $f(p) - p \geq m$. Now let q be such that $f(q) - q \geq k$. By $P(f(q) - q, q)$, it follows that

$f(f(q)) \geq f(f(q)-q)+f(f(q))-1$ which implies that $f(f(q)-q) \leq 1$. However, since f is non-decreasing

$$f(f(q)-q) \geq f(k)=k+c>c>1,$$

which is a contradiction.
Hence $f(n) \leq n+1$, which implies that $f(2007) \in \{1,2,\ldots,2008\}$.

Problem 3.3. Let $f:\mathbb{N} \to \mathbb{N}$ be a function such that $f(0)=1$ and

$$f(n)=f\left(\left[\frac{n}{2}\right]\right)+f\left(\left[\frac{n}{3}\right]\right), \text{ for all } n \geq 1.$$

Prove that

$$f(n-1)<f(n) \Leftrightarrow n=2^k3^h, \quad \text{for some} \quad k,h \in \mathbb{N}.$$

Solution. We will prove the result by induction on n in two steps. First, we will show by induction on n that f is a non-decreasing function, that is, $f(n) \geq f(n-1)$. Since we easily compute $f(1)=2f(0)=2>f(0)$, the base case is done. For the inductive step we simply note that for $n \geq 2$, we have $[\frac{n}{2}],[\frac{n}{3}]<n$ and hence

$$f(n)=f\left(\left[\frac{n}{2}\right]\right)+f\left(\left[\frac{n}{3}\right]\right) \geq f\left(\left[\frac{n-1}{2}\right]\right)+f\left(\left[\frac{n-1}{3}\right]\right)=f(n-1).$$

Now we prove the requested result by induction on n. Since $f(1)=2>f(0)$ and $1=2^03^0$, the base case is done. For the inductive step, note that since

$$f(n)=f\left(\left[\frac{n}{2}\right]\right)+f\left(\left[\frac{n}{3}\right]\right)$$

and

$$f(n-1)=f\left(\left[\frac{n-1}{2}\right]\right)+f\left(\left[\frac{n-1}{3}\right]\right)$$

and f is non-decreasing, we have $f(n)>f(n-1)$ if and only if either

$$f\left(\left[\frac{n}{2}\right]\right)>f\left(\left[\frac{n-1}{2}\right]\right) \quad \text{or} \quad f\left(\left[\frac{n}{3}\right]\right)>f\left(\left[\frac{n-1}{3}\right]\right).$$

If n is odd, then $[\frac{n}{2}] = [\frac{n-1}{2}]$ and clearly

$$f\left(\left[\frac{n}{2}\right]\right) = f\left(\left[\frac{n-1}{2}\right]\right).$$

If $n = 2m$ is even, then by the induction hypothesis

$$f(m) = f\left(\left[\frac{n}{2}\right]\right) > f\left(\left[\frac{n-1}{2}\right]\right) = f(m-1)$$

if and only if $m = 2^k 3^h$ for some $k, h \in \mathbb{N}$. Thus

$$f\left(\left[\frac{n}{2}\right]\right) > f\left(\left[\frac{n-1}{2}\right]\right)$$

if and only if $n = 2^{k+1} 3^h$ for some $k, h \in \mathbb{N}$.
Similarly, if n is not a multiple of 3, then $[\frac{n}{3}] = [\frac{n-1}{3}]$ and clearly

$$f\left(\left[\frac{n}{3}\right]\right) = f\left(\left[\frac{n-1}{3}\right]\right).$$

If $n = 3m$ is a multiple of 3, then by the induction hypothesis

$$f(m) = f\left(\left[\frac{n}{3}\right]\right) > f\left(\left[\frac{n-1}{3}\right]\right) = f(m-1)$$

if and only if $m = 2^k 3^h$ for some $k, h \in \mathbb{N}$. Thus

$$f\left(\left[\frac{n}{3}\right]\right) > f\left(\left[\frac{n-1}{3}\right]\right)$$

if and only if $n = 2^k 3^{h+1}$ for some $k, h \in \mathbb{N}$.
Combining these two calculation, we see that $f(n) > f(n-1)$ if and only if $n = 2^k 3^h$ for some $k, h \in \mathbb{N}$, completing the inductive step.

Problem 3.4. (**AMM 10728**) Find all functions $f : \mathbb{Z} \to \mathbb{R}$ satisfying

$$f(a^3 + b^3 + c^3) = f(a)^3 + f(b)^3 + f(c)^3,$$

whenever $a, b, c \in \mathbb{Z}$.

Solution. Set $f(x) = ax$ to get $a = 0, 1, -1$ and set $f(x) = c$ to get $c = 0, \pm\frac{1}{\sqrt{3}}$. We shall prove that these are the only solutions.
First, we shall find a way of computing $f(n)$ by induction on $|n|$. Since

$$f(0) = f(n^3 + (-n)^3 + 0^3) = f(n)^3 + f(-n)^3 + f(0)^3,$$

we compute $f(-n)$ from $f(n)$. Further, note that if

$$a^3 + b^3 + c^3 = m^3 + n^3 + p^3,$$

then

$$f(a)^3 + f(b)^3 + f(c)^3 = f(m)^3 + f(n)^3 + f(p)^3.$$

So if we can write n^3 as a sum of at most five cubes of numbers having absolute value less than n, we are done. Let us find such representations. Set $n = 2^m(2a+1)$. Note that

$$\begin{aligned} 5^3 &= 4^3 + 4^3 - 1^3 - 1^3, \\ 6^3 &= 3^3 + 4^3 + 5^3, \\ 7^3 &= 6^3 + 4^3 + 4^3 - 1^3. \end{aligned}$$

If $a = 1$ and $m \geq 1$, $a = 2$, or $a = 3$, then n is a multiple of $6, 5$ or 7 and we are done. For $a \geq 4$ we get the result by the identity

$$(2a+1)^3 = (2a-1)^3 + (a+4)^3 - (a-4)^3 - 5^3 - 1^3.$$

Finally,

$$2^{3m} = 2^{3(m-2)}(3^3 + 3^3 + 2^3 + 1^3 + 1^3)$$

and therefore this also holds if $a = 0$ and $m \geq 2$. So f is uniquely determined by $f(0)$ and $f(1)$, since then we compute

$$f(2) = f(1^3 + 1^3) \quad \text{and} \quad f(3) = f(1^3 + 1^3 + 1^3).$$

Now let us look at $f(0)$ and $f(1)$. By setting $a = b = c = 0$, we get

$$f(0) = 3f(0)^3.$$

Thus $f(0) = 0$ or $f(0) = \pm\frac{1}{\sqrt{3}}$. If $f(0) = 0$, then setting $a = 1$, $b = c = 0$ we get $f(1) = f(1)^3$. Hence $f(1) = 0, -1, 1$ and we obtain the solutions $f(x) = 0$, $f(x) = -x$ and $f(x) = x$.

Now assume that $f(0) \neq 0$. Without loss of generality, let $f(0) = \frac{1}{\sqrt{3}}$ (the second case is analogous since we could look at $-f$ instead of f). Then by setting $a = 1$, $b = c = 0$ we deduce $f(1) = f(1)^3 + 2f(0)^3$, which is a polynomial equation with respect to $f(1)$ with solutions $\frac{1}{\sqrt{3}}, \frac{-2}{\sqrt{3}}$. Similarly, by setting $a = -1$, $b = c = 0$, we deduce that $f(-1)$ is either $\frac{1}{\sqrt{3}}$ or $\frac{-2}{\sqrt{3}}$. But setting $a = 0$, $b = 1$ and $c = -1$ we find

$$\frac{1}{\sqrt{3}} = f(0) = f(0)^3 + f(1)^3 + f(-1)^3 = \frac{1}{3\sqrt{3}} + f(1)^3 + f(-1)^3.$$

A little checking shows that the only possibility is $f(1) = f(-1) = \frac{1}{\sqrt{3}}$, which makes $f(x) = \frac{1}{\sqrt{3}}$ for all x.

Problem 3.5. (**APMO 2008**) Consider the function $f : \mathbb{N} \to \mathbb{N}$ defined by the following conditions :

i) $f(0) = 0$;
ii) $f(2n) = 2f(n)$, for all $n \in \mathbb{N}$;
iii) $f(2n+1) = n + 2f(n)$, for all $n \in \mathbb{N}$.

a) Determine the three sets

$$L = \{n | f(n) < f(n+1)\},\ E = \{n | f(n) = f(n+1)\}, \text{ and}$$
$$G = \{n | f(n) > f(n+1)\};$$

b) For each $k \geq 0$, find a formula for $a_k = \max\{f(n) : 0 \leq n \leq 2^k\}$ in terms of k.

Solution. Let

$$L_1 := \{2k : k > 0\},\ E_1 := \{0\} \cup \{4k+1 : k \geq 0\},\ G_1 := \{4k+3 : k \geq 0\}.$$

We will show that $L_1 = L$, $E_1 = E$ and $G_1 = G$. It suffices to verify that $L_1 \subseteq L$, $E_1 \subseteq E$ and $G_1 \subseteq G$, because L_1, E_1, G_1 are mutually disjoint and $L_1 \cup E_1 \cup G_1 = \mathbb{N}$.

Firstly, if $k > 0$, then $f(2k) - f(2k+1) = -k < 0$ and therefore $L_1 \subseteq L$. Secondly, $f(1) = 0$ and

$$f(4k+1) = 2k + 2f(2k) = 2k + 4f(k)$$
$$f(4k+2) = 2f(2k+1) = 2(k + 2f(k)) = 2k + 4f(k),$$

for all $k \geq 0$. Thus $E_1 \subseteq E$.
In order to prove $G_1 \subseteq G$, we prove by induction on n that $f(n+1) - f(n) \leq n$, for all $n \in \mathbb{N}$. Notice that this is clearly true when $n = 2t$, as we have by definition of f that

$$f(2t+1) - f(2t) = t \leq n.$$

In particular, this proves the base case $n = 0$ and it suffices to consider the case when n is odd.

Assume now that the result holds for all integers up to some n, $n = 2t+1$. Then by the induction hypothesis we have

$$\begin{aligned} f(n+1) - f(n) &= f(2t+2) - f(2t+1) = 2f(t+1) - t - 2f(t) \\ &= 2(f(t+1) - f(t)) - t \leq 2t - t = t < n. \end{aligned}$$

This completes our induction. Now for all $k \geq 0$, we have

$$\begin{aligned} f(4k+4) - f(4k+3) &= f(2(2k+2)) - f(2(2k+1)+1) \\ &= 4f(k+1) - (2k+1+2f(2k+1)) \\ &= 4f(k+1) - (2k+1+2k+4f(k)) \\ &= 4(f(k+1) - f(k)) - (4k+1) \leq 4k - (4k+1) < 0. \end{aligned}$$

This proves $G_1 \subseteq G$ and completes the proof of part a).

b) First note that $a_0 = a_1 = f(1) = 0$.

Let $k \geq 2$ and let $N_k = \{0, 1, 2, \ldots, 2^k\}$. We first prove by induction that the maximum a_k occurs at the largest number in $G \cap N_k$, i.e. $a_k = f(2^k - 1)$.

The base case is $k = 2$ for which we have $a_2 = f(3) = f(2^2 - 1) = 1$.

Assume now that the result holds for all integers between 2 and $k-1$, for some $k \geq 3$. Then for every even number $2t$ with $2^{k-1} + 1 < 2t \leq 2^k$, by the induction hypothesis we have

$$f(2t) = 2f(t) \leq 2a_{k-1} = 2f(2^{k-1} - 1).$$

For every odd number $2t+1$ with $2^{k-1}+1 < 2t+1 < 2^k$ we have by induction hypothesis that

$$\begin{aligned} f(2t+1) &= t + 2f(t) \le 2^{k-1} - 1 + 2f(t) \\ &= 2^{k-1} - 1 + 2a_{k-1} = 2^{k-1} - 1 + 2f(2^{k-1} - 1). \end{aligned}$$

In either of the two cases above, the upper bound is at most

$$f(2^k - 1) = f(2(2^{k-1} - 1) + 1) = 2^{k-1} - 1 + 2f(2^{k-1} - 1),$$

so we have indeed that $a_k = f(2^k - 1)$, completing our induction.

From the above, we obtain the recursion

$$a_k = 2a_{k-1} + 2^{k-1} - 1, \forall k \ge 3.$$

We observe that the above formula also holds when $k = 0, 1, 2$. Then a simple induction yields that for all $k > 0$ we have $a_k = k2^{k-1} - 2^k + 1$. The inductive step reads

$$\begin{aligned} a_k &= 2a_{k-1} + 2^{k-1} - 1 \\ &= 2((k-1)2^{k-2} - 2^{k-1} + 1) + 2^{k-1} - 1 \\ &= k2^{k-1} - 2^k + 1. \end{aligned}$$

Problem 3.6. (**India 2000**) Suppose $f : \mathbb{Q} \to \{0, 1\}$ is a function with the property that for $x, y \in \mathbb{Q}$, if $f(x) = f(y)$, then $f(x) = f((x+y)/2) = f(y)$. If $f(0) = 0$ and $f(1) = 1$, show that $f(q) = 1$ for all rational number q greater than or equal to 1.

Solution. We first prove the following:

Lemma. Suppose that a and b are rational numbers. If $f(a) \ne f(b)$, then $f(n(b-a)+a) = f(b)$ for all positive integers n.

Proof. We prove the claim by strong induction on n. For $n = 1$, the claim is clear. Now assume that the claim is true for $n \le k$. Let

$$(x_1, y_1, x_2, y_2) = (b, k(b-a)+a, a, (k+1)(b-a)+a).$$

By the induction hypothesis, $f(x_1) = f(y_1)$. We claim that $f(x_2) \neq f(y_2)$. Otherwise, setting $(x, y) = (x_1, y_1)$ and $(x, y) = (x_2, y_2)$ in the given condition, we would have $f(b) = f((x_1 + y_1)/2)$ and $f(a) = f((x_2 + y_2)/2)$. However, this is impossible because $x_1 + y_1 = x_2 + y_2$. Therefore, $f(y_2)$ must equal the value in $\{0, 1\} \setminus \{f(a)\}$, namely $f(b)$. This completes the induction.

Applying the lemma with $a = 0$ and $b = 1$, we see that $f(n) = 1$ for all positive integers n. Thus, $f(1 + r/s) \neq 0$ for all natural numbers r and s, because otherwise applying the lemma with $a = 1$, $b = 1 + r/s$ and $n = s$ yields $f(1 + r) = 0$, which is a contradiction. Therefore, $f(q) = 1$ for all rational numbers $q \geq 1$.

Problem 3.7. Find all functions $f : \mathbb{Q}^+ \to \mathbb{Q}^+$, such that

$$f(x) + f\left(\frac{1}{x}\right) = 1, \quad \text{and} \quad f(1 + 2x) = \frac{1}{2} f(x),$$

for any $x \in \mathbb{Q}^+$.

Solution. From the first equation we deduce that $f(1) = \dfrac{1}{2}$.
Let us now prove that for m, n positive integers we have

$$f\left(\frac{m}{n}\right) = \frac{1}{1 + \frac{m}{n}}.$$

We shall prove the result by induction on $m + n$.
By the above, the statement holds when $m + n = 2$.
Let us show that if the statement holds for $m + n \leq k$, $k \in \mathbb{N}$, $k \geq 2$, then it also holds for $m + n = k + 1$.

We consider the following cases:

Case 1. If m and n are odd numbers.

1.1) If $m > n$, then

$$f\left(\frac{m}{n}\right) = f\left(1 + 2 \cdot \frac{\frac{m}{n} - 1}{2}\right) = \frac{1}{2} f\left(\frac{\frac{m}{n} - 1}{2}\right) = \frac{1}{2} f\left(\frac{\frac{m-n}{2}}{n}\right)$$
$$= \frac{1}{2} \cdot \frac{1}{1 + \frac{\frac{m-n}{2}}{n}} = \frac{1}{1 + \frac{m}{n}},$$

as $\frac{m-n}{2}+n=\frac{k+1}{2}\leq k$.

1.2) If $m=n$, we have that

$$f\left(\frac{m}{n}\right)=f(1)=\frac{1}{2}=\frac{1}{1+\frac{m}{n}}.$$

1.3) If $m<n$, then

$$f\left(\frac{m}{n}\right)=1-f\left(\frac{n}{m}\right)=1-\frac{1}{1+\frac{n}{m}}=\frac{1}{1+\frac{m}{n}}.$$

Case 2. If m and n are even numbers, then

$$f\left(\frac{m}{n}\right)=f\left(\frac{\frac{m}{2}}{\frac{n}{2}}\right)=\frac{1}{1+\frac{\frac{m}{2}}{\frac{n}{2}}}=\frac{1}{1+\frac{m}{n}},$$

as $\frac{m}{2}+\frac{n}{2}=\frac{k+1}{2}\leq k$.

Case 3. If m is even and n is odd, then let us consider the following pairs of positive integers: (m_0,n_0), (m_1,n_1),…,(m_p,n_p),…, where $m_0=m$, $n_0=n$ and

$$(m_i,n_i)=\begin{cases}\left(\dfrac{m_{i-1}}{2}+n_{i-1},\dfrac{m_{i-1}}{2}\right) & \text{if } 4\nmid m_{i-1}\\[2ex] \left(\dfrac{m_{i-1}}{2},\dfrac{m_{i-1}}{2}+n_{i-1}\right) & \text{if } 4\mid m_{i-1},\end{cases}\qquad \text{where}\quad i=1,2,\ldots$$

Note that m_i is an even number, n_i is an odd number and $m_i+n_i=m_{i-1}+n_{i-1}$, where $i=1,2,\ldots$

Since there are only finitely many pairs (m,n) with fixed sum $m+n$, the sequence (m_0,n_0), (m_1,n_1),… must eventually cycle. In general, the cycle need not begin with the first term sequence, but in this case we can step the recursion backwards. Indeed, if $m_i>n_i$, then $m_{i-1}=2n_i$, $n_{i-1}=m_i-n_i$ and if $m_i<n_i$, then $m_{i-1}=2m_i$, $n_{i-1}=n_i-m_i$. Therefore, the sequence (m_0,n_0), (m_1,n_1),… is periodic starting from its first element.

Let p be such that the pairs (m_p, n_p) and (m_0, n_0) coincide. We have that

$$f\left(\frac{m_0}{n_0}\right) = 1 - f\left(\frac{n_0}{m_0}\right) = 1 - 2f\left(\frac{n_0}{\frac{m_0}{2}} + 1\right) = 1 - 2f\left(\frac{n_1}{m_1}\right),$$

or

$$f\left(\frac{m_0}{n_0}\right) = 1 - 2f\left(\frac{m_1}{n_1}\right).$$

On the other hand,

$$1 - 2f\left(\frac{n_1}{m_1}\right) = -1 + 2f\left(\frac{m_1}{n_1}\right).$$

Thus,

$$f\left(\frac{m_{i-1}}{n_{i-1}}\right) = 1 - 2f\left(\frac{m_i}{n_i}\right),$$

or

$$f\left(\frac{m_{i-1}}{n_{i-1}}\right) = -1 + 2f\left(\frac{m_i}{n_i}\right),$$

where $i = 1, 2, \ldots, p+1$. Let

$$x_i = f\left(\frac{m_{i-1}}{n_{i-1}}\right), \quad \text{where} \quad i = 1, 2, \ldots, p+1.$$

Hence, we obtain that

$$\begin{cases} x_1 = \varepsilon_1 - 2\varepsilon_1 x_2 \\ x_2 = \varepsilon_2 - 2\varepsilon_2 x_3, \\ \ldots \\ x_p = \varepsilon_p - 2\varepsilon_p x_{p+1}, \\ x_{p+1} = x_1, \end{cases}$$

where $|\varepsilon_i| = 1$, $i = 1, 2, \ldots, p$.

Solving the last system by plugging in the corresponding values of the variables, we deduce that $x_1 = a + bx_1$, where b is an even number (in fact $b = \pm 2^p$). The last equation has a unique solution, hence the system has a unique solution.

Note that

$$x_i = \frac{1}{1+\frac{m_{i-1}}{n_{i-1}}},$$

where $i = 1, 2, \ldots, p+1$ is a solution of the given system of equations. Therefore,

$$f\left(\frac{m}{n}\right) = f\left(\frac{m_0}{n_0}\right) = x_1 = \frac{1}{1+\frac{m_0}{n_0}} = \frac{1}{1+\frac{m}{n}}.$$

Case 4. If m is odd and n is even, then

$$f\left(\frac{m}{n}\right) = 1 - f\left(\frac{n}{m}\right) = 1 - \frac{1}{1+\frac{n}{m}} = \frac{1}{1+\frac{m}{n}}.$$

Hence,

$$f\left(\frac{m}{n}\right) = \frac{1}{1+\frac{m}{n}}.$$

Thus, the function $f(x) = \dfrac{1}{1+x}$ is the only function satisfying the hypotheses of the problem.

Problem 3.8. Find all functions $f : [0, +\infty) \to [0, 1]$, such that for any $x \geq 0$, $y \geq 0$

$$f(x)f(y) = \frac{1}{2}f(yf(x)).$$

Solution. We will establish some properties of f which will allow us to solve the problem. We begin by showing the following:

P1. For any $x \geq 0$ we have $f(x) \leq \dfrac{1}{2}$.

Proof. We prove this by contradiction. Assume that $f(x_1) > \dfrac{1}{2} + a$, for some $x_1 > 0$, where $a > 0$. Under this assumption, we prove that there exists a sequence (x_n), such that for $n = 1, 2, \ldots$

$$f(x_n) > \frac{1}{2} + na. \tag{1}$$

We do this by induction on n. For $n = 1$, the result holds by our assumption.

We now show that if the statement holds for $n = k$, $k \in \mathbb{N}$, then it also holds for $n = k + 1$.

We have that $f(x_k) > \dfrac{1}{2} + ka$. Let $x_{k+1} = x_k f(x_k)$. Then

$$f(x_{k+1}) = f(x_k f(x_k)) = 2(f(x_k))^2 > 2\Big(\frac{1}{2} + ka\Big)^2 \geq \frac{1}{2} + (k+1)a.$$

This completes the induction step.

From (1) we obtain that the sequence $\left(\dfrac{1}{2} + na\right)$ is bounded, which is a contradiction. Hence $f(x) \leq \frac{1}{2}$, for all $x \geq 0$.

We now prove the following:

P2. For any $x \geq 0$ either $f(x) = 0$ or $f(x) = \dfrac{1}{2}$.

Proof. Assume that $0 < f(x_1) < \dfrac{1}{2}$, for some $x_1 \leq 0$. From hypothesis we have that

$$f(x_1) f\Big(\frac{x_1}{f(x_1)}\Big) = \frac{1}{2} f(x_1).$$

Hence $f(x_2) = \dfrac{1}{2}$, where $x_2 = \dfrac{x_1}{f(x_1)}$. Then

$$f(x_1) f\Big(\frac{x_2}{f(x_1)}\Big) = \frac{1}{2} f(x_2) = \frac{1}{4}.$$

Therefore, $f\Big(\dfrac{x_2}{f(x_1)}\Big) > \dfrac{1}{2}$. This contradicts what we established in **P1** above.

Let

$$f(x) = \frac{1}{2}, \quad \text{for} \quad x \in A \quad \text{and} \quad f(x) = 0, \quad \text{for } x \in [0, +\infty) \setminus A. \tag{2}$$

Note that if $x \in A$, then $f(x)f(x) = \dfrac{1}{2} f(xf(x))$. Hence $\dfrac{x}{2} \in A$. We also have

$$f(x) f\Big(\frac{x}{f(x)}\Big) = \frac{1}{2} f(x),$$

so $2x \in A$ as well. On the other hand, if $f(0) = \frac{1}{2}$, then $f(x) = \frac{1}{2}$ for any non-negative value of x (it is enough to take $y = 0$ in the hypothesis).

It is now an easy task to verify that if $A = \emptyset$, $A = [0, +\infty)$ or A is such that $0 \notin A$ and

$$x \in A \Rightarrow \frac{x}{2} \in A, 2x \in A,$$

then the functions in (2) satisfy the conditions of the problem.

Problem 3.9. (**China 2013**) Prove that there exists only one function $f : \mathbb{N}^* \to \mathbb{N}^*$ satisfying the following two conditions:

i) $f(1) = f(2) = 1$;

ii) $f(n) = f(f(n-1)) + f(n - f(n-1))$ for $n \geq 3$.

For each integer $m \geq 2$, find the value of $f(2^m)$.

Solution. For the first part of the question, we will prove by induction that for any $n \geq 2$ we have

$$f(n) - f(n-1) \in \{0, 1\} \quad \text{and} \quad \frac{n}{2} \leq f(n) \leq n.$$

Notice that this suffices to establish the first part of the question, since from $\frac{n}{2} \leq f(n) \leq n$ we know that $f(n-1) \leq n-1$, and from the given relation $f(n) = f(f(n-1)) + f(n - f(n-1))$ for $n \geq 3$, $f(n)$ is uniquely determined by the values of $f(f(n-1))$ and $f(n - f(n-1))$.

Our base case is $n = 2$, which is true, since we are given $f(2) = 1$. Assume now that the statement holds for all numbers between 2 and some integer $n \geq 2$. Let us prove it for $n + 1$. We have

$$f(n+1) - f(n) = f(f(n)) - f(f(n-1)) + f(n+1-f(n)) - f(n-f(n-1)).$$

From the induction hypothesis, we know that $f(n) \leq n$, $f(n-1) \leq n-1$ and $f(n) - f(n-1) \in \{0, 1\}$. We distinguish two cases:

Case 1. If $f(n) = f(n-1)$, then $f(f(n)) = f(f(n-1))$, so we have

$$\begin{aligned} f(n+1) - f(n) &= f(n+1-f(n)) - f(n-f(n-1)) \\ &= f(n+1-f(n)) - f(n-f(n)) \in \{0, 1\}. \end{aligned}$$

Case 2. If $f(n) = f(n-1)+1$, then $f(n+1-f(n)) = f(n-f(n-1))$, and we have

$$f(n+1)-f(n) = f(f(n))-f(f(n-1)) = f(f(n-1)+1)-f(f(n-1)) \in \{0,1\}.$$

So we obtained in both cases that $f(n+1)-f(n) \in \{0,1\}$. Since $f(n) \leq n$, we have that $f(n+1) \leq n+1$, from what we just proved. Furthermore,

$$f(n+1) = f(f(n)) + f(n+1-f(n)) \geq \frac{f(n)}{2} + \frac{n+1-f(n)}{2} = \frac{n+1}{2},$$

which completes the proof of our induction.

For the second part of the question, we will prove by induction that $f(2^m) = 2^{m-1}$, for all $m \geq 1$. The base case is verified, since we are given that $f(2) = 1$. Assume now that the result holds for all integers between 1 and some positive integer $m \geq 1$. We shall prove the result for $m+1$:

We prove by induction on k that if $2^m \leq k < 2^{m+1}$, then $f(k) \leq 2^m$. If $k = 2^m$, the statement is true from the induction hypothesis. Assuming it for k, if $k+1 < 2^{m+1}$, then

$$f(k) \leq 2^m \quad \text{and} \quad k+1-f(k) \leq k+1-\frac{k}{2} = \frac{k+2}{2} \leq \frac{2^{m+1}}{2} = 2^m.$$

From what we established in the first part of the question, we know that f is increasing, so $f(k) \leq 2^m$ and $k+1-f(k) \leq 2^m$ imply that $f(f(k)) \leq f(2^m)$ and $f(k+1-f(k)) \leq f(2^m)$, hence

$$f(k+1) = f(f(k)) + f(k+1-f(k)) \leq f(2^m) + f(2^m) = 2^{m-1} + 2^{m-1} = 2^m.$$

We now have $f(2^{m+1}-1) \leq 2^m$ and $f(2^{m+1}-1) \geq \frac{2^{m+1}-1}{2}$. So it must be that $f(2^{m+1}-1) = 2^m$. Finally,

$$\begin{aligned} f(2^{m+1}) &= f(f(2^{m+1}-1)) + f(2^{m+1} - f(2^{m+1}-1)) \\ &= f(2^m) + f(2^{m+1} - 2^m) \\ &= 2^{m-1} + 2^{m-1} \\ &= 2^m, \end{aligned}$$

completing our proof.

Problem 3.10. (**Silk Road MC**) Find all functions $f : \mathbb{N}^* \to \mathbb{N}^*$ which satisfy

$$2f(mn) \geq f(m^2+n^2) - f(m)^2 - f(n)^2 \geq 2f(m)f(n).$$

Solution. We will show that $f(n) = n^2$ for all n. For this, we will need two famous facts from number theory. The first is that if $q \equiv 3 \pmod 4$ is a prime, then q does not divide any number of the form a^2+1. The second is that if $p \equiv 1 \pmod 4$ is a prime, then there are positive integers a and b such that $p = a^2 + b^2$.

First note that comparing the outer expressions, we see that $f(mn) \geq f(m)f(n)$.

Taking $m = n = 1$, we get $2f(1) \geq f(2) - 2f(1)^2 \geq 2f(1)^2 \geq 2f(1)$. Hence we must have equality at each step and thus $f(1) = 1$ and $f(2) = 4$.

Taking $m = 1$, we get $2f(n) \geq f(n^2+1) - f(n)^2 - f(1)^2 \geq 2f(n)$. Hence again we must have equality throughout and $f(n^2+1) = (f(n)+1)^2$.

We will now show by a somewhat subtle induction on n that $f(n) \geq n^2$ for all n. By the inequality $f(mn) \geq f(m)f(n)$ it suffices to prove this for n a prime. Since we already saw that $f(2) = 4$, the base case is proven, and we may assume n is an odd prime. Define the complexity of a prime p to be p if $p \equiv 1 \pmod 4$ and p^2 if $p \equiv 3 \pmod 4$. We will prove the inequality by induction on the complexity.

Suppose we have proved the inequality for all primes whose complexity is less than the complexity of p (hence for all products of such primes).

If $p \equiv 1 \pmod 4$, then there are a, b with $p = a^2 + b^2$. Since $p > a^2, b^2$, every prime divisor of a or b has complexity less than that of p. Therefore by the inductive hypothesis $f(a) \geq a^2$ and $f(b) \geq b^2$. Hence taking $m = a$ and $n = b$ in the given inequality gives

$$f(p) \geq f(a)^2 + f(b)^2 + 2f(a)f(b) \geq a^4 + b^4 + 2a^2b^2 = (a^2+b^2)^2 = p^2.$$

If $p \equiv 3 \pmod 4$, then every prime divisor of p^2+1 is either 2 or is 1 modulo 4. Since there must be at least one factor of 2, these prime factors must be smaller than p^2. Hence by induction we conclude that

$$(f(p)+1)^2 = f(p^2+1) \geq (p^2+1)^2,$$

and hence $f(p) \geq p^2$.

Thus by induction we see that $f(n) \geq n^2$ for all n.

We saw above that $f(1) = 1$. Using the inequality, we have two rules which would let us conclude that $f(n) = n^2$ for some n.

Rule 1. Suppose $f(mn) = (mn)^2$. Then from

$$(mn)^2 = f(mn) \geq f(m)f(n) \geq m^2n^2,$$

we conclude that $f(m) = m^2$ and $f(n) = n^2$, and hence from

$$2(mn)^2 + m^4 + n^4 = 2f(mn) + f(m)^2 + f(n)^2 \geq f(m^2 + n^2) \geq (m^2 + n^2)^2$$

we further conclude that $f(m^2 + n^2) = (m^2 + n^2)^2$.

Rule 2. Suppose $f(m^2 + n^2) = (m^2 + n^2)^2$. Then from

$$(m^2 + n^2)^2 = f(m^2 + n^2) \geq (f(m) + f(n))^2 \geq (m^2 + n^2)^2$$

we conclude that both $f(m) = m^2$ and $f(n) = n^2$.

Let B be the set of all n for which we can prove that $f(n) = n^2$ starting with the result above that $f(1) = 1$ and applying some sequence of the two rules above. One can quickly prove that quite a few integers are in B using these rules. For example $1 \in B$ by definition. Since $2 = 1^2 + 1^2$, we conclude $2 \in B$ by Rule 1. Similarly $5 = 2^2 + 1^2$ and $26 = 5^2 + 1^2$, show by Rule 1 that $5, 26 \in B$. Then $13 = 26/2$, $170 = 13^2 + 1^2$, $10 = 170/17$, $17 = 170/10$, show by Rule 1 that $10, 13, 17 \in B$. Hence $10 = 3^2 + 1^2$ and $17 = 4^2 + 1^2$ show by Rule 2, that $3, 4 \in B$.

The trick is to organize what we know to show that B is in fact every positive integer. Here is one method that works (though it is somewhat computational at the start). Note that we already saw that using only Rule 1, we have $1, 2, 5, 26, 13$, and $170 \in B$. Continuing, in this way we get $34 = 170/5$, $1157 = 34^2 + 1^2$, $1338650 = 1157^2 + 1^2$, and $25 = 1338650/53546 \in B$. Thus we have shown $25 \in B$ using only Rule 1.

Now notice that Rule 1 has an extra feature. If we apply Rule 1 to k^2mn, then we can conclude that k^2m, k^2n and $k^2(m^2+n^2)$ are in B. Thus iterating the sequence of implications above, we conclude that $5^n \in B$ for all n.

Rule 2 is a little weaker than Rule 1 with respect to multiples. If we apply Rule 2 to $k^2(m^2+n^2)$, we can only conclude that km and kn are in B. However since arbitrarily high powers of 5 are in B, we can still conclude that if $n \in B$, then $5n \in B$ (Suppose a chain of implications that starts from 1 and shows that $n \in B$ has r steps that use Rule 2. Then starting from 5^{2^r} and applying these same steps with extra powers of 5, we will conclude that $5n \in B$.)

Now finally the conclusion is easy. From $n \in B$, we conclude that $n^2+1 \in B$ and hence by the long discussion above that $5(n^2+1) \in B$. Writing $5(n^2+1) = (n+2)^2 + (2n-1)^2$, we conclude using Rule 2 that $n+2 \in B$. Thus an easy induction starting with $1, 2 \in B$ shows that every positive integer is in B.

Problem 3.11. (**Turkey**) Find all functions $f : \mathbb{Q}^+ \to \mathbb{Q}^+$ such that

$$f\left(\frac{x}{x+1}\right) = \frac{f(x)}{x+1}, \quad f(x) = x^3 f\left(\frac{1}{x}\right), \quad \text{for all} \quad x \in \mathbb{Q}^+.$$

Solution. We prove that the only functions which satisfy the conditions of the problem are those of the form $f_a(\frac{n}{d}) = a \cdot \frac{n^2}{d}$, where d and n are positive integers with $(d, n) = 1$. One can easily check that these functions satisfy the functional equation. For the converse, we proceed by induction on $k = \max\{n, d\}$ (or alternatively, on $n + d$). Let $f(1) = a$.

Assume that whenever $k < m$ (for some $m \geq 2$), we have $f\left(\frac{n}{d}\right) = f_a\left(\frac{n}{d}\right)$ (we know the base case $k = 1$ holds by our construction). Consider the case $k = m$. Notice that since $\gcd(n, d) = 1$, and $\max(n, d) = m \geq 2$, we have $n \neq d$. We distinguish two cases:
If $n < d$, then $\gcd(n, d-n) = 1$ and $\max\{n, d-n\} < k$, so

$$f\left(\frac{n}{d}\right) = f\left(\frac{\frac{n}{d-n}}{\frac{n}{d-n}+1}\right) = \frac{f\left(\frac{n}{d-n}\right)}{\frac{d}{d-n}} = a \cdot \frac{n^2}{d}.$$

If $d < n$, then applying the previous case gives

$$f\left(\frac{n}{d}\right) = f\left(\frac{1}{\frac{n}{d}}\right) = \frac{f\left(\frac{d}{n}\right)}{\frac{d^3}{n^3}} = a \cdot \frac{n^2}{d}.$$

In either case, we obtain the result for the inductive step, completing our proof.

Problem 3.12. Find all functions $f : \mathbb{Z} \to \mathbb{Z}$ that satisfy

$$f(m+n) + f(mn-1) = f(m)f(n) + 2,$$

for all integers m, n.

Solution. We begin by calculating a few small values for f. Plugging in $m = n = 0$ gives

$$f(0) + f(-1) = f^2(0) + 2.$$

Setting $m = 1$, $n = 0$ gives

$$f(1) + f(-1) = f(1)f(0) + 2.$$

If we set $m = -1$, $n = 0$, we get

$$2f(-1) = f(-1)f(0) + 2.$$

Hence $f(-1)(2 - f(0)) = 2$. So we must have that $f(-1) \in \{\pm 1, \pm 2\}$. Combining this with $f(0) + f(-1) = f^2(0) + 2$, we obtain that the only integer value for $f(0)$ is obtained when $f(-1) = 2$ and in this case we get $f(0) = 1$.

Now, to find $f(1)$, we first plug in $m = n = -1$ which gives

$$f(-2) + f(0) = f^2(-1) + 2,$$

so $f(-2) = 5$. Then, for $m = 1$ and $n = -1$ we obtain

$$f(0) + f(-2) = f(1)f(-1) + 2,$$

which gives $f(1) = 2$. Further, $m = n = 1$ gives $f(2) + f(0) = f^2(1) + 2$, so $f(2) = 5$.

Notice that for the first small values we always had $f(m) = m^2 + 1$. Also notice that from the functional equation, for $n = -1$ we have

$$f(m-1) + f(-m-1) = f(-1)f(m) + 2,$$

so it would suffice to prove that $f(m) = m^2 + 1$ for $m \geq 0$, as it would then imply it for all integers.

We claim that for each $m \geq 1$, $f(m)$ has only one possible value and it must be m^2+1, and we shall prove this using strong induction.

For the base cases, $m=1$ and $m=2$, we have found that $f(1)=2$ and $f(2)=5$, so the base cases are true.

For the inductive step, we assume that for all $1 \leq m \leq k$, $f(m)$ has one possible value, m^2+1, and we want to show that $f(m+1)$ has only one possible value, $(m+1)^2+1$. First of all, since $f(m)$ and $f(m-1)$ are fixed, $f(m+1)$ is fixed. We have

$$\begin{aligned} f(m+1) &= 2f(m)+2-f(m-1) \\ &= 2(m^2+1)+2-(m-1)^2-1 \\ &= m^2+2m+2. \end{aligned}$$

We have obtained that $f(m+1)=(m+1)^2+1$, so the induction is complete.

Problem 3.13. (**Estonia 2000**) Find all functions $f:\mathbb{N}\to\mathbb{N}$ such that

$$f(f(f(n)))+f(f(n))+f(n)=3n \quad \text{for all} \quad n\in\mathbb{N}.$$

Solution. Observe that if $f(a)=f(b)$, then setting $n=a$ and $n=b$ into the given equation yields $3a=3b$, so $a=b$. Therefore, f is injective.

We now prove by induction on $n \geq 0$ that $f(n)=n$. The base case $n=0$ follows immediately by substituting 0 in the given relation and using the fact that $f(0)\geq 0$, as the image of f lies in $\mathbb{N}$.

For the induction step, assume that we have proved the result for all integers n less than some n_0, with $n_0 \geq 1$. We want to show that $f(n_0)=n_0$. Since f is injective, if $n \geq n_0 > k$, then $f(n) \neq f(k) = k$. Thus

$$f(n) \geq n_0, \quad \text{for all} \quad n \geq n_0. \tag{*}$$

In particular, (*) holds for $n=n_0$, i.e. $f(n_0)\geq n_0$. Then (*) holds for $n=f(n_0)$ and similarly for $f(f(n_0))$ as well. Substituting $n=n_0$ in the given equation, we find that

$$3n_0 = f(f(f(n_0)))+f(f(n_0))+f(n_0) \geq n_0+n_0+n_0.$$

Therefore, equality must occur, so $f(n_0)=n_0$, completing the induction step and the proof.

Problem 3.14. Prove that there exists a unique function $f : \mathbb{Q}^+ \to \mathbb{Q}^+$ satisfying all the following conditions:

a) If $0 < q < \frac{1}{2}$, then $f(q) = 1 + f\left(\frac{q}{1-2q}\right)$;

b) If $1 < q \le 2$, then $f(q) = 1 + f(q-1)$;

c) $f(q)f\left(\frac{1}{q}\right) = 1$ for all $q \in \mathbb{Q}^+$.

Solution. We show that there is a unique f satisfying these properties by showing that f is uniquely determined at every positive rational number $\frac{a}{b}$. We do this by strong induction on $a + b$.

We start with the base case, when $a + b = 2$ with $a = b = 1$ and $a/b = 1$. Note that $f(1)f(1) = 1$, so $f(1) = 1$ by c).

Now suppose $f(\frac{a}{b})$ is uniquely determined at all positive rational numbers $\frac{a}{b}$ with $a + b \le N$, and consider $\frac{a'}{b'}$ with $a' + b' = N + 1$. We distinguish the following cases:

Case 1. If $1 < \frac{a'}{b'} \le 2$, we have from b) that

$$f\left(\frac{a'}{b'}\right) = 1 + f\left(\frac{a'-b'}{b'}\right).$$

Since $(a' - b') + b' = a' < a' + b' = N + 1$, the right hand side is determined uniquely from the induction hypothesis, hence so is $f\left(\frac{a'}{b'}\right)$.

Case 2. If $1/2 \le \frac{a'}{b'} < 1$, using c) we have $f(\frac{a'}{b'}) = f(\frac{b'}{a'})^{-1}$ and now we can determine $f(\frac{b'}{a'})$, since $1 < \frac{b'}{a'} \le 2$.

Case 3. If $\frac{a'}{b'} < 1/2$ we have by a) that

$$f\left(\frac{a'}{b'}\right) = 1 + f\left(\frac{\frac{a'}{b'}}{1 - \frac{2a'}{b'}}\right) = 1 + f\left(\frac{a'}{b - 2a'}\right).$$

Since $b' + (a' - 2b') = a' - b' < a' + b' = N + 1$, $f\left(\frac{a'}{b-2a'}\right)$ is uniquely determined from the induction hypothesis, hence so is $f\left(\frac{a'}{b'}\right)$.

Case 4. If $\frac{a'}{b'} > 2$, then by c) we can write $f\left(\frac{a'}{b'}\right) = f\left(\frac{b'}{a'}\right)^{-1}$ and we can find $f\left(\frac{b'}{a'}\right)$, since $\frac{b'}{a'} < \frac{1}{2}$.

This covers all the possible cases and completes the induction step. Thus there is at most one function f satisfying the given conditions. We still need to show that there is such an f. Notice that for each pair of a rational $\frac{a}{b}$ and its reciprocal $\frac{b}{a}$, we found exactly one equation relating f at that point to f at a point $\frac{a'}{b'}$ with $a' + b' < a + b$. Thus we can define f inductively without ever trying to assign a value twice (and thus there is no danger of an inconsistency). This completes our proof.

4 Inequalities

Problem 4.1. There are $n \geq 1$ real numbers with non-negative sum written on a circle. Prove that one can enumerate them $a_1, a_2, \ldots, a_n$ such that they are consecutive on the circle and $a_1 \geq 0$, $a_1+a_2 \geq 0, \ldots, a_1+a_2+\ldots+a_{n-1} \geq 0$, $a_1 + a_2 + \ldots + a_n \geq 0$.

Solution. The proof is by induction on n. For one number, everything is clear.

Assume now that the result holds for $n-1$ numbers, some $n \geq 2$, and we show that it also holds for n numbers. As the total sum is non-negative, there are non-negative numbers on the circle. If they are all non-negative, there is nothing to prove. Otherwise, there is a negative one, say $a_n < 0$. Then by applying the induction hypothesis to $a_1, a_2, \ldots, a_{n-2}, a_{n-1} + a_n$ we can find a j such that $a_j, a_j + a_{j+1}, \ldots, a_j + \ldots + a_{n-2}, a_j + \ldots + a_{n-1} + a_n, \ldots$ are all non-negative. But as $a_n < 0$, we conclude that $a_j + \ldots + a_{n-1}$ is also non-negative. Therefore listing the numbers starting from a_j satisfies our condition. This completes our proof.

Problem 4.2. Let $a_1, a_2, \ldots, a_n$ be positive real numbers. Prove that

$$\frac{a_1}{(1+a_1)^2} + \frac{a_2}{(1+a_1+a_2)^2} + \cdots + \frac{a_n}{(1+a_1+\cdots+a_n)^2} < \frac{a_1+\cdots+a_n}{1+a_1+\cdots+a_n}.$$

Solution. We will prove the result by induction. For $n = 1$, the result is equivalent to

$$\frac{a_1}{(1+a_1)^2} < \frac{a_1}{1+a_1},$$

which holds, since $1 + a_1 > 1$.

Assume now that the result holds for $n-1$ variables, with $n \geq 2$. To prove the result for n variables, it suffices to show that

$$\frac{s_n - a_n}{1 + s_n - a_n} + \frac{a_n}{(1+s_n)^2} \leq \frac{s_n}{1+s_n},$$

where $s_n = a_1 + a_2 + \cdots + a_n$. After bringing expressions to common denominator and cross multiplying, the above inequality is equivalent to

$$(s_n - a_n)(1+s_n)^2 \leq (s_n + s_n^2 - a_n)(1 + s_n - a_n).$$

Expanding both sides and simplifying yields $a_n^2 \geq 0$, which is clear. This establishes our induction step and completes the proof.

Problem 4.3. Show that $2(\sqrt{n+1}-1) < 1 + \frac{1}{\sqrt{2}} + \ldots + \frac{1}{\sqrt{n}} < 2\sqrt{n}$.

Solution. The base case is $n = 1$, for which we have to prove that

$$2(\sqrt{2}-1) < 1 < 2.$$

Assuming $P(n)$ for some $n \geq 1$, for the induction step it suffices to show

$$2(\sqrt{n+1}-\sqrt{n}) < \frac{1}{\sqrt{n}} < 2(\sqrt{n}-\sqrt{n-1}).$$

However, this is obvious if we rewrite it as

$$\frac{2}{\sqrt{n+1}+\sqrt{n}} < \frac{1}{\sqrt{n}} < \frac{2}{\sqrt{n}-\sqrt{n-1}}.$$

Problem 4.4. Let x be a real number. Prove that for all positive integers n,

$$|\sin nx| \leq n|\sin x|.$$

Solution. We prove the given inequality by induction on $n \geq 1$. The base case $n = 1$ is clear.

Assume that our inequality is true for $n \geq 1$:

$$|\sin nx| \leq n|\sin x|.$$

To prove it for $n+1$, observe the following

$$|\sin(n+1)x| = |\sin nx \cos x + \cos nx \sin x| \leq |\sin nx| + |\sin x| \leq (n+1)|\sin x|,$$

and we are done.

Problem 4.5. Let $n > 2$. Find the least constant k such that for any $a_1, \ldots, a_n > 0$ with product 1 we have

$$\frac{a_1a_2}{(a_1^2+a_2)(a_2^2+a_1)} + \frac{a_2a_3}{(a_2^2+a_3)(a_3^2+a_2)} + \ldots + \frac{a_na_1}{(a_n^2+a_1)(a_1^2+a_n)} \leq k.$$

Solution. We shall prove that the answer is $n-2$.

Notice that if we put $a_1 = a_2 = .. = a_{n-1} = x, a_n = \frac{1}{x^{n-1}}$, then as x tends to 0 we get arbitrarily close to $n-2$. To prove the other part, observe that $(a_1^2+a_2)(a_2^2+a_1) \geq (a_1^2+a_1)(a_2^2+a_2)$. We use this inequality for every fraction and so it suffices to prove

$$\sum \frac{1}{(a_i+1)(a_{i+1}+1)} \leq n-2.$$

We prove the above inequality by induction on $n \geq 3$. For the base case, since $a_1a_2a_3 = 1$, we can write

$$a_1 = \frac{x}{y},\ a_2 = \frac{y}{z},\ a_3 = \frac{z}{x}.$$

Then

$$\begin{aligned}
&\frac{1}{(a_1+1)(a_2+1)} + \frac{1}{(a_2+1)(a_3+1)} + \frac{1}{(a_3+1)(a_1+1)} \\
&= \frac{yz}{(x+y)(y+z)} + \frac{xz}{(x+z)(y+z)} + \frac{xy}{(x+y)(x+z)} \\
&< \frac{xy}{xy+yz+zx} + \frac{yz}{xy+yz+zx} + \frac{xz}{xy+yz+zx} = 1.
\end{aligned}$$

Assume now that the result holds for $n-1$ variables, $n \geq 4$ and we prove it for n. Notice that unless $a_1 = \ldots = a_n = 1$ (in which case the inequality is immediate), there must be $1 \leq j \leq n$ such that $a_j > 1 > a_{j+1}$. As the inequality is cyclic, we can assume without loss of generality that this happens for $j = 2$. Now we can replace a_2 and a_3 with a_2a_3 and prove that the expression decreases by at most 1.

Indeed, we compute that

$$\frac{1}{(a_1+1)(a_2+1)} - \frac{1}{(a_1+1)(a_2a_3+1)} = \frac{-a_2(1-a_3)}{(a_1+1)(a_2+1)(a_2a_3+1)} \leq 0$$

and

$$\begin{aligned}
\frac{1}{(a_3+1)(a_4+1)} - \frac{1}{(a_2a_3+1)(a_4+1)} &= \frac{a_3(a_2-1)}{(a_3+1)(a_4+1)(a_2a_3+1)} \\
&\leq \frac{a_3(a_2-1)}{(a_3+1)(a_2a_3+1)}.
\end{aligned}$$

Thus the decrease is at most

$$\frac{1}{(a_2+1)(a_3+1)} + \frac{a_3(a_2-1)}{(a_3+1)(a_2a_3+1)} = \frac{1-a_3+a_2a_3+a_2^2a_3}{(a_2+1)(a_3+1)(a_2a_3+1)} < 1.$$

This shows that we can reduce the problem to $n-1$ variables and by the induction hypothesis, we are done.

Problem 4.6. Let $a_1, a_2, \ldots, a_n$ be integers, not all zero, such that $a_1 + a_2 + \cdots + a_n = 0$. Prove that

$$|a_1 + 2a_2 + \cdots + 2^{k-1}a_k| > \frac{2^k}{3},$$

for some $k \in \{1, 2, \ldots n\}$.

Solution. Assume that

$$\left|a_1 + 2a_2 + \cdots + 2^{k-1}a_k\right| \le \frac{2^k}{3} \text{ for all } k \in \{1, 2, \ldots, n\},$$

with a_i integers. We prove by induction on $1 \le k \le n$ that $a_1 = a_2 = \ldots = a_n = 0$.

For $k = 1$, the result is immediate, since $|a_1| \le \frac{2}{3} < 1$ necessarily implies $a_1 = 0$.

For the induction step, if the result is true for $i = 1, 2, \ldots, k-1$, then

$$\frac{2^k}{3} \ge \left|a_1 + 2a_2 + \cdots + 2^{k-1}a_k\right| = 2^{k-1}|a_k|,$$

yielding $|a_k| \le \frac{2}{3} < 1$, and again $a_k = 0$. This completes our induction step.

Therefore, $a_1 = a_2 = \ldots = a_n = 0$, which which contradicts the hypothesis of the problem. The result follows.

Problem 4.7. Let $n \ge 2$ and $a_1, a_2, \ldots, a_n \in (0, 1)$ with $a_1a_2\ldots a_n = A^n$. Show that

$$\frac{1}{1+a_1} + \ldots + \frac{1}{1+a_n} \le \frac{n}{1+A}.$$

Solution. We prove the result by induction on n.

When $n = 2$, the inequality is equivalent to

$$\frac{1}{1+x^2} + \frac{1}{1+y^2} \leq \frac{2}{1+xy},$$

which after clearing denominators is equivalent to

$$(2 + x^2 + y^2)(1 + xy) - 2(1 + x^2)(1 + y^2) \leq 0$$

or

$$(x - y)^2(xy - 1) \leq 0,$$

which is true.

For the induction step, set $b = \dfrac{a_n a_{n+1}}{A}$. Then $a_1 a_2 \ldots a_{n-1} b = A^n$, so by the induction hypothesis we have

$$\frac{1}{1+a_1} + \ldots + \frac{1}{1+a_{n-1}} + \frac{1}{1+b} \leq \frac{n}{1+A}$$

Therefore, it suffices to show that

$$\frac{1}{1+a_1} + \ldots + \frac{1}{1+a_{n-1}} + \frac{1}{1+b} + \frac{1}{1+A} \geq \frac{1}{1+a_1} + \ldots + \frac{1}{1+a_n} + \frac{1}{1+a_{n+1}},$$

or

$$\frac{1}{1+a_n} + \frac{1}{1+a_{n+1}} \leq \frac{1}{1+A} + \frac{A}{A + a_n a_{n+1}}.$$

By clearing denominators, this reduces to

$$(a_n - A)(a_{n+1} - A)(1 - a_n a_{n+1}) \leq 0.$$

We have from hypothesis that $a_n, a_{n+1} < 1$, so the above inequality holds if and only if one of a_n, a_{n+1} is greater than or equal to A and the other is less than or equal to A. Notice that there must be some $1 \leq j \leq n + 1$ such that $a_{j-1} \leq A$ and $a_j \geq A$ and since the original inequality is cyclic, we can assume without loss of generality that this happens when $j = n + 1$. This establishes the above inequality and thus the induction step, completing our proof.

Problem 4.8. **(APMO 1999)** Let $a_1, a_2, \ldots$ be a sequence of real numbers satisfying $a_{i+j} \leq a_i + a_j$ for all $i, j = 1, 2, \ldots$. Prove that

$$a_1 + \frac{a_2}{2} + \frac{a_3}{3} + \cdots + \frac{a_n}{n} \geq a_n$$

for each positive integer n.

Solution. We will prove this by induction on n. Note that the inequality holds when $n = 1$.

For the induction step, assume that the inequality holds for $1 \leq k \leq n$, so that we have

$$\begin{aligned} & a_1 \geq a_1, \\ & a_1 + \frac{a_2}{2} \geq a_2, \\ & a_1 + \frac{a_2}{2} + \frac{a_3}{3} \geq a_3, \\ & \vdots \\ & a_1 + \frac{a_2}{3} + \frac{a_3}{3} + \ldots + \frac{a_n}{n} \geq a_n. \end{aligned}$$

By summing all these inequalities we get

$$na_1 + (n-1)\frac{a_2}{2} + \cdots + \frac{a_n}{n} \geq a_1 + a_2 + \cdots + a_n.$$

We now add $a_1 + \ldots + a_n$ to both sides and we obtain

$$(n+1)\left(a_1 + \frac{a_2}{2} + \ldots + \frac{a_n}{n}\right) \geq (a_1+a_n)+(a_2+a_{n-1})+\ldots+(a_n+a_1) \geq na_{n+1}.$$

Finally, by dividing both sides by $n+1$ we get

$$a_1 + \frac{a_2}{2} + \ldots + \frac{a_n}{n} \geq \frac{na_{n+1}}{n+1},$$

i.e.

$$a_1 + \frac{a_2}{2} + \frac{a_3}{3} + \cdots + \frac{a_{n+1}}{n+1} \geq a_{n+1}.$$

Problem 4.9. (**Tuymaada 2000**) Let $n \geq 2$ be a positive integer and $x_1, \ldots, x_n$ be real numbers such that $0 < x_k \leq \frac{1}{2}$, for all $k = 1, 2, \ldots, n$. Prove that

$$\left(\frac{n}{x_1 + x_2 + \ldots + x_n} - 1\right)^n \leq \left(\frac{1}{x_1} - 1\right) \cdot \ldots \cdot \left(\frac{1}{x_n} - 1\right).$$

Solution. We begin by proving that if x and y are positive reals with $x+y \leq 1$, then

$$\left(\frac{1}{x} - 1\right)\left(\frac{1}{y} - 1\right) \geq \left(\frac{2}{x+y} - 1\right)^2. \tag{1}$$

Indeed, we can rewrite the inequality in the form

$$\frac{1-x-y}{xy} + 1 \geq \frac{1-x-y}{\left(\frac{x+y}{2}\right)^2} + 1,$$

which follows from $xy \leq \left(\frac{x+y}{2}\right)^2$.
Using the inequality (1), one can easily prove by induction on $m \geq 1$, that whenever $N = 2^m$ we have

$$\left(\frac{N}{x_1 + x_2 + \ldots + x_N} - 1\right)^N \leq \left(\frac{1}{x_1} - 1\right) \cdot \ldots \cdot \left(\frac{1}{x_N} - 1\right). \tag{2}$$

Let now $n \geq 2$ be arbitrary and take m sufficiently large so that $n < N = 2^m$. Let $x_1 + \ldots + x_n = nd$, for some positive real d. We set $x_{n+1} = \ldots = x_N = d$. Notice that we have $x_1 + \ldots + x_N = Nd$. Then the inequality (2) can be rewritten as

$$\left(\frac{N}{Nd} - 1\right)^N \leq \left(\frac{1}{x_1} - 1\right) \cdot \ldots \cdot \left(\frac{1}{x_N} - 1\right),$$

which is further equivalent to

$$\left(\frac{1}{d} - 1\right)^N \leq \left(\frac{1}{x_1} - 1\right) \cdot \ldots \cdot \left(\frac{1}{x_n} - 1\right) \cdot \left(\frac{1}{d} - 1\right)^{N-n}.$$

We derive that

$$\left(\frac{1}{d} - 1\right)^n = \left(\frac{n}{x_1 + \ldots + x_n} - 1\right)^n \leq \left(\frac{1}{x_1} - 1\right) \cdot \ldots \cdot \left(\frac{1}{x_n} - 1\right),$$

which is what we wanted.

Problem 4.10. **(China)** Let $a_1, \ldots, a_n$ be real numbers. Prove that the following two statements are equivalent:

a) $a_i + a_j \geq 0$ for all $i \neq j$;

b) If $x_1, \ldots, x_n$ are non-negative real numbers whose sum is 1, then

$$a_1x_1 + \ldots + a_nx_n \geq a_1x_1^2 + \ldots + a_nx_n^2.$$

Solution. We first assume that b) holds. For given $1 \leq i < j \leq n$, we take $x_i = x_j = \frac{1}{2}$ and we set $x_k = 0$ for all the other possible values. Then the inequality in b) reads

$$\frac{a_i + a_j}{2} \geq \frac{a_i + a_j}{4},$$

from which we deduce a).

Conversely, assuming that a) holds, we shall prove b) by induction on n. We start with the base case $n = 2$. Notice that from $x_1 + x_2 = 1$ we have

$$a_1x_1 + a_2x_2 - a_1x_1^2 - a_2x_2^2 = (a_1 + a_2)x_1x_2 \geq 0,$$

and we are done.

Assume now that the result holds for some $n \geq 2$ and let $x_1, \ldots, x_{n+1}$ be non-negative reals with sum 1. If $x_{n+1} = 1$, there is nothing to prove. Otherwise, we have

$$\sum_{i=1}^{n} \frac{x_i}{1 - x_{n+1}} = 1.$$

From the induction hypothesis we know that

$$\sum_{k=1}^{n} \frac{a_kx_k}{1 - x_{n+1}} \geq \sum_{k=1}^{n} a_k \left(\frac{x_k}{1 - x_{n+1}} \right)^2,$$

or

$$(1 - x_{n+1}) \sum_{k=1}^{n} a_kx_k \geq \sum_{k=1}^{n} a_kx_k^2.$$

We obtain that

$$\begin{aligned}\sum_{k=1}^{n+1} a_k x_k &= (1-x_{n+1})\sum_{k=1}^{n} a_k x_k + x_{n+1}\sum_{k=1}^{n} a_k x_k \\ &\quad + (1-x_{n+1})a_{n+1}x_{n+1} + a_{n+1}x_{n+1}^2 \\ &\geq \sum_{k=1}^{n+1} a_k x_k^2 + x_{n+1}\sum_{k=1}^{n} a_k x_k + a_{n+1}x_{n+1}\sum_{k=1}^{n} x_k \\ &\geq \sum_{k=1}^{n+1} a_k x_k^2,\end{aligned}$$

where the last inequality follows from the fact that $(a_k + a_{n+1})x_k x_{n+1} \geq 0$, for all k, by a). This completes our proof.

Problem 4.11. (**USAMO 2000**) Let $a_1, b_1, a_2, b_2, \ldots, a_n, b_n$ be non-negative real numbers. Prove that

$$\sum_{i,j=1}^{n} \min\{a_i a_j, b_i b_j\} \leq \sum_{i,j=1}^{n} \min\{a_i b_j, a_j b_i\}.$$

Solution. Define

$$L(a_1, b_1, \ldots, a_n, b_n) = \sum_{i,j} \left(\min\{a_i b_j, a_j b_i\} - \min\{a_i a_j, b_i b_j\}\right).$$

We shall prove that $L(a_1, b_1, \ldots, a_n, b_n) \geq 0$ by induction on n. The base case $n = 1$ is clear.

We have the following identities, which are easy to check by hand:

- $L(a_1, 0, a_2, b_2, \ldots) = L(0, b_1, a_2, b_2, \ldots) = L(a_2, b_2, \ldots)$;
- $L(x, x, a_2, b_2, \ldots) = L(a_2, b_2, \ldots)$;
- $L(a_1, b_1, a_2, b_2, a_3, b_3, \ldots) = L(a_1 + a_2, b_1 + b_2, a_3, b_3, \ldots)$ if $a_1/b_1 = a_2/b_2$;
- $L(a_1, b_1, a_2, b_2, a_3, b_3, \ldots) = L(a_2 - b_1, b_2 - a_1, a_3, b_3, \ldots)$ if $a_1/b_1 = b_2/a_2$ and $a_1 \leq b_2$.

These would be enough to perform the induction step, unless we are in the following case:

1. All of the a_i and b_i are non-zero;
2. For $i = 1, \ldots, n$ we have $a_i \neq b_i$;
3. For $i \neq j$, $a_i/b_i \neq a_j/b_j$ and $a_i/b_i \neq b_j/a_j$.

For $i = 1, \ldots, n$, let $r_i = \max\{a_i/b_i, b_i/a_i\}$. Without loss of generality, we may assume that $1 < r_1 < \ldots < r_n$ and $a_1 < b_1$. Notice that

$$f(x) = L(a_1, x, a_2, b_2, \ldots, a_n, b_n)$$

is a *linear* function of x in the interval $[a_1, r_2 a_1]$. Explicitly,

$$\begin{aligned} f(x) &= \min\{a_1 x, x a_1\} - \min\{a_1^2, x^2\} + L(a_2, b_2, \ldots, a_n, b_n) \\ &\quad + 2\sum_{j=2}^{n} (\min\{a_1 b_j, x a_j\} - \min\{a_1 a_j, x b_j\}) \\ &= (x - a_1)(a_1 + 2\sum_{j=2}^{n} c_j) + L(a_2, b_2, \ldots, a_n, b_n), \end{aligned}$$

where $c_j = -b_j$ if $a_j > b_j$ and $c_j = a_j$ if $a_j < b_j$.
In particular, because f is linear, we have

$$f(x) \geq \min\{f(a_1), f(r_2 a_1)\}.$$

Note that $f(a_1) = L(a_1, a_1, a_2, b_2, \ldots) = L(a_2, b_2, \ldots)$ and

$$\begin{aligned} f(r_2 a_1) &= L(a_1, r_2 a_1, a_2, b_2, \ldots) \\ &= \begin{cases} L(a_1 + a_2, r_2 a_1 + b_2, a_3, b_3, \ldots) & \text{if } r_2 = b_2/a_2, \\ L(a_2 - r_2 a_1, b_2 - a_1, a_3, b_3, \ldots) & \text{if } r_2 = a_2/b_2. \end{cases} \end{aligned}$$

Therefore, we deduce the desired inequality from the induction hypothesis in all cases.

Problem 4.12. (**Romania TST 1981**) Let $n \geq 1$ be a positive integer and let $x_1, x_2, \ldots, x_n$ be real numbers such that $0 \leq x_n \leq x_{n-1} \leq \ldots \leq x_3 \leq x_2 \leq$

x_1. We consider the sums

$$s_n = x_1 - x_2 + \ldots + (-1)^n x_{n-1} + (-1)^{n+1} x_n;$$
$$S_n = x_1^2 - x_2^2 + \ldots + (-1)^n x_{n-1}^2 + (-1)^{n+1} x_n^2.$$

Show that $s_n^2 \leq S_n$.

Solution. We first prove by induction on n that $0 \leq s_n$:
When $n = 1$, the result is clear, as $x_1 \geq 0$.
For $n = 2$ we have $s_2 = x_1 - x_2 \geq 0$, since $x_1 \geq x_2$.
Assume now that the result holds for $n-1$ and n, where $n \geq 2$.
If n is even, then

$$s_{n+1} = s_n + (-1)^{n+2} x_{n+1} = s_n + x_{n+1} \geq 0, \quad \text{since} \quad s_n \geq 0, x_{n+1} \geq 0.$$

If n is odd, then

$$s_{n+1} = s_{n-1} + x_n - x_{n+1} \geq 0, \quad \text{as} \quad s_{n-1} \geq 0, \quad \text{and} \quad x_{n+1} \leq x_n.$$

In both cases, the result holds for $n+1$, which completes the induction step.
We now prove that $s_n^2 \leq S_n$ by induction of step 2.
For $n = 1$, we have $s_1^2 = x_1^2 = S_1$, so the result holds.
For $n = 2$,

$$s_1^2 = (x_1 - x_2)^2 \leq (x_1 - x_2)(x_1 + x_2) = x_1^2 - x_2^2 = S_2.$$

Assume now that the result holds for all values up to some $n \geq 2$.
We prove that it also holds for $n+2$:
If $n = 2k+1$, $k \geq 0$, then

$$\begin{aligned}
s_{2k+3}^2 &= (s_{2k+1} - x_{2k+2} + x_{2k+3})^2 \\
&= s_{2k+1}^2 - 2s_{2k+1}(x_{2k+2} - x_{2k+3}) + (x_{2k+2} - x_{2k+3})^2 \\
&\leq S_{2k+1} - 2s_{2k+1}(x_{2k+2} - x_{2k+3}) - x_{2k+2}^2 + x_{2k+3}^2 + 2x_{2k+2}^2 - 2x_{2k+2}x_{2k+3} \\
&= S_{2k+1} - x_{2k+2}^2 + x_{2k+3}^2 - 2(x_{2k+2} - x_{2k+3})(s_{2k+1} - x_{2k+2}) \\
&= S_{2k+3} - 2(x_{2k+2} - x_{2k+3})s_{2k+2} \leq S_{2k+3}.
\end{aligned}$$

If $n = 2k$, then using what we have established above, we have

$$\begin{aligned} s_{2k+2}^2 &= (s_{2k+1} - x_{2k+2})^2 \\ &= s_{2k+1}^2 - 2s_{2k+1}x_{2k+2} + x_{2k+2}^2 \\ &\leq S_{2k+1} - x_{2k+2}^2 - 2s_{2k+1}x_{2k+2} + 2x_{2k+2}^2 \\ &= S_{2k+2} - 2x_{2k+2}s_{2k+2} \\ &\leq S_{2k+2}, \end{aligned}$$

as required.

Problem 4.13. Let $1 = x_1 \leq x_2 \leq \ldots \leq x_{n+1}$ be non-negative integers. Prove that

$$\frac{\sqrt{x_2 - x_1}}{x_2} + \ldots + \frac{\sqrt{x_{n+1} - x_n}}{x_{n+1}} < 1 + \frac{1}{2} + \ldots + \frac{1}{n^2}.$$

Solution. We prove the inequality by induction on n.
For $n = 1$, the result is clear, as

$$\frac{\sqrt{x_2 - x_1}}{x_2} \leq \frac{1}{2} < 1.$$

Assume now that the statement holds for some $n \geq 1$.
Let $1 = x_1 \leq x_2 \leq \ldots \leq x_{n+1} \leq x_{n+2}$ be positive integers. Then

$$\frac{\sqrt{x_2 - x_1}}{x_2} + \frac{\sqrt{x_3 - x_2}}{x_3} + \ldots + \frac{\sqrt{x_{n+2} - x_{n+1}}}{x_{n+2}}$$

$$\leq \frac{x_2 - x_1}{x_2} + \frac{x_3 - x_2}{x_3} + \ldots + \frac{x_{n+2} - x_{n+1}}{x_{n+2}}$$

$$\leq \left(\frac{1}{x_1 + 1} + \ldots + \frac{1}{x_2}\right) + \ldots + \left(\frac{1}{x_{n+1} + 1} + \ldots + \frac{1}{x_{n+2}}\right)$$

$$\leq \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{x_{n+2}}.$$

We now have two cases:

Case 1. If $x_{n+2} \le (n+1)^2$, then

$$\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{x_{n+2}} \le \frac{1}{2}+\ldots+\frac{1}{(n+1)^2} < 1+\frac{1}{2}+\ldots+\frac{1}{(n+1)^2}.$$

So the statement also holds for $n+1$ in this case.

Case 2. If $x_{n+2} > (n+1)^2$, then

$$\begin{aligned}\frac{\sqrt{x_{n+2}-x_{n+1}}}{x_{n+2}} &\le \frac{\sqrt{x_{n+2}-1}}{x_{n+2}}\\ &= \sqrt{\frac{1}{x_{n+2}}-\Big(\frac{1}{x_{n+2}}\Big)^2}\\ &< \sqrt{\frac{1}{(n+1)^2}-\frac{1}{(n+1)^4}}\\ &= \frac{\sqrt{n^2+2n}}{(n+1)^2}.\end{aligned}$$

Hence, we have that

$$\begin{aligned}&\frac{\sqrt{x_2-x_1}}{x_2}+\ldots+\frac{\sqrt{x_{n+1}-x_n}}{x_{n+1}}+\frac{\sqrt{x_{n+2}-x_{n+1}}}{x_{n+2}}\\ &\le 1+\frac{1}{2}+\ldots+\frac{1}{n^2}+\frac{\sqrt{x_{n+2}-x_{n+1}}}{x_{n+2}}\\ &\le 1+\frac{1}{2}+\ldots+\frac{1}{n^2}+\frac{\sqrt{n^2+2n}}{(n+1)^2}\\ &< 1+\frac{1}{2}+\ldots+\frac{1}{n^2}+\frac{2n+1}{(n+1)^2}\\ &= 1+\frac{1}{2}+\ldots+\frac{1}{n^2}+\underbrace{\frac{1}{(n+1)^2}+\ldots+\frac{1}{(n+1)^2}}_{\text{2n+1}}\\ &< 1+\frac{1}{2}+\ldots+\frac{1}{(n+1)^2}.\end{aligned}$$

So we have established the induction step in the second case as well, completing the proof.

Remark. If $1 = x_1 \leq x_2 \leq \ldots \leq x_{n+1}$ are non-negative integers, then with a little more effort it can be shown that

$$\frac{\sqrt{x_2 - x_1}}{x_2} + \ldots + \frac{\sqrt{x_{n+1} - x_n}}{x_{n+1}} \leq \left(\sum_{i=1}^{n^2} \frac{1}{i}\right) - \frac{1}{2}.$$

Problem 4.14. Prove that

$$\sum_{i=0}^{n} |\sin(2^i x)| \leq 1 + \frac{\sqrt{3}}{2} n,$$

where n is a non-negative integer.

Solution. We begin by proving that

$$2|\sin x| + |\sin 2x| \leq \frac{3\sqrt{3}}{2}.$$

We have that

$$\begin{aligned} 2|\sin x| + |\sin 2x| &= \frac{2}{\sqrt{3}} \cdot \sqrt{(3 - 3|\cos x|)(1 + |\cos x|)^3} \\ &\leq \frac{2}{\sqrt{3}} \cdot \sqrt{\left(\frac{6}{4}\right)^4} = \frac{3\sqrt{3}}{2}, \end{aligned}$$

where the inequality follows from AM-GM.
We now return to the original problem. An easy induction on $n \geq 1$ shows that

$$\frac{2}{3}|\sin x| + \sum_{i=1}^{n-1} |\sin 2^i x| + \frac{1}{3}|\sin 2^n x| \leq \frac{\sqrt{3}}{2} n.$$

The base case $n = 1$ is just the previous inequality divided by 3. The inductive step follows by just adding the $n = 1$ case with x replaced by $2^n x$.
To finish the problem, we simply add to the inequality above the easy inequality

$$\frac{1}{3}|\sin x| + \frac{2}{3}|\sin 2^n x| \leq 1.$$

Problem 4.15. Prove the following inequality

$$2(a^{2012}+1)(b^{2012}+1)(c^{2012}+1) \geq (1+abc)(a^{2011}+1)(b^{2011}+1)(c^{2011}+1),$$

where $a > 0$, $b > 0$, $c > 0$.

Solution. We begin by proving the following lemma:

Lemma. If $a > 0$ and n is a positive integer, then

$$2(1+a^{n+1})^3 \geq (1+a^3)(1+a^n)^3.$$

Proof. We prove the statement by induction on n. For $n = 1$, we need to prove that

$$2(1+a^2)^3 \geq (1+a^3)(1+a)^3.$$

We have that

$$2(1+a^2)^3 - (1+a^3)(1+a)^3 = (a-1)^4(a^2+a+1),$$

hence

$$2(1+a^2)^3 \geq (1+a^3)(1+a)^3.$$

Assume now that the statement holds for $n = k$, where k is some positive integer. We show that it also holds for $n = k+1$.
From the induction hypothesis,

$$2(1+a^{k+1})^3 \geq (1+a^3)(1+a^k)^3.$$

On the other hand,

$$(1+a^{k+2})(1+a^k) \geq (1+a^{k+1})^2.$$

Thus,

$$\begin{aligned}
2(1+a^{k+2})^3 &= 2(1+a^{k+1})^3\left(\frac{1+a^{k+2}}{1+a^{k+1}}\right)^3 \\
&\geq (1+a^k)^3(1+a^3)\left(\frac{1+a^{k+1}}{1+a^k}\right)^3 \\
&= (1+a^3)(1+a^{k+1})^3.
\end{aligned}$$

Therefore,

$$2(1 + a^{k+2})^3 \geq (1 + a^3)(1 + a^{k+1})^3.$$

This completes the induction step and hence the proof of our lemma.
For a, b, c positive reals we have

$$(1 + a^3)(1 + b^3)(1 + c^3) \geq (1 + abc)^3$$

(one way to prove this for example is to expand both sides, reduce the common terms and then use AM-GM twice). Combining this fact with the lemma we proved above, we have that

$$\begin{aligned}
&(2(a^{2012} + 1)(b^{2012} + 1)(c^{2012} + 1))^3 \\
&= 2(a^{2012} + 1)^3 \cdot 2(b^{2012} + 1)^3 \cdot 2(c^{2012} + 1)^3 \\
&\geq (1 + a^3)(a^{2011} + 1)^3 \cdot (1 + b^3)(b^{2011} + 1)^3 \cdot (1 + c^3)(c^{2011} + 1)^3 \\
&\geq (1 + abc)^3(a^{2011} + 1)^3(b^{2011} + 1)^3(c^{2011} + 1)^3.
\end{aligned}$$

We deduce that

$$2(a^{2012} + 1)(b^{2012} + 1)(c^{2012} + 1) \geq (1 + abc)(a^{2011} + 1)(b^{2011} + 1)(c^{2011} + 1),$$

which is what we wanted.

Problem 4.16. Prove that for any $n \in \mathbb{Z}$, $n \geq 14$ and any $x \in \left(0, \dfrac{\pi}{2n}\right)$ the following inequality holds:

$$\frac{\sin 2x}{\sin x} + \frac{\sin 3x}{\sin 2x} + \ldots + \frac{\sin(n+1)x}{\sin nx} < 2\cot x.$$

Solution. We have that

$$\frac{\sin(k+1)x}{\sin kx} = \cos x + \sin x \cdot \cot kx,$$

for $k = 1, \ldots, n$. Hence, it suffices to prove that

$$\sin x \cdot (\cot x + \ldots + \cot nx) < \cos x \cdot \left(\frac{2}{\sin x} - n\right).$$

Note that $kx \in \left(0, \dfrac{\pi}{2}\right)$, for $k = 1, \ldots, n$.
We now use the fact that if $0 < \alpha < \dfrac{\pi}{2}$, then $\sin\alpha < \alpha < \tan\alpha$. Therefore,

$$\begin{aligned}\sin x \cdot (\cot x + \cot(2x) + \ldots + \cot(nx)) &< x \cdot \left(\frac{1}{x} + \frac{1}{2x} + \ldots + \frac{1}{nx}\right)\\ &= \frac{1}{1} + \frac{1}{2} + \ldots + \frac{1}{n}.\end{aligned}$$

On the other hand,

$$0 < x < \frac{\pi}{2n} < \frac{\pi}{12}$$

and

$$\frac{\sqrt{3}}{2} = \cos\frac{\pi}{6} = 2\cos^2\frac{\pi}{12} - 1 > 2 \cdot \left(\frac{27}{28}\right)^2 - 1 = \frac{337}{392}.$$

Thus,

$$\cos\frac{\pi}{12} > \frac{27}{28}.$$

We obtain that

$$\cos x \cdot \left(\frac{2}{\sin x} - n\right) > \cos\frac{\pi}{12} \cdot \frac{4-\pi}{\pi} \cdot n > \frac{27}{28} \cdot \frac{4-\pi}{\pi} \cdot n > \frac{9}{35}n.$$

So it suffices to prove that for $n = 14, 15, \ldots$

$$1 + \ldots + \frac{1}{n} < \frac{9}{35}n.$$

We do this by mathematical induction on $n \geq 14$.
For the base case, we show that the statement holds for $n = 14$:

$$1+\frac{1}{2}+\ldots+\frac{1}{14} < 1+\frac{1}{2}+\ldots+\frac{1}{6}+\frac{1}{7}\cdot 7+\frac{1}{14} = 3+\frac{1}{4}+\frac{1}{5}+\frac{1}{14} < 3.6 = \frac{9}{35}\cdot 14.$$

Assume now that the statement holds for $n = k \geq 14$, $k \in \mathbb{Z}$. We prove that it also holds for $n = k + 1$.
We have that

$$1 + \frac{1}{2} + \ldots + \frac{1}{k} + \frac{1}{k+1} < \frac{9}{35} \cdot k + \frac{1}{k+1} \leq \frac{9}{35} \cdot k + \frac{1}{15} < \frac{9}{35} \cdot (k+1).$$

This completes the induction step and thus the proof of the question.

Problem 4.17. Let $A_n = \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$, $n \geq 2$. Prove that

$$e^{A_n} > \sqrt[n]{n!} \geq 2^{A_n}.$$

Solution. We first prove the inequality $e^{A_n} > \sqrt[n]{n!}$:
By the AM-GM inequality we have

$$\sqrt[n]{n!} < \frac{1+2+\cdots+n}{n} = \frac{n+1}{2}.$$

Since

$$\left(1+\frac{1}{n}\right)^n < e \iff 1+\frac{1}{n} < e^{\frac{1}{n}} \iff \frac{n+1}{n} < e^{\frac{1}{n}},$$

for any positive integer n, we have

$$\prod_{k=2}^{n} e^{\frac{1}{k}} > \prod_{k=2}^{n} \frac{k+1}{k} \iff e^{A_n} > \frac{n+1}{2} \implies e^{A_n} > \frac{n+1}{2} > \sqrt[n]{n!}.$$

We now prove that $\sqrt[n]{n!} \geq 2^{A_n}$. Note that

$$n+1 > 2^{(n+1)A_{n+1} - nA_n}, \quad n \geq 2.$$

Indeed, since

$$(n+1)\,A_{n+1} - nA_n = (n+1)\,A_n + 1 - nA_n = 1 + A_n$$

then

$$n+1 > 2^{(n+1)A_{n+1}-nA_n} \iff n+1 > 2^{1+A_n} \iff 2^{A_n} < \frac{n+1}{2},\ n \geq 2. \quad (1)$$

We will prove inequality (1) by induction on $n \geq 2$:
If $n = 2$, then $2^{A_2} = \sqrt{2}$ and $\sqrt{2} < \frac{3}{2} \iff 8 < 9$.
For the induction step, note that $\frac{n+2}{n+1} > 2^{\frac{1}{n+1}}$. Indeed, by the AM-GM Inequality

$$\sqrt[n+1]{2} = \sqrt[n+1]{2 \cdot 1 \cdot 1 \cdots 1} < \frac{2 + n \cdot 1}{n+1} = \frac{n+2}{n+1}.$$

So $2^{\frac{1}{n+1}} < \dfrac{n+2}{n+1}$ and by the induction hypothesis, $2^{A_n} < \dfrac{n+1}{2}$, hence we have

$$2^{A_n} \cdot 2^{\frac{1}{n+1}} < \frac{n+1}{2} \cdot \frac{n+2}{n+1} \Rightarrow 2^{A_{n+1}} < \frac{n+2}{2}.$$

We now prove that $\sqrt[n]{n!} \geq 2^{A_n}$ for all $n \geq 2$, by induction on n:
For $n = 2$, we have $2! = 2^{2A_2}$.
Assuming the result for some $n \geq 2$, we have that $n! \geq 2^{nA_n}$. Using that

$$2^{A_n} < \frac{n+1}{2},$$

we have

$$\begin{aligned}(n+1)! &= (n+1)\, n! \geq (n+1)\, 2^{nA_n} \\ &> 2^{(n+1)A_{n+1} - nA_n} \cdot 2^{nA_n} \\ &= 2^{(n+1)A_{n+1}}.\end{aligned}$$

Then

$$n! \geq 2^{nA_n} \iff \sqrt[n]{n!} \geq 2^{A_n}, \text{ for any } n \geq 2.$$

Problem 4.18. Show that if $x_1, x_2, \ldots, x_n \in (0, 1/2)$, then

$$\frac{x_1 x_2 \ldots x_n}{(x_1 + x_2 + \ldots + x_n)^n} \leq \frac{(1-x_1)\ldots(1-x_n)}{((1-x_1) + \ldots + (1-x_n))^n}.$$

Solution. We prove the result using Cauchy induction. We first prove the result when $n = 2^k$, by induction on k.

If $n = 2$, we must prove

$$\frac{x_1 x_2}{(x_1 + x_2)^2} \leq \frac{(1-x_1)(1-x_2)}{(2 - x_1 - x_2)^2},$$

or

$$x_1 x_2 \left[4 - 4(x_1 + x_2) + (x_1 + x_2)^2\right] \leq (x_1 + x_2)^2 \left[1 - (x_1 + x_2) + x_1 x_2\right].$$

This in turn gives $(x_1 - x_2)^2 (1 - x_1 - x_2) \geq 0$, which is true.

To pass from n to $2n$ we see that

$$\begin{aligned}
&\frac{x_1x_2\ldots x_nx_{n+1}\ldots x_{2n}}{(x_1+\ldots+x_n+x_{n+1}+\ldots+x_{2n})^{2n}}\\
=\;&\frac{x_1x_2\ldots x_n}{(x_1+x_2+\ldots+x_n)^n}\cdot\frac{x_{n+1}\ldots x_{2n}}{(x_{n+1}+\ldots+x_{2n})^n}\\
&\cdot\frac{(x_1+\ldots+x_n)^n(x_{n+1}+\ldots+x_{2n})^n}{(x_1+\ldots+x_n+x_{n+1}+\ldots+x_{2n})^{2n}}\\
=\;&\frac{x_1x_2\ldots x_n}{(x_1+x_2+\ldots+x_n)^n}\cdot\frac{x_{n+1}\ldots x_{2n}}{(x_{n+1}+\ldots+x_{2n})^n}\\
&\cdot\left[\frac{\dfrac{x_1+\ldots+x_n}{n}\,\dfrac{x_{n+1}+\ldots+x_{2n}}{n}}{\left(\dfrac{x_1+\ldots+x_n}{n}+\dfrac{x_{n+1}+\ldots+x_{2n}}{n}\right)^2}\right]^n\\
\le\;&\frac{(1-x_1)\ldots(1-x_n)}{[(1-x_1)+\ldots+(1-x_n)]^n}\,\frac{(1-x_{n+1})\ldots(1-x_{2n})}{[(1-x_{n+1})+\ldots+(1-x_{2n})]^n}\\
&\cdot\left[\frac{\left(1-\dfrac{x_1+\ldots+x_n}{n}\right)\left(1-\dfrac{x_{n+1}+\ldots+x_{2n}}{n}\right)}{\left[\left(1-\dfrac{x_1+\ldots+x_n}{n}\right)+\left(1-\dfrac{x_{n+1}+\ldots+x_{2n}}{2n}\right)\right]^2}\right]^n\\
=\;&\frac{(1-x_1)\ldots(1-x_{2n})}{[(1-x_1)+\ldots+(1-x_{2n})]^{2n}}.
\end{aligned}$$

Assume now that our assumption holds for n and let us prove it for $n-1$. To do this, take $x_n=\dfrac{x_1+\ldots+x_{n-1}}{n-1}$. Then

$$\begin{aligned}
&\frac{x_1x_2\ldots x_{n-1}\dfrac{x_1+\ldots+x_{n-1}}{n-1}}{\left(x_1+x_2+\ldots+x_{n-1}+\dfrac{x_1+\ldots+x_{n-1}}{n-1}\right)^n}\\
\le\;&\frac{(1-x_1)(1-x_2)\ldots(1-x_{m-1})(1-\dfrac{x_1+\ldots+x_{n-1}}{n-1})}{\left[(1-x_1)+\ldots+(1-x_{n-1})+\left(1-\dfrac{x_1+\ldots+x_{n-1}}{n-1}\right)\right]^n},
\end{aligned}$$

which after reducing the common terms in the numerators and denominators becomes exactly the inequality we wanted.

Problem 4.19. (**ELMO 2013 shortlist**) Let n be a fixed positive integer. Initially, n 1's are written on a blackboard. Every minute, David picks two numbers x and y written on the blackboard, erases them, and writes the number $(x+y)^4$ on the blackboard. Show that after $n-1$ minutes, the number written on the blackboard is at least $2^{\frac{4n^2-4}{3}}$.

Solution. We prove the result by induction on n. When $n = 1$, the claim is obviously true.

Suppose now that the result holds for all $n \le N$, with $N \ge 1$, and consider a blackboard with $N+1$ 1's written on it. After $N-1$ minutes, we are left with two numbers a_1 and a_2 on the blackboard.

Let S_1 be the set of 1's (from the original set consisting of $N+1$ 1's) which were involved in the computation of a_1, and S_2 be the set of 1's involved in the computation of a_2. Put $s_1 = |S_1|$ and $s_2 = |S_2|$. Note $s_1, s_2 > 0$ and $s_1 + s_2 = N+1$.

By the induction hypothesis,

$$a_1 \ge 2^{\frac{4s_1^2-4}{3}} \quad \text{and} \quad a_2 \ge 2^{\frac{4s_2^2-4}{3}}.$$

By the convexity of the function $f(x) = 2^x$, we have

$$\begin{aligned}
(a_1+a_2)^4 &= \left(2^{\frac{4s_1^2-4}{3}} + 2^{\frac{4s_2^2-4}{3}}\right)^4 \\
&\ge \left(2^{\frac{4s_1^2+4s_2^2-8}{6}+1}\right)^4 \\
&\ge \left(2^{\frac{(s_1+s_2)^2-1}{3}}\right)^4 = 2^{\frac{4(N+1)^2-4}{3}},
\end{aligned}$$

as desired.

5 Sequences and Recurrences

Problem 5.1. The sequence $(a_n)_{n\geq 1}$ is defined by $a_1 = 2$ and

$$a_{n+2} = \frac{2+a_n}{1-2a_n}, \quad n \geq 1.$$

Prove that all its terms are nonzero.

Solution. The recurrence of (a_n) resembles the formula for sum of tangents. Indeed, if $2 = \tan t, a_n = \tan b_n$ then

$$a_{n+1} = \frac{\tan(b_n) + \tan(t)}{1 - \tan(b_n)\tan(t)} = \tan(b_n + t).$$

Thus we can assume $b_{n+1} = b_n + t$. This gives by induction on n the relation $a_n = \tan nt$. So we reduced our question to showing that $\tan nt$ is never zero.

We prove by induction on n that there exist polynomials $p_n, q_n \in \mathbb{Z}[X]$ such that

$$\tan(nt) = \frac{p_n(\tan(t))}{q_n(\tan(t))}.$$

For $n = 1$, we take $p_1 = t, q_1 = 1$. Next,

$$\tan{(n+1)t} = \tan(nt+t) = \frac{t + \frac{p_n(t)}{q_n(t)}}{1 - t\frac{p_n(t)}{q_n(t)}} = \frac{tq_n(t) + p_n(t)}{q_n(t) - tp_n(t)}.$$

So we may take $p_{n+1}(t) = tq_n(t)+p_n(t), q_{n+1}(t) = q_n(t)-tp_n(t)$. Moreover, the same recursion and an easy induction show that $p_n(2) \equiv 2^n \pmod 5$ and $q_n(2) \equiv 2^{n-1} \pmod 5$.

In particular, $p_n(2)$ is nonzero and therefore $a_n \neq 0$. This completes our proof.

Problem 5.2. Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ be sequences defined by

$$a_{n+3} = a_{n+2} + 2a_{n+1} + a_n, \ n = 0, 1, \ldots, \ a_0 = 1, \ a_1 = 2, \ a_2 = 3$$

and

$$b_{n+3} = b_{n+2} + 2b_{n+1} + b_n, \ n = 0, 1, \ldots, \ b_0 = 3, \ b_1 = 2, \ b_2 = 1.$$

How many integers do the sequences have in common?

Solution. Clearly $a_3 = b_3 = 8$, while $a_4 = 16$, $a_5 = 35$, $a_6 = 75$, and $b_4 = 12$, $b_5 = 29$, $b_6 = 61$. Note that for $n = 4, 5, 6$, $a_n > b_n > a_{n-1}$, and an easy induction shows that for any $n \geq 3$,

$$\begin{aligned} a_{n+3} &= a_{n+2} + 2a_{n+1} + a_n \\ &> b_{n+3} = b_{n+2} + 2b_{n+1} + b_n \\ &> a_{n+1} + 2a_n + a_{n-1} = a_{n+2}. \end{aligned}$$

Therefore, since this inequality shows that both sequences are strictly increasing, no b_n for $n \geq 4$ may appear in (a_n), and the only values that appear in both sequences are $\{1, 2, 3, 8\}$, while the only n's for which $a_n = b_n$ are $n = 1$ and $n = 3$ with $a_1 = b_1 = 2$ and $a_3 = b_3 = 8$.

Problem 5.3. (**India 1996**) Define a sequence $(a_n)_{n\geq 1}$ by $a_1 = 1$ and $a_2 = 2$ and $a_{n+2} = 2a_{n+1} - a_n + 2$ for $n \geq 1$. Prove that for any $m \geq 1$, $a_m a_{m+1}$ is also a term in this sequence.

Solution. We first prove by induction on $n \geq 1$ that $a_n = (n-1)^2 + 1$.

The base cases $n = 1$ and $n = 2$ are satisfied from the hypotheses.

For the inductive step, we assume that the result holds for all values up to some $n \geq 2$ and we prove it for $n + 1$. We have

$$\begin{aligned} a_{n+1} &= 2a_n - a_{n-1} + 2 \\ &= 2(n-1)^2 + 2 - (n-2)^2 - 1 + 2 \\ &= (n-1)^2 + 2(n-1) + 2 = n^2 + 1, \end{aligned}$$

as required.

Therefore,

$$\begin{aligned} a_m a_{m+1} &= [(m-1)^2 + 1][m^2 + 1] \\ &= [m(m-1)]^2 + 2m^2 - 2m + 2 = [m(m-1) + 1]^2 + 1 \\ &= a_{m(m-1)+2}. \end{aligned}$$

Problem 5.4. (**Russia 2000**) Let $a_1, a_2, \ldots, a_n$ be a sequence of non-negative real numbers, not all zero. For $1 \leq k \leq n$, let

$$m_k = \max_{1\leq i\leq k} \frac{a_{k-i+1} + a_{k-i+2} + \ldots + a_k}{i}.$$

Prove that for any $\alpha > 0$, the number of integers k which satisfy $m_k > \alpha$ is less than $\frac{a_1+a_2+\ldots+a_n}{\alpha}$.

Solution. We prove the statement by induction on n. For $n = 1$, we have $m_1 = a_1$. If $\alpha > a_1$, then there are no k with $m_k > \alpha$, so the claim holds trivially. If $\alpha < a_1$, then there is exactly one such k and $1 < a_1/\alpha$. This establishes the base case.

Assume now that the result holds for all integers smaller than some $n > 1$. Let r be the number of integers k for which $m_k > \alpha$.

If $m_n \leq \alpha$, then the sequence $a_1, a_2, \ldots, a_{n-1}$ also contains r values of k for which $m_k > \alpha$. By the induction hypothesis,

$$r < \frac{a_1 + a_2 + \ldots + a_{n-1}}{\alpha} \leq \frac{a_1 + a_2 + \ldots + a_n}{\alpha},$$

as desired.

If instead $m_n > \alpha$, then there is some $1 \leq i \leq n$ such that

$$\frac{a_{n-i+1} + a_{n-i+2} + \ldots + a_n}{i} > \alpha.$$

Fix such an i. The sequence $a_1, \ldots, a_{n-i}$ contains at least $r - i$ values of k for which $m_k > \alpha$, so by the induction hypothesis we have

$$r - i < \frac{a_1 + a_2 + \ldots + a_{n-i}}{\alpha}.$$

Then

$$(a_1 + a_2 + \ldots + a_{n-i}) + (a_{n-i+1} + \ldots + a_n) > (r - i)\alpha + i\alpha = r\alpha.$$

Now we divide by α and we obtain the desired inequality. This completes our proof.

Problem 5.5. (**USAMO 2003**) Let $n \neq 0$. For every sequence of integers

$$A = a_0, a_1, a_2, \ldots, a_n \quad \text{satisfying} \quad 0 \leq a_i \leq i, \quad \text{for} \quad i = 0, \ldots, n,$$

define another sequence

$$t(A) = t(a_0), t(a_1), t(a_2), \ldots, t(a_n)$$

by setting $t(a_i)$ to be the number of terms in the sequence A that precede the term a_i and are different from a_i. Show that, starting from any sequence A as above, fewer than n applications of the transformation t lead to a sequence B such that $t(B) = B$.

Solution. Consider some sequence $C = c_0, \ldots, c_n$ as the image of A after t has been applied some finite number of times.

Lemma 1. If $t(c_k) = c_k = j$, then $c_j = \cdots = c_{j+i} = \cdots = c_k = j$ $(0 \leq i \leq k - j)$.

Proof. Since the j terms $c_0, \ldots, c_{j-1}$ are all less than j, no other terms that precede c_k can be different from c_k.

Lemma 2. If $t(c_k) = c_k = j$, then $t^2(c_k) = t(c_k)$ (here $f^2(x) = f(f(x))$).

Proof. Since only i terms precede the term c_i, we will have $t^m(c_i) \leq i$, for any integer m. This means that we will always have $t^m(c_0), \ldots, t^m(c_{j-1}) < j$. This means that $c_j = t(c_j) = \cdots = c_k = t(c_k)$, and the lemma follows by iteration.

Thus we may regard a term c_k as stable if $t(c_k) = c_k$. We will call a sequence stable if all of its terms are stable.

Lemma 3. If $c_k = j$ is not stable, then $t(c_k) > c_k$.

Proof. We have $c_0, \ldots, c_{j-1} < j$, so we always have $t(c_k) \geq c_k$. Equality implies that c_k is stable.

We will now prove the problem by induction on $n \geq 1$. When $n = 1$, we have one of the sequences $0, 0$ or $0, 1$, both of which are stable.

Now, suppose that $t^{n-2}(a_0), \ldots, t^{n-2}(a_{n-1})$ are stable. We must then have $t^{n-2}(a_n) \in \{n-2, n-1, n\}$. If $t^{n-2}(a_n) = n$, then the sequence is already stable. If $t^{n-2}(a_n) = n - 1$, then either it is already stable or $t^{n-1}(a_n) = n$, which is stable. If $t^{n-2}(a_n) = t^{n-2}(a_{n-1})$, then $t^{n-2}(a_n)$ must already be stable.

The only possibilities remaining are $t^{n-2}(a_n) = n - 2, t^{n-2}(a_{n-1}) = n - 1$ and $t^{n-2}(a_n) = n - 2, t^{n-2}(a_{n-1}) < n - 2$. In the first case, we must have $t^{n-1}(a_n)$ equal to $n - 1$ or n, both of which will make it stable. In the second case, we must have $t^{n-2}(a_{n-2}) = t^{n-2}(a_{n-1}) < n - 2$, giving us $t^{n-1}(a_n) = n$, which will make it stable. This completes our induction.

Problem 5.6. (**Russia 2000**) Let $a_1, a_2, \dots$ be a sequence with $a_1 = 1$ satisfying the recursion

$$a_{n+1} = \begin{cases} a_n - 2 & \text{if } a_n - 2 \notin \{a_1, a_2, \dots, a_n\} \quad \text{and} \quad a_n - 2 > 0; \\ a_n + 3 & \text{otherwise.} \end{cases}$$

Prove that for every positive integer $k > 1$, we have $a_n = k^2 = a_{n-1} + 3$ for some n.

Solution. We use induction to prove that for all non-negative n, $a_{5n+1} = 5n + 1$, $a_{5n+2} = 5n + 4$, $a_{5n+3} = 5n + 2$, $a_{5n+4} = 5n + 5$ and $a_{5n+5} = 5n + 3$. The base case $n = 0$ can be verified easily from the recursion.

Now assume that the claim is true for all integers less than some $n \geq 1$. Observe that by the induction hypothesis, $(a_1, a_2, \dots, a_{5n})$ is a permutation of $1, 2, \dots, 5n$. Thus $a_{5n} - 2 = 5n - 4$ is included in this set, and hence $a_{5n+1} = a_{5n} + 3 = 5n + 1$. Similarly, $a_{5n+2} = a_{5n+1} + 3 = 5n + 4$. On the other hand, $a_{5n+2} - 2 = 5n + 2$ is not in $\{a_1, a_2, \dots, a_{5n+2}\}$, so $a_{5n+3} = 5n + 2$. Continuing in this fashion, we find that $a_{5n+4} = a_{5n+3} + 3 = 5n + 5$ and $a_{5n+5} = a_{5n+4} - 2 = 5n + 3$. This completes the induction.

Each positive integer greater than 1 is included in the sequence $a_2, a_3, \dots$ exactly once. Also, all squares are congruent to either $0, 1$ or $4 \pmod 5$, respectively. From above, for any $n > 1$, in one of these residue classes we have $a_n = a_{n-1} + 3$, and this completes the proof.

Problem 5.7. Let $(x_n)_{n \geq 1}$ be defined by the relations $x_1 = 1$, and

$$x_{n+1} = \frac{x_n}{n} + \frac{n}{x_n}, \; n \geq 1.$$

Prove that $\lfloor x_n^2 \rfloor = n$, for all $n \geq 4$.

Solution. Let

$$f_n(x) = \frac{x}{n} + \frac{n}{x}.$$

Notice that f is decreasing on $[0, n]$.

We have to prove that $\sqrt{n} < x_n < \sqrt{n+1}$ for $n \geq 4$. However, we cannot prove the statement in this form by induction, because $\sqrt{n} < \sqrt{x_n} < \sqrt{n+1}$ implies

$$f_n(\sqrt{n+1}) < \sqrt{n} < f_n(\sqrt{n})$$

or

$$\frac{n}{\sqrt{n+1}} + \frac{\sqrt{n+1}}{n} < x_{n+1} < \sqrt{n} + \frac{1}{\sqrt{n}},$$

and while

$$\sqrt{n+1} < \frac{n}{\sqrt{n+1}} + \frac{\sqrt{n+1}}{n},$$

we actually have

$$\sqrt{n} + \frac{1}{\sqrt{n}} > \sqrt{n+2}$$

(just square to see this relation).

So we should strengthen the relation $x_n > \sqrt{n}$ to perform the induction step. How do we strengthen it? We see that the other endpoint of the interval gives a better lower bound for x_{n+1}: $\frac{\sqrt{n+1}}{n} + \frac{n}{\sqrt{n+1}}$ instead of $\sqrt{n+1}$. If we replace $n+1$ by n we get

$$x_n > \frac{\sqrt{n}}{n-1} + \frac{n-1}{\sqrt{n}} = \frac{n}{\sqrt{n}} - \frac{1}{\sqrt{n}} + \frac{\sqrt{n}}{n-1} = \sqrt{n} + \frac{1}{(n-1)\sqrt{n}}.$$

We therefore try to prove by induction that

$$\sqrt{n} + \frac{1}{(n-1)\sqrt{n}} \leq x_n < \sqrt{n+1}.$$

As

$$f_n(\sqrt{n+1}) = \sqrt{n+1} + \frac{1}{n\sqrt{n+1}},$$

one part of the induction is done. So we need to prove the other part, namely

$$f_n\left(\sqrt{n} + \frac{1}{(n-1)\sqrt{n}}\right) < \sqrt{n+2},$$

which can be rewritten as

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

or

$$(\sqrt{n}+\frac{1}{\sqrt{n}}-\sqrt{n+2})+\frac{1}{n(n-1)\sqrt{n}}<\frac{\sqrt{n}}{n^2-n+1}.$$

In order to evaluate differences of radicals, we shall use the identity

$$a-b=\frac{a^2-b^2}{a+b}.$$

So we get

$$\sqrt{n}+\frac{1}{\sqrt{n}}-\sqrt{n+2}=\frac{\left(\sqrt{n}+\frac{1}{\sqrt{n}}\right)^2-n-2}{\sqrt{n+2}+\sqrt{n}+\frac{1}{\sqrt{n}}}<\frac{1}{2n\sqrt{n}}.$$

We also have $\dfrac{1}{n(n-1)\sqrt{n}}\leq\dfrac{1}{2n\sqrt{n}}$ for $n\geq 3$. So

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

and the induction step is proved.

We are now left to check the base cases. We have

$$x_1=1,\ x_2=2,\ x_3=2,\ x_4=2+\frac{1}{6}.$$

So $x_4-\sqrt{4}=\frac{1}{6}=\frac{1}{(4-1)\sqrt{4}}$ and also $x_4=2+\frac{1}{6}<\sqrt{5}$. This completes our proof.

Problem 5.8. (**Russia 2000**) For any odd integer $a_0>5$, consider the sequence $a_0,a_1,a_2,\ldots$, where

$$a_{n+1}=\begin{cases}a_n^2-5 & \text{if } a_n \quad \text{is odd;}\\ \frac{a_n}{2} & \text{if } a_n \quad \text{is even}\end{cases},$$

for all $n\geq 0$. Prove that this sequence is not bounded.

Solution. We use induction on n to show that a_{3n} is odd and that $a_{3n} > a_{3n-3} > \cdots > a_0 > 5$, for all $n \geq 1$. The base case $n = 0$ is true from the hypotheses of the question.

Assume now that the result holds for all integers less than some $n \geq 1$. Because a_{3n} is odd, $a_{3n} \equiv 1 \pmod 8$ and hence $a_{3n+1} = a_{3n}^2 - 5 \equiv 4 \pmod 8$. Thus a_{3n+1} is divisible by 4 but not by 8, which implies that $a_{3(n+1)} = \frac{a_{3n+1}}{4}$ is indeed odd. Moreover, $a_{3n} > 5$ by the induction hypothesis, which implies that $a_{3n}^2 > 5a_{3n} > 4a_{3n} + 5$. This shows that $a_{3(n+1)} = \frac{1}{4}(a_{3n}^2 - 5) > a_{3n}$, which completes the induction and shows that the sequence is unbounded.

Problem 5.9. (**China 1997**) Let $(a_n)_{n\geq 1}$ be a sequence of non-negative real numbers satisfying $a_{n+m} \leq a_n + a_m$ for all positive integers m, n. Prove that if $n \geq m$, then

$$a_n \leq ma_1 + \left(\frac{n}{m} - 1\right) a_m.$$

Solution. Note that a simple induction on k shows that if we set $a_0 = 0$, then for any non-negative integers r and k we have

$$a_{km+r} \leq a_r + ka_m.$$

The base case $k = 0$ is trivial and the induction step is simply

$$a_{(k+1)m+r} \leq a_m + a_{km+r} \leq a_m + (a_r + ka_m) = a_r + (k+1)a_m.$$

In particular, taking $m = 1$ and $r = 0$, we get $a_k \leq ka_1$.

To solve the problem, we write $n = km + r$ with $k \geq 1$ and $0 \leq r \leq m - 1$. Note that $r < m$ gives $k > n/m - 1$. Then

$$\begin{aligned} a_n = a_{km+r} \leq ka_m + a_r &\leq \left(\frac{n}{m} - 1\right) a_m + \left(k + 1 - \frac{n}{m}\right) a_m + a_r \\ &\leq \left(\frac{n}{m} - 1\right) a_m + (mk + m - n)a_1 + ra_1 = ma_1 + \left(\frac{n}{m} - 1\right) a_m. \end{aligned}$$

Problem 5.10. Let $n \geq 2$ be an integer. Show that there exist $n+1$ numbers $x_1, x_2, \ldots, x_{n+1} \in \mathbb{Q} \setminus \mathbb{Z}$, so that $\{x_1^3\} + \{x_2^3\} + \ldots + \{x_n^3\} = \{x_{n+1}^3\}$, where $\{x\}$ is the fractional part of x.

Solution. Notice that if $y_1 < y_2 < \ldots < y_{n+2}$ are positive integers such that

$$y_1^3 + y_2^3 + \cdots + y_{n+1}^3 = y_{n+2}^3, \tag{1}$$

then the numbers $x_k = \frac{y_k}{y_{n+2}}$ for $1 \leq k \leq n$ and $x_{n+1} = -\frac{y_{n+1}}{y_{n+2}}$ satisfy the required conditions.

We now show by induction of step 2 that numbers satisfying (1) exist for every $n \geq 2$.

The equalities $3^3 + 4^3 + 5^3 = 6^3$ and $3^3 + 15^3 + 21^3 + 36^3 = 39^3$ settle the cases $n = 2$ and $n = 3$, respectively.

For the induction step, notice that if $3 < y_2 < \ldots < y_{n+1} < y_{n+2}$ are $n+2$ integers satisfying (1), then $3 < 4 < 5 < 2y_2 < \ldots < 2y_{n+1} < 2y_{n+2}$ are $n+4$ integers satisfying (1) as well.

Problem 5.11. Let $(a_n)_{n\geq 1}$ be a sequence of real numbers such that

$$a_1 = a_2 = 1 \quad \text{and} \quad a_{n+2} = a_{n+1} + \frac{a_n}{3^n}, \quad \text{for} \quad n \geq 1.$$

Prove that $a_n \leq 2$ for any $n \geq 1$.

Solution. We will show more strongly that

$$a_n \leq 2 - \frac{1}{3^{n-2}}$$

for $n \geq 2$ by induction on n. Note that this implies in particular that $a_n \leq 2$ as desired and this simpler inequality also holds for $n = 1$. The base case $n = 2$ follows from the hypotheses. For the induction step, suppose the inequality holds up to $n - 1$, then we have

$$a_n = a_{n-1} + \frac{a_{n-2}}{3^{n-2}} \leq 2 - \frac{1}{3^{n-3}} + \frac{2}{3^{n-2}} = 2 - \frac{1}{3^{n-2}}.$$

Note that since we only used the weaker bound on a_{n-2}, the argument for the inductive step is valid even for $n = 3$.

Problem 5.12. (**IMO 1995**) The positive real numbers $x_0, x_1, \ldots, x_{1995}$ satisfy $x_0 = x_{1995}$ and

$$x_{i-1} + \frac{2}{x_{i-1}} = 2x_i + \frac{1}{x_i},$$

for $i = 1, 2, \ldots, 1995$. Find the maximum value that x_0 can have.

Solution. The given condition is equivalent to $(2x_i - x_{i-1})(x_i x_{i-1} - 1) = 0$, so either $x_i = \frac{1}{2}x_{i-1}$ or $x_i = \frac{1}{x_{i-1}}$.

We shall show by induction on n that for any $n \geq 0$, $x_n = 2^{k_n} x_0^{e_n}$, for some integer k_n, where $|k_n| \leq n$ and $e_n = (-1)^{n-k_n}$. Indeed, this is true for $n = 0$. If it holds for some n, then $x_{n+1} = \frac{1}{2}x_n = 2^{k_n - 1}x_0^{e_n}$ (hence $k_{n+1} = k_n - 1$ and $e_{n+1} = e_n$) or $x_{n+1} = \frac{1}{x_n} = 2^{-k_n}x_0^{-e_n}$ (hence $k_{n+1} = -k_n$ and $e_{n+1} = -e_n$).

Thus $x_0 = x_{1995} = 2^{k_{1995}} x_0^{e_{1995}}$. Note that $e_{1995} = 1$ is impossible, since in that case k_{1995} would be odd, although it should equal 0. Therefore $e_{1995} = -1$, which gives $x_0^2 = 2^{k_{1995}} \leq 2^{1994}$, so the maximal value that x_0 can have is 2^{997}. This value is attained in the case $x_i = 2^{997-i}$ for $i = 0, \ldots, 1994$ and $x_{1995} = x_{1994}^{-1} = 2^{997}$.

Problem 5.13. (**INMO 2010**) Define a sequence $(a_n)_{n \geq 0}$ by $a_0 = 0$, $a_1 = 1$ and

$$a_n = 2a_{n-1} + a_{n-2}, \quad \text{for} \quad n \geq 2.$$

a) For every $m > 0$ and $0 \leq j \leq m$, prove that $2a_m$ divides $a_{m+j} + (-1)^j a_{m-j}$.
b) Suppose that 2^k divides n for some natural numbers n and k. Prove that 2^k divides a_n.

Solution. a) We have $2a_m = a_{m+1} - a_{m-1}$.
We will prove by induction on $j \geq 1$ that $2a_m | \left(a_{m+j} + (-1)^j a_{m-j}\right)$.
Suppose that this is true for all $j \leq k$. Then we have that

$$2a_m \mid \left(2a_{m+k} + 2(-1)^k a_{m-k}\right).$$

But

$$2a_{m+k} + 2(-1)^k a_{m-k} = a_{m+k+1} - a_{m+k-1} + (-1)^k a_{m-k+1} + (-1)^{k+1} a_{m-k-1},$$

so

$$2a_m \mid \left(a_{m+k+1} + (-1)^{k+1} a_{m-k-1} - \left(a_{m+k-1} + (-1)^{k-1} a_{m-k+1}\right)\right),$$

hence $2a_m \mid \left(a_{m+k+1} + (-1)^{k+1} a_{m-k-1}\right)$.

b) For a fixed m, we prove by induction on $n \geq 0$ that

$$a_{m+n} = a_{m-1}a_n + a_m a_{n+1}.$$

When $n = 0$, the identity holds since $a_0 = 0$ and $a_1 = 1$. Also, since $a_2 = 2a_1 = 2$, we have that $a_{m+1} = a_{m-1} + 2a_m$, which is true from the hypothesis.

Assume now that the identity holds for all values up to some $n \geq 1$. Then, using the inductive hypothesis for n and $n-1$ we have

$$\begin{aligned} a_{m+n+1} &= 2a_{m+n} + a_{m+n-1} \\ &= 2a_{m-1}a_n + 2a_m a_{n+1} + a_{m-1}a_{n-1} + a_m a_n \\ &= a_{m-1}(2a_n + a_{n-1}) + a_m(2a_{n+1} + a_n) \\ &= a_{m-1}a_{n+1} + a_m a_{n+2}, \end{aligned}$$

so the statement is proved for $n+1$. Since m was arbitrary, the identity holds for any $m, n \geq 0$. Notice that this implies in particular that if $m|n$ then $a_m|a_n$.

So if we prove that $2^k|a_{2^k}$, then the problem is solved because $a_{2^k}|a_n$.

From $a_{m+n} = a_{m-1}a_n + a_m a_{n+1}$, for $m = n$ we have:

$$a_{2m} = a_m(a_{m-1} + a_{m+1}) = 2a_m(a_{m-1} + a_m).$$

Then $2a_m|a_{2m}$, and since $a_2 = 2$, an easy induction gives now that $2^k|a_{2^k}$, which is what we wanted.

Problem 5.14. (**IMO 2013 shortlist**) Let n be a positive integer and let $a_1, \dots, a_{n-1}$ be arbitrary real numbers. Define the sequences $u_0, \dots, u_n$ and $v_0, \dots, v_n$ inductively by $u_0 = u_1 = v_0 = v_1 = 1$, and

$$u_{k+1} = u_k + a_k u_{k-1}, \; v_{k+1} = v_k + a_{n-k}v_{k-1} \text{ for } k = 1, \dots, n-1.$$

Prove that $u_n = v_n$.

Solution. We prove by induction on k that

$$u_k = \sum_{\substack{0<i_1<\ldots<i_t<k, \\ i_{j+1}-i_j \geq 2}} a_{i_1} \dots a_{i_t}. \tag{1}$$

Note that we have one trivial summand equal to 1 (which corresponds to $t = 0$ and the empty sequence, whose product is 1).

For $k = 0, 1$, the sum on the right-hand side contains only the empty product, so (1) holds due to $u_0 = u_1 = 1$. For $k \geq 1$, assuming the result is true for $0, 1, \ldots, k$ we have

$$\begin{aligned} u_{k+1} &= \sum_{\substack{0<i_1<\ldots<i_t<k,\\ i_{j+1}-i_j\geq 2}} a_{i_1}\ldots a_{i_t} + \sum_{\substack{0<i_1<\ldots<i_t<k-1,\\ i_{j+1}-i_j\geq 2}} a_{i_1}\ldots a_{i_t}\cdot a_k \\ &= \sum_{\substack{0<i_1<\ldots<i_t<k+1,\\ i_{j+1}-i_j\geq 2,\\ k\notin\{i_1,\ldots,i_t\}}} a_{i_1}\ldots a_{i_t} + \sum_{\substack{0<i_1\ldots<i_t<k+1,\\ i_{j+1}-i_j\geq 2,\\ k\in\{i_1,\ldots,i_t\}}} a_{i_1}\ldots a_{i_t} \\ &= \sum_{\substack{0<i_1<\ldots<i_t<k+1,\\ i_{j+1}-i_j\geq 2}} a_{i_1}\ldots a_{i_t}, \end{aligned}$$

as required.
Applying (1) to the sequence $b_1, \ldots, b_n$ given by $b_k = a_{n-k}$ for $1 \leq k \leq n$, we get

$$v_k = \sum_{\substack{i_1<\ldots<i_t<k,\\ i_{j+1}-i_j\geq 2}} b_{i_1}\ldots b_{i_t} = \sum_{\substack{n>i_1>\ldots>i_t>n-k,\\ i_j-i_{j+1}\geq 2}} a_{i_1}\ldots a_{i_t}. \tag{2}$$

For $k = n$, the expressions (1) and (2) coincide, so indeed $u_n = v_n$.

Problem 5.15. (**IMO 2006 shortlist**) The sequence of real numbers $a_0, a_1, a_2, \ldots$ is defined recursively by

$$a_0 = -1, \quad \sum_{k=0}^{n} \frac{a_{n-k}}{k+1} = 0, \quad \text{for} \quad n \geq 1.$$

Show that $a_n > 0$ for $n \geq 1$.

Solution. We prove the statement by induction on $n \geq 1$.

For the base case, we note that $a_1 = 1/2$.

Now assume that $a_1, \ldots, a_{n-1}$ are all positive for some $n \geq 2$. We note that a_n is positive if and only if $\sum_{k=1}^{n} \frac{a_{n-k}}{k+1}$ is negative. Now, since $a_1, \ldots, a_{n-1}$

are all positive, we know

$$-\frac{a_0}{n+1} = \frac{n}{n+1} \cdot \left(-\frac{a_0}{n}\right) = \frac{n}{n+1} \sum_{k=0}^{n-2} \frac{a_{n-1-k}}{k+1}$$
$$> \sum_{k=0}^{n-2} \frac{k+1}{k+2} \cdot \frac{a_{n-1-k}}{k+1} = \sum_{k=0}^{n-2} \frac{a_{n-1-k}}{k+2}.$$

This implies that

$$\sum_{k=1}^{n} \frac{a_{n-k}}{k+1} = \frac{a_0}{n+1} + \sum_{k=1}^{n-1} \frac{a_{n-k}}{k+1} = \frac{a_0}{n+1} + \sum_{k=0}^{n-2} \frac{a_{n-1-k}}{k+2} < \frac{a_0}{n+1} - \frac{a_0}{n+1} = 0,$$

which is what we wanted.

Problem 5.16. Let $(a_n)_{n\geq 0}$ be a sequence defined by

$$a_0 = a_1 = 47 \quad \text{and} \quad 2a_{n+1} = a_n a_{n-1} + \sqrt{(a_n^2 - 4)(a_{n-1}^2 - 4)}.$$

Prove that $a_n + 2$ is a perfect square for any $n \geq 0$.

Solution. We prove the statement by induction on n. For $n = 0$, $n = 1$ and $n = 2$ we have that $a_0 + 2 = a_1 + 2 = 49$ and $a_2 + 2 = 47^2$, so the result holds in these cases.

Assume now that the result holds for $n-1$ and n, for some $n \geq 1$. Then

$$2a_{n+1} = a_n a_{n-1} + \sqrt{(a_n^2 - 4)(a_{n-1}^2 - 4)}$$

gives

$$4a_{n+1}^2 - 4a_{n+1}a_n a_{n-1} + a_n^2 a_{n-1}^2 = a_n^2 a_{n-1}^2 - 4a_n^2 - 4a_{n-1}^2 + 16,$$

which simplifies further to

$$a_{n+1}^2 + a_n^2 + a_{n-1}^2 = a_{n+1}a_n a_{n-1} + 4. \tag{1}$$

By completing the square on the right side, we get

$$(a_{n+1} + a_n + a_{n-1})^2 = 2a_{n-1}a_n + 2a_{n-1}a_{n+1} + 2a_n a_{n+1} + a_{n-1}a_n a_{n+1} + 4. \tag{2}$$

We want to relate $a_{n-1}+2$ and a_n+2 to $a_{n+1}+2$ to be able to use our inductive hypothesis. Also, if we look at the right side of (2), we notice that it looks very similar to the factorization

$$\begin{aligned}(a_{n-1}+2)(a_n+2)(a_{n+1}+2) &= 4(a_{n-1}+a_n+a_{n+1}) \\ &\quad + 4(a_{n-1}a_n + a_{n-1}a_{n+1} \\ &\quad + a_na_{n+1}) + a_{n-1}a_na_{n+1} + 8.\end{aligned}$$

So we add $4(a_{n-1}+a_n+a_{n+1})+4$ to both sides of (1) and we obtain that

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

Since we assumed that $a_{n-1}+2$ and a_n+2 are both perfect squares, from the above relation we would be done if we knew that a_{n+1} was an integer. To prove this, we make use of (2). By writing the relation for n and $n+1$ we have

$$a_{n+1}^2 + a_n^2 + a_{n-1}^2 = a_{n+1}a_na_{n-1} + 4a_{n+2}^2 + a_{n+1}^2 + a_n^2 = a_{n+2}a_{n+1}a_n + 4.$$

Subtracting these two relations we get

$$a_{n+2}^2 - a_{n+1}^2 = a_{n+1}a_n(a_{n+2}-a_{n-1}) \Leftrightarrow a_{n+2}+a_{n-1} = a_{n+1}a_n.$$

We computed at the beginning that a_0, a_1 and a_2 were all integers. So assuming that a_{n-1}, a_n and a_{n+1} are all integers for some $n \geq 1$, from the above relation we have that a_{n+2} must also be an integer, by induction. This completes our proof.

Problem 5.17. Let $a_0, a_1, a_2, \ldots$ be an increasing sequence of non-negative integers such that every non-negative integer can be expressed uniquely in the form $a_i + 2a_j + 4a_k$, where i, j and k are not necessarily distinct. Determine a_{1998}.

Solution. We will prove that this sequence is unique, showing that a_n is uniquely determined, by induction on n.

Clearly a_0=0, so the base case holds.

Now, suppose we have proven that

$$\{a_0, a_1, \ldots, \} \cap \{1, 2, \ldots, n\}$$

is uniquely determined. If $n+1$ can be written as $x + 2y + 4z$ where $x, y, z \in \{a_0, a_1, \ldots, \} \cap \{1, 2 \ldots, n\}$ then it cannot belong to the sequence due to the uniqueness of this representation, while if not, it definitely must because such a representation should exist. So, whether $n+1$ belongs or not to the sequence depends only on the terms of the sequence smaller that n. As these terms are uniquely determined, the induction step is done.

Therefore, it suffices to find an example of such a sequence. The expression $a_i + 2a_j + 4a_k$ is strongly related to base 8 expansion and with this idea in mind, we easily find that the sequence consists of those non-negative integers whose digits in base 8 are only 0 or 1. Then we check that a_n is obtained by writing n in base 2 and reading the result in base 8. In particular,

$$1998 = 2^{10} + 2^9 + 2^8 + 2^7 + 2^6 + 2^3 + 2^2 + 2^1,$$

so

$$a_{1998} = 8^{10} + 8^9 + 8^7 + 8^3 + 8^2 + 8 = 1,227,096,648.$$

Problem 5.18. A sequence of integers $a_1, a_2, a_3, \ldots$ is defined as follows: $a_1 = 1$ and for $n \geq 1$, a_{n+1} is the least integer greater than a_n such that $a_i + a_j \neq 3a_k$ for any i, j and k in $\{1, 2, 3, \ldots, n+1\}$, not necessarily distinct. Determine a_{1998}.

Solution. The first few terms of the sequence are

$$1, 3, 4, 7, 10, 12, 13, 16, 19, 21.$$

We notice that the above sequence consists only of numbers of the form $3k+1$ and $3(3k+1)$. So let us try to prove by induction on n that $x \leq n$ belongs to the sequence if and only if $x \equiv 1 \pmod 3$ or $x \equiv 3 \pmod 9$. The base cases for $n \leq 21$ were proven above.

For the induction step, we distinguish the following cases:

a) $n = 3k+1$. Then n must belong to the sequence, as otherwise we would find a triple $a_i + a_j = 3a_k$ such that the largest of a_i, a_j, a_k is n. Clearly

$a_k \neq n$. Thus we may assume $a_i = n$, so $a_j = 3a_k - n \equiv 2 \pmod 3$, which contradicts the induction hypothesis.

b) $n = 3(3k+1)$. n must belong to the sequence, otherwise we would, like above, find $a_j, a_k \leq n$ such that $a_j = 3a_k - n$. As $3a_k \equiv 3 \pmod 9$ or $3a_k \equiv 0 \pmod 9$, we find $a_j \equiv 0 \pmod 9$ or $a_j \equiv 6 \pmod 9$, contradicting the induction assumption.

c) $n = 3k+2$. Then n does not belong to the sequence, because one of the numbers $k+1, k+2, k+3$ is of the form $3m+1$, thus belongs to the sequence and we get $3k+2+1 = 3(k+1)$ or $3k+2+4 = 3(k+2)$ or $3k+2+7 = 3(k+3)$

d) $n = 3(3k+2)$. Then one of the numbers $3(k+1), 3(k+2), 3(k+3)$ is of form the form $3m+1$, thus belongs to the sequence and we get $3(3k+2)+3 = 3(3(k+1))$ or $3(3k+2)+12 = 3(3(k+2))$ or $3(3k+2)+21 = 3(3(k+3))$.

e) $n = 9k$. Then n does not belong to the sequence, as $9k+3 = 3(3k+1)$.

From what we have proven it is clear that $a_{4k} = 9k-2$, $a_{4k+1} = 9k+1$, $a_{4k+2} = 9k+3$, $a_{4k+3} = 9k+4$. As $1998 = 4 \cdot 499 + 2$, we obtain

$$a_{1998} = 9 \cdot 499 + 3 = 4494.$$

Problem 5.19. Let k be a positive integer. The sequence $a_1, a_2, a_3, \ldots$ is defined by $a_1 = k+1$, and $a_{n+1} = a_n^2 - ka_n + k$, $n \geq 1$. Prove that a_m and a_n are coprime (for $m \neq n$).

Solution. We rewrite the condition as $a_{n+1} - k = a_n(a_n - k)$ and from here $a_{n+1} - k = a_1 a_2 \ldots a_n$. So if $m < n$ we get that $a_m | a_n - k$ so it only remains to prove that $(a_m, k) = 1$.

We do this by induction on $m \geq 1$. The base case $m = 1$ is verified from the hypothesis.

For the induction step, notice that $a_{n+1} \equiv a_n^2 \pmod k$, which shows that $(a_n, k) = 1 \Rightarrow (a_{n+1}, k) = 1$, completing our proof.

Problem 5.20. (**Bulgaria TST 2011**) Let $(x_n)_{n\geq 1}$ be a sequence defined by $x_1 = \frac{2}{3}$ and

$$x_{n+1} = \frac{3x_n + 2}{3 - 2x_n}, \quad \text{for all} \quad n \geq 1.$$

Is this sequence eventually periodic?

Solution. The answer is no. Assume by contradiction that there exist N, T such that for all $i \geq N$ we have $x_{i+T} = x_i$. Observe that

$$x_{n+1} = \frac{3x_n + 2}{3 - 2x_n} \quad \text{implies} \quad x_n = \frac{3x_{n+1} - 2}{3 + 2x_{n+1}},$$

so if $x_{n+1+T} = x_{n+1}$, we must also have $x_{n+T} = x_n$. This shows that we can take $N = 1$.

Now let us assume that $x_n = \frac{p_n}{q_n}$ in its reduced form. Then

$$x_{n+1} = \frac{p_{n+1}}{q_{n+1}} = \frac{3p_n + 2q_n}{3q_n - 2p_n}.$$

If we denote $d = \gcd(3p_n + 2q_n, 3q_n - 2p_n)$, we observe that we must have $d = 1$ or $d = 13$. Now we define two sequences a_n and b_n by $a_1 = 2$, $b_1 = 3$ and $a_{n+1} = 3a_n + 2b_n$, $b_{n+1} = 3b_n - 2a_n$. One can check by induction that $a_n \equiv 9 \cdot 6^n \pmod{13}$ and $b_n \equiv 7 \cdot 6^n \pmod{13}$. Thus none of the terms are divisible by 13. Therefore, we have that $p_{n+1} = 3p_n + 2q_n$ and $q_{n+1} = 3q_n - 2p_n$.

One can now prove inductively that

$$p_{n+1}^2 + q_{n+1}^2 = 13(p_n^2 + q_n^2) = \ldots = 13^n.$$

By periodicity, there must be some n such that $x_{n+1} = x_1 = \frac{2}{3}$. But then $x_n = 0$, so $13 \mid p_n$ and then $13 \mid q_n$, which is a contradiction to p_n and q_n being coprime. Therefore, the sequence is not periodic.

Problem 5.21. (**St. Petersburg**) Let $(x_n)_{n\geq 1}$ and $(y_n)_{n\geq 1}$ be two sequences given by $x_1 = \frac{1}{10}$, $y_1 = \frac{1}{8}$ and

$$x_{n+1} = x_n + x_n^2, \quad y_{n+1} = y_n + y_n^2 \quad \text{for} \quad n \geq 1.$$

Prove that for any positive integers m and n, we cannot have $x_n = y_m$.

Solution 1. Notice that $x_1 = 0.1$ and $y_1 = 0.125$. We prove by induction on $n \geq 1$ that the last non-trivial decimal of x_n is 1, while the last non-trivial decimal of y_n is 5. This holds for $n = 1$ from hypothesis.

Assuming the result for some $n \geq 1$, we have $x_n = \frac{x}{10^N}$, for some positive integer x whose last digit is 1. Then $x_{n+1} = \frac{10^N x + x^2}{10^{2N}}$. Notice that the last digit of $10^N x + x^2$ is 1, which establishes the claim for x_{n+1}. A similar proof works for y_{n+1}, showing that x_n and y_m can never be equal.

Solution 2. Notice that $x_2 = 0.11$, $x_3 = 0.1221$, $x_4 = 0.1221+0.1221^2 > 0.13$, so $x_3 < 0.125 = y_1 < x_4$. We prove by induction on n that $x_{n+2} < y_n < x_{n+3}$. We have just proved the base case.

Assuming the result for some $n \geq 1$, we have $x_{n+2} < y_n < x_{n+3}$. An easy induction shows that all x_n's and y_n's are positive, so by squaring our inequality we obtain $x_{n+2}^2 < y_n^2 < x_{n+3}^2$. Adding this to the original inequality, we obtain that

$$x_{n+3} < y_{n+1} < x_{n+4},$$

completing our proof.

Problem 5.22. (**Taiwan 2000**) Let $f : \mathbb{N}^* \to \mathbb{N}$ be defined recursively by $f(1) = 0$ and

$$f(n) = \max_{1 \leq j \leq \lfloor \frac{n}{2} \rfloor} \{f(j) + f(n-j) + j\},$$

for all $n \geq 2$. Determine $f(2000)$.

Solution. For each positive integer n, we consider the binary representation of n. Consider the substrings of the representation formed by removing at least one digit from the left side of the representation, such that the substring so formed begins with a 1. We call the decimal values of these substrings the *tail-values* of n. Also, for each 1 that appears in the binary representation of n, if it represents the number 2^k, let $2^k \cdot \frac{k}{2}$ be a *place-value* of n.

Let $g(n)$ be the sum of the tail- and place-values of n. We prove by induction on n that $f(n) = g(n)$. For convenience, let $g(0) = 0$. It is clear that $g(1) = 1$. It will therefore suffice to show that $g(n)$ satisfies the same recurrence as $f(n)$. We first prove that

$$g(n) \geq g(j) + g(n-j) + j, \tag{1}$$

for all n, j such that $0 \leq j \leq \lfloor \frac{n}{2} \rfloor$. The relation is definitely true for $j = 0$, since we defined $g(0) = 0$. Now we induct on the number of (binary) digits of $n - j$. For the base case (when $n - j$ has 1 binary digit), we can only have $n - j = 1$. In this case, $(n, j) = (2, 1)$ or $(n, j) = (1, 0)$, in which cases (1) is easily seen to be true. Now we prove the induction step by distinguishing two cases:

Case 1. If $n-j$ and j have the same number of digits, say $k+1$. Let a and b be the numbers formed by taking off the leftmost 1's (which represent 2^k) from $n-j$ and j. We want to show that

$$g(n) = g(a+b+2^{k+1}) \geq g(2^k+a) + g(2^k+b) + (2^k+b).$$

Subtracting the inequality $g(a+b) \geq g(a) + g(b) + b$ (which is true by the induction hypothesis), we see that it suffices to show that

$$g(a+b+2^{k+1}) - g(a+b) \geq g(2^k+a) - g(a) + g(2^k+b) - g(b) + 2^k. \quad (2)$$

On the right hand side, $g(2^k+a)$ equals $g(a)$ plus the place-value $2^a \cdot \frac{k}{2}$ and the tail-value a. Similarly, $g(2^k+b) = g(b) + 2^k \cdot \frac{k}{2} + b$. Hence, the right hand side equals

$$2^k \cdot \frac{k}{2} + a + 2^k \cdot \frac{k}{2} + b + 2^k = 2^{k+1} \cdot \frac{k+1}{2} + a + b.$$

As for the left hand side of (2), because $a < 2^k$ and $b < 2^k$, the binary representation of $a+b+2^{k+1}$ is simply the binary representation of $a+b$, with an additional 1 in the 2^{k+1} position. Hence, $g(a+b+2^{k+1})$ equals $g(a+b)$ plus the additional place value $2^{k+1} \cdot \frac{k+1}{2}$. Thus, $g(a+b+2^{k+1}) - g(a+b)$ equals the right hand side, proving the inequality in (2).

Case 2. If $n-j$ has more digits than j, then let $n-j$ have $k+1$ digits and, as before, let $a = n-j = 2^k$. We need to prove that

$$g(a+j+2^k) \geq g(a+2^k) + g(j) + j.$$

We know by the induction hypothesis that

$$g(a+j) \geq g(a) + g(j) + \min\{a, j\}.$$

Subtracting, we see that it suffices to prove that

$$g(a+j+2^k) - g(a+j) \geq g(a+2^k) - g(a) + j - \min\{a, j\}. \quad (3)$$

We find as in Case 1 that on the right hand side,

$$g(a+2^k) - g(a) = 2^k \cdot \frac{k}{2} + a.$$

Hence, the right hand side equals

$$2^k \cdot \frac{k}{2} + a + j - \min\{a, j\} = 2^k \cdot \frac{k}{2} + \max\{a, j\}.$$

On the left hand side of (3), if $a + j < 2^k$ (i.e. so that the 2^k digits do not carry in the sum $(a+j) + 2^k$), then $g(a+j+2^k)$ equals $g(a+j)$ plus the additional place-value $2^k \cdot \frac{k}{2}$ and the additional tail-value $a+j$. Hence, the left hand side of (3) is indeed greater than or equal to the right hand side. Otherwise, if the 2^k digits *do* carry in the sum $(a+j)+2^k$, then $g(a+j+2^k)$ equals $g(a+j)$ plus the additional place-value $2^{k+1} \cdot \frac{k+1}{2}$, minus the original place-value $2^k \cdot \frac{k}{2}$. Thus, the left hand side equals

$$2^{k+1} \cdot \frac{k+1}{2} - 2^k \cdot \frac{k}{2} = 2^k \cdot \frac{k}{2} + 2^k > 2^k \cdot \frac{k}{2} + \max\{a, j\},$$

so again (3) is true. This completes the induction.

Hence, $g(n) \geq \max\limits_{1 \leq j \leq \lfloor \frac{n}{2} \rfloor} \{g(j) + g(n-j) + j\}$ for all n. We now prove that in fact equality holds, by showing that $g(n) = g(j) + g(n-j) + j$ for some j. Let 2^k be the largest power of 2 less than n, and set $j = n - 2^k$. Then $g(n)$ equals $g(n-2^k)$ plus the additional place-value $g(2^k) = g(n-j)$ and the additional tail-value $n - 2^k = j$.

It follows that $f(n) = g(n)$ for all n. Hence, by finding the place- and tail-values of 2000 (with binary representation 11111010000), we may compute that $f(2000) = 10864$.

Problem 5.23. (**Taiwan 1997**) Let $n > 2$ be an integer. Suppose that $a_1, a_2, \ldots, a_n$ are positive real numbers such that $k_i = \frac{a_{i-1}+a_{i+1}}{a_i}$ is a positive integer for all i (here $a_0 = a_n$ and $a_{n+1} = a_1$). Prove that

$$2n \leq k_1 + k_2 + \ldots + k_n \leq 3n.$$

Solution. Notice that

$$k_1 + \ldots + k_n = \sum_{i=1}^{n} \left(\frac{a_i}{a_{i+1}} + \frac{a_{i+1}}{a_i} \right).$$

So the first inequality follows from AM-GM.

For the second inequality, we prove by induction on n that $k_1 + \ldots + k_n \leq 3n$ for $n \geq 3$.

The base case is $n = 3$. By symmetry, we can assume without loss of generality that a_1 is the maximum among a_1, a_2, a_3. Then $k_1 a_1 = a_2 + a_3 \leq 2a_1$ so k_1 is 1 or 2. If $k_1 = 2$, then we get $a_1 = a_2 = a_3$ hence $k_1 = k_2 = k_3 = 2$ and the inequality follows. If $k_1 = 1$ then $a_1 = a_2 + a_3$. Then from $k_2 a_2 = a_1 + a_3$ and $k_3 a_3 = a_1 + a_2$ we get $2a_2 = (k_3 - 1)a_3$ and $2a_3 = (k_2 - 1)a_2$, so by multiplying these two we get that $(k_3 - 1)(k_2 - 1) = 4$. As k_2, k_3 are positive integers, we have one of the cases $(k_2 = 2,\ k_3 = 5)$, $(k_2 = 5,\ k_3 = 2)$ or $(k_2 = 3, k_3 = 3)$. In all of them we have $k_1 + k_2 + k_3 \leq 8$.

Assume now that the result holds for $n - 1$ numbers, for some $n \geq 4$. We prove that it also holds for n numbers subject to the conditions in the hypothesis.

Let $a_j = \max\{a_i : 1 \leq i \leq n\}$. Then $k_j a_j = a_{j-1} + a_{j+1} \leq 2a_j$. So $k_j \leq 2$ and since k_j is a positive integer we have that $k_j = 1$ or $k_j = 2$.

If $k_j = 1$, then $a_j = a_{j-1} + a_{j+1}$. We want to remove the a_j from our sequence and use induction. Now $k_{j-1} a_{j-1} = a_{j-2} + a_j = a_{j-2} + a_{j-1} + a_{j+1}$, so $(k_{j-1} - 1)a_{j-1} = a_{j-2} + a_{j+1}$. Similarly we have that $(k_{j+1} - 1)a_{j+1} = a_{j-1} + a_{j+2}$.

So we by removing a_j from our sequence and replacing the numbers $k_1, \ldots, k_n$ by $k_1, \ldots, k_{j-1} - 1, k_{j+1} - 1, \ldots k_n$ we obtain a sequence of $n - 1$ numbers which satisfies the same hypotheses. Applying the inductive hypothesis we have that $k_1 + \ldots + (k_{j-1} - 1) + (k_{j+1} - 1) + \ldots + k_n \leq 3(n - 1)$, so $k_1 + \ldots + k_n \leq 3n$, as required.

If $k_j = 2$, then from $2a_j = a_{j-1} + a_{j+1}$ and the fact that a_j is the maximum of the numbers, we have that $a_j = a_{j-1} = a_{j+1}$. We recursively obtain that the sequence is constant, which implies that $k_1 = \ldots = k_n = 2$, so the inequality clearly holds. This finishes the induction step and hence the solution.

Problem 5.24. **(Zeckendorf)** Prove that any positive integer N can be represented uniquely as a sum of distinct and non-consecutive terms of the Fibonacci sequence:

$$N = \sum_{j=1}^{m} F_{i_j}, \quad i_j - i_{j-1} \geq 2.$$

Solution. We will prove the existence by induction on N. It is straightforward to check that the property holds for $N \leq F_4 = 3$.

Assume now that such a sum exists for all numbers up to F_n, $n \geq 4$. We prove that this suffices to show that it also holds for all N with $F_n < N \leq F_{n+1}$. If $N = F_{n+1}$, we are done. Otherwise, $N = F_n + (N - F_n)$ and $N - F_n < F_{n+1} - F_n = F_{n-1}$, so $N - F_n$ can be written as

$$N - F_n = F_{t_1} + \ldots + F_{t_r}, \quad t_{i+1} \leq t_i - 2, \quad t_1 \leq n - 2.$$

Therefore, we have that $N = F_n + F_{t_1} + \ldots + F_{t_r}$.

To prove the uniqueness, we use induction again. Notice that if $F_n \leq N < F_{n+1}$, then F_n is part of the sum in the representation of N. Indeed, this follows from the fact that a sum of Fibonacci terms with $i_j - i_{j-1} \geq 2, j = 1, \ldots, r-1$ and $i_1 \geq 2$ is at most

$$\begin{aligned} F_{i_{r-1}} + F_{i_{r-1}-2} + \ldots &= (F_{i_{r-1}+1} - F_{i_{r-1}-1}) + (F_{i_{r-1}-1} - F_{i_{r-1}-3}) + \ldots \\ &= F_{i_{r-1}+1} - 1. \end{aligned}$$

Therefore, if $N = F_n$, then $N = F_n$ is the unique representation of N, and if $F_n < N < F_{n+1}$ then any representation of N contains F_n and $N - F_n < F_{n-1}$. Now, using the induction hypothesis for $N - F_n$, we are done.

Problem 5.25. Let $a > 1$ be a real number which is not an integer. Prove that the sequence $(a_n)_{n \geq 0}$ defined by

$$a_n = [a^{n+1}] - a[a^n]$$

is not periodic. Here, $[x]$ denotes the integer part of x.

Solution. We assume by contradiction that the given sequence is periodic. So let T and n_0 be positive integers such that $a_{n+T} = a_n$ for any positive integer $n \geq n_0$. In particular, for $n \geq n_0$ we have

$$[a^{n+T+1}] - [a^{n+T}] \cdot a = [a^{n+1}] - [a^n] \cdot a.$$

Therefore,

$$a([a^{n+T}] - [a^n]) = [a^{n+T+1}] - [a^{n+1}]. \tag{3}$$

We now show that $[a^{n_0+T}] - [a^{n_0}] = 0$. If this does not hold then setting $n = n_0$ in (3) we obtain that a is a rational number. Further, by mathematical induction with respect to m, one can prove that

$$a^m([a^{n+T}] - [a^n]) = [a^{n+m+T}] - [a^{n+m}], \tag{4}$$

where $m, n \in \mathbb{N}$ and $n \geq n_0$.

Let $a = \dfrac{p}{q}$, where $p, q \in \mathbb{N}$, $(p, q) = 1$, $q > 1$. Then we have that $(p^m, q^m) = 1$. Hence, using (4) we have that for any positive integer m

$$q^m \mid \left([a^{n_0+T}] - [a^{n_0}]\right).$$

We have that $q^m \geq 2^m \geq m + 1$, so from the condition $q^m \mid \left([a^{n_0+T}] - [a^{n_0}]\right)$, one can deduce that $[a^{n_0+T}] = [a^{n_0}]$.

Without loss of generality, we can assume that $T > \dfrac{1}{(a-1)a^{n_0}}$ (this can be achieved by example by taking a sufficiently large integer multiple of a given period). But now

$$\begin{aligned} a^{n_0+T} - a^{n_0} &= a^{n_0}(a^T - 1) \\ &= a^{n_0}((1 + a - 1)^T - 1) \\ &\geq a^{n_0}(1 + (a-1)T - 1) \\ &= a^{n_0}(a-1)T > 1. \end{aligned}$$

This gives a contradiction. Here we have used that for $n \in \mathbb{N}$ and $h \geq -1$ we have $(1 + h)^n \geq 1 + nh$, which is Bernoulli's inequality.

Remark. Notice that in the above solution we have proved something stronger than the question asked, namely that our sequence is not even eventually periodic.

Problem 5.26. (**China 2004**) For a given real number a and a positive integer n, prove that:

i) There exists exactly one sequence of real numbers $x_0, x_1, \ldots, x_n, x_{n+1}$ such that

$$\begin{cases} x_0 = x_{n+1} = 0, \\ \frac{1}{2}(x_i + x_{i+1}) = x_i + x_i^3 - a^3, \ i = 1, 2, \ldots, n. \end{cases}$$

ii) the sequence $x_0, x_1, \ldots, x_n, x_{n+1}$ in i) satisfies $|x_i| \leq |a|$ where $i = 0, 1, \ldots, n+1$.

Solution. Notice first that since the recurrence $\frac{1}{2}(x_i + x_{i+1}) = x_i + x_i^3 - a^3$ holds for $i = 1, 2, \ldots, n$, we cannot use induction in the form $P(k) \Rightarrow P(k+1)$ since we cannot prove the base case when $k = 1$. So we have to come up with a different type of induction:

We prove both parts of the question by induction by showing that $P(k) \Rightarrow P(k-1)$.

For the first part, we are given $x_{n+1} = 0$. Having constructed x_{k+1}, for some $1 \leq k \leq n$, we have the equation

$$\frac{1}{2}(x_k + x_{k+1}) = x_k + x_k^3 - a^3,$$

which is equivalent to

$$2x_k^3 + x_k = x_{k+1} + 2a^3. \tag{1}$$

Since the value of x_{k+1} is determined, let us consider the function

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = 2x^3 + x - x_{k+1} - 2a^3.$$

Since $f'(x) = 6x^2 + 1$, this function is increasing on $\mathbb{R}$. In particular, the equation $f(x) = 0$ has a unique real solution. Hence x_k is uniquely determined by the equation (1) and we are done by induction.

For the second part, first notice that since $x_{n+1} = 0$, we clearly have $|x_{n+1}| \leq |a|$.

Assume now that $|x_k| \le |a|$, for some $2 \le k \le n+1$. Then using (1), we have

$$|x_{k-1}|(2|x_{k-1}|^2+1) = |2x_{k-1}^3 + x_{k-1}| = |x_k + 2a^3| \le |x_k| + 2|a^3| \le |a| + 2|a^3|,$$

which implies that

$$2|x_{k-1}^3| + |x_{k-1}| - 2|a^3| - |a| \le 0.$$

This can be rewritten equivalently as

$$(|x_{k-1}| - |a|)(2|x_{k-1}|^2 + 2|x_{k-1}||a| + 2|a|^2 + 1) \le 0,$$

which gives that $|x_{k-1}| \le |a|$, showing that $P(k) \Rightarrow P(k-1)$. Finally, notice that $x_0 = 0$, so $|x_k| \le |a|$, for all $k = 0, \dots, n+1$, as we wanted.

Problem 5.27. (**IMO 2010 Shortlist**) A sequence $x_1, x_2, \dots$ is defined by

$$x_1 = 1 \quad \text{and} \quad x_{2k} = -x_k, \quad x_{2k-1} = (-1)^{k+1} x_k, \quad \text{for all} \quad k \ge 1.$$

Prove that for all $n \ge 1$, we have

$$x_1 + x_2 + \ldots + x_n \ge 0.$$

Solution. Let $S_n = x_1 + x_2 + \ldots + x_n$. We also define $S_0 = 0$. We split our solution into a series of claims which we will prove by induction:

Claim 1. $|x_k| = 1$ for every $k \ge 1$.

Proof. We do induction on k. The statement is true for $k = 1$ from the hypothesis. Let $n > 1$ and suppose that the claim holds for every $0 < k < n$. If $n = 2m$, we have $0 < m < n$, then $|x_m| = 1$, and so $|x_n| = |x_m| = 1$.
If $n = 2m-1$, we have that $0 < m < n$, then $|x_m| = 1$, and so $|x_n| = |x_m| = 1$. This establishes the claim.

Claim 2. $S_{4m} = 2S_m$ for every $m \ge 0$.

Proof. Notice first that for every $m \geq 1$, from the hypothesis we have:

$$\begin{aligned} x_{4m} &= -x_{2m} = x_m; \\ x_{4m-1} &= -x_{2m} = x_m; \\ x_{4m-2} &= -x_{2m-1} = (-1)^m x_m; \\ x_{4m-3} &= x_{2m-1} = (-1)^{m+1} x_m. \end{aligned}$$

We now prove the claim by induction on m. It is certainly true for $m = 0$. Now let $m > 0$ and suppose it is true for $m - 1$. Then

$$\begin{aligned} S_{4m} &= S_{4m-4} + x_{4m-3} + x_{4m-2} + x_{4m-1} + x_{4m} \\ &= S_{4m-4} + 2x_m = 2S_{m-1} + 2x_m = S_m. \end{aligned}$$

This completes the inductive proof of our claim.
Now we prove the problem by induction on n:
A quick check shows that $S_n \geq 0$ for $n = 0, 1, 2, 3, 4$.
Let $n > 4$ and suppose that $S_k \geq 0$ for every $0 \leq k < n$. Let $m = \lceil n/4 \rceil$. This implies that $1 < m < n$.
Then $S_{m-1} \geq 0$ and $S_m \geq 0$ by induction hypothesis. We distinguish two possible cases:

Case 1. If $S_{m-1} = 0$ then 0 is the sum of $m - 1$ numbers that are ± 1 and so m is odd. Also $x_m = 1$ since $S_m \geq 0$.
Then $x_{4m-3} = 1, x_{4m-2} = -1, x_{4m-1} = 1, x_{4m} = 1$.
We also have that $S_{4m-4} = S_{m-1} = 0$.
Now $S_{4m-3}, S_{4m-2}, S_{4m-1}, S_{4m}$ are all ≥ 0.
Then $S_n \geq 0$, since $n \in \{4m - 3, 4m - 2, 4m - 1, 4m\}$.

Case 2. If $S_{m-1} > 0$, then $S_{m-1} \geq 1$, and so $S_{4m-4} \geq 2$.

$$\begin{aligned} S_{4m-3} &= S_{4m-4} + x_{4m-3} \geq 2 - 1 = 1; \\ S_{4m-2} &= S_{4m-4} + x_{4m-3} + x_{4m-3} = S_{4m-4} \geq 2; \\ S_{4m-1} &= S_{4m-2} + x_{4m-1} \geq 2 - 1 = 1; \\ S_{4m} &= S_{4m-2} + x_{4m-1} + x_{4m} \geq 2 - 1 - 1 = 0. \end{aligned}$$

Therefore, $S_{4m-3}, S_{4m-2}, S_{4m-1}, S_{4m}$ are all ≥ 0.
Hence $S_n \geq 0$, since $n \in \{4m-3, 4m-2, 4m-1, 4m\}$.
In both cases we have that $S_n \geq 0$. Then, by induction $S_n \geq 0$ for every $n \geq 0$.

Problem 5.28. Let a_n be the number of strings of length n which contain only the digits 0 and 1 and such that no two 1's can be distance two apart. Find a formula for a_n in closed form.

Solution. We begin by first finding a suitable recursion for a_n. Notice that if a string starts with 0, we have a_{n-1} ways to continue the string. If it starts with 1, it can either start with 100 or 1100. For the former, there are a_{n-3} ways to continue the strings, while for the latter we have a_{n-4} possibilities. Thus

$$a_n = a_{n-1} + a_{n-3} + a_{n-4}.$$

The first few values for a_n can be readily compute to be $a_1 = 2$, $a_2 = 4$, $a_3 = 6$ and $a_4 = 9$. Let us now study the first few terms of the sequences a_n and F_n, where $(F_n)_{n\geq 0}$ is the Fibonacci sequence:

n	F_n	a_n
1	1	$2 = 2 \cdot 1$
2	1	$4 = 2^2$
3	2	$6 = 2 \cdot 3$
4	3	$9 = 3^2$
5	4	$15 = 3 \cdot 5$

From the above table, it seems that $a_{2n} = F_{n+2}^2$ and $a_{2n+1} = F_{n+2} \cdot F_{n+3}$. We will prove this by induction on n. The base cases are given in the table. Assume now that the statement holds for all values less than $2n$. Then

$$\begin{aligned} a_{2n} &= a_{2n-1} + a_{2n-3} + a_{2n-4} \\ &= F_{n+1}F_{n+2} + F_nF_{n+1} + F_n^2 \\ &= F_{n+1}F_{n+2} + F_nF_{n+2} = F_{n+2}^2, \end{aligned}$$

and

$$\begin{aligned}a_{2n+1} &= a_{2n} + a_{2n-2} + a_{2n-3} \\ &= F_{n+2}^2 + F_{n+1}^2 + F_n F_{n+1} \\ &= F_{n+2}^2 + F_{n+1}F_{n+2} = F_{n+2}F_{n+3},\end{aligned}$$

as required.

Alternative solution. First consider the easier (and famous) problem of counting the number b_n of strings of length n such that no two 1's are adjacent. It is easy to check that $b_0 = 1$ and $b_1 = 2$. Notice that if a string starts with 0 we have b_{n-1} ways to continue the string. If a string starts with a 1, then it must start with 10, and hence we have b_{n-2} ways to continue the string. Thus we see that for $n \geq 2$ we have

$$b_n = b_{n-1} + b_{n-2}.$$

It follows by induction on n that $b_n = F_{n+2}$ is the Fibonacci sequence.

Now for the actual problem, suppose we take a string of length n with no two 1's two apart and split it into two strings, one with every odd character and the other with every even character. Then we get strings of length $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$ with no consecutive 1's. Conversely, given any two such strings we can interleave them to get a string of length n with no 1's two apart. Thus

$$a_n = b_{\lceil n/2 \rceil} b_{\lfloor n/2 \rfloor} = F_{\lceil n/2 \rceil + 2} F_{\lfloor n/2 \rfloor + 2},$$

or equivalently $a_{2n} = F_{n+2}^2$ and $a_{2n+1} = F_{n+2}F_{n+3}$.

Problem 5.29. (**IMO 2009 shortlist**) Let n be a positive integer. Given a sequence $\varepsilon_1, \ldots, \varepsilon_{n-1}$ with $\varepsilon_i = 0$ or $\varepsilon_i = 1$ for each $i = 1, \ldots, n-1$, the sequences $a_0, \ldots, a_n$ and $b_0, \ldots, b_n$ are constructed by the following rules:

$$a_0 = b_0 = 1, \quad a_1 = b_1 = 7,$$

$$a_{i+1} = \begin{cases} 2a_{i-1} + 3a_i, & \text{if } \varepsilon_i = 0, \\ 3a_{i-1} + a_i, & \text{if } \varepsilon_i = 1, \end{cases} \qquad \text{for each } i = 1, \ldots, n-1,$$

$$b_{i+1} = \begin{cases} 2b_{i-1} + 3b_i, & \text{if } \varepsilon_{n-i} = 0, \\ 3b_{i-1} + b_i, & \text{if } \varepsilon_{n-i} = 1, \end{cases} \qquad \text{for each } i = 1, \ldots, n-1.$$

Prove that $a_n = b_n$.

Solution. For a binary word $\omega = \sigma_1 \dots \sigma_n$ of length n and a letter $\sigma \in \{0,1\}$, let $\omega\sigma = \sigma_1 \dots \sigma_n \sigma$ and $\sigma\omega = \sigma\sigma_1 \dots \sigma_n$. Moreover, let $\overline{\omega} = \sigma_n \dots \sigma_1$ and let $\emptyset$ be the empty word (of length 0 and with $\overline{\emptyset} = \emptyset$. Let (u,v) be a pair of two real numbers. For binary words ω, we define recursively the numbers $(u,v)^\omega$ as follows:

$$(u,v)^\emptyset = v, \ (u,v)^0 = 2u + 3v, \ (u,v)^1 = 3u + v,$$

$$(u,v)^{\omega\sigma\varepsilon} = \begin{cases} 2(u,v)^\omega + 3(u,v)^{\omega\sigma}, & \text{if } \varepsilon = 0, \\ 3(u,v)^\omega + (u,v)^{\omega\sigma}, & \text{if } \varepsilon = 1. \end{cases}$$

It easily follows by induction on the length of ω that for all real numbers $u_1, v_1, u_2, v_2, \lambda_1$ and λ_2

$$(\lambda_1 u_1 + \lambda_2 u_2, \lambda_1 v_1 + \lambda_2 v_2)^\omega = \lambda_1 (u_1, v_1)^\omega + \lambda_2 (u_2, v_2)^\omega \tag{1}$$

and that for $\varepsilon \in \{0,1\}$

$$(u,v)^{\varepsilon\omega} = (v, (u,v)^\varepsilon)^\omega . \tag{2}$$

Obviously, for $n \geq 1$ and $\omega = \varepsilon_1 \dots \varepsilon_{n-1}$ we have $a_n = (1,7)^\omega$ and $b_n = (1,7)^{\overline{\omega}}$. Thus it suffices to prove that

$$(1,7)^\omega = (1,7)^{\overline{\omega}} \tag{3}$$

for each binary word ω. We proceed by induction on the length of ω. The assertion is obvious if ω has length 0 or 1. Now let $\omega\sigma\varepsilon$ be a binary word of length $n \geq 2$ and suppose that the assertion is true for all binary words of lengths at most $n-1$.

Note that $(2,1)^\sigma = 7 = (1,7)^\emptyset$ for $\sigma \in \{0,1\}$, $(1,7)^0 = 23$ and $(1,7)^1 = 10$.

First let $\varepsilon = 0$. Then in view of the induction hypothesis and the equalities (1) and (2), we obtain

$$\begin{aligned} (1,7)^{\omega\sigma 0} &= 2(1,7)^\omega + 3(1,7)^{\omega\sigma} \\ &= 2(1,7)^{\overline{\omega}} + 3(1,7)^{\sigma\overline{\omega}} \\ &= 2(2,1)^{\sigma\overline{\omega}} + 3(1,7)^{\sigma\overline{\omega}} \\ &= (7,23)^{\sigma\overline{\omega}} = (1,7)^{0\sigma\overline{\omega}} \end{aligned}$$

Now let $\varepsilon = 1$. Analogously, we obtain

$$\begin{aligned}(1,7)^{\omega\sigma 1} &= 3(1,7)^{\omega} + (1,7)^{\omega\sigma}\\ &= 3(1,7)^{\overline{\omega}} + (1,7)^{\sigma\overline{\omega}}\\ &= 3(2,1)^{\sigma\overline{\omega}} + (1,7)^{\sigma\overline{\omega}}\\ &= (7,10)^{\sigma\overline{\omega}} = (1,7)^{1\sigma\overline{\omega}}.\end{aligned}$$

Thus the induction step is complete, so (3) and hence also $a_n = b_n$ are proved.

Problem 5.30. (**IMO 2008 shortlist**) Let a_0, a_1, a_2, ... be a sequence of positive integers such that the greatest common divisor of any two consecutive terms is greater than the preceding term; in symbols, $\gcd(a_i, a_{i+1}) > a_{i-1}$. Prove that $a_n \geq 2^n$ for all $n \geq 0$.

Solution. Since $a_i \geq \gcd(a_i, a_{i+1}) > a_{i-1}$, the sequence is strictly increasing. In particular, $a_0 \geq 1$ and $a_1 \geq 2$. For each $i \geq 1$ we also have $a_{i+1} - a_i \geq \gcd(a_i, a_{i+1}) > a_{i-1}$ and consequently $a_{i+1} \geq a_i + a_{i-1} + 1$. Hence $a_2 \geq 4$ and $a_3 \geq 7$. The equality $a_3 = 7$ would force equalities in the previous estimates, leading to $\gcd(a_2, a_3) = \gcd(4, 7) > a_1 = 2$, which is false. Thus $a_3 \geq 8$. So we established that the result is valid for $n = 0, 1, 2, 3$. These are the base cases for our proof by induction:

Take an $n \geq 3$ and assume that $a_i \geq 2^i$ for $i = 0, 1, \ldots, n$. We must show that $a_{n+1} \geq 2^{n+1}$. Let $\gcd(a_n, a_{n+1}) = d$. We know that $d > a_{n-1}$. The induction claim is reached immediately in the following cases:

$$\begin{array}{ll} \text{if } a_{n+1} \geq 4d & \text{then } a_{n+1} > 4a_{n-1} \geq 4 \cdot 2^{n-1} = 2^{n+1};\\ \text{if } a_n \geq 3d & \text{then } a_{n+1} \geq a_n + d \geq 4d > 4a_{n-1} \geq 4 \cdot 2^{n-1} = 2^{n+1};\\ \text{if } a_n = d & \text{then } a_{n+1} \geq a_n + d = 2a_n \geq 2 \cdot 2^n = 2^{n+1}.\end{array}$$

The only remaining possibility is that $a_n = 2d$ and $a_{n+1} = 3d$, which we assume for the sequel. So $a_{n+1} = \frac{3}{2}a_n$.

Let now $\gcd(a_{n-1}, a_n) = d'$. Then $d' > a_{n-2}$. We write $a_n = md'$, for some integer m. Keeping in mind that $d' \leq a_{n-1} < d$ and $a_n = 2d$, we get that $m \geq 3$. Also $a_{n-1} < d = \frac{1}{2}md'$, $a_{n+1} = \frac{3}{2}md'$. Again, we single out the cases which imply the induction claim immediately:

$$\begin{array}{ll} \text{if } m \geq 6 & \text{then } a_{n+1} = \frac{3}{2}md' \geq 9d' > 9a_{n-2} \geq 9 \cdot 2^{n-2} > 2^{n+1}; \\ \text{if } 3 \leq m \leq 4 & \text{then } a_{n-1} < \frac{1}{2} \cdot 4d' \text{ and hence } a_{n-1} = d', \\ & \text{so } a_{n+1} = \frac{3}{2}ma_{n-1} > \frac{3}{2} \cdot 3a_{n-1} > \frac{9}{2} \cdot 2^{n-1} > 2^{n+1}. \end{array}$$

So we are left with the case $m = 5$, which means that $a_n = 5d'$, $a_{n+1} = \frac{15}{2}d'$, $a_{n-1} < d = \frac{5}{2}d'$. The last relation implies that a_{n-1} is either d' or $2d'$. Anyway, $a_{n-1} \mid 2d'$.

The same pattern repeats once more. We denote $\gcd(a_{n-2}, a_{n-1}) = d''$. Then $d'' > a_{n-3}$. Because d'' is a divisor of a_{n-1}, hence also of $2d'$, we may write $2d' = m'd''$, for some integer m'. Since $d'' \leq a_{n-2} < d'$, we get $m' \geq 3$. Also, $a_{n-2} < d' = \frac{1}{2}m'd''$, $a_{n+1} = \frac{15}{2}d' = \frac{15}{4}m'd''$. As before, we consider the cases:

$$\begin{array}{ll} \text{if } m' \geq 5 & \text{then } a_{n+1} = \frac{15}{4}m'd'' \geq \frac{75}{4} > \frac{75}{4}a_{n-3} \geq \frac{75}{4} \cdot 2^{n-3} > 2^{n+1}; \\ \text{if } 3 \leq m' \leq 4 & \text{then } a_{n-2} < \frac{1}{2} \cdot 4d'', \text{ and hence } a_{n-2} = d'', \\ & a_{n+1} = \frac{15}{4}m'a_{n-2} \geq \frac{15}{4} \cdot 3a_{n-2} \geq \frac{45}{4} \cdot 2^{n-2} > 2^{n+1}. \end{array}$$

Both of them have produced the induction claim. At this stage we have covered all the cases, so the induction is complete and the inequality $a_n \geq 2^n$ holds for all n.

6 Number Theory

Problem 6.1. Show that for any positive integer n, 3^n divides $\underbrace{11\ldots11}_{3^n\text{ones}}$.

Solution. For the base case $n = 1$ we have to prove $3|111$, which is true, as $111 = 3 \cdot 37$.

Assume now that the result holds for some $n \geq 1$. For the induction step, note that

$$\underbrace{11\ldots11}_{3^{n+1}\text{digits}} = \underbrace{11\ldots11}_{3^n\text{digits}} \cdot 10\ldots010\ldots01,$$

where we have $3^n - 1$ zeroes between every two consecutive ones in the second factor. We know that $\underbrace{11\ldots11}_{3^n\text{digits}}$ is divisible by 3^n by the induction hypothesis and $10\ldots010\ldots01$ is divisible by 3 because the sum of its digits is 3, therefore $\underbrace{11\ldots11}_{3^{n+1}\text{digits}}$ is divisible by $3^n \cdot 3 = 3^{n+1}$.

Problem 6.2. Prove that for any two positive integers a and m , the sequence

$$a, a^a, a^{a^a}, \ldots$$

is eventually constant modulo m.

Solution. We prove the result by induction on m. For $m = 1$, everything is clear. We can also suppose that $a > 1$, since the conclusion for $a = 1$ is immediate.

Assume now that the result holds for all positive integers which are less than some $m > 1$.

Assume first that a, m have a common prime factor p. Write $m = p^e k$, where $\gcd(p, k) = 1$. Since $a > 1$, the sequence of exponents

$$1, a, a^a, a^{a^a}, \ldots$$

is increasing, hence all the terms after a certain point will be divisible by p^e. By the induction hypothesis, since $k < m$, the sequence is eventually constant modulo k. It is also eventually constant 0 modulo p^e, so by the

Chinese Remainder Theorem, the sequence is also eventually constant modulo $p^e k = m$.

We are left to treat the case when a and m are coprime. We know from Euler's Theorem that $a^{\phi(m)} \equiv 1 \pmod{m}$. Notice that starting from the second term, the exponents of a in the given sequence are $1, a, a^a, a^{a^a}, \ldots$, which is our original sequence. So it suffices to look at this sequence modulo $\phi(m)$. Since $\phi(m) < m$, we are done by our induction hypothesis. This completes our proof for this case as well.

Problem 6.3. (**Poland 1998**) Let x, y be real numbers such that the numbers $x+y$, x^2+y^2, x^3+y^3, and x^4+y^4 are all integers. Prove that for all positive integers n, the number x^n+y^n is an integer.

Solution. We will prove the result by induction on $n \geq 1$. First notice that

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

so $2xy$ is an integer. But if $2xy$ is odd, the left hand side of

$$2(x+y)^4 = 2(x^4+y^4) + 4(2xy)(x^2+y^2) + 3(2xy)^2$$

is an even integer, but the right hand side is odd. Thus $2xy$ is even and xy is an integer.

Now, for $n = 1$ and $n = 2$, the conclusion holds from the hypothesis.

Assume that the result is true for all integers $n \leq k$, $k \geq 3$ and let us prove it for $n = k+1$. We have

$$x^{k+1} + y^{k+1} = (x^k + y^k)(x+y) - xy(x^{k-1} + y^{k-1}).$$

As the right hand side is an integer, the conclusion follows.

Problem 6.4. (**Ibero American 2012**) Let a, b, c, d be integers such that the number $a-b+c-d$ is odd and it divides the number $a^2-b^2+c^2-d^2$. Show that for every positive integer n, $a-b+c-d$ divides $a^n-b^n+c^n-d^n$.

Solution. We prove the question by induction on $n \geq 1$. The following observation is an immediate consequence of the modulo congruences properties and will be useful throughout the proof:

If $(x-y) \mid (z_1 - t_1)$ and $(x-y) \mid (z_2 - t_2)$, then $(x-y) \mid (z_1 z_2 - t_1 t_2)$.
Back to our original problem, notice $(a+c)-(b+d)$ divides

$$(a+c)^2 - (b+d)^2 = [(a^2+c^2) - (b^2+d^2)] + 2(ac - bd).$$

Since our assumption is that $(a+c)-(b+d)$ divides $(a^2+c^2)-(b^2+d^2)$, it follows that it also divides $2(ac-bd)$. As we are told that $(a+c)-(b+d)$ is odd, it must divide $ac - bd$.

The base cases for our induction are $n = 1$ and $n = 2$ which are both immediate from the hypotheses.

For the induction step, we let $n \geq 3$ and suppose that $(a+c)-(b+d)$ divides both $(a^{n-1}+c^{n-1}) - (b^{n-1}+d^{n-1})$ and $(a^{n-2}+c^{n-2}) - (b^{n-2}+d^{n-2})$. From the above observation, we have that

$$((a+c)-(b+d)) \mid \left((a+c)(a^{n-1}+c^{n-1}) - (b+d)(b^{n-1}+d^{n-1})\right).$$

Further,

$$\begin{aligned}&(a+c)(a^{n-1}+c^{n-1}) - (b+d)(b^{n-1}+d^{n-1})\\ &= (a^n+c^n) - (b^n+d^n) + ac(a^{n-2}+c^{n-2}) - bd(b^{n-2}+d^{n-2}).\end{aligned}$$

The induction hypothesis together with the above observation imply that $ac(a^{n-2}+c^{n-2}) - bd(b^{n-2}+d^{n-2})$ is divisible by $(a+c)-(b+d)$. Hence $(a+c)-(b+d)$ divides $(a^n+c^n)-(b^n+d^n)$, as we wanted.

Problem 6.5. (**AoPS**) Let m and n be positive integers with $\gcd(m,n) = 1$. Compute

$$\gcd(5^m + 7^m, 5^n + 7^n).$$

Solution. We claim that $\gcd(5^m+7^m, 5^n+7^n)$ is 2 when $2 \mid mn$, and 12 otherwise. We prove this by induction on $n+m$. By symmetry, we can assume throughout the proof that $m \leq n$.
The base cases are $m = 0$, $n = 1$ for which we have $\gcd(5^m+7^m, 5^n+7^n) = 2$ and $m = n = 1$, for which $\gcd(5^m+7^m, 5^n+7^n) = 12$, respectively. Thus our claim holds in these situations.

Assume now that the result holds for some positive integers m, n with $m \le n$ and $m + n \ge 2$. Notice that

$$\begin{aligned}\gcd(5^m + 7^m, 5^n + 7^n) &= \gcd(5^m + 7^m, 7^{n-m}(5^m + 7^m) - (5^n + 7^n))\\ &= \gcd(5^m + 7^m, 7^{n-m}5^m - 5^n)\\ &= \gcd(5^m + 7^m, 7^{n-m} - 5^{n-m}).\end{aligned}$$

We distinguish two cases:
If $n \ge 2m$, then

$$\begin{aligned}&\gcd(5^m + 7^m, 7^{n-m} - 5^{n-m})\\ &= \gcd(5^m + 7^m, 7^{n-2m}(5^m + 7^m) - (7^{n-m} - 5^{n-m}))\\ &= \gcd(5^m + 7^m, 7^{n-2m}5^m + 5^{n-m})\\ &= \gcd(5^m + 7^m, 7^{n-2m} + 5^{n-2m})\end{aligned}$$

If $2m \ge n$, then

$$\begin{aligned}&\gcd(5^m + 7^m, 7^{n-m} - 5^{n-m})\\ &= \gcd(5^m + 7^m, (5^m + 7^m) - 7^{2m-n}(7^{n-m} - 5^{n-m}))\\ &= \gcd(5^m + 7^m, 5^m + 5^{n-m}7^{2m-n})\\ &= \gcd(5^m + 7^m, 5^{2m-n} + 7^{2m-n})\end{aligned}$$

Since $m+n-2m < m+n$ (unless $m = 0$, a base case) and $m+2m-n < m+n$ (unless $m = n = 1$, a base case), by the induction hypothesis, $\gcd(5^m+7^m, 5^n+7^n)$ is 2 when $2 \mid mn$, and 12 otherwise. As $2 \mid mn \Leftrightarrow 2 \mid m(2m-n)$, this completes our proof.

Problem 6.6. Let n be a non-negative integer. Prove that the numbers $0, 1, 2, \ldots, n$ can be rearranged into a sequence $a_0, a_1, \ldots, a_n$ such that $a_i + i$ is a perfect square for all $0 \le i \le n$.

Solution. We will prove the statement by induction on $n \ge 0$. We think of the sequence as giving us pairings of two copies of the numbers $0, 1, \ldots, n$. We check below the first 4 base cases:

For $n = 0$ we have $0 + 0$.

For $n = 1$, we pair them $0 + 1$ and $1 + 0$.

For $n = 2$, we pair them $0 + 1, 1 + 0, 2 + 2$.

When $n = 3$, we pair them $0 + 0$, $1 + 3$, $2 + 2$, $3 + 1$.

For $n = 4$, we pair them $0 + 4$, $1 + 3$, $2 + 2$, $3 + 1$, $4 + 0$.

Assume now that the result holds for all values smaller than some $n > 4$ and we prove that it also holds for n. Since $n > 4$, there exists a unique $k \geq 2$ such that $k^2 \leq n < (k + 1)^2$. Then $0 < (k + 1)^2 - n < n$, and we can use the pairings

$$((k + 1)^2 - n) + n,\ ((k + 1)^2 - n + 1) + (n - 1),\ \ldots,\ n + ((k + 1)^2 - n).$$

By the induction hypothesis, since $((k+1)^2 - n) - 1 < n$, we have suitable pairings for the numbers $0, 1, \ldots, ((k + 1)^2 - n) - 1$, which completes our induction step.

Problem 6.7. We call a positive n triangular if it can be written as $n = \frac{a(a+1)}{2}$, for some positive integer a. Show that $11\ldots1_9$ is triangular, where the subscript denotes the fact that $11\ldots1$ is written in base 9.

Solution. We prove the result by induction on the number k of ones in the representation of $11\ldots1_9$ in base 9. The base case $k = 1$ works, since $1_9 = 1$ and $1 = \frac{1\cdot2}{2}$.

Now assume that

$$\underbrace{11\ldots1_9}_{k \text{ ones}} = \frac{t(t+1)}{2}.$$

Then

$$\begin{aligned}\underbrace{11\ldots1_9}_{k+1 \text{ ones}} &= 9 \cdot \underbrace{11\ldots1_9}_{k \text{ ones}} + 1 = \frac{9t(t+1)}{2} + 1\\ &= \frac{9t^2 + 9t + 2}{2} = \frac{(3t+1)(3t+2)}{2},\end{aligned}$$

so $\underbrace{11\ldots1_9}_{k+1 \text{ ones}}$ is a triangular number, too.

Problem 6.8. Let a, b, m be positive integers such that $\gcd(b, m) = 1$. Prove that the set $\{a^n + bn : n = 1, 2, \ldots, m^2\}$ contains a complete set of residues modulo m.

Solution. We prove the result by induction on m. We distinguish two cases:

Case 1. If $\gcd(m, ab) = 1$. The statement is clear for $m = 1$.

Assume that the result is true for all integers less than some $m \geq 2$ and write $m = p_1^{\alpha_1} \cdot \ldots \cdot p_t^{\alpha_t}$, where $p_1 < p_2 < \ldots < p_t$ are distinct prime numbers. Let $m_0 = \frac{m}{p_t}$. By the induction hypothesis, for all integers r there exists a number $k \leq m_0^2$ such that we have

$$a^k + bk = r + m_0 q_0, \quad \text{some} \quad q_0 \in \mathbb{Z}.$$

Since $\gcd(b(p_1 - 1) \cdot \ldots \cdot (p_t - 1), p_t) = 1$, there exists a non-negative integer $q < p_t$ such that

$$qb(p_1 - 1) \cdot \ldots \cdot (p_t - 1) \equiv -q_0 \pmod{p_t}.$$

Set $n = k + m_0 q(p_1 - 1) \cdot \ldots \cdot (p_t - 1)$. Since $\phi(m) \mid m_0(p_1 - 1) \cdot \ldots \cdot (p_t - 1)$, we have

$$\begin{aligned} a^n + bn &\equiv a^k + b(k + m_0 q(p_1 - 1) \cdot \ldots \cdot (p_t - 1)) \\ &\equiv r + m_0(q_0 + qb(p_1 - 1) \cdot \ldots \cdot (p_t - 1)) \\ &\equiv r \pmod{m}. \end{aligned}$$

Finally, notice that

$$\begin{aligned} 0 < k \leq n &\leq m_0^2 + m_0(p_1 - 1) \cdot \ldots \cdot (p_t - 1) \cdot (p_t - 1) \\ &< m_0^2 + m_0(p_t - 1)m = m_0^2(1 + p_t^2 - p_t) < m_0^2 p_t^2 = m^2. \end{aligned}$$

Case 2. If $\gcd(m, a) > 1$, set $m = uv$, where $u > 1$, such that a is divisible by all prime divisors of u and $\gcd(v, a) = 1$. Let r be an arbitrary integer. Then there exists $s \in \{0, 1, \ldots, u - 1\}$ such that

$$bs \equiv r \pmod{u}.$$

We also know that $\gcd(a^u, v) = 1$, hence $\gcd\left((a^{-1})^s bu, v\right) = 1$. Using the previous case, there exists $k \le v^2$ such that

$$(a^u)^k + (a^{-1})^s buk \equiv (a^{-1})^s(r - bs) \pmod{v}.$$

Now set $n = uk + s$. Then $n \le uv^2 + u - 1 < u(v^2 + 1) \le u^2v^2 = m^2$ and

$$a^{ku+s} + buk \equiv r - bs \pmod{v}.$$

This implies that $a^n + bn \equiv r \pmod{v}$. Finally, for any p dividing a we know that $\nu_p(u) < u$, since $p^u \ge 2^u \ge 1 + u$, hence

$$a^n + bn = a^{uk+s} + b(uk + s) \equiv 0 + bs \equiv r \pmod{u},$$

and our proof is complete.

Problem 6.9. (**Kömal**) Let a and n be two positive integers such that $a^n - 1$ is divisible by n. Prove that the numbers $a + 1$, $a^2 + 2$, $\ldots$, $a^n + n$ are distinct modulo n.

Solution. We will proceed by induction on n. The base case $n = 1$ is clear.

Assume now that the result holds for all integers less than some $n \ge 2$. Let k be the order of a modulo n, that is, the least value such that $a^k \equiv 1 \pmod{n}$. We know by Euler that $k < n$ and from hypothesis we have that $k \mid n$. Since $k < n$, by the induction hypothesis we have that $a + 1, a^2 + 2, \ldots, a^k + k$ are distinct modulo k. Now let us prove that for $1 \le x, y \le n$ we have $a^x + x \not\equiv a^y + y \pmod{n}$. Write $x = kz + t$ and $y = ku + v$, where $1 \le t, v \le k$ and $0 \le z, u < \frac{n}{k}$. Then $a^x \equiv a^t \pmod{n}$ and $a^y \equiv a^v \pmod{n}$. We distinguish two cases:

Case 1. If $t \ne v$, then $a^x + x \equiv a^t + t$ and from the induction hypothesis we have $a^t + t \not\equiv a^v + v \equiv a^y + y \pmod{k}$.

Case 2. If $t = v$, then $z \ne u$, hence

$$\begin{aligned} a^x + x &\equiv a^t + kz + t = a^v + ku + v + k(z - u) \\ &\equiv a^y + y + k(z - u) \not\equiv a^y + y \pmod{n}. \end{aligned}$$

This completes our proof.

Problem 6.10. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = ax^2 + bx + c$, where a, b, c are positive integers. Let n be some given positive integer. Prove that for any positive integer m, there exist n consecutive integers $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that each of the numbers $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)$ has at least m distinct prime factors.

Solution. We prove the result by induction on m. For $m = 1$, the statement clearly holds.

Assume now that the result holds for some $m \geq 1$ and let $\alpha_1, \ldots, \alpha_n$ be such that each of the numbers $f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)$ has at least m distinct prime factors. Take

$$A = f(\alpha_1)^2 \cdot f(\alpha_2)^2 \cdot \ldots \cdot f(\alpha_n)^2.$$

For each $j = 1, \ldots, n$, let $\beta_j = A + \alpha_j$. Notice that the integers $\beta_1, \ldots, \beta_n$ are also consecutive and

$$f(\beta_j) = f(\alpha_j) \cdot (1 + C_j f(\alpha_j)),$$

for some integer C_j. Hence the term $1 + C_j f(\alpha_j)$ is coprime to $f(\alpha_j)$, so $f(\beta_j)$ has at least one prime factor more than $f(\alpha_j)$. This completes our proof.

Problem 6.11. (**Iran 2005**) Find all $f : \mathbb{N}^* \to \mathbb{N}^*$ such that for every $m, n \in \mathbb{N}^*$,

$$f(m) + f(n) \mid m + n.$$

Solution. We will prove by strong induction that $f(n) = n$. Letting $m = n = 1$, we obtain that $2f(1) \mid 2$, hence $f(1) \mid 1 \Rightarrow f(1) = 1$. This establishes the base case.

Assume now that the result holds for all positive integers which are less than some $n > 1$. By Chebyshev's theorem, we know that there exists a prime number between n and $2n$, so there exists $m < n$ such that $m+n = p$ is prime.

Since $f(m) + f(n)$ divides p, we have that $f(m) + f(n)$ is either 1 or p. But $f(m) + f(n) \geq 1 + 1 = 2$, so it cannot be 1. Therefore $f(m) + f(n) = p$. Since $f(m) = m$ from the induction hypothesis, we obtain $f(n) = p - m = n$. This completes our proof.

Problem 6.12. (**Kvant M2252**) Prove that

$$1 + 3^3 + 5^5 + \ldots + (2^n - 1)^{2^n - 1} \equiv 2^n \pmod{2^{n+1}}, \quad \text{for} \quad n \geq 2.$$

Solution. We will use in our proof the following two facts

i) If k is odd, then $k^{2^n} \equiv 1 \pmod{2^{n+2}}$;

ii) $(k + 2^n)^k \equiv k^k(1 + 2^n) \pmod{2^{n+2}}$.

i) can be proved either using the identity

$$k^{2^n} - 1 = (k-1)(k+1)(k^2+1)\ldots(k^{2^{n-1}} + 1),$$

or by induction on n. Fact ii) is an easy consequence of the binomial theorem.

Back to our original problem, let us define

$$S_n = 1 + 3^3 + 5^5 + \ldots + (2^n - 1)^{2^n - 1}.$$

Then $S_{n+1} = S_n + R_n$, where $R_n = (2^n + 1)^{2^n + 1} + \ldots + (2^{n+1} - 1)^{2^{n+1} - 1}$. Notice that all of the $2^n - 1$ terms in R_n are of the form $m = 2^n + k$, with $k < 2^n$. Therefore, we can write

$$m^m \equiv m^{2^n} \cdot m^k \equiv m^k \equiv k^k(1 + 2^n) \pmod{2^{n+2}},$$

using facts i) and ii) presented above. This implies that

$$R_n \equiv (1 + 2^n)\left(1 + 3^3 + 5^5 + \ldots + (2^n - 1)^{2^n - 1}\right) \equiv (1 + 2^n)S_n \pmod{2^{n+2}}.$$

Thus, $S_{n+1} \equiv 2S_n(1 + 2^{n-1}) \pmod{2^{n+2}}$. We now use induction on n. For $n = 2$, we see that indeed $S_2 = 28 \equiv 4 \pmod 8$.

Assume now that the result holds for some $n \geq 2$. Then we have $S_n = 2^{n+1}k + 2^n$, for some $k \in \mathbb{Z}$. By what we have established above, we obtain that

$$S_{n+1} \equiv (2^{n+2}k + 2^{n+1})(1 + 2^{n-1}) \equiv 2^{n+1} \pmod{2^{n+2}},$$

which completes our proof.

Problem 6.13. (**GMA 2013**) Prove that for any positive integer n and any prime p, the sum

$$S_n = \sum_{k=0}^{\lfloor \frac{n}{p} \rfloor} (-1)^k \binom{n}{kp}$$

is divisible by $p^{\lfloor \frac{n-1}{p-1} \rfloor}$.

Solution. We begin by proving the following

Lemma. For all $n \geq p$ we have

$$S_n - \binom{p}{1} S_{n-1} + \binom{p}{2} S_{n-2} - \ldots + \binom{p}{p-1} S_{n-p+1} = 0.$$

Proof. Let ζ be a primitive p-th root of unity. We know that

$$\sum_{j=0}^{p-1} \zeta^i = \begin{cases} 0 & \text{if } p \nmid i \\ p & \text{if } p \mid i \end{cases}.$$

Using this identity, one sees that

$$S_n = \frac{1}{p} \sum_{i=0}^{p-1} (1 - \zeta^i)^n.$$

Since ζ^i, $i = 1, 2, \ldots, p-1$ are roots of the polynomial $\frac{x^p - 1}{x-1}$, we have that $1 - \zeta^i$ are roots of the polynomial

$$\frac{1 - (1-x)^p}{x} = x^{p-1} - \binom{p}{1} x^{p-2} + \ldots + \binom{p}{p-1}.$$

By setting $x = 1 - \zeta^i$, $i = 1, \ldots, p-1$ and adding up, we see that for every $n \geq p$, one has

$$S_n - \binom{p}{1} S_{n-1} + \binom{p}{2} S_{n-2} + \ldots + \binom{p}{p-1} S_{n-p+1} = 0,$$

proving the lemma.

Back to the original problem, we can now prove the result by induction on n. Things are clear if $n = 1, 2, \ldots, p-1$.

Assume now that $n \geq p$ and that the result holds for all positive integers less than n. We also know that for all $1 \leq j \leq p-1$, we have $p \mid \binom{p}{j}$. By the induction hypothesis, we know that for $1 \leq j \leq p-1$, $p^{\lfloor \frac{n-j-1}{p-1} \rfloor}$ divides S_{n-j}. Furthermore, for all $1 \leq j \leq p-1$, we have

$$\left\lfloor \frac{n-j-1}{p-1} \right\rfloor + 1 = \left\lfloor \frac{n+p-j-2}{p-1} \right\rfloor \geq \left\lfloor \frac{n-1}{p-1} \right\rfloor .$$

Therefore, all the terms of the sum

$$\binom{p}{1} S_{n-1} - \binom{p}{2} S_{n-2} + \ldots - \binom{p}{p-1} S_{n-p+1}$$

are divisible by $p^{\lfloor \frac{n-1}{p-1} \rfloor}$, hence the same is true for S_n by our lemma. This completes our proof.

Problem 6.14. (**Bulgaria 1996**) Let $k \geq 3$ be an integer. Show that there exist odd positive integers x and y with $2^k = 7x^2 + y^2$.

Solution. We prove the result by induction, showing that for each $k \geq 3$ there are odd positive integers x_k, y_k such that

$$2^k = 7x_k^2 + y_k^2.$$

For $k = 3$, choose $x_3 = y_3 = 1$. For the inductive step, we use the identity

$$2(7a^2 + b^2) = 7\left(\frac{a \pm b}{2}\right)^2 + \left(\frac{7a \mp b}{2}\right)^2 .$$

If $2^k = 7x_k^2 + y_k^2$ with x_k, y_k odd, then either $(x_k + y_k)/2$ or $|x_k - y_k|/2$ is odd, as their sum is x_k or y_k, both of which are odd. If $(x_k + y_k)/2$ is odd, then take $x_{k+1} = (x_k + y_k)/2$ and $y_{k+1} = |7x_k - y_k|/2$.

If $(x_k + y_k)/2$ is even, then $|x_k - y_k|/2$ is odd and define $x_{k+1} = |x_k - y_k|/2$ and $y_{k+1} = (7x_k + y_k)/2$. This completes the induction step and therefore the proof.

Problem 6.15. (**USAMO 1998**) Prove that for each $n \geq 2$, there is a set S of n integers such that $(a-b)^2$ divides ab for every distinct $a, b \in S$.

Solution. We will prove by induction on n, that we can find such a set S_n consisting of non-negative integers.

The base case is $n = 2$ for which we take $S_n = \{0, 1\}$.

Now suppose that for some $n \geq 2$, the desired set S_n of n non-negative integers exists. Let L be the least common multiple of those numbers $(a-b)^2$ and ab which are non-zero, with (a, b) ranging over all pairs of distinct elements from S_n. We define:

$$S_{n+1} = \{L + a : a \in S_n\} \cup \{0\}.$$

Notice that S_{n+1} consists of $n + 1$ non-negative integers, since $L > 0$. Moreover, if $i, j \in S_{n+1}$ and either i or j is zero, then $(i-j)^2$ divides ij.

Finally, notice that if $L + a, L + b \in S_{n+1}$ with a, b, distinct elements of S_n, then

$$(L+a)(L+b) \equiv ab \equiv 0 \pmod{(a-b)^2},$$

so $((L+a) - (L+b))^2$ divides $(L+a)(L+b)$, thus completing our induction step.

Problem 6.16. (**Brazil 2011**) Prove that there exist positive integers $a_1 < a_2 < \ldots < a_{2011}$ such that for all $1 \leq i < j \leq 2011$ we have $\gcd(a_i, a_j) = a_j - a_i$.

Solution. First notice that the condition $\gcd(a_i, a_j) = a_j - a_i$ is equivalent to $a_j - a_i \mid a_i$. Indeed, if $\gcd(a_i, a_j) = a_j - a_i$, then $a_j - a_i \mid a_i$; conversely, if $a_j - a_i \mid a_i$, then $a_j - a_i \mid a_j$, so $a_j - a_i \mid \gcd(a_i, a_j)$ and from Euclid we know that $\gcd(a_i, a_j) \leq a_j - a_i$, so we are done.

We will now construct a sequence $a_1 < a_2 < \ldots < a_n$ with the required properties by induction on n. For the base case $n = 2$ we can take for example $a_1 = 2$ and $a_2 = 4$.

Assume now that we have constructed $n-1$ numbers $a_1 < a_2 < \ldots < a_{n-1}$ which satisfy the required properties. Consider the n numbers

$$a_0 < a_0 + a_1 < a_0 + a_2 < \ldots < a_0 + a_{n-1},$$

where a_0 will be chosen later. By the observation we made at the beginning, the condition $\gcd(a_i + a_0, a_j + a_0) = (a_j + a_0) - (a_i + a_0)$ is equivalent to $a_j - a_i \mid a_0 + a_i$, and we already have $a_j - a_i \mid a_i$, so this is equivalent to $a_j - a_i \mid a_0$. Also the condition $\gcd(a_0, a_i + a_0) = a_i$ is equivalent to $a_i \mid a_0$. Therefore, taking a_0 to be the least common multiple of all the numbers $a_1, \ldots, a_{n-1}$ and the numbers $a_j - a_i$, $1 \le i < j \le n-1$, we obtain the construction for n. This completes our proof.

Problem 6.17. (**Bulgaria TST**) Let $a, m \ge 2$ and let $\mathrm{ord}_m^a = k$ (i.e. $a^k \equiv 1 \pmod m$ and $a^s \not\equiv 1 \pmod m$ for any $0 < s < k$). Prove that if t is an odd number such that every prime that divides t also divides m and $\gcd\left(t, \frac{a^k-1}{m}\right) = 1$, then $\mathrm{ord}_{mt}^a = kt$.

Solution. We prove the statement by induction on the number of prime factors of t, taking into account also their multiplicity. We start with the case when t is a prime. We then have $t \mid m$. Let $d = \frac{a^k-1}{m}$, so that we have $a^k = 1 + md$. Then

$$a^{kt} = (1+md)^t \equiv 1 \pmod{mt},$$

thus $s = \mathrm{ord}_{mt}^a \mid kt$. We know that $mt \mid a^s - 1$, so $m \mid a^s - 1$, and since $k = \mathrm{ord}_m^a$ we must have $k \mid s \mid kt$, thus $s = k$ or $s = kt$. If $s = k$, then $1 + md = a^k \equiv 1 \pmod{mt}$, which implies that $t \mid d$, which is a contradiction. Hence $s = kt$.

Now assume that t has at least two prime factors (not necessarily distinct) and write $t = rt_0$, where r is prime and $t_0 > 1$. Since r is prime, by the base case we proved above we have that $\mathrm{ord}_{mr}^a = kr$.

We first prove that $\gcd\left(t_0, \frac{a^{kr}-1}{mr}\right) = 1$. If a prime r_0 divides t_0, then it also divides t, so $r_0 \mid m \mid mr$. Now

$$\begin{aligned} d_0 = \frac{a^{kr}-1}{mr} &= \frac{a^k-1}{m} \cdot \frac{a^{k(r-1)} + a^{k(r-2)} + \ldots + 1}{r} \\ &= d \cdot \frac{a^{k(r-1)} + a^{k(r-2)} + \ldots + 1}{r}. \end{aligned}$$

Since $a^k \equiv 1 \pmod m$, we have $a^k \equiv 1 \pmod r$, so $\frac{a^{k(r-1)}+a^{k(r-2)}+\ldots+1}{r}$ is an integer.

If $r \neq r_0$, then $a^k \equiv 1 \pmod{r_0}$, so we also have

$$\frac{a^{k(r-1)} + a^{k(r-2)} + \ldots + 1}{r} \equiv 1 \pmod{r_0}.$$

Thus $\gcd\left(r_0, \frac{a^{kr}-1}{mr}\right) = 1$.
If $r = r_0$, then we can set $a^k \equiv 1 + br \pmod{r^2}$ for some integer b. Therefore, for all $j = 0, 1, \ldots, r-1$, by using the binomial theorem we find that $a^{kj} \equiv 1 + jbr \pmod{r^2}$, hence

$$a^{k(r-1)} + a^{k(r-2)} + \ldots + 1 \equiv r + br(1 + 2 + \ldots + r - 1) \equiv r + \frac{b}{2} r^2 (r-1) \pmod{r^2}.$$

Since t is odd, so must be r, so we obtain that

$$\frac{a^{k(r-1)} + a^{k(r-2)} + \ldots + 1}{r} \equiv 1 \pmod{r}$$

and again $\gcd\left(r_0, \frac{a^{kr}-1}{mr}\right) = 1$.
Combining these two cases we see that $\gcd\left(t_0, \frac{a^{kr}-1}{mr}\right) = 1$ and hence we can apply the induction hypothesis, completing our proof.

Problem 6.18. (**China TST**) Prove that for all positive integers m, n there exists an integer k such that $2^k - m$ has at least n distinct prime divisors.

Solution. We prove the result by induction on n, the base case $n = 1$ being clear.

Assume now that there is some integer k_n such that $A_n = 2^{k_n} - m$ has at least n distinct prime divisors. Without loss of generality, assume that A_n is odd (we can take out the power of 2 dividing m and work with the rest, so we can assume m is odd). By Euler's theorem we then have $2^{\phi(A_n^2)} \equiv 1 \pmod{A_n^2}$, so

$$2^{k_n + \phi(A_n^2)} - m \equiv A_n \pmod{A_n^2}.$$

We deduce that $\frac{2^{k_n+\phi(A_n^2)}-m}{A_n} \equiv 1 \pmod{A_n}$. Therefore, there exists a prime $p \nmid A_n$ such that $p \mid \frac{2^{k_n+\phi(A_n^2)}-m}{A_n}$. Taking $k_{n+1} = k_n + \phi(A_n^2)$, we complete the induction step and hence the proof.

Problem 6.19. (**Serbia**) Prove that for all positive integers m there exists a positive integer $k \geq 2$ such that $3^k - 2^k - k$ is divisible by m.

Solution. We define a sequence $(x_n)_{n\geq 0}$ by $x_0 = 2$ and $x_{n+1} = 3^{x_n} - 2^{x_n}$. We prove that for all positive integers d, there exists some n for which $x_{n+1} \equiv x_n \pmod{d}$ (notice that this implies the statement of the problem by taking $d = m$ and $k = x_n$). We proceed by induction on d. The base case $d = 1$ is clear.

Assume now that the result holds for all integers less than some $d \geq 2$. Then $\phi(d) < d$ and since the statement holds for $\phi(d)$, there is some n such that $x_{n+1} \equiv x_n \pmod{\phi(d)}$. Then by Euler's theorem we have

$$3^{x_{n+1}} \equiv 3^{x_n} \pmod{d} \quad \text{and} \quad 2^{x_{n+1}} \equiv 2^{x_n} \pmod{d},$$

proving our result.

Problem 6.20. Let k be a positive integer. Prove that for all non-negative integers m there exists a positive integer n with at least m prime factors (not necessarily distinct) such that $2^{kn^2} + 3^{kn^2}$ is divisible by n^3.

Solution. We begin by proving the following:

Lemma. If a and b are positive integers with $\gcd(a, b) = 1$ and p an odd prime such that $\nu_p(a + b) = s \geq 1$, then $\nu_p(a^p + b^p) = s + 1$.

Proof. Let $a + b = x$ be an integer divisible by p. Then

$$\begin{aligned}\frac{a^p + b^p}{a + b} &= \frac{(x - b)^p + b^p}{x} \\ &= x^{p-1} - bpx^{p-2} + \ldots - \binom{p}{2} b^{p-2}x + pb^{p-1} \\ &\equiv pb^{p-1} \pmod{p^2}.\end{aligned}$$

As a and b are coprime, it follows that $\nu_p\left(\frac{a^p+b^p}{a+b}\right) = 1$. As

$$\nu_p(a^p + b^p) = \nu_p(a + b) + \nu_p\left(\frac{a^p + b^p}{a + b}\right),$$

we obtain the result.

We will now prove the statement of the original question by induction on $m \geq 0$. For $m = 0$, we can take $n = 1$.

Assume now that $m \geq 1$ and we have found an integer n_{m-1} which has at least $m-1$ prime factors such that n_{m-1}^3 divides $2^{kn_{m-1}^2} + 3^{kn_{m-1}^2}$. We distinguish two cases:

Case 1. If there exists a prime p such that $p \nmid n_{m-1}$, but $p \mid 2^{kn_{m-1}^2} + 3^{kn_{m-1}^2}$. Since $p \mid 2^{kn_{m-1}^2} + 3^{kn_{m-1}^2}$, by the above lemma we have that $p^3 \mid 2^{kp^2 n_{m-1}^2} + 3^{kp^2 n_{m-1}^2}$. So we can take $n_m = pn_{m-1}$ and we are done.

Case 2. If n_{m-1} and $2^{kn_{m-1}^2} + 3^{kn_{m-1}^2}$ have the same prime factors, then we use the fact that $n_{m-1} \neq 2^{kn_{m-1}^2} + 3^{kn_{m-1}^2}$ (since $2^{kn_{m-1}^2} + 3^{kn_{m-1}^2} > 3^{n_{m-1}^2} > n_{m-1}^3$. Therefore, there exists a prime q such that $\nu_q(n_{m-1}^3) = \alpha$ and $\nu_q\left(2^{kn_{m-1}^2} + 3^{kn_{m-1}^2}\right) = \beta \geq \alpha + 1$. By taking $n_m = qn_{m-1}$, we have $q^{\alpha+3} \mid q^{\beta+2} = \nu_q\left(2^{kn_m^2} + 3^{kn_m^2}\right)$. This completes our proof.

Problem 6.21. Prove that for all positive integers k there exists an integer n which has exactly k prime divisors and $n^3 \mid (2^{n^2} + 1)$.

Solution. We first prove two important lemmas:

Lemma 1. Let a be a positive integer and p an odd prime. Then the following statements are equivalent:

a) $\nu_p(a+1) = s \geq 1$;

b) $\nu_p(a^p + 1) = s + 1$.

Proof. If $\nu_p(a+1) = s \geq 1$, then writing $a^p = (a+1-1)^p$ and using the binomial theorem we obtain that $\frac{a^p+1}{a+1} \equiv p \pmod{p^2}$. Since

$$\nu_p(a^p + 1) = \nu_p(a+1) + \nu_p\left(\frac{a^p+1}{a+1}\right),$$

we are done. Conversely, if $p \mid a^p + 1$, using Fermat's Little Theorem we find that $a+1$ is divisible by p and by the identity we used above, we are done.

Lemma 2. Let a be a positive integer. Then there exists a prime q such that $q \mid \frac{a^p+1}{a+1}$, but $q \nmid (a+1)$, except for the cases $a = 2$ and $p = 3$.

Proof. Assume the contrary. Then any prime q dividing $\frac{a^p+1}{a+1}$ also divides $a+1$. Notice that $\gcd\left(\frac{a^p+1}{a+1}, a+1\right) = 1$ or p. So we must have $q = p$ and hence that both $\frac{a^p+1}{a+1}$ and $a+1$ are powers of p. We know from the previous lemma that if $p \mid (a+1)$, then $\nu_p\left(\frac{a^p+1}{a+1}\right) = 1$. So we must have that $\frac{a^p+1}{a+1} = p$, while

$$\frac{a^p+1}{a+1} > a^2 - a + 1 > a + 1 \geq p,$$

except for the case $a = 2$ and $p = 3$. This completes the proof of the lemma.

Back to our original problem, we will prove the result by induction on k. We know from Lemma 1 that if $p \mid 2^m + 1$, then $p^3 \mid 2^{mp^2} + 1$. For $k = 1$ we can thus take $n_1 = p_1 = 3$. Since $2^9 + 1 = 513 = 27 \cdot 19$, for $k = 2$ we can take $n_2 = p_1 p_2$, where $p_2 = 19$.

Assume now that for some $k \geq 2$ we have constructed the corresponding $n_k = p_1 p_2 \dots p_k$ such that $n_k \mid 2^{n_{k-1}} + 1$ and $n_k^3 \mid 2^{n_k^2} + 1$. By Lemma 2, there exists a prime p_{k+1} such that $p_{k+1} \nmid 2^{p_1 p_2 \cdots p_{k-1}} + 1 = a + 1$ (and hence p_{k+1} is not one of $p_1, \dots, p_k$), but $p_{k+1} \mid 2^{p_1 p_2 \cdots p_k} + 1 = a^{p_k} + 1$. Now set $n_{k+1} = p_{k+1} n_k$. Since $n_k^3 \mid 2^{n_k^2} + 1 \mid 2^{n_k^2 p_{k+1}^2} + 1$ and by Lemma 1 we have $p_{k+1}^3 \mid 2^{n_k^2 p_{k+1}^2} + 1$, we deduce that $n_k^3 p_{k+1}^3 = n_{k+1}^3 \mid 2^{n_k^2 p_{k+1}^2} + 1 = 2^{n_{k+1}^2} + 1$. This completes our proof.

Problem 6.22. (Polish Training Camp) Let k be a positive integer. The sequence $(a_n)_{n \geq 1}$ is given by

$$\sum_{d \mid n} d a_d = k^n, \quad \text{for all} \quad n \geq 1.$$

Prove that every term of the sequence is an integer.

Solution. We shall prove the result by induction on n. The base case $n = 1$ is clear.

Assume now that $n \geq 2$ and that the result is true for all integers less than n. From the hypothesis we have that

$$n a_n + \sum_{d \mid n,\, d < n} d a_d = k^n.$$

We shall prove that $n \mid \left(k^n - \sum\limits_{d|n,\, d<n} da_d\right)$. Let p be a prime dividing n and let $n = p^r x$, for some x coprime to p. From the induction hypothesis we know that

$$\begin{aligned} k^n - \sum_{d|n,\, d<n} da_d &\equiv k^n - \sum_{d|p^{r-1}x} da_d \\ &= k^{p^r x} - k^{p^{r-1}x} \\ &= k^{p^{r-1}x}\left(k^{p^{r-1}(p-1)x} - 1\right) \pmod{p^r}. \end{aligned}$$

If $p \mid k$, then we are done. Otherwise, by Euler's Theorem, we have that

$$p^r \mid \left(k^{p^{r-1}(p-1)x} - 1\right).$$

In either case, we obtain that p^r divides $\left(k^n - \sum\limits_{d|n,\, d<n} da_d\right)$. As this holds for all $p \mid n$, we obtain the desired conclusion.

Problem 6.23. Prove that there is a permutation $(a_1, a_2, \ldots, a_n)$ of $(1, 2, 3, \ldots, n)$ such that none of the numbers $a_1, a_1+a_2, \ldots, a_1+a_2+\ldots+a_{n-1}$ is a perfect square.

Solution. We prove the statement by strong induction on n. For the base cases $n = 1, 2, 3$ we can take the permutations (1), $(2, 1)$, and $(3, 2, 1)$, respectively.

Now assume we have proved the assertion for $n < k$, some $k \geq 4$. If $1 + 2 + \ldots + (k-1)$ is not a perfect square, then we can take the permutation to be $(a_1, a_2, \ldots, a_{k-1}, k)$, where $(a_1, a_2, \ldots, a_{k-1})$ is the suitable permutation of $(1, 2, \ldots, k-1)$ from the induction hypothesis. If $1 + 2 + \ldots + (k-1)$ is a perfect square, then $1 + 2 + \ldots + (k-2)$ and $1 + 2 + \ldots + (k-2) + k$ are not, so we can take the permutation to be $(a_1, a_2, \ldots, a_{k-2}, k, k-1)$, where $(a_1, a_2, \ldots, a_{k-2})$ is a suitable permutation of $(1, 2, \ldots, k-2)$.

Problem 6.24. Prove that any integer can be represented in infinitely many ways as $\pm 1^2 \pm 2^2 \pm 3^2 \pm \ldots \pm t^2$, for a convenient t and a suitable choice of the signs $+$ and $-$.

Solution. We divide our proof into two steps. We first show that each number can be written in the given form in at least one way. Then we prove that if a number can be written in at least one way in the given form, it can be written in infinitely many ways. The key ingredient for our proof is the identity

$$(t+1)^2 - (t+2)^2 - (t+3)^2 + (t+4)^2 = 4.$$

This shows that if n can be represented as a sum of t terms, then $n+4$ can be written as a sum of $t+4$ terms. This will be our induction step (from n to $n+4$) for the first part of the proof. We are left to check the base cases $n=0$, $n=1$, $n=2$, $n=3$. For these we have

$$\begin{aligned}
&0 = 1^2 + 2^2 - 3^2 + 4^2 - 5^2 - 6^2 + 7^2;\\
&1 = 1^2;\\
&2 = -1^2 - 2^2 - 3^2 + 4^2;\\
&3 = -1^2 + 2^2.
\end{aligned}$$

This completes our induction, establishing the first part of the proof.
For the second part of the proof, notice that from the identity

$$(t+1)^2 - (t+2)^2 - (t+3)^2 + (t+4)^2 = 4,$$

we immediately deduce the identity

$$(t+1)^2-(t+2)^2-(t+3)^2+(t+4)^2-(t+5)^2+(t+6)^2+(t+7)^2-(t+8)^2 = 0.$$

Therefore, if n can be represented as a sum of t squares, then n can be represented also as a sum of $t+8$ squares. This completes the second part of our proof, establishing the result.

Problem 6.25. (**Romania TST 2013**) Find all positive integers n that can be written as

$$n = \frac{(a_1^2 + a_1 - 1)(a_2^2 + a_2 - 1)\cdots(a_k^2 + a_k - 1)}{(b_1^2 + b_1 - 1)(b_2^2 + b_2 - 1)\cdots(b_k^2 + b_k - 1)},$$

for some positive integers $a_i, b_i \in \mathbb{N}^*$ and some $k \in \mathbb{N}^*$.

Solution. First, notice that $a^2 = a - 1$ is always odd, hence any such n must be odd. Next, notice that for any odd prime $p \neq 5$ with $p \mid (a^2 + a - 1)$ we have $p \mid ((2a+1)^2 - 5)$, so 5 is a quadratic residue mod p, hence $p \equiv 1, 4 \pmod{5}$.

We now claim that all odd integers with all prime factors to 0, 1 or 4 modulo 5 can be written in this manner. We call such a number *good.*

We can prove this by induction. The base cases are covered by $1 = \frac{1^2+1-1}{1^2+1-1}$ and $5 = \frac{2^2+2-1}{1_1^2-1}$. Suppose we have written all *good* numbers strictly less than $k > 5$, which we assumed to be *good.* If k is not prime, then it is a product of two *good* numbers and these can be written in the given form, so their product is also *good.* Otherwise, if k is prime and *good*, then we can find a $0 < a \leq k - 1$ such that $k \mid a^2 - 5$. Since we also have $k \mid (k-a)^2 - 5$, after possibly replacing a by $k - a$, we may assume $a = 2b + 1$ is odd. Thus $k \mid b^2 + b - 1$. Since $\frac{b^2+b-1}{k} < k$ and obviously it is a good number, we know that it can be written in the required form and hence so can its reciprocal $\frac{k}{b^2+b-1}$. Thus writing $k = \frac{k}{b^2+b-1} \cdot \frac{1^2+1-1}{b^2+b-1}$ finishes this case, hence the induction step and the proof.

Problem 6.26. (**USA TST 2006**) Let n be a positive integer. Find, with proof, the least positive integer d_n which cannot be expressed in the form

$$\sum_{i=1}^{n} (-1)^{a_i} 2^{b_i},$$

where a_i and b_i are non-negative integers for each i.

Solution. The answer is $d_n = 2\frac{4^n-1}{3} + 1$. We first show that d_n cannot be obtained. For any p, let $t(p)$ be the minimum number n of terms required to express p in the desired form and call any realization of this minimum a *minimal representation* of p. If p is even, any sequence of b_i that can produce p must contain an even number of zeros. If this number is non-zero, then cancelling one against another or replacing two with a $b_i = 1$ term would reduce the number of terms in the sum. Thus a minimal representation cannot contain a $b_i = 0$ term, and by dividing each term by two we see that $t(2m) = t(m)$. If p is odd, there must be at least one $b_i = 0$ and removing it gives a sequence that produces either $p - 1$ or $p + 1$. Therefore, we obtain

$$t(2m-1) = 1 + \min\{t(2m-2), t(2m)\} = 1 + \min\{t(m-1), t(m)\}.$$

With d_n as defined above and $c_n = \frac{2^{2n}-1}{3}$, we have $d_0 = c_1 = 1$, so $t(d_0) = t(c_1) = 1$ and

$$t(d_n) = 1 + \min\{t(d_{n-1}, t(c_n))\} \quad \text{and} \quad t(c_n) = 1 + \min\{t(d_{n-1}), t(c_{n-1})\}.$$

Therefore, by induction, $t(c_n) = n$ and $t(d_n) = n+1$ and d_n cannot be obtained by a sum with n terms.

Next we show by induction on n that any positive integer less than d_n can be obtained with n terms. By the inductive hypothesis and symmetry about zero, it suffices to show that by adding one summand we can reach every p in the range $d_{n-1} \le p < d_n$ from an integer q in the range $-d_{n-1} < q < d_{n-1}$. Suppose that $c_n + 1 \le p \le d_n - 1$. By using a term 2^{2n-1}, we see that $t(p) \le 1 + t(|p - 2^{2n-1}|)$. Since

$$d_n - 1 - 2^{2n-1} = 2^{2n-1} - (c_n + 1) = d_{n-1} - 1,$$

it follows from the induction hypothesis that $t(p) \le n$. Now suppose that $d_{n-1} \le p \le c_n$. By using a term 2^{2n-2}, we see that $t(p) \le 1 + t(|p - 2^{2n-2}|)$. Since

$$c_n - 2^{2n-2} = 2^{2n-2} - d_{n-1} = c_{n-1} < d_{n-1},$$

it follows again that $t(p) \le n$.

Problem 6.27. (USAMO 2003) Prove that for every positive integer n there exists an n-digit number divisible by 5^n, all of whose digits are odd.

Solution. We will use induction on $n \ge 1$.

For the base case $n = 1$ we can take the number 5.

For the induction step, assume that we have constructed a number A of k odd digits, which is divisible by 5^k. Notice that each of the numbers 1, 3, 5,7, 9 gives a different residue modulo 5. Since there are just 5 residues modulo 5, they must cover all residues modulo 5. Hence there exists $c \in \{1, 3, 5, 7, 9\}$ such that

$$c \cdot 2^k \equiv -\frac{A}{5^k} \pmod 5 \quad \text{or} \quad c \cdot 10^k \equiv -A \pmod{5^{k+1}}.$$

Then the number $10^k c + A$ is the desired number. So the statement is also true for $n = k + 1$, which completes our proof.

Problem 6.28. (IMO 2004) We call a positive integer *alternating* if every two consecutive digits in its decimal representation are of different parity. Find all positive integers n such that n has a multiple which is alternating.

Solution. The answer depends on the very essence on the decimal number system, thus on the powers of two dividing n. For example, if 10 divides n, then the last digit of any multiple of n is 0, so if an alternating number is divisible by n its ten's digit must be odd, implying that n is not divisible by 20. Let us prove that this is also a sufficient condition, namely that the numbers which are not multiples of 20 have a multiple which is alternating.

The numbers not divisible by 20 split into four groups:

$$m, 2^k \cdot m, 5^k \cdot m, 2 \cdot 5^k \cdot m, \quad \text{where} \quad \gcd(m, 10) = 1.$$

We first treat the case of m with $\gcd(m, 10) = 1$. Among the numbers of the form $101010101\dots01$ there are two who give the same remainder upon division on m. Subtracting the smaller from the larger and cutting the irrelevant zeroes, we get a number of the form $1010\dots01$ (which is alternating) and divisible by m. The method above can be generalized to construct a number of the form $100\dots100\dots0\dots100\dots01$ divisible by m, where between any two consecutive unities there are k zeroes. This number is not alternating, but we shall need it for the next cases.

Turning to the other cases, we first prove that 2^k has an alternating multiple of $k+1$ digits, by induction on k. The base case is clear.

Now assume that we have constructed an alternating multiple M of 2^k with $k+1$ digits. If M is divisible by 2^{k+1}, just add 10^{k+1} or $2 \cdot 10^{k+1}$ to it to make it into an alternating multiple of 2^{k+1} with $k+2$ digits. If it is not divisible by 2^{k+1}, then we may add 2 or subtract 2 from the digit representing 10^{k-1} in the decimal representation of M (if the digit is larger than 1, subtract 2, otherwise add) and make it into an alternating number divisible by 2^{k+1}. Again we add 10^{k+1} or $2 \cdot 10^{k+1}$ to it to ensure the induction step.

By induction on k, we can also prove that $2 \cdot 5^k$ (hence also 5^k) has an alternating multiple of $k+1$ digits. The base case is clear. Now let us consider a multiple M of $2 \cdot 5^k$ with $k+1$ digits. If M is also divisible by 5^{k+1}, add 1 or 2 in front of it to do the induction. Otherwise, look at the second digit of M

(i.e. the one accompanying 10^k). Only its parity matters, so we can change it into one of the other four possible values, and it is clear exactly one of these changes will turn M into a multiple of 5^{k+1}. Again, by adding 1 or 2 in front of the resulting number we complete the induction.

We have showed that we can find an alternating multiple of 2^k, and $2 \cdot 5^k$ having $k+1$ digits. We may assume that it has an even number of digits, as otherwise we can add one more digit at the beginning such that the multiple is still alternating. Thus its leading digit is odd. Now if it has d digits, then multiplying it by a multiple of m of the form $10^{k(d+1)} + 10^{(k-1)(d+1)} + \ldots + 1$ (which we proved above exists), we construct an alternating multiple of $2^k m$ or $2 \cdot 5^k m$.

Thus for n of the form $2^k \cdot m$, $2 \cdot 5^k \cdot m$ or $5^k \cdot m$, an alternating multiple exists. So the answer to our question is all numbers except for the multiples of 20.

Problem 6.29. (IMO shortlist 2000) Does there exist a positive integer n such that n has exactly 2000 prime divisors and n divides $2^n + 1$?

Solution. We prove the existence of an odd positive integer n divisible by 3, having exactly k prime divisors and dividing $2^n + 1$, by induction on k.

The base case is $k = 1$ for which we take $n = 3$.

Now assume that we have found n with k prime divisors and $n \mid (2^n + 1)$. As $(x^2 - x + 1, x + 1) \mid 3$ (because $x^2 - x + 1 - (x+1)(x-2) = 3$), we deduce that either $2^{2n} - 2^n + 1$ is a power of 3 or $2^{2n} - 2^n + 1$ has a prime divisor p which does not divide $2^n + 1$ and thus does not divide n.

The first case is impossible, as $2^{2n} - 2^n + 1$ is not divisible by 9 since $2^{2n} = 2^{6k} \equiv 1 \pmod 9$, $2^n = 2^{6m+3} \equiv -1 \pmod 9$, so $2^{2n} - 2^n + 1 \equiv 3 \pmod 9$. Thus there is a prime divisor $p \mid 2^{2n} - 2^n + 1$ which does not n. Then $p \mid (2^{3n} + 1)$ and $3 \mid (2^{2n} - 2^n + 1)$, so $3np \mid (2^n + 1)(2^{2n} - 2^n + 1)$, hence $3np \mid (2^{3n} + 1)$ and since $(2^{3n} + 1) \mid (2^{3np} + 1)$, we have that $3np$ is the desired number.

Problem 6.30. (**IMO shortlist 2002**) Let $p_1, p_2, \ldots, p_n$ be distinct primes greater than 3. Show that $2^{p_1 p_2 \cdots p_n} + 1$ has at least 4^n divisors.

Solution. We will prove the result by induction on n. Before we do that let us prove some preliminary results:

Lemma 1. If $\gcd(a,b)=1$ and a,b are odd, then $\gcd(2^a+1,2^b+1)=3$.

Proof. Let $d=\gcd(2^a+1,2^b+1)$. $3|d$ and

$$d|\gcd(2^{2a}-1,2^{2b}-1)=2^{\gcd(2a,2b)}-1=2^2-1=3,$$

so $d=3$.

Lemma 2. If a is an odd positive integer and $3\nmid a$, then $9\nmid(2^a+1)$.

Proof. If $a=3k+1$ we have $2^a+1=2(2^{3k}+1)-1=9t-1$, and if $a=3k+2$ we have $2^a+1=4(2^{3k}+1)-3=9t-3$.

Back to our original problem, let $N=p_1\cdot p_2\cdot\ldots\cdot p_n$ and $N'=\frac{N}{p_n}$.

The base case $n=1$ is easy to prove, since 2^p+1 has the divisors $1,3,\frac{2^p+1}{3},2^p+1$ and we have $3\neq\frac{2^p+1}{3}$ because $p>3$.

Assume that we have proved the result for $n-1$, where $n\geq 2$. This means that $2^{N'}+1$ has at least 4^{n-1} divisors. Notice that

$$\left(2^{N'}+1\right)\mid\left(2^N+1\right),2^{p_n}+1\mid\left(2^N+1\right),$$

so by applying Lemma 1 to $a=p_n,b=N'$, we get

$$\frac{2^{p_n}+1}{3}\mid\frac{2^N+1}{2^{N'}+1}.$$

On the other hand, we clearly have

$$p_n<\frac{2^{p_n}+1}{3}\quad\text{and}\quad p_n\left(\frac{2^{p_n}+1}{3}\right)^2<\frac{2^N+1}{2^{N'}+1},$$

which means that $\frac{2^N+1}{2^{N'}+1}$ has at least four divisors:

$$d_1=1,\ d_2=\frac{2^{p_n}+1}{3},\ d_3=\frac{3}{2^{p_n}+1}\cdot\frac{2^N+1}{2^{N'}+1},\text{ and }d_4=\frac{2^N+1}{2^{N'}+1}$$

and the inequalities give $p_nd_k<d_{k+1}$. For each $d\mid\left(2^{N'}+1\right)$, each of d_id is a divisor of 2^N+1. Since

$$\begin{aligned}\gcd(2^{N'}+1,\frac{2^N+1}{2^{N'}+1})&=\gcd(2^{N'}+1,2^{N'(p_n-1)}-2^{N'(p_n-2)}+\cdots+1)\\&=\gcd(2^{N'}+1,p_n)\mid p_n,\end{aligned}$$

we cannot have $d_i d = d_j d'$ with $i < j$ as $\gcd(d_j, d) \mid p_n$ would force $d_j \mid d_i p_n$, and hence $d_j \leq d_i p_n$, contrary to our construction. Thus the divisors $d_i d$ are distinct.

Hence $2^N + 1$ has at least $4 \cdot 4^{n-1} = 4^n$ divisors.

Problem 6.31. (**IMO 1988**) Show that if a, b and $q = \dfrac{a^2+b^2}{ab+1}$ are non-negative integers, then $q = \gcd(a,b)^2$.

Solution. We will prove the result using induction on ab. For $ab = 0$, things are clear.

For $ab > 0$, by symmetry, we can assume without loss of generality that $a \leq b$. Assume that the result holds for all possible values less than ab. To use the induction hypothesis, we look for an integer $0 \leq c < b$ such that

$$q = \frac{a^2+c^2}{ac+1}, \quad 0 \leq c < b.$$

We can find the value of c explicitly, since

$$\frac{a^2+b^2}{ab+1} = \frac{a^2+c^2}{ac+1} = q,$$

so

$$\frac{b^2-c^2}{ab-ac} = q \Leftrightarrow \frac{b+c}{a} = q \Leftrightarrow c = aq - b.$$

Since $c = aq - b$ and $q = (a,b)$, we have that $q = (a,c)$ as well. Therefore, from the inductive hypothesis for $q = (a,c)^2$ we have the result for $q = (a,b)^2$. The only thing left to show is that we indeed have $0 \leq c < b$. For this, notice that

$$q = \frac{a^2+b^2}{ab+1} < \frac{a^2+b^2}{ab} = \frac{a}{b} + \frac{b}{a},$$

so

$$aq < \frac{a^2}{b} + b \leq \frac{b^2}{b} + b = 2b \Rightarrow aq - b < b \Rightarrow c < b.$$

Moreover,

$$q = \frac{a^2+c^2}{ac+1} \Rightarrow ac + 1 > 0 \Rightarrow c > \frac{-1}{a} \Rightarrow c \geq 0.$$

This completes our proof.

Problem 6.32. (**IMO 1999 shortlist**) Prove that there exist two strictly increasing sequences (a_n) and (b_n) such that $a_n(a_n+1)$ divides b_n^2+1 for every natural n.

Solution. We begin our solution by noting that it suffices to find positive integers c_n, d_n such that $a_n \mid (c_n^2+1)$, $(a_n+1) \mid (d_n^2+1)$, because then, since $(a_n, a_n+1)=1$, we can find $b_n \equiv c_n \pmod{a_n}$, $b_n \equiv d_n \pmod{a_n+1}$.

Let us show that $a_n = 5^{2n}$ works.We have $a_n+1 = 5^{2n}+1 = (5^n)^2+1$, so all we need to show is that there is some c_n such that $5^{2n} \mid (c_n^2+1)$.

We can show something more general: if p is a prime of the form $4k+1$, then there is a t_n such that $p^n \mid (t_n^2+1)$.

We construct t_n inductively. Since $p \equiv 1 \pmod 4$, the Legendre symbol $\left(\frac{-1}{p}\right)=1$, so t_1 exists.

For the induction step, assuming that $p^{n-1} \mid (t_{n-1}^2+1)$ we have to find $k \in \{0,1,\ldots,p-1\}$ such that $p^n \mid \left((t_{n-1}+kp^{n-1})^2+1\right)$. This reduces to $p \mid \left(\frac{t_{n-1}^2+1}{p^{n-1}} + 2kt_{n-1}\right)$, and it is clear that we can find such a k.

Problem 6.33. Prove that for any two positive integers n and m, we have

$$\gcd(F_n, F_m) = F_{\gcd(n,m)}.$$

Solution. We first note that two special cases are trivial. If $n=m$, then the identity becomes $\gcd(F_m, F_m) = F_m$, which is obvious. If $n = m+1$, then the identity becomes $\gcd(F_{m+1}, F_m) = 1$. This follows easily by induction on m. The base case reads $\gcd(F_2, F_1) = \gcd(1,1) = 1$. The inductive step follows since

$$\gcd(F_{m+1}, F_m) = \gcd(F_{m+1} - F_m, F_m) = \gcd(F_{m-1}, F_m) = 1.$$

We will prove the general result by induction on $n+m$.

If $n+m=2$, then, since n and m are positive integers, we must have $n=m=1$, and this case was treated above.

Now let k be a positive integer, and assume that the relation $\gcd(F_n, F_m) = F_{\gcd(n,m)}$ holds for any two positive integers n and m with $n+m<k$. We will then show that it also holds for any two positive integers n and m with $n+m=k$.

Let n and m be two such integers. The case $n = m$ was handled above, so without loss of generality, we can assume that $n > m$. Then, $n - m$ is a positive integer. Now, applying Cassini's identity: $F_{m+n} = F_{m+1}F_n + F_mF_{n-1}$ to the positive integers $n - m$ and m, we get

$$F_n = F_{(n-m)+m} = F_{n-m}F_{m+1} + F_{(n-m)-1}F_m.$$

Thus

$$\gcd(F_n, F_m) = \gcd(F_{n-m}F_{m+1} + F_{n-m-1}F_m, F_m) = \gcd(F_{n-m}F_{m+1}, F_m).$$

But we saw above that F_{m+1} and F_m are relatively prime, so this gives

$$\gcd(F_n, F_m) = \gcd(F_{n-m}, F_m) = F_{\gcd(n-m,m)} = F_{\gcd(n,m)},$$

where the second inequality is the induction hypothesis.

Problem 6.34. Let n be a positive integer which is not divisible by 3. Show that $x^3+y^3 = z^n$ has at least one solution (x, y, z) with x, y, z positive integers.

Solution. We prove the result by induction on n, when $n \not\equiv 0 \pmod 3$. The base cases are $n = 1$ and $n = 2$. For $n = 1$, we have $1^3 + 2^3 = 9$. For $n = 2$, we have $1^3 + 2^3 = 3^2$.

For the induction step, assume that we have

$$x_n^3 + y_n^3 = z_n^n, \quad \text{for} \quad n \not\equiv 0 \pmod 3.$$

We then take $x_{n+3} = z_nx_n$, $y_{n+3} = z_ny_n$, $z_{n+3} = z_n$ and we have

$$\begin{aligned} x_{n+3}^3 + y_{n+3}^3 &= z_n^3(x_n^3 + y_n^3) \\ &= z_n^3z_n^n = z_n^{n+3} = z_{n+3}^{n+3}. \end{aligned}$$

This completes our proof.

Remark. There is also a short proof for the above question which does not use induction: If $n = 3r + 1$, take $x = a(a^3 + b^3)^r$, $y = b(a^3 + b^3)^r$, and $z = a^3 + b^3$. If $n = 3r + 2$, take $x = a(a^3 + b^3)^{2r+1}$, $y = b(a^3 + b^3)^{2r+1}$, and $z = (a^3 + b^3)^2$.

Problem 6.35. Let n be a positive integer. What is the largest number of elements that one can choose from the set $A = \{1, 2, \ldots, 2n\}$ such that the sum any two chosen numbers is composite?

Solution. We prove by induction on n that if we choose $n+1$ numbers from A, then there exist two numbers whose sum is prime.

When $n = 1$, there is nothing to prove, as $1 + 2 = 3$ is prime.

Assume now that the statement holds for all integers $1 \le n \le m - 1$, $m \ge 2$, and we show that it also holds for $n = m$.

By Chebyshev's theorem, there is a prime number in the interval $(2m, 4m)$, call it p. Let $p = 2m + k$, for some $k > 0$. Then k is odd and $k - 1 \le 2(m-1)$. Notice that if we choose $m + 1$ numbers from the set $\{1, 2, \ldots, 2m\}$, we either choose more than half of elements in the set $\{1, 2, \ldots, k-1\}$ or by Pigeonhole Principle, we choose both elements of one of the following pairs:

$$(k, 2m), (k, 2m-1), \ldots, \left(\frac{k+2m-1}{2}, \frac{k+2m+1}{2}\right)$$

In the first case, the conclusion follows from the induction hypothesis applied to $n = (k-1)/2$, while in the second case the result holds because the sum of elements in each of the above pairs is p.

This completes our induction.

To finish the question, notice that an example with n elements is given by choosing the elements $2, 4, 6, \ldots, 2n$.

Problem 6.36. (**Bulgaria 1999**) Find the number of positive integers n, $4 \le n \le 2^k - 1$, whose binary representations do not contain three equal consecutive digits.

Solution. Let a_k be the number of numbers having k digits in binary and not containing three consecutive digits. Then the answer to our question is $a_3 + a_4 + \ldots + a_k$. To establish a recurrence for a_k, we use a method which is very helpful for many similar problems: we split the sequence into more sequences which are easier to investigate.

More precisely, we define $a(00)_n$, $a(01)_n$, $a(10)_n$, $a(11)_n$ to be the number of numbers of n binary digits that do not contain three consecutive equal digits and end up in 00, 01, 10 and 11, respectively (in binary). By looking at the third digit from the end we establish the following relations:

$$\begin{aligned} a(00)_n &= a(10)_{n-1}; \\ a(01)_n &= a(00)_{n-1} + a(10)_{n-1}; \\ a(10)_n &= a(11)_{n-1} + a(01)_{n-1}; \\ a(11)_n &= a(01)_{n-1}. \end{aligned}$$

From these, we get that

$$a(00)_n + a_(11)_n = a(10)_{n-1} + a(01)_{n-1}$$

and

$$a(10)_n + a(01)_n = a(00)_{n-1} + a_(11)_{n-1} + a(10)_{n-1} + a(01)_{n-1} = a_{n-1}.$$

Hence $a(00)_n + a_(11)_n = a_{n-2}$, so

$$a(00)_n + a(11)_n + a(10)_n + a(01)_n = a_{n-1} + a_{n-2}.$$

Since $a_3 = 3$, $a_4 = 5$, we conclude that $a_n = F_{n+1}$ by induction on n. So the desired number is $F_4 + F_5 + \ldots + F_k$. Now the relation $F_1 + F_2 + \ldots + F_k = F_{k+2} - 1$ (which follows by a simple induction on k), gives us that the answer to our problem is $F_{k+2} - 1 - F_1 - F_2 - F_3 = F_{k+2} - 5$.

Problem 6.37. Prove that every positive integer is a sum of one or more numbers of the form $2^r 3^s$, where r and s are non-negative integers and no summand divides another (for example, $23 = 9 + 8 + 6$).

Solution. We prove the statement using strong induction. The base cases $n = 1$, $n = 2$ and $n = 3$ are clear.

For the induction step, assume that the result holds for all integers less than some $n \geq 4$. We distinguish two cases:

If n is even, then $\frac{n}{2}$ can be properly represented. By multiplying each of the powers in the representation by 2, we get a representation of n.

If n is odd, then let $3^k \leq n < 3^{k+1}$. If $n = 3^k$ the representation is clear. If $n > 3^k$, then $\frac{n-3^k}{2}$ is an integer which can be represented as a desired sum.

Multiply each term in the representation of $\frac{n-3^k}{2}$ by 2 and add 3^k. We claim we get a good representation for n. Clearly no summand coming from the representation of $\frac{n-3^k}{2}$ can divide another such summand. Also since all such summands are even they cannot divide 3^k. Since $\frac{n-3^k}{2}$ is less than 3^k, no summand coming from the representation of $\frac{n-3^k}{2}$ can be divisible by 3^k.

Problem 6.38. Let $p \geq 3$ be a prime and let $a_1, a_2, \ldots, a_{p-2}$ be a sequence of positive integers such that p does not divide either a_k or $a_k^k - 1$ for all $k = 1, 2, \ldots, p-2$. Prove that the product of some elements of the sequence is congruent to 2 modulo p.

Solution. Clearly 2 is not really relevant modulo p, so we may suspect that any non-zero residue modulo p may be represented as the product of some elements of the sequence. A natural approach by induction would be the following statement:

The set of products of elements of the set $\{1, a_1, a_2, \ldots, a_k\}$ contains at least $k+1$ distinct residues modulo p (we added 1 to our set because we need a set of $k+1$ numbers to build $k+1$ different residues and the condition $p \nmid a_k^k - 1$ implies that none of a_i's is equal to 1).

The base case $k = 1$ is true as a_1 and 1 have different residues modulo p.

Now assume that there are j elements $b_1, b_2, \ldots, b_j$ with different non-zero remainders modulo p represented as the product of some elements of the set $\{1, a_1, \ldots, a_{j-1}\}$. The numbers $b_1 a_j, b_2 a_j, \ldots, b_j a_j$ also have different non-zero remainders modulo p. These remainders cannot be a permutation of $b_1, b_2, \ldots, b_j$, because otherwise by multiplying we would get

$$b_1 b_2 \ldots b_j \equiv (a_j b_1)(a_j b_2) \ldots (a_j b_j) \pmod{p},$$

thus $a_j^j \equiv 1 \pmod{p}$, which is a contradiction. So one of the numbers $a_j b_i$ does not give the same residue modulo p as any of the elements in the set $\{b_1, b_2, \ldots, b_j\}$. Hence the set $\{b_1, b_2, \ldots, b_j, a_j b_i\}$ is the desired set of $j+1$ numbers. This establishes the induction step.

According to the above result, the product of some elements of $\{1, a_1, \ldots, a_{p-2}\}$ is congruent to 2 modulo p, because all the $p-1$ different nonzero residues modulo p were proved to be represented as such a product.

We can remove 1 from the set since it is irrelevant for the products, and we get that the product of some numbers from $\{a_1, a_2, , \ldots, a_{p-2}\}$ is 2 modulo p.

Problem 6.39. Let n be a positive integer. Prove that the number of ordered pairs (a, b) of relatively prime positive divisors of n is equal to the number of divisors of n^2.

Solution. We will prove the result by induction on the number of distinct prime factors of n. As usual, we denote the cardinality of a set S by $|S|$, and the number of positive divisors of an integer n by $\tau(n)$.

The base case $n = 1$ is immediate. Now suppose n has only one prime factor, $n = p^k$. Then $\tau(n^2) = \tau(p^{2k}) = 2k + 1$, and

$$|\{(a, b) : \ a \mid n, b \mid n, \ \gcd(a, b) = 1\}|$$

$$= |\{(p^i, 1) : \ 1 \leq i \leq k\}| + |\{(1, p^i); \ 1 \leq i \leq k\}| + |\{(1, 1)\}| = 2k + 1.$$

For the induction step, suppose that $n = mp^k$, where the prime p is not a factor of m. Since the function τ is multiplicative and $\gcd(m^2, p^{2k}) = 1$, we have $\tau(n^2) = \tau(m^2)(2k + 1)$. By the induction hypothesis we obtain

$$|\{(a, b) : \ a \mid m, b \mid m, \gcd(a, b) = 1\}| = \tau(m^2).$$

We now see that

$$\begin{aligned}
|\{(c, d) : \ c \mid n, d \mid n, \gcd(c, d) = 1\}| &= |\{(a, b) : \ a \mid m, b \mid m, \gcd(a, b) = 1\}| \\
&+ |\{(ap^i, b) : \ a \mid m, b \mid m, \gcd(a, b) = 1, 1 \leq i \leq k\}| \\
&+ |\{(a, bp^i) : \ a \mid m, b \mid m, \gcd(a, b) = 1, 1 \leq i \leq k\}| \\
&= \tau(m^2)(2k + 1),
\end{aligned}$$

which completes the proof.

Problem 6.40. Prove that for each integer $n \geq 3$, there are n pairwise distinct positive integers such that each of them divides the sum of the remaining $n-1$.

Solution. We proceed by induction on $n \geq 3$. The base case is $n = 3$, for which we can take $(3, 6, 9)$ as an example.

Assume now that the statement holds for some $n \geq 3$ and let $(x_1, x_2, \ldots, x_n)$ be n pairwise distinct positive integers with this property, and let s be their sum. Then the statement also holds for $(x_1, x_2, \ldots, x_n, s)$, since from the induction hypothesis we have

$$x_i \mid \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} x_j + s \right) = \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} x_j + x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} x_j \right), \quad \forall i \in \{1, \ldots, n\}.$$

The check for s is immediate. This completes our induction step and thus the proof.

Problem 6.41. (**USAMO 2008**) Show that for all positive integers n we can find distinct positive integers $a_1, a_2, \ldots a_n$ such that $a_1 \cdot a_2 \cdots a_n - 1$ is the product of two consecutive integers.

Solution. We prove the result by induction on n. For $n = 1$, choose $a_1 = 3$ and we have $a_1 - 1 = 1(1+1)$. For $n = 2$ choose $a_1 = 1$, $a_2 = 7$ and we have that $a_1 \cdot a_2 - 1 = 2(2+1)$.

Assume now that the result holds for some $n \geq 2$, that is, there are distinct positive integers $a_1, \ldots a_n$ such that $a_1 \cdot a_2 \cdots a_n - 1 = k(k+1)$, for some positive integer k. This implies that

$$a_1 \cdot a_2 \cdots a_n = k^2 + k + 1,$$

so

$$(k^2 - k + 1) \cdot a_1 \cdot a_2 \cdots a_n = (k^2 - k + 1)(k^2 + k + 1) = k^4 + k^2 + 1.$$

So we choose $a_{n+1} = k^2 - k + 1$ and we have that

$$a_1 \cdot a_2 \cdots a_{n+1} - 1 = k^2(k^2 + 1).$$

Also, from $a_1 \cdot a_2 \cdots a_n - 1 = k(k+1)$, we have that none of $a_1, \ldots, a_n$ can be equal to $k^2 - k + 1$, since $2(k^2 - k + 1) > k^2 + k$ for any $k \geq 3$. Hence $a_1, \ldots, a_{n+1}$ are all distinct. This completes our induction.

Problem 6.42. We begin with a triple (a, b, c) of positive reals a, b, c, such that $a \leq b \leq c$. Starting with the given triple, at each step, we perform the transformation

$$(x, y, z) \rightarrow (|x-y|, |y-z|, |z-x|).$$

Prove that we can eventually reach a 0 in one of the triples if and only if there exist positive integers $n \geq k \geq 0$ such that and

$$nb = ka + (n-k)c.$$

Solution. Let

$$nb = ka + (n-k)c, \tag{5}$$

where $n \in \mathbb{N}$, $k \in \mathbb{Z}$ and $n \geq k \geq 0$.
We prove by induction on n that we can eventually reach 0 in one of the triples.
If $n = 1$, then $b = a$ or $b = c$, thus $|a-b| = 0$ or $|b-c| = 0$.
For the induction step we show that if the statement holds true for $1 \leq n \leq S-1$, where $S \geq 2$, $S \in \mathbb{N}$, then it also holds for $n = S$.
Let $Sb = pa + (S-p)c$, $S \geq p \geq 0$, $S \in \mathbb{N}$, $p \in \mathbb{Z}$.
If $0 < p < \dfrac{S}{2}$, we have that

$$(S-p)(a-b) = p(b-c) + (S-2p)(a-c).$$

Therefore, 0 will be in one of the triples obtained from the triple

$$(|a-b|, |b-c|, |c-a|).$$

If $\dfrac{S}{2} \leq p < S$, the conclusion follows from the fact that

$$p(b-c) = (S-p)(a-b) + (2p-S)(a-c).$$

If $p = 0$ or $p = S$ then we clearly have the result as well.
Conversely, let us now prove that, if 0 is in one of the triples, then (5) holds. We notice that, if for some triple (5) holds, then it also holds for any triple obtained before this triple: assume that we have obtained the triple $(x-y, y-z, x-z)$

from the triple (x,y,z), where $x \geq y \geq z \geq 0$ and let $y-z \geq x-y$. We have that

$$n(y-z) = k(x-y) + (n-k)(x-z).$$

Therefore,

$$(n+k)y = nx + kz.$$

Hence, if 0 is in one of the triples obtained from the triple (a,b,c), then the triple (d,d,e), where $1 \cdot d = 1 \cdot d + 0 \cdot e$, is the triple that precedes the triple containing 0 in the sequence of transformations. So (5) holds for this triple and therefore, (5) holds for (a,b,c) too. This completes the proof of the question.

Problem 6.43. (**Poland 2000**) A sequence $p_1, p_2, \ldots$ of prime numbers satisfies the following condition: for $n \geq 3$, p_n is the greatest prime divisor of $p_{n-1} + p_{n-2} + 2000$. Prove that the sequence is bounded.

Solution. Let $b_n = \max\{p_n, p_{n+1}\}$ for $n \geq 1$.
We first prove that $b_{n+1} \leq b_n + 2002$ for all such n. Certainly $p_{n+1} \leq b_n$, so it suffices to show that $p_{n+2} \leq b_n + 2002$. If either p_n or p_{n+1} equals 2, then we have $p_{n+2} \leq p_n + p_{n+1} + 2000 = b_n + 2002$. Otherwise, p_n and p_{n+1} are both odd, so $p_n + p_{n+1} + 2000$ is even. Because $p_{n+2} \neq 2$ divides this number, we have

$$p_{n+2} \leq \frac{p_n + p_{n+1} + 2000}{2} = \frac{p_n + p_{n+1}}{2} + 1000 \leq b_n + 1000.$$

This proves the claim.
Choose k large enough so that $b_1 \leq k \cdot 2003! + 1$. We prove by induction that $b_n \leq k \cdot 2003! + 1$ for all n. If this statement holds for some n, then

$$b_{n+1} \leq b_n + 2002 \leq k \cdot 2003! + 2003.$$

If $b_{n+1} > k \cdot 2003! + 1$, then let $m = b_{n+1} - k \cdot 2003!$. We have $1 < m \leq 2003$, implying that $m \mid 2003!$. Hence, m is a proper divisor of $k \cdot 2003! + m = b_{n+1}$, which is impossible because b_{n+1} is prime. Thus $p_n \leq b_n \leq k \cdot 2003! + 1$ for all n, which establishes the result we wanted to prove.

Problem 6.44. Prove that for all positive integers m, there exists an integer n such that $\phi(n) = m!$.

Solution. For $k \geq 1$, let p_k be the k-th prime number. We will prove by induction on m that if $p_k \leq m < p_{k+1}$ then there is a number n whose prime factors are $p_1, \ldots, p_k$ with some (positive) multiplicities and such that $\phi(n) = m$.

The base cases are $1! = \phi(1)$, $2! = \phi(4)$, $3! = \phi(18)$, $4! = \phi(72)$.

Suppose now the result holds for all values smaller than m. If $m = p_k$, then by the induction hypothesis, we have

$$(p_k - 2)! = \phi(p_1^{e_1} \cdot p_2^{e_2} \cdot \ldots \cdot p_{k-1}^{e_{k-1}}),$$

for some positive integers $e_1, \ldots, e_{k-1}$. By the multiplicative property of ϕ, we have

$$\begin{aligned} p_k! = p_k(p_k - 1)(p_k - 2)! &= \phi(p_k^2)(p_k - 2)! \\ &= \phi(p_k^2)\phi(p_1^{e_1} \cdot p_2^{e_2} \cdot \ldots \cdot p_{k-1}^{e_{k-1}}) \\ &= \phi(p_1^{e_1} \cdot p_2^{e_2} \cdot \ldots \cdot p_{k-1}^{e_{k-1}} \cdot p_k^2). \end{aligned}$$

This proves the statement for $m = p_k$.

Assume now that $p_k < m < p_{k+1}$. Then $m = \prod_{i=1}^{k} p_i^{\alpha_i}$. Also by the induction hypothesis, we have

$$\phi(p_1^{\beta_1} \cdot p_2^{\beta_2} \cdot \ldots \cdot p_k^{\beta_k}) = (m-1)!,$$

for some positive integers $\beta_1, \ldots, \beta_k$. Since the β_i are positive we have

$$\begin{aligned} m! = m \cdot (m-1)! &= p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \ldots \cdot p_k^{\alpha_k} \phi(p_1^{\beta_1} \cdot p_2^{\beta_2} \cdot \ldots \cdot p_k^{\beta_k}) \\ &= \phi(p_1^{\beta_1+\alpha_1} \cdot p_2^{\beta_2+\alpha_2} \cdot \ldots \cdot p_k^{\beta_k+\alpha_k}). \end{aligned}$$

This completes the inductive step and the proof.

Problem 6.45. (**Bulgaria 2012**) Let p be an odd prime and let $a_1, \ldots, a_{2p-1}$ be distinct integers lying in the interval $[1, p^2]$ such that their sum is divisible by p. Prove that there exist positive integers $b_1, \ldots, b_{2p-1}$ none of which is divisible by p, such that their representation in base p contains only 1 and 0 and such that $\sum_{j=1}^{2p-1} a_j b_j$ is divisible by p^{2012}.

Solution. We shall prove the statement for any exponent $n \geq 1$ of p, not just 2012. We begin by showing the following:

Lemma. If $x_1, \ldots, x_{p-1}$ are positive integers, none of which is divisible by p, then for all $1 \leq r \leq p-1$ we can choose some of the x_i's such that their sum is r modulo p.

Proof. Let $1 \leq k \leq p-2$ be such that the numbers $r_1, \ldots, r_k$ have different residues modulo p. Adding x_{k+1} to each term, we obtain the sequence

$$x_{k+1} + r_1, x_{k+1} + r_2, \ldots, x_{k+1} + r_k.$$

Notice that no two terms in the new sequence will give the same residue modulo p. Also notice that the residues given by the second sequence cannot be a permutation of the residues from the first sequence, as we would then have

$$r_1 + r_2 + \ldots + r_k \equiv (r_1 + x_{k+1}) + \ldots + (r_k + x_{k+1}) \pmod{p},$$

giving $k \cdot x_{k+1} \equiv 0 \pmod{p}$, which contradicts the fact that $p \nmid x_{k+1}$. Therefore, we obtain at least one new residue by this construction and inductively, we can obtain all the non-zero residues. This proves our lemma.
Back to our problem, we prove by induction on n that we can choose numbers b_j with at most n digits in base p and all base p digits either 0 or 1, such that the sum $\sum\limits_{j=1}^{2p-1} a_j b_j$ is a multiple of p^n.

The base case is $n = 1$. For this, we know that $p \mid \sum\limits_{i=1}^{2p-1} a_i$, so we can choose $b_1 = b_2 = \ldots = b_{2p-1} = 1$.
For the induction step, assume that there are n-digit positive integers $b_1, \ldots, b_{2p-1}$, none of which is divisible by p, such that their representation in base p contains only 1 and 0 and such that

$$\sum_{j=1}^{2p-1} a_j b_j = p^n \cdot A, \quad \text{where} \quad A \in \mathbb{N}^*.$$

We can assume that A is not divisible by p (as otherwise we would be done). Since $a_1, \ldots, a_{2p-1} \in [1, p^2]$, there are at least $p-1$ numbers in this sequence which are not divisible by p. Without loss of generality, let them be $a_1, \ldots, a_p$. By our lemma, we can find $a_1, \ldots, a_k$ such that

$$a_1 + a_2 + \ldots + a_k + A \equiv 0 \pmod{p}.$$

Now set $b_i' = b_i + p^n$, for $i = 1, \ldots, k$ and $b_j' = b_j$, for $j = k+1, \ldots, 2p-1$. These numbers satisfy all the required conditions, and by above we have that $\sum\limits_{l=1}^{2p-1} a_l b_l'$ is divisible by p^{n+1}, which completes our proof.

Problem 6.46. (**Poland 2010**) Prove that there exists a set A consisting of 2010 positive integers such that for any nonempty subset $B \subseteq A$, the sum of the elements in B is a perfect power greater than 1.

Solution. We prove by induction on t that for every $0 \le t \le 2^{2010} - 1$ there exists a set U_t of 2010 positive integers such that at least t nonempty subsets have their sums of elements a perfect power greater than 1. The base case $t = 0$ is trivial since we can just take any set of 2010 positive integers.

Assume now that we have constructed a set U_t for some $0 \le t < 2^{2010} - 1$. There are t subsets of U_t whose sums are perfect powers and let the sums of these subsets be $n_1^{m_1}, \ldots, n_t^{m_t}$. Let V be any other subset of U_t and let c be the sum of elements of V. We define U_{t+1} to be the set consisting of all elements of U_t multiplied by $c^{\operatorname{lcm}(m_1,\ldots,m_t)}$, i.e.

$$U_{t+1} = \{a \cdot c^{\operatorname{lcm}(m_1,\ldots,m_t)} :\ a \in U_t\}.$$

Notice that the t subsets of U_t whose sums were perfect powers give subsets of U_{t+1} whose sums are perfect powers. Moreover, the sum of elements in the subset of U_{t+1} which corresponds to V is now $c^{1+\operatorname{lcm}(m_1,\ldots,m_t)}$, hence a perfect power. This completes the inductive step.

The set $U_{2^{2010}-1}$ will be the desired set A.

Problem 6.47. (**AMM**) For a positive integer m, we let $\sigma(m)$ be the sum

$$\sigma(m) = \sum_{\substack{1 \le d < m \\ d \mid m}} d.$$

Prove that for every integer $t \geq 1$ there exists an m such that

$$m < \sigma(m) < \sigma(\sigma(m)) < \ldots < \sigma^t(m),$$

where $\sigma^k = \underbrace{\sigma \circ \sigma \circ \ldots \circ \sigma}_{k \text{ terms}}$.

Solution. We prove the following, stronger statement: For all positive integers t, there exists a positive integer n_t, such that for all positive integers m for which $n_t \mid m$ and $\gcd(n_t, \frac{m}{n_t})$ we have

$$m < \sigma(m) < \sigma(\sigma(m)) < \ldots < \sigma^t(m).$$

We prove this by induction on t. The base case is $t = 1$, for which we can take $n_1 = 12$.
Assume now that we have constructed n_t for some $t \geq 1$. Choose a prime p which does not divide n_t and let $k \geq 1$ be such that

$$p^k \equiv 1 \pmod{n_t^2(p-1)}.$$

(one can take for example $k = \phi(n_t^2(p-1)))$. We claim that $n_{t+1} = p^{k-1} \cdot n_t$ satisfies the required properties for the induction step.
Indeed, let m be such that $n_{t+1} \mid m$ and $\gcd(n_{t+1}, \frac{m}{n_{t+1}}) = 1$.
Also let $A(m) = m + \sigma(m)$. Notice that $\nu_p(m) = k - 1$ and $A(p^{k-1}) \mid A(m)$, so we have $n_t^2 \mid \frac{p^k-1}{p-1} \mid A(m)$. Hence

$$\sigma(m) = A(m) - m \equiv -m \pmod{n_t^2} \Rightarrow \gcd(n_t, \frac{\sigma(m)}{n_t}) = 1.$$

From the induction hypothesis, $\sigma(m)$ satisfies the conditions of the problem, so we have

$$\sigma(m) < \sigma^2(m) < \ldots < \sigma^t(\sigma(m)) = \sigma^{t+1}(m).$$

Since $\gcd(n_t, \frac{m}{n_t}) = 1$, we also have $m < \sigma(m)$ and we are done.

Problem 6.48. (**Moscow 2013**) For a positive integer m, we denote by $S(m)$ the sum of digits of m. Prove that for any positive integer n, there exists an integer k such that

$$S(k) = n, \quad S(k^2) = n^2, \quad S(k^3) = n^3.$$

Solution. We prove by induction on n, the following, stronger statement: For any positive integer n, there exists a positive integer k, such that the decimal representation of k consists only of 1's and 0's, the decimal representation of k^2 consists only of 0's, 1's and 2's, and

$$S(k) = n, \quad S(k^2) = n^2, \quad S(k^3) = n^3.$$

For $n = 1$ we simply take $k = 1$. Assume now that the result holds for some $n \geq 1$ and let k_n be the corresponding desired number. Let d be the number of digits of k_n. Now take $k_{n+1} = k_n + 10^{10d}$. We prove that this number satisfies the conditions for $n+1$. We clearly have $S(k_{n+1}) = 1 + S(k_n) = n + 1$ and

$$S(k_{n+1}^2) = S(k_n^2 + 2k_n 10^{10d} + 10^{20d}) = S(k_n^2) + 2S(k_n) + 1 = (n+1)^2,$$

since $k_n^2 < 10^{2d}$, $2k_n 10^{10d} < 10^{20d}$, so the numbers k_n^2, $2k_n \cdot 10^{10d}$ and 10^{20d} do not share common digits which can affect their addition. Moreover, since k_n^2 consists only of the digits 0, 1 and 2, the same holds for k_{n+1}^2. By a similar argument we have that the number

$$k_{n+1}^3 = k_n^3 + 3k_n^2 \cdot 10^{10d} + 3k_n 10^{20d} + 10^{30d}$$

satisfies $S(k_{n+1}^3) = (n+1)^3$. This completes our proof.

Problem 6.49. Establish whether there exist positive integers $a_1 < a_2 < \ldots < a_n < \ldots$ such that every number in the sequence a_1^2, $a_1^2 + a_2^2, \ldots$, $a_1^2 + a_2^2 + \ldots + a_n^2, \ldots$ is the square of a positive integer.

Solution. Let us take $a_1 = 3$ and for $k \geq 1$ we set

$$a_{k+1} = \frac{a_1^2 + a_2^2 + \ldots + a_k^2 - 1}{2}.$$

We first show that for $n \geq 2$, a_n is an even number. We do this by induction on n.
For $n = 2$, we have $a_2 = \frac{3^2-1}{2} = 4$.
Assume now that the result holds for all numbers $a_2, \ldots a_n$, where $n \geq 2$.
By definition we have

$$a_{n+1} = \frac{a_1^2 + a_2^2 + \ldots + a_n^2 - 1}{2}.$$

As all of $a_2, \ldots, a_n$ are divisible by 2, we have that $a_2^2 + \ldots + a_n^2$ is divisible by 4. So by adding $a_1^2 - 1 = 8$, we obtain a number divisible by 4. Thus

$$a_{n+1} = \frac{a_1^2 + a_2^2 + \ldots + a_n^2 - 1}{2}$$

is even, as we wanted.
Note that $a_{k+1} > a_k$ since

$$a_{k+1} - a_k = \frac{a_1^2 - 2 + \ldots + a_{k+1}^2 + (a_k - 1)^2}{2} > 0,$$

for any $k \geq 1$.
On the other hand, we have that

$$a_1^2 + a_2^2 + \ldots + a_k^2 + a_{k+1}^2 = \left(\frac{a_1^2 + a_2^2 + \ldots + a_k^2 + 1}{2}\right)^2,$$

and by what we proved above, $\dfrac{a_1^2 + a_2^2 + \ldots + a_k^2 + 1}{2}$ is a positive integer for any $k \geq 1$.
Therefore, the sequence we constructed satisfies all the conditions of the problem.

Problem 6.50. For which pairs (a, b) of positive integers do there exist only finitely many positive integers n such that $n^2 \mid a^n + b^n$?

Solution. Let (a, b) be a pair of positive integers for which there exist only finitely many positive integers n such that $n^2 \mid a^n + b^n$.

We begin by proving that if a and b have different parity, then $a = 2$ and $b = 1$ (or vice-versa).

So let us assume that $a \geq 3$. We show by induction on k that the sequence $(n_k)_{k\geq 1}$ defined by $n_1 = 1$ and

$$n_{k+1} = \frac{a^{n_k} + b^{n_k}}{n_k}, \quad \forall k \geq 1$$

satisfies $n_k \in \mathbb{N}$, $n_k^2 \mid a^{n_k} + b^{n_k}$, and $n_1 < n_2 < \ldots$.

If $k = 1$, the result clearly holds, as $n_1 = 1$ and $1 \mid (a + b)$.

For the induction step, we prove that if the result holds for $k \geq 1$, then it also holds for $k + 1$.

From the induction hypothesis we have that $n_k \in \mathbb{N}$,

$$n_k^2 \mid a^{n_k} + b^{n_k} \quad \text{and} \quad n_1 < n_2 < \ldots < n_k.$$

Note that $a^{n_k} + b^{n_k}$ is an odd number and therefore n_k is odd as well. Let $a^{n_k} + b^{n_k} = n_k^2 \cdot m_k$, where m_k is an odd number. We have that

$$n_{k+1} = \frac{a^{n_k} + b^{n_k}}{n_k} > \frac{3^{n_k}}{n_k} > n_k :$$

Indeed, the number of terms in the sequence $\sqrt{3}, \sqrt{3}^2, \ldots, \sqrt{3}^{n_k}$ is equal to n_k and

$$\sqrt{3}^{m+1} - \sqrt{3}^m = \sqrt{3}^m(\sqrt{3} - 1) \geq \sqrt{3}(\sqrt{3} - 1) > 1,$$

where $m \in \mathbb{N}$. Thus $\sqrt{3}^{n_k} > n_k$, hence $3^{n_k} > n_k^2$.

It follows that

$$m_k = \frac{a^{n_k} + b^{n_k}}{n_k^2} > 1.$$

We have that m_k is an odd number, hence

$$\begin{aligned}
a^{n_{k-1}} + b^{n_{k-1}} &= a^{n_k \cdot m_k} + b^{n_k \cdot m_k} \\
&= (a^{n_k} + b^{n_k})((a^{n_k})^{m_k - 1} - (a^{n_k})^{m_k - 2} b^{n_k} + \ldots + (b^{n_k})^{m_k - 1}) \\
&= n_k n_{k+1}(n_{k+1} \cdot A_k + m_k (b^{n_k})^{m_k - 1}) \\
&= n_{k+1}^2 (A_k \cdot n_k + (b^{n_k})^{m_k - 1})),
\end{aligned}$$

where $A_k \in \mathbb{Z}$. Thus, $n_{k+1}^2 \mid (a^{n_{k+1}} + b^{n_{k+1}})$ and $n_1 < n_2 < \ldots < n_k < n_{k+1}$.

This shows that we cannot have $a \geq 3$. The same proof applies if we assume $b \geq 3$. It follows that $a = 2$, $b = 1$ (or vice-versa). When $a = 2$, $b = 1$ we obtain $n^2 \mid (2^n + 1)$. It is a nice exercise to show that this can only happen when $n = 1$ or $n = 3$.

We now treat the case when a and b are even numbers. We consider the sequence $(n_k)_{k\geq 1}$ given by $n_k = 2^k$. From Bernoulli's inequality, it follows that $2^{k-1} \geq k$. Hence $(2^k)^2 \mid \left(a^{2^k} + b^{2k}\right)$, for any $k \geq 1$.

In a similar way one can prove that if $p \in \mathbb{N}$ and $p > 1$, $p \mid a$, $p \mid b$, then $(p^k)^2 \mid a^{p^k} + b^{p^k}$, for any $k \geq 1$.

This leaves us with the case when a and b are odd and coprime. Let $a + b = 2^k \cdot m$, where $k \in \mathbb{N}$ and m is odd.

First suppose $m > 1$. We prove by induction on l, that if we consider the sequence $n_1 = 1$,

$$n_{l+1} = \frac{a^{n_l} + b^{n_l}}{2^k \cdot n_l}, \quad \text{for} \quad l \geq 1,$$

then $n_l \in \mathbb{N}$, $2^k \cdot (n_l)^2 \mid (a^{n_l} + b^{n_l})$, $2^{k+1} \nmid (a^{n_l} + b^{n_l})$ and $n_1 < n_2 < \ldots < n_l$.

If $l = 1$, then $2^k \cdot 1^2 \mid a + b$ and $2^{k+1} \nmid a + b$.

For the induction step we show that if the statement holds for l, then it also holds for $l + 1$. We have that $a + b \geq 3 \cdot 2^k$, hence

$$n_{l+1} = \frac{a^{n_l} + b^{n_l}}{2^k \cdot n_l} \geq \frac{\left(\frac{a+b}{2}\right)^{n_l}}{2^{k-1} \cdot n_l} \geq \frac{2^{k-1} \cdot 3^{n_l}}{2^{k-1} \cdot n_l} > n_l.$$

On the other hand, we have that $n_{l+1} > 1$ is odd. Let $n_{l+1} = n_l \cdot m_l$, where $m_l > 1$ is odd. It follows that

$$a^{n_{l+1}} + b^{n_{l+1}} = (a^{n_l})^{m_l} + (b^{n_l})^{m_l}$$

$$= 2^k n_l n_{l+1}((2^k n_l n_{l+1} - b^{n_l})^{m_l - 1} - (2^k n_l n_{l+1} - b^{n_l})^{m_l - 2} \cdot b^{n_l} + \ldots + (b^{n_l})^{m_l - 1})$$
$$= 2^k n_l n_{l+1}(2^k n_l n_{l+1} A_l + m_l b^{n_l(m_l - 1)}),$$

where $A_l \in \mathbb{Z}$. Therefore, the number

$$a^{n_{l+1}} + b^{n_{l+1}} = 2^k n_{l+1}^2 (2^k n_l^2 A_l + b^{n_l(m_l-1)})$$

is divisible by $2^k \cdot n_{l+1}^2$ and it is not divisible by 2^{k+1}.

This proves that we must have $m = 1$. Thus, $a + b = 2^k$, a and b are relatively prime.

Let $n > 1$ and
$$n^2 \mid a^n + b^n. \tag{6}$$

Note that n is odd, as for n even, we would get $4 \nmid (a^n + b^n)$, while both a^n and b^n are 1 modulo 4.

Let p be the least prime divisor of n. Then p is odd and from (6) we deduce that $p \nmid a$, $p \nmid b$ (because if it divides one, it must divide the other as well).

By Fermat's little theorem, we obtain that

$$p \mid (a^{p-1} - 1) - (b^{p-1} - 1) = a^{p-1} - b^{p-1}.$$

On the other hand, we have that $a^n + b^n \mid a^{2n} - b^{2n}$. Thus, $p \mid a^{2n} - b^{2n}$. If $a = b$, then $a = 1$, $b = 1$ and these values satisfy the conditions of the problem.

If $a > b$, then

$$p \mid (a^{2n} - b^{2n}, a^{p-1} - b^{p-1}) = a^{(2n,p-1)} - b^{(2n,p-1)} = a^2 - b^2.$$

We also have that $a + b = 2^k$, therefore $p \mid a - b$ and $p \mid (a^n - b^n)$. On the other hand, $p \mid a^n + b^n$. Hence, $p \mid a$ and $p \mid b$. This leads to a contradiction.

Thus, up to permuting a and b, the pairs (a, b) satisfying the conditions of the problem are $(2, 1)$, $(1, 1)$ and $(2^k - 2l + 1, 2l - 1)$, where $k \in \mathbb{N}$, $k \geq 2$ and $l = 1, 2, \ldots, 2^{k-2}$.

7 Combinatorics

Problem 7.1. (**IMO 2002**) Let n be a positive integer and S the set of points $(x,\, y)$ in the plane, where x and y are non-negative integers such that $x + y < n$. The points of S are colored in red and blue so that if (x, y) is red, then (x', y') is red as long as $x' \leq x$ and $y' \leq y$. Let A be the number of ways to choose n blue points such that all their x-coordinates are different and let B be the number of ways to choose n blue points such that all their y-coordinates are different. Prove that $A = B$.

Solution. Let a_k be the number of blue points with x-coordinate equal to k and b_k the number of blue points with y-coordinate equal to k. Using $n-1$ times the law of the product, we have that $A = a_0 a_1 \cdots a_{n-1}$. In the same way we have that $B = b_0 b_1 \cdots b_{n-1}$.

We are going to prove that the numbers $a_0, a_1, \ldots, a_{n-1}$ are a permutation of the numbers $b_0, b_1, \ldots, b_{n-1}$ using strong induction. If we do this, we establish that their product is the same, so $A = B$.

If $n = 1$, S consists of only one point. Then a_0 and b_0 are both 1 or 0 depending on whether the point is painted blue or red, so $a_0 = b_0$.

Now assume that the assertion holds for every $k \leq n$ and we want to prove it for $n + 1$.

There are two cases:

1) Every point (x, y) with $x + y = n$ is blue. If this occurs, we let S' be the set of points (x, y) such that $x + y < n$ and let a'_k be the number of points in S' with k in the x-coordinate and b'_k the number of points in S' with k in the y-coordinate. Then $a'_0, a'_1, \ldots, a'_{n-1}$ is a permutation of $b'_0, b'_1, \ldots, b'_{n-1}$ from the induction hypothesis. Moreover, we know that $a_n = b_n = 1$ and $a_k = a'_k + 1$, $b_k = b'_k + 1$ for every $k < n$. So $b_0, b_1, \ldots, b_n$ is a permutation of $a_0, a_1, \ldots, a_n$.

2) At least one of the points (x, y) with $x + y = n$ is red. Suppose that the point $(k, n-k)$ is red. Let S_1 be the set of points (x, y) with $x + y < n + 1$, $x < k$ and $y > n - k$. Notice that $a_0, a_1, \ldots, a_{k-1}$ and $b_{n-k+1}, b_{n-k+2}, \ldots, b_n$ are the a_i and b_i that would be associated with S_1. Therefore, from the induction hypothesis, one must be a permutation of the other. In a similar

manner we obtain that $a_{k+1}, a_{k+2}, \ldots, a_n$ is a permutation of $b_0, b_1, \ldots, b_{n-k+1}$. Finally, $a_k = b_{n-k} = 0$, so putting all together, we have that $a_0, a_1, \ldots, a_n$ is a permutation of $b_0, b_1, \ldots, b_n$.

Problem 7.2. (**IMO 2013 shortlist**) Let n be a positive integer. Find the smallest integer k with the following property: Given any real numbers $a_1, \ldots, a_d$ such that $a_1 + a_2 + \ldots + a_d = n$ and $0 \leq a_i \leq 1$ for $i = 1, 2, \ldots, d$, it is possible to partition these numbers into k groups (some of which may be empty) such that the sum of the numbers in each group is at most 1.

Solution. The answer to the problem is $k = 2n - 1$. If $d = 2n - 1$ and $a_1 = \ldots = a_{2n-1} = \frac{n}{2n-1}$, then each group in such a partition can contain at most one number, since $\frac{2n}{2n-1} > 1$. Therefore, $k \geq 2n - 1$. It remains to show that a suitable partition into $2n - 1$ groups always exists.

We proceed by induction on d. For $d \leq 2n - 1$, the result is trivial. If $d \geq 2n$, then, since

$$(a_1 + a_2) + \ldots + (a_{2n-1} + a_{2n}) \leq n,$$

we may find two numbers a_i, a_{i+1} such that $a_i + a_{i+1} \leq 1$. We "merge" these two numbers into one new number $a_i + a_{i+1}$. By the induction hypothesis, a suitable partition exists for the $d - 1$ numbers $a_1, \ldots, a_{i-1}, a_i + a_{i+1}, a_{i+2}, \ldots, a_d$. This induces a suitable partition for $a_1, \ldots, a_d$.

Problem 7.3. (**TOT 2002**) The spectators are seated in a row with no empty places. Each is in a seat which does not match the spectator's ticket. An usher can order two spectators in adjacent seats to trade places unless one of them is already seated correctly. Is it true that from any initial arrangement, the usher can place all the spectators in their correct seats?

Solution. The answer is yes. We prove this by induction on the number n of spectators.

The case $n = 2$ is clear, since a single switch fixes everything. Assume now that the result holds for all numbers up to some positive integer $n \geq 2$ and consider the situation when we have $n + 1$ spectators. Denote by S_k the spectator which has the ticket with number k.

Assume S_{n+1} is seated in the seat m for some $m \leq n$. If the spectators in seats m to $n+1$ are in order $S_{n+1}, S_m, S_{m+1}, \ldots, S_n$, then we can fix them all by successively swapping S_{n+1} until he reaches his correct position. We are then done by the induction hypothesis applied to the first $m-1$ seats. Otherwise, there is a least $l > m$ such that the seat l is occupied by some S_x with $x \neq l-1$. By construction, for $m < k < l$, the seat k is occupied by S_{k-1} so we have in seats m to l the spectators $S_{n+1}, S_m, \ldots, S_{l-2}, S_x$. We perform a chain of switches to move S_x from the seat l to seat $m+1$ and we obtain another derangement $S_{n+1}, S_x, S_m, \ldots, S_{l-2}$ where now S_k is in the seat $k+2$ for $m \leq k < l-1$. This brings S_x to the seat $m+1$ and now we can swap him with S_{n+1}, bringing S_{n+1} one seat closer to his actual place without putting anyone else in his correct seat. We repeat this process until S_{n+1} is in seat $n+1$ and then use the induction hypothesis.

Problem 7.4. (**USSR 1991**) Several (more than two) consecutive positive integers $1, 2, \ldots, n$ are written on a blackboard. In one move, it is permitted to erase any pair of numbers, say p and q and replace them by the numbers $p+q$ and $|p-q|$ instead. In several moves, a student was able to make all numbers on the blackboard equal to k. Find all possible values of k.

Solution. The answer to the question is that k can be any number of the form 2^s for any s satisfying the inequality $2^s \geq n$.

First notice that after any move, there are only non-negative integers on the board. If the sum and the difference of two integers are both divisible by an odd number d, then these numbers themselves are also divisible by d. Since 1 was among the initial numbers, k has no odd prime divisors. Thus $k = 2^s$. Since at any move, the maximum of the numbers on blackboard does not decrease, $k = 2^s \geq n$.

To prove that each number $k = 2^s \geq n$ can be obtained, we shall use induction. Observe that if at some stage zero appears on the blackboard, then each number can be doubled in two moves:

$$(0, a) \to (a, a) \to (0, 2a).$$

Thus, if the following zeros and powers of 2 are on the blackboard:

$$0, \ldots, 0, 2^{k_1}, \ldots, 2^{k_m},$$

with at least one zero, then after several moves, it is possible to obtain

$$0, 2^k, \ldots, 2^k, \quad k = \max(k_1, \ldots, k_m).$$

Lemma. There is a sequence of moves that transforms the numbers $1, 2, \ldots, n$, where $n \geq 3$ into the numbers $0, 2^{s+1}, \ldots, 2^{s+1}$, where s is the greatest integer satisfying the inequality $2^s < n$.

Proof. It is easy to check that the lemma holds for $3 \leq n \leq 6$. For example,

$$1, 2, 3, 4, 5 \to 1, 2, 2, 4, 8 \to 0, 1, 4, 4, 8 \to 0, 8, 8, 8, 8.$$

Now assume that $n > 6$. Suppose that for $n' < n$ the statement of the lemma is true. Let us represent n in the form $n = 2^s + b$, where $0 < b \leq 2^s$. If $b = 2^s$, then $n = 2^{s+1}$ and the inductive step is trivial, since by the induction hypothesis $1, 2, \ldots, n-1$ can be transformed into $0, 2^{s+1}, \ldots, 2^{s+1}$. Suppose that $0 < b < 2^s$ (note that this implies either $n = 7$ or $n > 8$) and divide $1, 2, \ldots, n$ into four groups:

a) $1, 2, \ldots, 2^s - b - 1$;
b) 2^s;
c) $2^s - 1, \ldots, 2^s - b$;
d) $2^s + 1, \ldots, 2^s + b$.

After b moves involving the pairs $(2^s + i, 2^s - i)$, we get the following four groups:

a) $1, 2, \ldots, 2^s - b - 1$;
b) 2^s;
c) $2, 4, 6, \ldots, 2b$;
d) $2^{s+1}, 2^{s+1}, \ldots, 2^{s+1}$.

Since $b = 3$ for $n = 7$ and $(2^s - b - 1) + b = 2^s - 1 \geq 7$ for $n > 8$, in the first or third group, there are more than two numbers, so the induction hypothesis can be applied to this group giving at least one zero. For the other group, either it has at least three elements and we can also apply the induction hypothesis or this group contains only powers of 2. In both cases we get only zeros and powers of 2 on the blackboard. This proves the lemma.

Let us now turn the original problem. Suppose that we want to obtain the numbers $2^m, 2^m, \ldots, 2^m$ on the blackboard, where $2^m \geq n$. First, using the lemma, we obtain the numbers $0, 2^{s+1}, \dots, 2^{s+1}$. Then, by doubling, if necessary, we can obtain $0, 2^m, \dots, 2^m$, and finally $2^m, 2^m, \dots, 2^m$.

Problem 7.5. (**USSR 1990**) We are given $4m$ coins, among which exactly half of the coins are counterfeit. All genuine coins have equal weights, all counterfeit coins also have equal weights, but a counterfeit coin is lighter than a genuine one. How can one determine all counterfeit coins in no more than $3m$ weighings, using a balance without weights?

Solution. We prove by induction the following, stronger statement: for an even number n, if n coins are given and if it is known how many coins among them are counterfeit, then it is possible to determine all counterfeit coins in $[3n/4]$ weighings:

For $n = 2$ the statement is true, since everything can be determined in one weighing. Suppose that $n \geq 4$. Let us compare two arbitrary coins. If they have different weights, we can classify them and since

$$\left[\frac{3(n-2)}{4}\right] + 1 \leq \left[\frac{3n}{4}\right],$$

the problem reduces to the same problem for $n - 2$ coins.

So suppose instead that the compared coins have equal weights. Then we shall compare this pair of coins with another pair. If the weights differ, then we shall compare the two coins of the second pair and all four coins after that will be classified. Since

$$\left[\frac{3(n-4)}{4}\right] + 3 \leq \left[\frac{3n}{4}\right],$$

the problem reduces to the problem for $n-4$ coins. If the two pairs have equal weights, we compare these four coins with another foursome. If these two sets of coins differ in weight, then the first can be classified and the problem again reduces to $n - 4$ coins. If the two sets have equal weights, we compare these eight coins with another eight and so on.

If at some stage of this procedure a group of 2^m coins differs from another group of 2^m coins, then the coins of the first group (which are identical) can be classified. Since

$$\left\lceil \frac{3(n-2^m)}{4} \right\rceil + (m-1) \le \left\lceil \frac{3n}{4} \right\rceil,$$

the problem reduces to the problem for $n - 2^m$ coins. Finally, if there are not enough coins to form another group of 2^m coins, then $2^m > n/2$ and the 2^m coins can be classified by choosing $n - 2^m$ of them and comparing them with the remaining $n - 2^m$ coins. This completes our proof.

Problem 7.6. (**IMO 2006 shortlist**) An (n, k)-tournament is a contest with n players held in k rounds such that:

i) Each player plays in each round, and every two players meet at most once.

ii) If player A meets player B in round i, player C meets player D on round i, and player A meets player C in round j, then player B meets player D in round j.

Determine all pairs (n, k) for which there exists an (n, k)-tournament.

Solution. Let t be the greatest integer such that 2^t divides n. We will show that there exists an (n, k)-tournament if and only if $k \le 2^t - 1$.

We first prove that if $k \le 2^t - 1$, then there exists an (n, k)-tournament. Since we may partition our n players into $\frac{n}{2^t}$ different groups of size 2^t, it suffices to prove this for $n = 2^t$.

For convenience, let us label the 2^t players with the elements of $(\mathbb{Z}/2\mathbb{Z})^t$, and label the different rounds with distinct non-zero elements of $(\mathbb{Z}/2\mathbb{Z})^t$. In the round with label j, let player a meet player $a + j$. We then have a $(2^t, k)$ tournament, for if $a - b = c - d = i$, then $a - c = b - d$, for all $a, b, c, d \in (\mathbb{Z}/2\mathbb{Z})^t$.

We next prove that if $k > 2^t - 1$, then there is no (n, k)-tournament. For this, we first prove an intermediate result:

Lemma. The number of players in any minimal sub-tournament of an (n, k)-tournament is a power of 2.

Proof. We induct on k. The case $k = 0$ is trivial.

Suppose now that all minimal sub-tournaments of any $(n, k-1)$-tournament have sizes that are powers of 2. Then in any (n, k)-tournament, the minimal sub-tournaments, ignoring the last round, are powers of 2. Let S be the set of players for such a minimal tournament, and for any player a in the (n, k)-tournament, let $K(a)$ be the player whom a meets in round k. Then either $K(S) = S$, or $K(S)$ and S are disjoint; furthermore, K induces a bijection from S to $K(S)$, so S and $K(S)$ have the same cardinality. It follows that the minimal sub-tournament of the (n, k)-tournament containing any element of S has size either $|S|$ or $2|S|$, completing the inductive step and proving the lemma.

To finish the question, we show that we cannot have an (n, k)-tournament if $k > 2^t - 1$. Assume the contrary. Then every minimal sub-tournament of our tournament has a multiple of 2^{t+1} players, since each must have a size that is a power of 2, and no two players meet more than once. It follows that 2^{t+1} divides n, which is a contradiction. This completes the proof of our initial claim.

Problem 7.7. In an $m \times n$ rectangular board of mn unit squares, *adjacent* squares are ones with a common edge, and a *path* is a sequence of squares in which any two consecutive squares are adjacent. Each square of the board can be colored black or white. Let N denote the number of colorings of the board such that there exists at least one black path from the left edge of the board to its right edge, and let M denote the number of colorings in which there exist at least two non-intersecting black paths from the left edge to the right edge. Prove that $N^2 \geq 2^{mn} M$.

Solution. Suppose that a two-sided $m \times n$ board T is considered, where exactly k of the squares are transparent. A transparent square is colored only on one side (then it looks the same from the other side), while a non-transparent one needs to be colored on both sides, not necessarily in the same color.

Let $C = C(T)$ be the set of colorings of the board in which there exist two black paths from the left edge to the right edge, one on top and one underneath, not intersecting at any transparent square. If $k = 0$, then $|C| = N^2$. We prove by induction on k that $2^k |C| \leq N^2$. This inequality will imply the statement of the problem, as $|C| = M$ for $k = mn$.

Let q be a fixed transparent square. Consider any coloring B in C. If q is converted into a non-transparent square, a new board T' with $k-1$ transparent square is obtained, so by the induction hypothesis, $2^{k-1}|C(T')| < N^2$. Since B contains two black paths at most one of which passes through q, coloring q in either color on the other side will result in coloring C'. Hence $|C(T')| \geq 2|C(T)|$, implying $2^k|C(T)| \leq N^2$ and finishing the induction.

Problem 7.8. (**IMO shortlist 1998**) Let $U = \{1, 2, \ldots, n\}$, where $n \geq 3$. A subset S of U is said to be *split* by an arrangement of the elements of U if an element not in S occurs in the arrangement somewhere between two elements of S. For example, 13542 splits $\{1, 2, 3\}$ but not $\{3, 4, 5\}$. Prove that for any $n-2$ subsets of U, each containing at least 2 and at most $n-1$ elements, there is an arrangement of the elements of U which splits all of them.

Solution. The proof is by induction on n. The base case $n = 3$ is clear.

For the induction step, assume that the result holds for all integers which are less than some $n \geq 4$. Consider the subsets $U_1, U_2, \ldots, U_{n-2}$ of S. We want to remove an element x from S and apply the induction hypothesis for some $n-3$ of $U_1 \backslash \{x\}, \ldots, U_{n-2} \backslash \{x\}$. Assume k of the sets have $n-1$ elements. We wish that after the removal, they are left with $n-2$ elements, so that we could use the induction hypothesis. This gives us k forbidden values for x, one for each set. Next, we would like to ensure no set of two elements becomes a singleton after removal. Unfortunately this is not always possible, but we can ensure that at most one becomes a singleton. This is because we have $n-k$ elements and at most $n-2-k$ two-element sets, so we cannot have each element belong to at least two of these sets.

So perform the removal we described above. Next, we ignore one of the obtained sets. Let it be the singleton or an arbitrary set if there is no singleton. Suppose we have removed $U_{n-2} \backslash \{x\}$. By the induction hypothesis, there is an arrangement of $\{1, 2, \ldots, n\} \backslash \{x\}$ that splits $U_1 \backslash \{x\}, U_2 \backslash \{x\}, \ldots, U_{n-3} \backslash \{x\}$. It is clear that however we insert x into this arrangement, the sets $U_1, U_2, \ldots, U_{n-3}$ will be split. So it remains to ensure U_{n-2} is split. But this is simple to achieve: if $x \in U_{n-2}$ then we place x either at the beginning or at the end of the arrangement, and one of these two variants will satisfy our requirement; if $x \notin U_{n-2}$, then we place x between two elements of U_{n-2} and we are done.

Problem 7.9. Let $n \neq 4$ be a positive integer. Consider a set $S \subseteq \{1, 2, \ldots, n\}$ with $|S| > \left\lceil \frac{n}{2} \right\rceil$. Prove that there exist $x, y, z \in S$ with $x + y = 3z$.

Solution. We proceed by induction on n. For $n \leq 15$ we can verify our statement directly.

For the induction step, assume that the result holds for all integers less than some $n \geq 16$. Let A be a set that has no triples x, y, z with $x + y = 3z$. We need to prove that $|A| \leq \left\lceil \frac{n}{2} \right\rceil$.

First suppose A contains some element t with $\frac{n+5}{3} < t \leq \frac{2n}{3}$. Then $3t - n - 1 \geq 5$, so by the induction hypothesis $A \cap \{1, 2, \ldots, 3t - n - 1\}$ contains at most $\left\lceil \frac{3t-n-1}{2} \right\rceil$ elements. Group the numbers larger than $3t - n - 1$ into pairs $(3t - n, n)$, $(3t - n + 1, n - 1), \ldots$ such that the sum of the numbers in each pair is $3t$. Since $t \in A$, A can contain at most one element out of each pair (and if t is even then A cannot contain $3t/2$). Thus A contains at most $\left\lfloor \frac{2n+1-3t}{2} \right\rfloor$ elements larger than $3t - n - 1$. Hence

$$|A| \leq \left\lceil \frac{3t - n - 1}{2} \right\rceil + \left\lfloor \frac{2n + 1 - 3t}{2} \right\rfloor \leq \frac{3t - n}{2} + \frac{2n + 1 - 3t}{2} = \frac{n + 1}{2}.$$

Hence

$$|A| \leq \left\lceil \frac{n}{2} \right\rceil,$$

as desired.

Thus we may assume A does not contain any element in this range. Let

$$B = A \cap \left[1, \frac{n+5}{3}\right] \quad \text{and} \quad C = A \cap \left[\frac{2n+1}{3}, n\right].$$

Since $n \geq 16$, $\left\lfloor \frac{n+5}{3} \right\rfloor \geq 7$. Therefore by the induction hypothesis we have

$$|B| \leq \left\lceil \frac{1}{2} \left\lfloor \frac{n+5}{3} \right\rfloor \right\rceil.$$

Also we obviously have $|C| \leq n - \left\lfloor \frac{2n}{3} \right\rfloor$. Thus we get

$$|A| \leq \left\lceil \frac{1}{2} \left\lfloor \frac{n+5}{3} \right\rfloor \right\rceil + \left(n - \left\lfloor \frac{2n}{3} \right\rfloor\right).$$

This upper bound is only sufficient to solve the problem if $n = 6k+3$ in which case it becomes

$$|A| \leq (k+1) + (2k+1) = 3k+2 = \left\lceil \frac{n}{2} \right\rceil.$$

Still with a little case checking we can knock off the other possibilities. If $n \neq 6k+4$, then we find that the bound above gives

$$|A| \leq \left\lceil \frac{n}{2} \right\rceil + 1.$$

Thus we only need to show that we cannot have equality in all the bounds above. For the equality in the bound on $|C|$, we need C to contain every number larger than $\left\lfloor \frac{2n}{3} \right\rfloor$. But A cannot contain both x and $2x$ for any x (since $x + 2x = 3x$). Thus the largest element of B is at most

$$\left\lceil \frac{1}{2} \left(\left\lfloor \frac{2n}{3} \right\rfloor + 1 \right) \right\rceil - 1 = \left\lfloor \frac{n}{3} \right\rfloor.$$

Thus our upper bound on $|B|$ becomes $\left\lceil \frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor \right\rceil$ so the upper bound on $|A|$ becomes

$$|A| \leq \left\lceil \frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor \right\rceil + \left(n - \left\lfloor \frac{2n}{3} \right\rfloor \right),$$

which one can check implies

$$|A| \leq \left\lceil \frac{n}{2} \right\rceil \quad \text{for} \quad n \neq 6k+4.$$

That leaves only the case $n = 6k+4$. In this case $B = A \cap \{1, 2, \ldots, 2k+3\}$ and $C = A \cap \{4k+3, \ldots, 6k+4\}$. In this case our upper bound on $|A|$ reads

$$|A| \leq (k+2) + (2k+2) = 3k+4 = \left\lceil \frac{n}{2} \right\rceil + 2.$$

If $2k+3 \in B$, then $4k+6$ cannot be in C and the inequality improves to

$$|A| \leq \left\lceil \frac{n}{2} \right\rceil + 1.$$

Thus we are done unless we have equality in all the remaining places. In particular, C must contain all the numbers $4k+7,\ldots,6k+4$. Hence B cannot contain any of the numbers $3(2k+3)-(4k+7),\ldots,3(2k+3)-(6k+4)$. But these are all the numbers from 5 to $2k+2$. That leaves B with $k+2$ numbers that must come from just some subset of $\{1,2,3,4\}$ and $2k+3$. Except for $k=2$ and $n=16$ this is a clear contradiction (if 2 is in B, then 1 and 4 are not in B). For $n=16$, this leaves just the possibility $A=\{1,3,4,7,11,12,13,15,16\}$, which fails since $1+11=3\cdot 4$.
If $2k+3\notin B$, then by the induction hypothesis B has at most $k+1$ elements and we again improve to $|A|\le\left\lceil\frac{n}{2}\right\rceil+1$. Thus we are done unless we have equality in all the remaining places. Thus C must contain all the numbers $4k+3,\ldots,6k+4$ and hence B cannot contain $2k+2$. If $n\ne 16$, then the induction hypothesis implies $2k+1\in B$. But then B cannot contain any number of the form $3(2k+1)-c$ for $c\in C$, which leaves B with just $2k+1$, a contradiction. If $n=16$, then we get the extra possibility $A=\{1,3,4,11,12,13,14,15,16\}$, which again fails since $1+11=3\cdot 4$.

Problem 7.10. A word consists of n letters from the alphabet $\{a,b,c,d\}$. A word is called *convoluted* if it has two consecutive identical blocks of letters. For example, *caab* and *cababdc* are convoluted, but *abcab* is not. Prove that the number of non-convoluted words with n letters is greater than 2^n.

Solution. We prove the result by induction on n. For $n\le 3$, we can easily check the cases manually.

For the induction step, suppose $n\ge 4$ and that the result holds for all positive integers less than n. To solve the problem, let us first try to establish a recurrence for the number a_n of convoluted words. We also denote the number of non-convoluted (simple) words by b_n. We clearly have $a_n+b_n=4^n$. Since we want to use recursion, let us delete the last letter from a word of n letters. If the remaining word of $n-1$ letters is convoluted, so is the original one. If not, then the two consecutive blocks of letters must occur at the end so the word if of the form AXX. Moreover, AX is a simple word thus, we can obtain such a word in the following way: choose a simple word T with length l at least $\frac{n}{2}$ and add the last $n-l$ letters of the word at the end. Note, however that some convoluted words may be counted twice, as, for example *abacabcabc*, but every word is listed. For this reason, we can deduce only the inequality that

$$a_n \leq 4a_{n-1} + (b_{[\frac{n}{2}]} + \dots b_{n-1}).$$

We can rewrite this as

$$b_n \geq 4b_{n-1} - (b_{[\frac{n}{2}]} + \dots b_{n-1}).$$

Now it is an easy induction to show that $b_n \geq 2b_{n-1}$. The base case is easy since $b_1 = 4$ and $b_2 = 12$. For the inductive step, we note that by the induction hypothesis

$$b_{[\frac{n}{2}]} + \dots b_{n-1} \leq b_{n-1}\left(1 + \frac{1}{2} + \frac{1}{4} + \dots\right) = 2b_{n-1},$$

and hence

$$b_n \geq 4b_{n-1} - 2b_{n-1} = 2b_{n-1}.$$

From this, it is a trivial induction to show that $b_n > 2^n$.

Problem 7.11. (**IMO 2006 shortlist**) We have $n \geq 2$ lamps $L_1, \dots, L_n$ in a row, each of them being either on or off. Every second we simultaneously modify the state of each lamp as follows: if the lamp L_i and its neighbors (only one neighbor for $i = 1$ or $i = n$, two neighbors for other i) are in the same state, then L_i is switched off; otherwise, L_i is switched on.

Initially, all the lamps are off except the leftmost one which is on.

a) Prove that there are infinitely many integers n for which all the lamps will eventually be off.

b) Prove that there are infinitely many integers n for which the lamps will never be all off.

Solution. We order the lamps so that L_1 is the leftmost lamp.

For part a) of the question, we show that for $n = 2^k$, all lamps will be switched on in $n - 1$ steps and off in n steps. We do this by induction on k, with the base case $k = 1$ being clear.

Assume now that the result holds for some $k \geq 1$ and let $n = 2^{k+1}$. Let $A = \{L_1, \dots, L_{2^k}\}$ and $B = \{L_{2^k+1}, \dots, L_{2^{k+1}}\}$. Notice that the first $2^k - 1$

steps do not affect the state of the lamps in B. Hence, after $2^k - 1$ steps, the lamps in A will all be on and those in B will all be off. After the 2^k-th step, L_{2^k} and L_{2^k+1} are both one and the other lamps are off. The key observation is that from this point on, the lamps L_i and $L_{2^{k+1}-i}$, for $i = 1, \ldots, 2^k - 1$ will have the same state and there is no other interference between the lamps in A and those in B. As B starts with only the leftmost lamp on, by the induction hypothesis, it will have all its lamps off in 2^k steps. The same will be true for A by our previous observation. So there are $2^k + 2^k = 2^{k+1}$ steps in total, which completes the proof of this part of the question.

For part b), we show that if $n = 2^k + 1$, the lamps will never be all off. After the first step, only L_1 and L_2 will be on. Based on what we proved for a), after $2^k - 1$ steps, all lamps except L_n will be on, so after the 2^k-th step, all lamps will be off except for L_{n-1} and L_n. This position is symmetric with respect to the one we had after the first step, so by periodicity, we will never have all the lamps off.

Problem 7.12. (**IMO 2005 shortlist**) Let k be a fixed positive integer. A company has a special method to sell sombreros. Each customer can convince up to two persons to buy a sombrero after he/she buys one; convincing someone already convinced does not count. Each of these new customers can convince up to two others and so on. If a customer convinces two people to but sombreros and each of these two people in turn makes at least k persons buy sombreros (directly or indirectly), then that customer wins a free instructional video. Prove that if n persons bought sombreros, then at most $\frac{n}{k+2}$ of them got videos.

Solution. Suppose m persons receive videos. We wish to prove $n \geq m(k+2)$. Since this is clear for $m = 0$, let us assume without loss of generality that m is positive. Under this assumption, we will now prove the stronger bound

$$n \geq (m+1)(k+1) + m,$$

by induction on m.

We say a person A is a direct successor of B if B directly convinced A to buy a sombrero. We say A is an indirect successor of B if B caused A to buy a sombrero, directly or indirectly. We will call a person who receives a video a blossom. We will call direct successor of a blossom who is not a blossom him- or her-self a bud.

For the base case of our induction, when there is only one blossom, that blossom must have at least two buds, each of which must have at least k indirect successors. Hence $n \geq 2k+3 = (m+1)(k+1)+m$.

Now, suppose there are m blossoms. Since there are only finitely many blossoms, there exists at least one blossom which has no blossoms as indirect successors. We remove this blossom, one of its buds, and all indirect successors of that bud; we then make the disconnected bud a direct successor of whatever person was a direct successor of the blossom we removed, if there was such a person. All other blossoms stay blossoms; all other buds stay buds. We have thus removed at least $2+k$ people, and we have removed one blossom. If there are n' people remaining, then the inductive hypothesis now tells us

$$n-(2+k) \geq n' \geq (m)(k+1)+(m-1),$$

or

$$n \geq n'+k+2 \geq (m+1)(k+1)+m.$$

Therefore for all m, we have

$$n \geq (m+1)(k+1)+m = m(k+2)+k+1 > m(k+2),$$

as desired.

We note that the bound $n \geq (m+1)(k+1)+m$ is sharp, for it is possible to have blossoms $A_1, \ldots, A_m$, with A_{i+1} a direct successor of A_i, and all buds having exactly k indirect successors.

Problem 7.13. Let r be a positive integer. Consider an infinite collection of sets, each having r elements such that each two of these sets are not disjoint. Prove that there is a set with $r-1$ elements that is not disjoint from any member of the collection.

Solution. Assume this is not the case, so for any $r-1$ elements, there is a set from the collection containing none of its elements. Let us prove by induction on $j \leq r-1$ that we can find a finite number of sets $A_1, A_2, \ldots, A_m$

from our collection such that no set of j elements can meet all $A_1, A_2, \ldots, A_m$ simultaneously. The base case $j = 0$ is clear.

Now assume we have proven the statement for $j = k$, so we have found $A_1, A_2, \ldots, A_m$ such that no set of k elements meets all of them. If there is no set of $k+1$ meeting all of them, the induction step is proven. If there is such a set B with $|B| = k+1$, then $B \subset A_1 \cup A_2 \ldots \cup A_m$, as otherwise B would contain an element not contained in any of $A_1, A_2, \ldots, A_m$, so we could remove it and the remaining set of k elements would contradict the induction hypothesis. As $A_1 \cup A_2 \ldots \cup A_m$ has a finite number of subsets, there are a finite number of such sets: $B_1, B_2, \ldots, B_l$. But as $|B_i| \leq r-1$, there is a set C_i from the collection which does not intersect B_i. Now consider the sets $A_1, A_2, \ldots, A_m, C_1, C_2, \ldots, C_l$. If a set of $k+1$ elements would meet all these sets, then it would be one of $B_1, B_2, \ldots, B_l$, but then it would not meet the corresponding set from $C_1, C_2, \ldots, C_l$, which is impossible. Therefore, the induction step is verified by the collection $A_1, A_2, \ldots, A_m, C_1, C_2, \ldots, C_l$.

Thus we can find a finite numbers of sets $A_1, A_2, \ldots, A_m$ from our collection such that no set of $r-1$ elements meets them all. But then any set of the collection would be contained in $A_1 \cup A_2 \ldots \cup A_m$, as otherwise it would contain an element not belonging to any of $A_1, A_2, \ldots, A_m$ and removing it would produce a set of $r-1$ elements meeting $A_1, A_2, \ldots, A_m$. But $A_1 \cup A_2 \ldots \cup A_m$ is finite, so it has a finite number of subsets, which contradicts the fact that the collection in the problem is infinite.

Problem 7.14. We have n finite sets $A_1, A_2, \ldots, A_n$ such that the intersection of any collection of them has an even number of elements, except for the intersection of all subsets, which has an odd number of elements. Find the least possible number of elements that $A_1 \cup A_2 \cup \ldots \cup A_n$ can have.

Solution. Set $\overline{A_i} = (A_1 \cup A_2 \cup \ldots \cup A_n) \setminus A_i$. For a nonempty subset J of $\{1, 2, \ldots, n\}$ let

$$X_J = \bigcap_{j \in J} A_j \bigcap_{k \notin J} \overline{A_k}.$$

In words, X_J is all elements which are in the sets A_j for $j \in J$ and in no other A_k. From this it is clear that the sets X_J are pairwise disjoint. Also, any element of $\bigcap\limits_{i \in I} A_i$ must be in X_J for some superset J of I. In formulas, this says

$$\bigcup_{J \supseteq I} X_J = \bigcap_{i \in I} A_i.$$

Now let us prove by a downward induction on $|I|$ that $|X_I|$ is odd for all nonempty sets I. The base case $|I| = n$ is just the statement that the intersection of all the A_i has an odd number of elements. For the induction step, assume we know $|X_J|$ is odd for all sets J larger than I. The previous formula shows that $\bigcap_{i \in I} A_i$, which contains an even number of elements by hypothesis, is a disjoint union of $2^{n-|I|}$ of the sets X_J all of which have an odd number of elements by the induction hypothesis except X_I. Since $2^{n-|I|}$ is even, it follows that X_I also has an odd number of elements.

Thus each X_J has an odd number of elements and hence at least one element. Since every point in some A_i is in X_J for some J, we see that

$$\bigcup_{i=1}^{n} A_i = \bigcup_{J \neq \emptyset} X_J$$

has at least $2^n - 1$ elements.

The equality case can be achieved and the construction is inspired by the solution: if we let A_i be set of all numbers between 1 and $2^n - 1$ whose i-th binary digit is 1, then we can see that each of the sets X_J is a singleton. It is easy to check that this example satisfies all the required conditions.

Problem 7.15. (**USSR 1991**) A $k \times l$ minor of an $n \times n$ table consists of all cells which lie on the intersection of any k rows with any l columns. The number $k+l$ is called the semiperimeter of this minor. It is known that several minors, each of semiperimeter not less than n, jointly cover the main diagonal of the table. Prove that these minors jointly cover at least half of all cells.

Solution. We prove the result by induction on n. We denote the statement of the problem for an $n \times n$ table by $P(n)$. We will prove that $P(n)$ is true by proving $P(1)$ and $P(2)$ and the fact that $P(n-2)$ implies $P(n)$.

There is nothing to prove for $P(1)$ and $P(2)$, because for a 1×1 table the only cell is a diagonal one, and for a 2×2 table, the number of diagonal cells is exactly half of the number of all cells.

A cell in the table is characterized by a pair of integers (i,j), where i is the number of the row and j is the number of the column to which this cell belongs.

Observe that in one case we do not actually need the induction. Consider all pairs of cells which are symmetric about the main diagonal; they have coordinates (i,j) and (j,i) for some $i \neq j$. If in each such pair at least one of the two cells is covered, then the minors jointly cover at least $n+(n^2-n)/2 = (n^2+n)/2 > n^2/2$ cells and $P(n)$ is true.

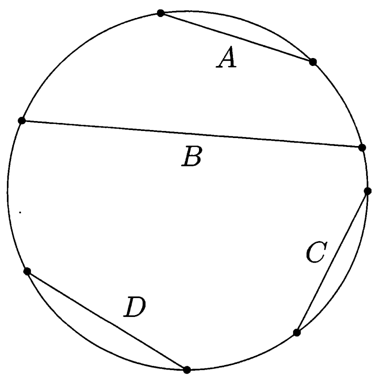
Let us now assume that $P(n-2)$ is true for some $n \geq 3$ and consider an $n \times n$ table. We only have to consider the case in which for some i,j neither cell (i,j) nor cell (j,i) is covered by the minors. Delete the i-th and j-th rows and the i-th and j-th columns and reduce the minors correspondingly if they contained some of these rows or columns. We obtain an $(n-2)\times(n-2)$ table and several reduced minors which again jointly cover the main diagonal. No reduced minor could contain both the i-th row and j-th column simultaneously (otherwise this minor would cover the cell (i,j)). Similarly, no reduced minor could contain simultaneously both the j-th row and the i-th column. Therefore, the semiperimeter of a reduced minor is not less than $n-2$. Now we can apply the induction hypothesis to show that the reduced minors cover at least half of all cells of the reduced table.

To complete the induction, it suffices to prove that among the deleted $4n-4$ cells, at least $2n-2$ were covered by the minors. Consider the minor that covers the cell (i,i). Since its semiperimeter is at least n, it covers at least $n-1$ of the deleted cells in the i-th row and i-th column. The minor that covers (j,j) (note that this could be the same minor) also covers $n-1$ deleted cells in the j-th row and in the j-th column. But since (i,j) and (j,i) are not covered, the two minors jointly cover at least $2(n-1)$ cells. This completes the proof.

Problem 7.16. (**IMO 2013 shortlist**) Let $n \geq 2$ be an integer. Consider all circular arrangements of the numbers $0, 1, \ldots, n$; the $n+1$ rotations of an arrangement are considered to be equal. A circular arrangement is called *beautiful* if, for any four distinct numbers $0 \leq a, b, c, d \leq n$ with $a+c=b+d$, the chord joining the numbers a and c does not intersect the chord joining numbers b and d.

Let M be the number of beautiful arrangements of $0, 1, \ldots, n$. Let N be the number of pairs (x, y) of positive integers such that $x + y \leq n$ and $gcd(x, y) = 1$. Prove that $M = N + 1$.

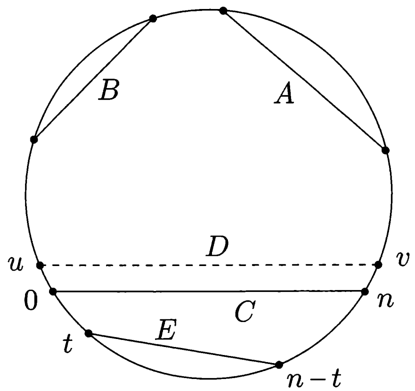
Solution. Given a circular arrangement of $[0, n] = \{0, 1, \ldots, n\}$, we define a k-chord to be a (possibly degenerate) chord whose (possibly equal) endpoints add up to k. We stay that three chords of a circle are *aligned* if one of them separates the other two. Say that $m \geq 3$ chords are aligned if any three of them are aligned. For instance, in the figure below, A, B, and C are aligned, while B, C, and D are not.



Claim. In a beautiful arrangement, the k-chords are aligned for any integer k.

Proof. We proceed by induction. For $n \leq 3$ the statement is trivial. Now let $n \geq 4$, and proceed by contradiction. Consider a beautiful arrangement S where the three k-chords A, B, C are not aligned. If n is not among the endpoints of A, B, and C, then by deleting n from S we obtain a beautiful arrangement $S \backslash \{n\}$ of $[0, n-1]$, where A, B, and C are aligned by the induction hypothesis. Similarly, if 0 is not among these endpoints, then deleting 0 and decreasing all the numbers by 1 gives a beautiful arrangement $S \setminus \{0\}$, where A, B, and C are aligned. Therefore both 0 and n are among the endpoints of these segments. If x and y are their respective partners, we have $n \geq 0 + x = k = n + y \geq n$. Thus 0 and n are the endpoints of one of the chords, say it is C.

Let D be the chord formed by the numbers u and v which are adjacent to 0 and n and on the same side of C as A and B, as show in the figure below:



Set $t = u + v$. If we had $t = n$, the n-chords A, B and D would not be aligned in the beautiful arrangement $S \setminus \{0, n\}$, contradicting the induction hypothesis. If $t < n$, then the t-chord from 0 to t cannot intersect D, so the chord C separates t and D. The chord E from t to $n - t$ is on the same side of C. But then the chords A, B and E are not aligned in $S \setminus \{0, n\}$, a contradiction. Finally, the case $t > n$ is equivalent to the case $t < n$ via the beauty-preserving relabelling $x \to n - x$ for $0 \leq x \leq n$, which sends t-chords to $(2n - t)$-chords. This proves the claim.

Having established the claim, we prove the desired result by induction. The case $n = 2$ is trivial. Now assume that $n \geq 3$. Let S be a beautiful arrangement of $[0, n]$ and delete n to obtain the beautiful arrangement T of $[0, n-1]$. The n-chords of T are aligned, and they contain every point except 0. Say T is of *Type 1* if 0 lies between two of these n-chords, and it is of *Type 2* otherwise, i.e., if 0 is aligned with these n-chords. We will show that each Type 1 arrangement of $[0, n-1]$ arises from a unique arrangement of $[0, n]$ and each Type 2 arrangement of $[0, n-1]$ arises from exactly two beautiful arrangements of $[0, n]$.

If T is of Type 1, let 0 lie between chords A and B. Since the chord from 0 to n must be aligned with A and B, n must be on the other arc between A and B. Therefore, S can be recovered uniquely from T. In the other direction, if T is of Type 1 and we insert n as above, then we claim that the resulting arrangement S is beautiful. For $0 < k \leq n$, the k-chords of S are also parallel by construction. There is an antisymmetry axis l such that x is symmetric to $n - x$ with respect to l for all x. If we had two k-chords which intersected for some $k > n$, then their reflections across l would be two $(2n-k)$-chords which intersect, where $0 < 2n - k < n$, a contradiction.

If T is of Type 2, there are two possible positions for n in S, on either side of 0. As above, we check that both positions lead to beautiful arrangements of $[0, n]$.

Hence, if we let M_n be the number of beautiful arrangements of $[0, n]$, and let L_n be the number of beautiful arrangements of $[0, n-1]$ of Type 2, we have

$$M_n = (M_{n-1} - L_{n-1}) + 2L_{n-1} = M_{n-1} + L_{n-1}.$$

It then remains to show that L_{n-1} is the number of pairs (x, y) of positive integers with $x + y = n$ and $\gcd(x, y) = 1$. Since $n \geq 3$, this number equals $\phi(n) = \#\{x : 1 \leq x \leq n, \gcd(x, n) = 1\}$.

To prove this, consider a Type 2 beautiful arrangement of $[0, n-1]$. Label the positions $0, \ldots, n-1 \pmod{n}$ clockwise around the circle, so that number 0 is in position 0. Let $f(i)$ be the number in position i; note that f is a permutation of $[0, n-1]$. Let a be the position such that $f(a) = n - 1$.

Since the n-chords are aligned with 0, and every point is in an n-chord, these chords are all parallel and

$$f(i) + f(-i) = n \quad \text{for all} \quad i.$$

Similarly, since the $(n-1)$-chords are aligned and every point is an $(n-1)$-chord, these chords are also parallel and

$$f(i) + f(a - i) = n - 1 \quad \text{for all} \quad i.$$

Therefore $f(a - i) = f(-i) - 1$ for all i. Since $f(0) = 0$, we get

$$f(-ak) = k \quad \text{for all} \quad k. \tag{1}$$

Recall that this is an equality modulo n. Since f is a permutation, we must have $(a, n) = 1$. Hence $L_{n-1} \leq \phi(n)$.

To prove the equality, it remains to observe that the labelling (1) is beautiful. To see this, consider four numbers w, x, y, z on the circle with $w+y = x+z$. Their positions around the circle satisfy $(-aw)+(-ay) = (-ax)+(-az)$, which means that the chord from w to y and the chord from x to z are parallel. Thus (1) is beautiful, and by construction it has Type 2. The desired result follows.

Problem 7.17. Prove that among any $2k+1$ integers, each having absolute value at most $2k-1$, one can always choose three that add up to 0.

Solution. We prove the result by induction on k. The base case $k = 1$ holds, since the numbers must be $-1, 0$, and 1 in this case and hence they add to 0.

Assume now that the result holds for some $k \geq 1$. Then consider a set of $2k+2$ integers in $[-(2k+1), 2k+1]$. We have to show that 3 of these integers sum to 0. Assume the contrary. If 0 is chosen, then we can pair off each number with its negative, and by the Pigeonhole Principle, we must pick 2 numbers from one of the pairs, which along with 0 yield a set which sums to 0. If 0 is not chosen, then at least 3 integers from the set $\{-(2k+1), -2k, 2k, 2k+1\}$ must be chosen, as otherwise at least $2k+1$ integers were chosen from $[-(2k-1), 2k-1]$ which by the induction hypothesis yields a set of 3 which sums to 0.

Therefore, we may assume without loss of generality that $2k$ and $2k+1$ were both chosen, else we could replace each number by its corresponding negative without changing the fact that the sum we are interested in is 0. We now distinguish 2 cases:

Case 1. If $-(2k+1)$ was chosen. Then 1 cannot be chosen because of the triple $\{-(2k+1), 1, 2k\}$. Pair off the negative integers not including $-(2k+1)$ so that in each pair the sum is $-(2k+1)$, and pair off the positive integers not including $2k+1$ and $2k$ so that in each pair the sum is $2k+1$. Then by Pigeonhole Principle, at least 2 numbers from some pair were chosen. This pair along with either $2k+1$ or $-(2k+1)$ yields a set which sums to 0.

Case 2. If $-(2k+1)$ was not chosen. Then $-2k$ must have been chosen, and therefore -1 cannot be chosen because of the triple $\{-2k, -1, 2k+1\}$. Pair off the negative numbers not including $-2k$ and $-(2k+1)$ so that in each pair the sum is $-2k-1$, and pair off the positive integers not including $2k$ and $2k+1$ into pairs which sum to $2k$ (with k being a singleton). By the Pigeonhole Principle, 2 distinct numbers from some pair must have been chosen. Thus we get a pair of distinct integers which sum to either $2k$ or $-(2k+1)$. This pair along with $-2k$ or $2k+1$ yields a set which sums to 0. This completes our proof.

Problem 7.18. Prove that a cube C can be divided into n cubes for all $n \geq 55$.

Solution. This is a three-dimensional version of the problem with dividing a square into n squares, and it is done in the same way. Only this time we have more cases to consider, and the examples are harder to construct.

Again, a cube can be divided into k^3 cubes for any $k \in \mathbb{N}$. In particular, if C is divided into n cubes we can divide one of them in 8 producing a division into $n+7$ cubes. Thus we need to check the base cases for $n = 55, 56, 57, 58, 59, 60, 61$.

$n = 55$: Divide C into 27 cubes, and four of these 27 into 8 cubes each, so we have a total of $27 + 4 \cdot 7 = 55$.

$n = 56$: Divide C into 8 cubes, and then divide the top four of these eight into 27 smaller cubes. The top face of C is now partitioned into a 6×6 grid. Looking at the grid as nine $2 \cdot 2$ squares, join eight of these nine squares with the squares below them into 8 cubes. We obtain a total of $8 + 4 \cdot 26 - 8 \cdot 7 = 56$ cubes.

$n = 57$: Divide C into 64 cubes, then join 8 of them to form a new one.

$n = 58$: Divide C into 27 cubes and join 8 of them into a new one. Then do the same thing with two cubes of the partition.

$n = 59$: Divide C into 64 cubes and join 27 of them into a new one. Then divide 3 of the remaining cubes into 8 each.

$n = 60$: Divide C into 8 cubes, then divide two of these into 27 each.

$n = 61$: Divide C into 27 cubes and join 8 of them into a new one. Then consider 4 of the remaining cubes that share a part of a face of C, let us call this part P, and divide them into 27 cubes each. Consider the 9 cubes having as one face the ninth part of P, obtained by joining 8 little cubes into a single one.

Problem 7.19. For a permutation $a_1, a_2, \ldots, a_n$ of the numbers $1, 2, \ldots, n$ one is allowed to change the places of any consecutive blocks, that is from

$$a_1, \ldots, a_i, a_{i+1}, \ldots, a_{i+p}, a_{i+p+1}, \ldots, a_{i+q}, a_{i+q+1}, \ldots, a_n$$

one can obtain

$$a_1, \ldots a_i, a_{i+p+1}, \ldots, a_{i+q}, a_{i+1}, \ldots, a_{i+p}, a_{i+q+1}, \ldots, a_n.$$

Find the least number of such changes after which one can obtain n, $n-1, \ldots$, 2, 1 from $1, 2, \ldots, n$.

Solution. One part here does not use induction, namely the estimation. We prove that we need at least $\left[\frac{n}{2}\right]+1$ moves. We call a pair of consecutive numbers in a permutation *normal* if the left number is smaller than the right number. Initially there are $n-1$ normal pairs and at the end zero. It is clear that the first and the last moves can decrease the number of normal pairs by at most 1. Now we will prove all the other moves can decrease the number of normal pairs by at most two. Indeed, if we pick consecutive pairs $(a, b), (c, d), (e, f)$ and swap the blocks $b \ldots c$ and $d \ldots e$, then three pairs will be modified: $(a, b), (c, d)$ and (e, f) will become $(a, d), (e, b), (c, f)$. It is not possible that all the initial pairs were normal while all the final pairs were not, because this would imply $a < b, c < d, e < f$ so $ace < bdf$ from one side, and $a > d, e > b, c > f$ so $ace > bdf$ from the other side. Therefore, if we did k moves, the number of normal pair has decreased by at most $1+2(k-2)+1 = 2k-2$. So $2k-2 \geq n-1$ and from here we get that $k \geq \left[\frac{n}{2}\right] + 1$.

Now we prove the part which uses induction: that we can reverse the order in $\left[\frac{n}{2}\right] + 1$ moves i.e. we can reverse $1, 2, \ldots, 2k-1$ in k moves and $1, 2, \ldots, 2k$ in $k+1$ moves. It suffices to prove the former, because the latter follows directly from it: if we reverse the first $2k-1$ numbers in k moves, by one more move we bring $2k$ in the front.

We do this by induction on k. When $k = 2$, it is simple. The case $k = 3$ can be exemplified by $(1, 2, 3, 4, 5) \to (1, 4, 5, 2, 3) \to (5, 2, 1, 4, 3) \to (5, 4, 3, 2, 1)$. Next, we prove by induction the following statement:

"For $k \geq 3$ there exists a sequence of k moves reversing a sequence of $2k-1$ numbers such that the first move swaps the blocks in positions $2, 3, \ldots, k$ and $k+1, k+2, \ldots, 2k-1$."

The base case was done above. Now let us prove the step $k-1 \to k$. Consider the first move, which transforms

$$1, 2, \ldots, 2k+1 \quad \text{to} \quad 1, k+2, k+3, \ldots, 2k+1, 2, \ldots, k, k+1.$$

Now we can "glue" 1 and $k+2$, i.e. considering them a a single entity in our next operations, and we can disregard the $k+1$ at the end. We get a sequence of $2k-1$ objects in which the first move was already done. So by the induction hypothesis, we can make $k-1$ more steps to completely reverse it, i.e. get $2k+1, 2k, \ldots, 2, [1, k+2]$. Then the operations will bring our initial sequence to $2k+1, 2k, \ldots, k+3, k, k-1, \ldots, 2, 1, k+2, k+1$. One more move takes $[k+2, k+1]$ to its place between $k+3$ and k. We are done.

Problem 7.20. Let m circles intersect at points A and B. We write numbers using the following algorithm: We write 1 at points A and B, at every midpoint of an open arc AB we write 2, then at the midpoint of the arc between every two points with numbers written we write the sum of the two numbers and so on repeating n times. Let $r(n,m)$ be the number of appearances of the number n after writing all of them on our m circles.

a) Determine $r(n,m)$;

b) For $n = 2006$, find the least m for which $r(n,m)$ is a perfect square.

Example. For a single arc AB the numbers written on successive rounds are: $1-1$; $1-2-1$; $1-3-2-3-1$; $1-4-3-5-2-5-3-4-1$; $1-5-4-7-3-8-5-7-2-7-5-8-3-7-4-5-1$.

Solution. Observe that it suffices to look at only one circle, and even at one of the arcs between A and B, and multiply the result by $2m$ to get $r(n,m)$. Even more, we can divide this arc into two halves and look at one of the halves, because the construction will be symmetric on these halves. Now let us investigate this case. Producing $a+b$ from a and b reminds us of an inverse of the Euclidean Algorithm. As the Euclidean Algorithm finds out the greatest common divisor, we might have an idea to look at the greatest common divisor of two consecutive numbers. As $(a+b, a) = (a+b, b) = (a, b)$ and initially $(1,1) = 1$ we conclude that two consecutive numbers on the circle are coprime, by induction on the number of steps. Even more, if at one step the numbers k, l with $k < l$ are neighbors on the circle, then it is clear that l was obtained by summing k with $l-k$ at the previous step. This allows us to work backwards from a given configuration. Thus, if at some point the pair (n,m) occurs on the circle, at the next point the pair $(n, n+m)$ occurs on the circle, and at the next point the next pair $(n, n+2m)$ and so on, after k steps the pair $(n, n+km)$ appearing on the circle. Vice-versa, if the pair $(n, n+km)$ appears on the circle, then the pair $(n, n+m)$ appears k steps earlier.

Let H_1 be the arc after one step $(1-2)$, H_2 after two steps $(1-3-2)$, and so on.

Now a careful observation allows us to conjecture the general result of the problem: after n steps, on the regarded half of the circle, all pairs (u, v) of numbers at most n such that $(u, v) = 1$ appear in exactly one of $H_1, H_2, \ldots, H_n$ and appear there exactly once.

The proof is by induction on n. The base case is clear. Firstly we establish the appearance part: if $(u, v) = 1$ and $u < v \leq n$, then $\max\{u, v-u\} < n$, thus by the induction hypothesis, $(u, v-u)$ appears on exactly one of $H_1, H_2, \ldots, H_{n-1}$. Hence (u, v) appears one exactly one of $H_2, H_3, \ldots, H_n$, and this settles the claim (notice that (u, v) cannot appear on H_1 because H_1 consists of one pair $1-2$ which cannot appear at higher level because all the newly added numbers will be greater than 2).

Hence according to the induction for any $0 < j < n$ with $(j, n) = 1$ the pair (n, j) will appear once on one H_k $(1 \leq k \leq n)$. Correspondingly, $(n, j + (n-k)n)$ will appear on H_n. So it contains a pair (n, m) where $m \equiv j \pmod{n}$. This pair should be unique, because each pair (n, m) has arisen from its prototype (n, j), as we have prove earlier, and (n, j) appears just once on $H_1, H_2, \ldots, H_n$. So there is exactly one pair (n, m) with $m \equiv j \pmod{n}$ for every residue j modulo n coprime with n, and no other pairs involving n. Hence the total number of them is $\phi(n)$ If n appears k then it will be involved in $2k$ pairs so $k = \frac{\phi(n)}{2}$. Therefore $r(n, m) = 4m\frac{\phi(n)}{2} = 2m\phi(n)$.

As $\phi(2006) = \phi(2 \cdot 17 \cdot 59) = 2^5 \cdot 29$, for $2m\phi(2006)$ be a perfect square we need m to be 29 times a perfect square so $m = 29$ is the least.

Problem 7.21. (**Hungary 2000**) Given a positive integer k and more than 2^k integers, prove that a set S of $k+2$ of these numbers can be selected such that for any positive integer $m \leq k+2$, all the m-element subsets of S have different sums of elements.

Solution. Let us introduce some terminology first: Given a positive integer m, we call a set *weakly m-efficient* if its m-element subsets have different sums of elements, and we call a set *strongly m-efficient* if it is weakly i-efficient for

$1 \le i \le m$. Also, given any set T of integers, let $\sigma(T)$ be the sum of the elements of T.

With this setup, we prove the result of our question by induction on k. For $k = 1$, it is easy to check that we may let S consist of any three of the given integers.

Now assuming that the claim is true for $k = n$, we prove that it is true for $k = n+1$. Given more than 2^{n+1} different integers $a_1, a_2, \ldots, a_t$, let 2^α be the largest power of 2 such that $a_1 \equiv a_i \pmod{2^\alpha}$ for each $i = 1, 2, \ldots, t$. Write $b_i = \frac{a_i - a_1}{2^\alpha}$ for $1 \le i \le t$, yielding t distinct integers $b_1, b_2, \ldots, b_t$. By the Pigeonhole Principle, applied to the b_i's there exist more than 2^n different integers of the same parity. By the induction hypothesis, from these integers we may choose an $n+2$-element, strongly $(n+2)$-efficient set S_1. Furthermore, there exists a b_{i_0} of the opposite parity because the b_i are not all of the same parity: $b_1 = 0$ is even, and by the maximality of α, at least one of the b_i is odd.

We claim that the $(n+3)$-element set $S_2 = S_1 \cup \{b_{i_0}\}$ is strongly $(n+3)$-efficient. Suppose, for the sake of contradiction that X and Y are two distinct m-element subsets of S with the same sums of elements, where $1 \le m \le n+3$. Because $X \neq Y$, $m > 1$. Notice that X and Y cannot both be subsets of S_1 because S_1 is weakly m-efficient. Nor can they both contain b_{i_0} because then $X \setminus \{b_{i_0}\}$ and $Y \setminus \{b_{i_0}\}$ would be two distinct $(m-1)$-element subsets of S_1 with the same sums of elements, which is impossible because S_1 is weakly $(m-1)$-efficient. Therefore, one of X and Y contains b_{i_0} and the other does not. This in turn implies that $\sigma(X)$ and $\sigma(Y)$ are of opposite parity, which is a contradiction.

Let Φ be the map which sends any set A of reals to $\{\frac{a-a_1}{2^\alpha} : a \in A\}$. There exists an $(n+3)$-element subset $S \subseteq \{a_1, a_2, \ldots, a_t\}$ such that $\Phi(S) = S_2$. Suppose that there existed $X, Y \subseteq S$ such that $X \neq Y$, $|X| = |Y| = m$, and $\sigma(X) = \sigma(Y)$. Then we would also have

$$\Phi(X), \Phi(Y) \subseteq \Phi(S) = S_2, \ \Phi(X) \neq \Phi(Y), \ |\Phi(X)| = m = |\Phi(Y)|,$$

and

$$\sigma(\Phi(X)) = \frac{\sigma(X) - ma_1}{2^\alpha} = \frac{\sigma(Y) - ma_1}{2^\alpha} = \sigma(\Phi(Y)).$$

However, this is impossible because S_2 is weakly m-efficient. Therefore, S is strongly $(n+3)$-efficient as well. This completes the induction step and our proof.

Problem 7.22. (**Russia 2000**) There is a finite set of congruent square cards, placed on a rectangular table with their sides parallel to the sides of the table. Each card is colored in one of k colors. For any k cards of different colors, it is possible to pierce two of them with a single pin. Prove that all the cards of some color can be pierced by $2k-2$ pins.

Solution. We prove the claim by induction on k. If $k=1$, we are told that given any set containing 1 card (of the single color), two cards in the set can be pierced with one pin. This does not make sense unless there are no cards to start with, in which case all the cards can be pierced by $0=2k-2$ pins.

Now assume that the claim holds when $k=n-1$, (for some $n \geq 2$) and consider a set of cards colored in n colors. We rotate the table such that the sides of the cards are horizontal and vertical. Let X be a card whose top edge has minimum distance to the top edge of the table. Because all of the cards are congruent and identically oriented, any card that overlaps with X must overlap either X's lower left corner or X's lower right corner. Pierce pins P_1 and P_2 through these two corners.

Now let S be the set of cards which are not pierced by either of these two pins and which are colored differently than X. None of the cards in S intersects X and they are each colored in one of $n-1$ colors. Given a set $T \subseteq S$ of $n-1$ cards of different colors, it is possible to pierce two of the cards in $T \cup \{X\}$ with a single pin. Because no card in T overlaps with X, this single pin actually pierces two cards in T. So we may apply the induction hypothesis to S and pierce all the cards of some color c in S with $2n-4$ pins. Combined with the pins P_1 and P_2, we find that all the cards of color c can be pierced with $2n-2$ pins. This completes our proof.

Problem 7.23. (**Russia 2000**) Each of the numbers $1, 2, \ldots, N$ is colored black or white. We are allowed to simultaneously change the colors of any three numbers in arithmetic progression. For which numbers N can we always make all the numbers white?

Solution. We clearly cannot always make all the numbers white if $N = 1$. Suppose that $2 \leq N \leq 7$, and suppose that only the number 2 is colored black. Call a number from $\{1, \ldots, N\}$ *heavy* if it is not congruent to 1 modulo 3. Let X be the number of heavy numbers which are black, where X changes as we change the colors. Suppose we change the colors of the numbers in $\{a - d, a, a + d\}$, where $1 \leq a - d < a < a + d \leq N$. If d is not divisible by 3, then $a-d$, a, $a+d$ are all distinct modulo 3, so exactly two of them are heavy. If instead d is divisible by 3, then $a - d$, a, $a + d$ must equal $1, 4, 7$, none of which are heavy. In either case, changing the colors of these three numbers changes the color of an even number of heavy numbers. Hence X is always an odd number, and we cannot make all the numbers white.

Next we show that for $N \geq 8$, we can always make all the numbers white. To do this, it suffices to show that we can invert the color of any single number n. We prove this by strong induction. If $n \in \{1, 2\}$, then we can invert the color of n by changing the colors of the numbers in $\{n, n + 3, n + 6\}$, $\{n+3, n+4, n+5\}$ and $\{n+4, n+5, n+6\}$. Now assuming that we can invert the color of $n - 2$ and $n - 1$ (where $3 \leq n \leq N$), we can invert the color of n by first inverting the colors of $n - 2$ and $n - 1$ then changing the colors of the numbers in $\{n - 2, n - 1, n\}$.

Hence, we can always make all the numbers white if and only if $N \geq 8$.

Problem 7.24. For $n \geq 1$, let O_n be the number of $2n$-tuples $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ with all entries being either 0 or 1 and for which the sum $x_1y_1 + \ldots + x_ny_n$ is odd, and E_n the number of $2n$-tuples of same type for which the sum is even. Prove that

$$\frac{O_n}{E_n} = \frac{2^n - 1}{2^n + 1}.$$

Solution. We prove the statement by induction on n. The base case is $n = 1$, for which we have $O_n = 1$, $E_n = 3$.

For the induction step, it suffices to find suitable recurrences for O_n and E_n. Any sequence from O_{n+1} has either $\sum_{i=1}^{n} x_iy_i$ even or odd. If it is even, then we are forced at the end to have $x_{n+1} = y_{n+1} = 1$. Otherwise we have at the end that $x_{n+1}y_{n+1} = 0$ and this can be done in three ways. Thus $O_{n+1} = E_n + 3O_n$.

A similar reasoning shows that $E_{n+1} = 3E_n + O_n$. Thus we have that

$$\frac{O_{n+1}}{E_{n+1}} = \frac{E_n + 3O_n}{3E_n + O_n}.$$

Now using the induction hypothesis, we are done.

Problem 7.25. (**TOT 2001**) In a row there are 23 boxes such that for $1 \le k \le 23$, there is a box containing exactly k balls. In one move, we can double the number of balls in any box by taking balls from another box which has more. Is it always possible to end up with k balls in the k-th box for $1 \le k \le 23$?

Solution. We prove more generally, that we can always achieve the task for any number $n \ge 1$ of boxes. We do this by induction on n. The result is clear when $n = 1$.

Assume now that the result holds for some $n \ge 1$. Then we can arrange the boxes in a line in increasing order (from left to right) of the number of balls in them, without regard to the box numbers. We now perform the following operation: we start with the rightmost box and we move balls from each box to the next one to its left, so that the box to its left will have twice as many balls as it initially did (this is possible by the ordering we chose among boxes). This is illustrated in the figure below

$$\begin{array}{ccccccc}
1 & 2 & 3 & \ldots & n-1 & n & n+1 \\
\hline
 & & & & & 2n & 1 \\
 & & & & 2n-2 & n+1 & \\
 & & & \ldots & n & & \\
 & & & & \vdots & & \\
 & & 6 & \ldots & & & \\
 & 4 & 4 & & & & \\
2 & 3 & & & & & \\
\hline
2 & 3 & 4 & \ldots & n & n+1 & 1
\end{array}$$

Notice that in this manner, we end up with a cyclic permutation of the number of balls from the initial arrangement that we made. This means that if we keep performing this operation (but now starting with the box with the highest number of balls and moving to its left), we must eventually end up having $(n+1)$ balls in the box with number $(n+1)$. Then from the induction hypothesis, we can sort the balls from the other n boxes remaining, and we are done.

Problem 7.26. Consider a $2^n \times 2^n$ square. Prove that after removing a 1×1 square from one of its corners, the remaining region can be tiled by "corners" (a corner is a 2×2 square with one of the four corner unit squares removed).

Solution. The base $n = 1$ is clear, as the obtained region is itself a corner.

Now assume that we can tile a $2^k \times 2^k$ region with one corner square removed. Consider a $2^{k+1} \times 2^{k+1}$ region. Without loss of generality, we can assume that the removed square is at the lower-right corner. If we draw the two lines parallel to the sides of the big square through its center, we will divide the $2^{k+1} \times 2^{k+1}$ region into four $2^k \times 2^k$ regions, call them S_1, S_2, S_3, S_4. The removed square is from S_4. Now we look at the central 2×2 square. It contains a unit square from each of S_1, S_2, S_3, S_4. So we can place a corner to cover the unit squares belonging to S_1, S_2 and S_3, respectively. This will leave S_1, S_2, S_3 without a unit square in the corner, too. Thus we can apply the induction hypothesis for S_1, S_2, S_3, S_4 to produce the desired tiling.

Problem 7.27. Consider 10×10 grid square, such that in some of its squares there are written ten ones, ten twos, ..., ten nines and one ten (such that in each square there is at most one number written). Prove that one can choose ten squares from different rows, such that in the chosen squares we have the numbers $1, 2, \ldots, 10$.

Solution. This problem could also be solved as an application of Hall's Theorem: the graph has vertices which are the rows and numbers, with a row joined to a number if the number occurs in that row; since any set of k rows has at least $10k-9$ numbers written in it, they must include at least k different values; thus, by Hall's Theorem, a matching exists and this is exactly what is requested. We give, however, a self-contained proof below, by proving the following more general result:

Consider a rectangular grid, such that each row has n squares. Prove that if in some of its squares there are written n ones, in some of its squares there are written n twos, ..., in some its squares there are written n $m-1$'s and in one square there is written m, $m \geq 2$, (in each square there is at most one number written), then one can choose m rows and from each row a number, such that in the m chosen squares we have the numbers $1, 2, \dots, m$.

We prove this statement by induction on $m \geq 2$. The base case is $m = 2$.

Note that in the row which contains the number 2 there are written at most $n-1$ ones. So there must be a different row containing a one. Hence, we can choose a square from this row which has the number 1.

Let us show that if the statement holds true for $m = 2, \dots, k$, $k \in \mathbb{N}$, then it also holds for $m = k+1$, $k \geq 2$.

From the whole grid consider only the squares where the numbers $2, 3, \dots, k, k+1$ are written. From the induction hypothesis on $m = k$, there exist k rows where we have the numbers $2, 3, .., k+1$. Let us denote the corresponding rows by $\langle 2\rangle, \langle 3\rangle, \dots, \langle k+1\rangle$. The total number of numbers written in the rows $\langle 2\rangle$, $\langle 3\rangle, \dots, \langle k+1\rangle$ is not more than nk and there are $nk+1$ numbers written in the whole grid. So apart from the rows $\langle 2\rangle, \langle 3\rangle, \dots, \langle k+1\rangle$, there must be at least one extra row which has a number written in one of its squares.

Let us denote by M_1 the set of numbers that are written in the rows different from $\langle 2\rangle$, $\langle 3\rangle, \dots, \langle k+1\rangle$. Obviously, if $1 \in M_1$, then the statement is proved. If $a \in M_1$ and in the row $\langle a\rangle$ we have the number 1, then the statement is proven. Consider all the rows $\langle a\rangle$, where $a \in M_1$. Denote by M_2 the set of numbers in all the rows $\langle a\rangle$ for $a \in M_1$ which not belong to M_1. We have that $1 \notin M_2$. Now, let us consider all $\langle a\rangle$ rows, where $a \in M_2$. Denote by M_3 the set of all numbers written in the considered rows that do not belong to the set $M_1 \cup M_2$.

Iterating this construction, either the statement is proven or we obtain the following sets $M_1, M_2, \dots, M_l$, where $M_{l+1} = \emptyset$ and $1 \notin M_i$, $i = 1, 2, \dots, l$.

Note that in the rows $\langle a\rangle$ where $a \in M_1 \cup M_2 \cup \dots \cup M_l$ are written only the elements of the set $M_1 \cup M_2 \cup \dots \cup M_l$.

Now, if we consider all rows $\langle a\rangle$, where $a \notin M_1 \cup M_2 \cup \dots \cup M_l$, then by the induction hypothesis one can choose different rows, in which are written all the numbers a, $a \notin M_1 \cup M_2 \cup \ldots \cup M_l$.

The choice for all the numbers a with $a \in M_1 \cup M_2 \cup \ldots \cup M_l$ is obvious. This ends the proof of the statement.

Problem 7.28. (**IMO 2009**) Let $a_1, a_2, \ldots a_n$ be different positive integers and M a set of $n-1$ positive integers not containing the number $s = a_1 + a_2 + \ldots + a_n$. A grasshopper is going to jump along the real axis. It starts at the point 0 and makes n jumps to the right of lengths $a_1, a_2, \ldots, a_n$ in some order. Prove that the grasshopper can organize its jumps in such a way that it never falls in any point of M.

Solution. We will prove the statement by strong induction on n. For $n = 1$, there is nothing to prove. For the inductive step, without loss of generality, we can order the steps as $a_1 < a_2 < \ldots < a_n$ and the elements of M as $b_1 < b_2 < \ldots < b_{n-1}$. Let $s' = a_1 + a_2 + \ldots + a_{n-1}$. If we remove a_n and b_{n-1}, there are two cases:

1) s' is not among the first $n-2$ elements of M. In this case, by induction, we can order the first $n-1$ jumps until we reach s'. If at any moment we fell on b_{n-1}, we change that last step for a_n and then we continue in any way to reach s. From the induction hypothesis we know that we have never fallen on $b_1, b_2, \ldots, b_{n-2}$. Also, if we had to use the change, since there are no elements of M after b_{n-1}, we do not have to worry about falling on a b_k in the rest of the jumps.

2) s' is one of the first $n-2$ elements of M. If this happens, then since $s' = s - a_n$ is in M, among the $2(n-1)$ numbers of the form $s - a_i, s - a_i - a_n$ with $1 \leq i \leq n-1$ there are at most $n-2$ elements of M. If we look at the pairs of numbers $(s - a_i, s - a_i - a_n)$, since we have $n-1$ of these pairs and they contain at most $n-2$ elements of M, there is a number a_i such that neither $s - a_i$, nor $s - a_i - a_n$ are in M. Notice that after $s - a_i - a_n$ we have s' and b_{n-1}, which are two elements of M. Therefore, there are at most $n-3$ elements of M before $s - a_i - a_n$. Then, by the induction hypothesis, we can use the other $n-2$ jumps to reach $s - a_i - a_n$, then use a_n and then use a_i to get to s without falling on a point of M.

This completes the inductive step and hence the solution to our problem.

Problem 7.29. (**IMO shortlist 2004**) For a finite graph G, let $f(G)$ be the number of triangles and $g(G)$ the number of tetrahedra formed by edges of G. Find the least constant c such that

$$g(G)^3 \le c \cdot f(G)^4,$$

for every graph G.

Solution. If $|G| = n$, we could try an induction on n. Notice first that if G is the complete graph then we get

$$g(G) = \binom{n}{4} \sim \frac{n^4}{24} \quad \text{and} \quad f(G) = \binom{n}{3} \sim \frac{n^3}{6},$$

so

$$\frac{g(G)^3}{f(G)^4} \sim \frac{6^4}{24^3} = \frac{3}{32}.$$

It is natural to suspect that this is the best possible c, because each tetrahedron contains four triangles, whereas a triangle can even belong to no tetrahedron, and the maximal number of tetrahedra a triangle belongs to is attained for a complete graph.

The most popular induction approach would be to take a vertex A, consider the subgraph G_1 induced by the neighbors of A and G_2 induced by all the vertices except A, and try to use induction step on G_1, G_2. But we confront with a problem as we cannot really know the relationship between the triangles and edges of G_1 that arises in our computations.

So let us lower the level and try at first to establish a relation between the number of edges $e(G)$ and the number of triangles $f(G)$ in a graph G. Reasoning as above, we can suspect that $f(G)^2 \le \frac{2}{9}e(G)^3$. The base cases $|G| \le 3$ are obvious.

Now let us perform the induction step. Pick up a vertex A and consider the subgraphs G_1 induced by the neighbors of A and G_2 induced by the non-neighbors of A. Let $|G_1| = k$, $e(G_1) = l$, $e(G_2) = s$, $f(G_2) = t$. We establish that $e(G) = k + s$, $f(G) = t + l$. Now we know from the induction hypothesis that $t^2 \le \frac{2}{9}s^3$, and it is also clear that $l \le \frac{k^2}{2}$, $l \le s$. So

$$f(G)^2 = (t + l)^2 = t^2 + 2tl + l^2.$$

Moreover, we also have

$$\frac{2}{9}e(G)^3 = \frac{2}{9}(s+k)^3 = \frac{2}{9}s^3 + \frac{2}{3}s^2k + \frac{2}{3}sk^2 + \frac{2}{9}k^3.$$

But $t^2 \leq \frac{2}{9}s^3$, hence

$$l \leq \sqrt{ls} \leq \frac{1}{\sqrt{2}}\sqrt{s}k,$$

so

$$2tl \leq 2\sqrt{\frac{2}{9}}s\sqrt{s}\frac{1}{\sqrt{2}}\sqrt{s}k = \frac{2}{3}s^2k$$

and also

$$l^2 \leq ls \leq \frac{sk^2}{2} \leq \frac{2}{3}sk^2.$$

So we conclude that $f(G)^2 \leq \frac{2}{9}e(G)^3$.
Now let us use this result to our main induction. We define G_1, G_2 as above. Set $f(G_2) = s$, $g(G_2) = t$, $e(G_1) = k$, $f(G_1) = l$. We know that

$$t^3 \leq \frac{3}{32}s^4,\ l^2 \leq \frac{2}{9}k^3,\ l \leq s.$$

Now we get $f(G) = s + k, g(G) = t + l$. Thus

$$g(G)^3 = (t+l)^3 = t^3 + 3t^2l + 3tl^2 + l^3,$$

$$\frac{3}{32}f(G)^4 = \frac{3}{32}(s+k)^4 = \frac{3}{32}s^4 + \frac{3}{8}s^3k + \frac{9}{16}s^2k^2 + \frac{3}{2}sk^3 + \frac{3}{32}k^4.$$

We break it down again:

$$t^3 \leq \frac{3}{32}s^4,\ l \leq s^{\frac{1}{3}}l^{\frac{2}{3}} \leq \left(\frac{2}{9}\right)^{\frac{1}{3}} s^{\frac{1}{3}}k.$$

It follows that

$$3t^2l \leq 3\left(\frac{3}{32}\right)^{\frac{2}{3}} s^{\frac{8}{3}} \left(\frac{2}{9}\right)^{\frac{1}{3}} s^{\frac{1}{3}}k = \frac{3}{8}s^3k,$$

then

$$3tl^2 \leq 3\cdot\left(\frac{3}{32}\right)^{\frac{1}{3}} s^{\frac{4}{3}} \left(\frac{2}{9}\right)^{\frac{2}{3}} s^{\frac{2}{3}}k^2 = \frac{1}{2}s^2k^2 \leq \frac{9}{16}s^2k^2,$$

and finally

$$l^3 \leq \frac{2}{9} sk^3 \leq \frac{3}{8} sk^3.$$

So

$$g(G)^3 \leq \frac{3}{32} f(G)^4.$$

The example of the complete graph shows us that $\frac{3}{32}$ is the best constant.

Remark. Using the method described above, one can prove by induction on $n+k$ the following generalization:

If we set $||G||_k$ be the number of complete subgraphs with k vertices of G, the we have

$$||G||_{k+1}^k \leq c||G||_k^{k+1},$$

where $c = \frac{(k!)^{k+1}}{(k+1)!^k} = \frac{k!}{k^k}$, and this constant is the best possible, as shown by the complete graph.

Problem 7.30. Prove that a graph with $\binom{n+k-2}{k-1}$ vertices contains either a K_n or a $\overline{K_k}$ i.e. either n mutually connected vertices or k mutually not connected vertices.

Solution. The proof is by induction on $n+k$. If $n+k=2$, then a graph with at least one vertex satisfies the condition, and the same holds if at least one of n or k is 0.

Assume now that the result holds whenever $n+k<m$, for some $m \geq 3$ and let $n+k=m$. Take a vertex A and let S_1 be the set of its neighbors, S_2 be the set of its non-neighbors. Then

$$|S_1| + |S_2| \geq \binom{n+k-2}{k-1} - 1 = \binom{n+k-3}{k-1} + \binom{n+k-3}{k-2} - 1,$$

therefore either $|S_1| \geq \binom{n+k-2}{k-1}$ or $|S_2| \geq \binom{n+k-2}{k-2}$. In the first case, by the induction hypothesis S_1 contains either a K_{n-1} or a $\overline{K_k}$. If it contains a $\overline{K_k}$ we are done, so are we if it contains a K_{n-1} because we can add A to it. In the second case, by the induction hypothesis S_2 contains either a K_n or a $\overline{K_{k-1}}$. Again, if it contains a K_n we are done, and if it contains a $\overline{K_{k-1}}$ we add A to it and establish the induction step.

Problem 7.31. In a simple graph with a finite number of vertices each vertex has degree at least three. Prove that the graph contains an even cycle.

Solution. We prove the result by induction on the number n of vertices. The base case $n = 4$ is true because we the have a cycle with length four.

Now assume that we have proven the problem for all graphs which have fewer than n vertices. Take a graph G with n vertices. Assume by contradiction that it does not satisfy our condition. Consider a minimal cycle $X_1X_2\dots X_{2k+1}$. Then the vertices X_i and X_j are not connected to each other unless they are consecutive on the cycle, so from each of them emerges at least one edge outside the cycle. These edges point to different vertices, because if, for example, X_i and X_j are connected with the same vertex Y then the two cycles $YX_iX_{i+1}\dots X_j$ and $YX_iX_{i-1}\dots X_j$ have sum of lengths $2k+5$ thus one of them has even length. Thus we can suppress $X_1, X_2, \dots, X_{2k+1}$ into one vertex X and join X by Y if and only Y is connected to one of the X_i in G. In this new graph G', X will have degree at least $2k+1$ and the other vertices will preserve their degrees from G. Thus by the induction hypothesis, G' must contain an even cycle. If X is not in this cycle, then the cycle appears also in G, so G contains an even cycle, contradiction. If X is in this cycle, say the cycle is $XA_1A_2\dots A_{2l-1}$, then A_1 and A_{2l-1} are connected to different X_i and X_j in G. Thus we have two cycles $X_jX_{j+1}\dots X_iA_1A_2\dots A_{2l-1}$ and $X_jX_{j-1}\dots X_iA_1A_2\dots A_{2l-1}$. The sum of their lengths is $2k+4l+1$, so one of them is even. This gives us the desired contradiction, completing the proof.

Problem 7.32. For a positive integer n, let S be a set of 2^n+1 elements. Let f be a function from the set of two-element subsets of S to $\{0, \dots, 2^{n-1}-1\}$. Assume that for any elements x, y, z of S, one of $f(\{x,y\}), f(\{y,z\}), f(\{z,x\})$ is equal to the sum of the other two. Prove that there exist a, b, c in S such that $f(\{a,b\}), f(\{b,c\}), f(\{c,a\})$ are all equal to 0.

Solution. We can reformulate the problem in a more comfortable language: the edges of a complete graph on 2^n+1 vertices are assigned numbers from the set $\{0, 1, \dots, 2^{n-1}-1\}$ such that in every triangle, the sum of numbers assigned to two edges of a triangle equals the sum of numbers assigned to the third. We must prove there is a triangle with all edges null. We shall do this by induction on $n \geq 1$. The base case $n = 1$ is immediate.

To use the induction hypothesis, we would like to find a subgraph with $2^{n-1}+1$ vertices whose all edges are assigned numbers from 0 to $2^{n-2}-1$. Unfortunately this approach fails. However, we can try another one: find a subgraph with $2^{n-1}+1$ vertices whose all edges are assigned even numbers. Then we could apply the induction hypothesis to it, because the numbers assigned to its edges are twice a number from the set $\{0, 1, \ldots, 2^{n-2}-1\}$. We can achieve this.

Indeed, we see that every triangle contains an even number of edges which are assigned an odd number. Then so does any cycle, because the edges of the cycle $A_1A_2\ldots A_n$ can be split into the edges of $n-2$ triangles

$$A_1A_2A_3, A_1A_3A_4, \ldots, A_1A_{n-1}A_n$$

(the edges $A_1A_3, A_3A_4, \ldots, A_1A_{n-1}$ are counted twice, so they do not affect the parity). In particular, there is no odd cycle such that the numbers assigned to its edges are all odd. Therefore, the graph obtained by joined two vertices if and only if the edge joining them in the original graph is assigned an odd number contains no odd cycle, hence it is bipartite by a famous graph theory result. Then one part contains at least $2^{n-1}+1$ vertices, and it is our desired subgraph.

Problem 7.33. Let G be a graph on n vertices such that there are no K_4 subgraphs in it. Prove that G contains at most $\left(\frac{n}{3}\right)^3$ triangles.

Solution. We prove the result by induction of step 3. For $n=2$, $n=3$, and $n=4$, the result is clear.

Assume now that the result holds for all integers less than some $n \geq 5$. Let G be a graph with n vertices, having no 4-clique. Observe that we can add edges to G until it contains a triangle ABC. Let $G \setminus \triangle ABC$ be the subgraph formed by the remaining $n-3$ vertices.

Using Turan's Theorem, since $G \setminus \triangle ABC$ contains no K_4 subgraphs, it has at most $\frac{(n-3)^2}{3}$ edges. From the induction hypothesis, we also know that it contains at most $\left(\frac{n-3}{3}\right)^3$ triangles. The remaining triangles in G are formed as a union of one vertex belonging to ABC and an edge from $G \setminus \triangle ABC$, or by a union of one edge from ABC and one vertex belonging to $G \setminus \triangle ABC$, or ABC itself.

Every edge from $G \setminus \triangle ABC$ forms a triangle with at most one vertex belonging to ABC. Each vertex from $G \setminus \triangle ABC$ can form a triangle with at most one pair of vertices that belongs to ABC. Otherwise, in both cases, we get a K_4 subgraph formed. Hence the total number of triangles is at most

$$\left(\frac{n-3}{3}\right)^3 + \frac{(n-3)^2}{3} + (n-3) + 1 = \left(\frac{n-3}{3}+1\right)^3 = \left(\frac{n}{3}\right)^3,$$

and our induction is complete. The equality case holds in the case of a 3-partite graph $K_{\frac{n}{3},\frac{n}{3},\frac{n}{3}}$.

Problem 7.34. (**Moscow 2000**) In a country, there is at least one road going out of each city (each road connects exactly two cities). We call a city *marginalized* if there is only one road going out of it. It is known that it is not possible to get out of one city and then to get back into it using a closed circuit. The cities were split into two sets so that no two cities belonging to the same set are connected by a road. Assuming that there are at least as many cities in the first set as in the second, show that the first set must contain a marginalized city.

Solution. We begin by showing that there is a marginalized city in the country. Indeed, considering the route which contains the largest number of cities and which passes through each of its cities exactly once, it is clear that the first and the last city of this route are marginalized.

We now prove the statement of the question by induction on the number n of the cities in the country. For $n = 2$, the conclusion is clear.

Assume now that we proved the statement for all values $n \leq m - 1$. We shall prove it for $n = m$ as well. Consider a marginalized city. If it belongs to the first set, then we are done. Otherwise, it must be in the second set. If we remove this city, in the country we obtain, there are two possible cases:

Case 1. The city we removed is connected to a city which is not marginalized in the country we obtain. Then in the country we obtain, in the second set there must be fewer cities than in the first set. From the induction hypothesis, there is a marginalized city in the first set. This is the marginalized city in the original country.

Case 2. The city we removed is connected to a city A, which is marginalized in the country we obtain. Then the city A is connected only to cities in the second set. We now also remove the city A. By this procedure, we removed a city from each set, hence the induction hypothesis applies to the country we obtain by doing these removals, so it must be that there is a marginalized city in the first set. But this city will also be marginalized in the country we started with. Therefore the induction step is proved in this case as well.

Problem 7.35. (**Moscow 2001**) There are 20 teams each belonging to a different city that play some football games among themselves, such that each team plays a home game and at most two away games. Prove that we can schedule the games in such a way that each team does not play more than one game per day and all games are played in three days.

Solution. We prove the statement by induction on the number n of the teams. If $n = 2$, everything is clear.

Assume first that there is a team A which only plays a home game, say against B. Then the hypotheses of the question are still true for the rest of the teams, if we remove the team A and its game against B. Therefore using the induction hypothesis, for the rest of the teams we can schedule the games so that they take only 3 days. From our assumption, team B could only play two games in the reduced schedule, so we can schedule the game between A and B for the day when B is not scheduled to play. So we proved the induction step for this case.

Assume now that every team played at least one away game. So it must be that each team played exactly one home game and one away game. This implies that we can partition the teams into groups such that if a group has k teams $A_1, A_2, \dots, A_k$, team A_1 plays at home against A_2, A_2 plays at home against $A_3, \dots$, A_{k_1} plays at home against A_k and A_k plays at home against A_1. Then for each group, we schedule the game between A_1 and A_2 in the first day, the game between A_2 and A_3 in the second day, the game between A_3 and A_4 in the first day and so on. If k is even, then all games can be scheduled in two days and if k is odd, then the last game between A_k and A_1 is scheduled for the third day. This completes the proof in this case.

Problem 7.36. (**Five color theorem**) Prove that the vertices of every planar graph can be colored with 5 colors such that each edge has its corresponding vertices of different colors.

Solution. To solve this problem we will need two corollaries of Euler's formula (Example 7.12) which states that if a connected planar graph with V vertices and E edges divides the plane into F faces, then $V-E+F=2$. The first fact we will need is that any planar graph must have a vertex of degree at most 5. Clearly, it suffices to prove this for connected planar graphs. Since each face is bounded by at least 3 edges and each edge is counted in at most 2 faces, we have $3F \leq 2E$. Thus from Euler's formula we deduce that $E \leq 3V-6$. If every vertex has degree at least 6, by the handshaking lemma ($\sum\limits_{v \in V(G)} \deg(v) = 2E$) we would obtain $E \geq 3V$. This is impossible, so there must be a vertex of degree at most 5. The second fact we will need is that the complete graph K_5 with 5 vertices cannot be planar. This follows from the bounds above since K_5 has 10 edges and 5 vertices, and these violate the inequality above since $10 > 3 \cdot 5 - 6 = 9$.

Note that as soon as we showed that there was a vertex of degree at most 5, we have an easy inductive proof of the six color theorem. We induct over the number of vertices. The base case is trivial. For the inductive step, we simply delete a vertex of degree at most 5. This results in a graph with fewer vertices, so by the induction hypothesis we can color it with six colors. Since the neighbors of the deleted vertex use only five of these colors, there is a color available for it.

The proof of the five color theorem is similar, but we will have to work a little to make sure a color is available. Again we proceed by induction on the number of vertices and the base case is trivial. Pick a vertex v with degree at most 5. If it has degree 4 or less, then we simply delete v. The resulting planar graph has fewer vertices, and hence can be colored with our five colors. Since v has at most 4 neighbors, there must be a color not used by any of its neighbors and we may color v this color.

Thus we may assume v has degree 5, say with neighbors A, B, C, D, and E in clockwise order. Note that we can draw an edge from A to B which closely follows the edges Av and vB and hence doesn't cross any other edges. Thus we may assume A and B are adjacent, and similarly for the other consecutive pairs. If the remaining pairs (A, C), $(B, D), \ldots, (E, B)$ were all adjacent, we would have a copy of the complete graph K_5 which we saw above is impossible. Thus without loss we may assume A and C are not adjacent. Note that after we delete v, we can join A and C by a rubber band that follows the edges Av and vC and hence misses all the other edges. Pulling this rubber band tight, we can collapse A and C into a single vertex, call it w. A vertex x is adjacent to w if and only if it was adjacent to either A or C in the original graph.

This new planar graph has fewer vertices, hence by the inductive hypothesis we can color it with 5 colors. When we mentally cut the rubber band joining A and C, we will get a coloring of every vertex except v such that A and C (the two halves of w) have the same color. Note that since A and C are not adjacent, this is a valid coloring of the graph with v removed. Since A and C have the same color, the neighbors of v use at most four of the five colors, hence there is a color available to color v. This completes the inductive step.

Problem 7.37. In a simple graph with a finite number of vertices, each vertex has degree at least three. Prove that the graph contains a cycle whose length is not divisible by 3.

Solution. We prove the result by induction on the number of vertices of a graph. The minimal possible number of vertices is 4 for a complete graph on four vertices, in which we have a cycle of length four.

Now assume we have proven the assertion for all graphs with less than k vertices (where $k \geq 5$), and let us prove it for all graphs with k vertices. Assume we do have a counter-example with k vertices. Pick up a minimal cycle $X_1X_2 \ldots X_m$ (we do have such a cycle because we have at least $\frac{3k}{2} > k-1$ edges). Then m is divisible by 3. Also no two vertices X_i and X_j are connected unless they are consecutive on the cycle, because otherwise they would split the cycle into two smaller cycles. Furthermore, no other vertex Y is connected with two different vertices X_i, X_j because then one of the cycles $YX_iX_{i+1} \ldots X_j, YX_jX_{j+1} \ldots X_i$ would have length not divisible by 3, as the sum of their lengths is $m + 4$, which is not divisible by 3. Collapse $X_1, X_2, \ldots, X_m$ into one vertex X and join X with other vertex Y from the original graph if and only if Y is connected with (only) one of $X_1, X_2, \ldots, X_m$. In this new graph, every vertex except X will maintain its degree, while X will have degree at least m (as from $X_1, X_2, \ldots, X_m$ there are at least m more edges emerging outside the cycle) so also at least 3. Finally, if this new graph contains a cycle with length not divisible by 3, then it must contain X, otherwise it would be a cycle with length not divisible by 3 in the original graph. If we let the cycle be $XA_1A_2\ldots A_{r-1}$ where r is not divisible by 3, then A_1 is connected to some vertex X_i in the original graph and A_{r-1} is connected to some vertex X_j in the original graph. Thus there are cycles $X_jX_{j+1}\ldots X_iA_1A_2\ldots A_{r-1}$ and $X_jX_{j-1}\ldots X_iA_1A_2\ldots A_{r-1}$ in the original graph. The sum of their length is $r+m$, not divisible by 3, so one of the cycles will have length not divisible by 3, which is a contradiction. We deduce that our newly constructed graph has less than k vertices and has no cycle whose length is divisible by 3. But this contradicts the induction hypothesis. Our proof is complete.

Problem 7.38. (**China 2000**) A table tennis club wishes to organize a doubles tournament, a series of matches where in each match one pair of players competes against a pair of two different players. Let a player's *match number* for a tournament be the number of matches he or she participates in. We are given a set $A = \{a_1, a_2, \ldots, a_k\}$ of distinct positive integers all divisible by 6. Find with proof the minimal number of players among whom we can schedule a doubles tournament such that

i) each participant belongs to at most 2 pairs;

ii) any two different pairs have at most 1 match against each other;

iii) if two participants belong to the same pair, they never compete against each other;

iv) the set of the participants' match numbers is exactly A.

Solution. We begin by proving the following auxiliary result:

Lemma. Suppose that $k \geq 1$ and $1 \leq b_1 < b_2 < \ldots < b_k$. Then there exists a graph of $b_k + 1$ vertices such that the set $\{b_1, b_2, \ldots, b_k\}$ consists of the degrees of the $b_k + 1$ vertices.

Proof. We prove the lemma by strong induction on k. If $k = 1$, the complete graph on b_1 vertices suffices. If $k = 2$, then take $b_2 + 1$ vertices, distinguish b_1 of these vertices, and connect two vertices by an edge if and only if one of the vertices is distinguished.

We now prove that the claim is true when $k = i \geq 3$ assuming that it is true when $k < i$. We construct a graph G of $b_i + 1$ vertices, forming the edges in two steps and thus "changing" the degrees of the vertices in each step. Take $b_i + 1$ vertices and partition them into three sets S_1, S_2, S_3 with $|S_1| = b_1$, $|S_2| = b_{i-1} - b_1 + 1$ and $|S_3| = b_i - (b_{i-1} + 1)$. By the induction hypothesis, we can construct edges between the vertices in S_2 such that the degrees of those vertices form the set $\{b_2 - b_1, \ldots, b_{i-1} - b_1\}$. Further, construct every edge which has some vertex in S_1 as an endpoint. Each vertex in S_1 now has degree b_i, each vertex in S_3 has degree b_1 and the degrees of the vertices in S_2 form the set $\{b_2, \ldots, b_{i-1}\}$. Hence, altogether, the degrees of the $b_i + 1$ vertices in G form the set $\{b_1, b_2, \ldots, b_i\}$. This completes the inductive step and the proof.

Back to our original problem, suppose that we have a doubles tournament among n players satisfying the given conditions. At least one player, call him/her X, plays in the maximal number of matches which we will denote by $\max(A)$. Let m be the number of different pairs of players X has played against. Each of X's matches involves two opponents for a total count of $2m$. Any player is counted at most twice in this manner, as any player belongs to at most two pairs. Hence, player X must have played against at least m different players. If X is in j pairs (with $j = 1$ or $j = 2$), then there are at least $m + j + 1$ players in total (X, the j players who played with X, and at least m players who played against X). Also, X plays in at most jm matches, implying that $jm \geq \max(A)$. Hence,

$$n \geq m + j + 1 \geq \max(A)/j + j + 1 \geq \min\{\max(A) + 2, \max(A)/2 + 3\}.$$

Since $\max(A) \geq 6$, we have $\max(A) + 2 > \max(A)/2 + 3$, implying that $n \geq \max(A)/2 + 3$.

We now prove that $n = \max(A)/2 + 3$ is attainable. From the lemma, we can construct a graph G of $\frac{\max(A)}{6} + 1$ vertices whose degrees form the set $\{\frac{a_1}{6}, \frac{a_2}{6}, \ldots, \frac{a_k}{6}\}$. Partition the n players into $\max(A)/6+1$ triples, and let two players be in a pair if and only if they are in the same triple. Assign each triple (and, at the same time, the three pairs formed by the corresponding players) to a vertex in G, and let two pairs compete if and only if their corresponding vertices are adjacent. Suppose that we have a pair assigned to a vertex v of degree $\frac{a_i}{6}$. For each of the $\frac{a_i}{6}$ vertices w adjacent to w, that pair competes against the three pairs assigned to w, for a total of $\frac{a_i}{2}$ matches. Each player assigned to v is in two pairs and hence has match number $2\frac{a_i}{2} = a_i$. Therefore, the set of participants' match numbers is $\{a_1, a_2, \ldots, a_k\}$, as we wanted.

Problem 7.39. (**Poland 2000**) Given a natural number $n \geq 2$, find the smallest integer k with the following property: Every set consisting of k cells of an $n \times n$ table contains a non-empty subset S such that in every row and in every column of the table there are an even number of cells belonging to S.

Solution. The answer is $2n$. To see that $2n - 1$ cells do not suffice, consider the "staircase" of cells consisting of the main diagonal and the diagonal immediately below it. Number these cells from upper-left to lower-right. For any subset S of the staircase, consider its lowest-numbered cell; this cell is either the only cell of S in its row or the only cell of S in its column, so S cannot have the desired property.

To see that $2n$ suffices, we use the following lemma:

Lemma. If a graph is drawn such that its vertices are the cells of an $m \times n$ grid, where two vertices are connected by an edge if and only if they lie in the same row or the same column, then any set T of at least $m + n$ vertices includes the vertices of some cycle whose edges alternate between horizontal and vertical.

Proof. We induct on $m + n$. If $m = 1$ or $n = 1$, the statement is clearly true. Otherwise, we construct a trail as follows. We arbitrarily pick a starting vertex in T and, if possible, proceed horizontally to another vertex of T. We then continue vertically to another vertex of T. We continue this process, alternating between horizontal and vertical travel. Eventually, we must either

(a) be unable to proceed further, or (b) return to a previously visited vertex. In case (a), we must have arrived at a vertex which is the only element of T in its row or its column; then remove this row or column from the grid, remove the appropriate vertex from T, and apply the induction hypothesis. In case (b), we have formed a cycle. If there are two consecutive horizontal edges (resp. vertical edges), as is the case if our cycle contains an odd number of vertices, then we replace these two edges by a single horizontal (resp. vertical) edge. We thus obtain a cycle which alternates between horizontal and vertical edges. By construction, our cycle does not visit any vertex twice.

To see that the result of the lemma solves our original problem, suppose that we have a set of $2n$ cells from our $n \times n$ grid. It then contains a cycle: let S be the set of vertices of this cycle. Consider any row of the grid. Every square of S in this row belongs to exactly one horizontal edge, so if the row contains m horizontal edges, then it contains $2m$ cells of S. Thus, every row (and similarly, every column) contains an even number of cells of S.

Problem 7.40. (**Austrian-Polish MO 2000**) We are given a set of 27 distinct points in the plane, no three collinear. Four points from this set are vertices of a unit square; the other 23 points lie inside this square. Prove that there exist three distinct points X, Y, Z in this set such that $[XYZ] \leq \frac{1}{48}$.

Solution. We prove by induction on n that, given $n \geq 1$ points inside the square (with no three collinear), the square may be partitioned into $2n + 2$ triangles, where each vertex of these triangles is either one of the n points or one of the vertices of the square. For the base case $n = 1$, because the square is convex, we may partition the square into 4 triangles by drawing line segments from the interior point to the vertices of the square.

For the induction step, assume that the claim holds for some $n \geq 1$. Then for $n + 1$ points, take n of the points and partition the square into $2n + 2$ triangles whose vertices are either vertices of the square or are among the n chosen points. Call the remaining point P. Because no three of the points in the set are collinear, P lies inside one of the $2n + 2$ partitioned triangles ABC. We may further divide this into the triangles APB, BPC and CPA. This yields a partition of the square into $2(n + 1) + 2 = 2n + 4$, completing the induction.

For the special case $n = 23$, we may divide the square into 48 triangles with total area 1. One of the triangles has area at most $\frac{1}{48}$, as desired.

Problem 7.41. (**Moscow 1999**) We consider in the plane a convex polygon such that each of its sides is a segment which is colored towards the outside of the polygon (i.e. we regard a part of the segment as colored, and the other one not). Inside the polygon we draw some diagonals which again, have a side that is colored and one which is not. Prove that one of the polygons that was formed while partitioning the initial polygon must also have its sides colored towards the outside.

Solution. We prove the result by induction on the number n of diagonals that were drawn. For $n = 0$, the statement holds from the hypothesis.

We now assume the statement for some $n \geq 0$. Before we draw the $(n+1)$-st diagonal, there is a polygon $\mathcal{P}$ which is colored towards the exterior from the induction hypothesis. When we draw the last diagonal, it either intersects $\mathcal{P}$ or not. If it does not intersect $\mathcal{P}$, then in the final configuration, $\mathcal{P}$ is still colored towards the exterior. If it intersects $\mathcal{P}$, then it partitions $\mathcal{P}$ into two convex polygons, and the one situated on the uncolored side of the diagonal is the one which will now be colored towards the exterior. This completes the induction step.

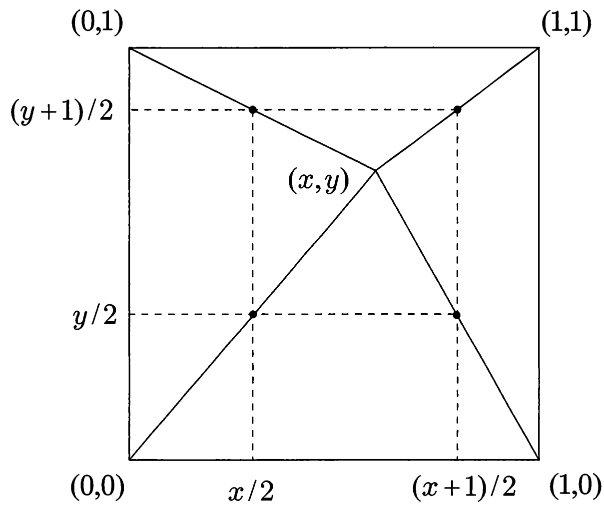
Problem 7.42. (**USSR 1989**) A fly and a spider are on a 1×1 meter square ceiling. In one second, the spider can jump from its position to the middle of any of the four segments which join it to the vertices of the ceiling. The fly does not move. Prove that in eight moves the spider can be within 1 centimeter of the fly.

Solution. We are going to prove a more general statement. Let us first introduce a rectangular system of coordinates of the ceiling, with the origin at one of its corners and the axes going along the corresponding edges. We take 1 meter as the unit of the length.

Let (x, y) be the coordinates of the initial position of the spider. Consider the set of points

$$A_k = \left\{ \left(\frac{x+i}{2^k}, \frac{y+j}{2^k} \right), i, j \in \mathbb{Z},\ 0 \leq i, j < 2^k \right\}.$$

This set gives us a network of cells whose sides have length equal to 2^{-k}. Obviously A_0 consists of just the single point (x, y) and the four points in A_1 are shown in the diagram below.



We see that when jumping from a point (x, y), the spider can change the first coordinate x of its location to either $x/2$ or $(x+1)/2$ and independently, it can change the second coordinate to either $y/2$ or $(y+1)/2$. We prove by induction on k that the spider can jump to any point of the set A_k after k jumps.

For $k = 0$, this is clear. Suppose that the spider can jump on any point of A_k after k jumps. Take an arbitrary point

$$M = \left(\frac{x+i}{2^{k+1}}, \frac{y+j}{2^{k+1}}\right)$$

of A_{k+1}. We have to find a point $N = (x_1, y_1)$ of A_k from which the spider can jump to M. Take

$$x_1 = \begin{cases} \frac{x+i}{2^k} & \text{if } i < 2^k \\ \frac{x+i}{2^k} - 1 & \text{if } i \geq 2^k \end{cases} \quad \text{and similarly} \quad y_1 = \begin{cases} \frac{y+j}{2^k} & \text{if } j < 2^k \\ \frac{y+j}{2^k} - 1 & \text{if } i \geq 2^k \end{cases}.$$

The point N which has these coordinates lies in A_k and the spider can jump from it to M.

Observe now that for any point P on the ceiling there exists a point in A_k whose distance from P is not greater than $\sqrt{2} \cdot 2^{-k}$. In particular, for $k = 8$ we have $\sqrt{2} \cdot 2^{-8} < \frac{1}{100}$, which solves the original problem.

Problem 7.43. Prove that if the plane is divided into parts ("countries") by lines and circles, then the obtained map can be painted in two colors so that the parts separated by an arc or a segment are of distinct colors.

Solution. We prove the statement by induction on the total number of lines and circles. For one line or circle, the statement is obvious.

Now suppose that it is possible to paint any map given by n lines and circles in the required way and we want to show how to paint a map given by $n + 1$ lines and circles. Let us delete one of these lines (or circles) and paint the map given by the remaining n lines and circles using our inductive hypothesis. Then we retain the colors of all the parts lying on one side of the deleted line (or circle) and replace the colors of all the parts lying on the other side of the deleted line (or circle) with opposite ones. In this way we achieve a coloring for $n + 1$ lines and circles, which finishes our induction and hence the question.

Problem 7.44. (**IMO 2006 shortlist**) Let S be a finite set of points in the plane such that no three of them are on a line. For each convex polygon P whose vertices are in S, let $a(P)$ be the number of vertices of P, and let $b(P)$ be the number of points of S which are outside P. A line segment, a point and the empty set are considered as convex polygons of 2, 1 and 0 vertices, respectively. Prove that for every real number x we have

$$\sum_P x^{a(P)}(1-x)^{b(P)} = 1,$$

where the sum is taken over all convex polygons with vertices in S.

Solution. We use strong induction on $|S|$, where the base case $|S| = 0$ is clear.

Consider the convex hull of S, and say it has n points $Q_1, Q_2, \ldots, Q_n$. Then by the inductive hypothesis (letting $S' = S - \{Q_{i_1}, Q_{i_2}, \ldots, Q_{i_k}\}$),

$$f(S') = \sum_{P \in S'} x^{a(P)}(1-x)^{b_{S'}(P)} = 1.$$

Using the Inclusion-Exclusion Principle to sum up $x^{a(P)}(1-x)^{b_S(P)}$ over all convex polygons $P \in S$ excluding at least one point on the convex hull $Q_1Q_2\ldots Q_n$ (note that the k points excluded are outside P), we obtain

$$f(S)-x^n(1-x)^0 = -\sum_{k=1}^{n}(-1)^k\binom{n}{k}(1-x)^k f(S') = -((1-(1-x))^n-(1-x)^0)$$

and thus $f(S) = 1$, as desired.

Problem 7.45. (**IMO 2006 shortlist**) A holey triangle is an upward equilateral triangle of side length n with n upward unit triangular holes cut out. A diamond is a $60° - 120°$ unit rhombus. Prove that a holey triangle T can be tiled with diamonds if and only if the following condition holds: Every upward equilateral triangle of side length k in T contains at most k holes, for $1 \le k \le n$.

Solution. Every diamond contains an upward and a downward triangle with a common edge. Now in every upward triangle T with sidelength k we have $\frac{k^2+k}{2}$ upward triangles and $\frac{k^2-k}{2}$ downward triangles. The downward triangles are adjacent only with upward triangles from T, so if more than k upward triangles are removed, there are not enough of them to match with the downward ones. This proves the easier part of the problem. Now let us prove the converse by induction on n:

Call a sequence $2k+1$ of consecutive upward and downward triangles (starting and ending with upward triangles) from the same row a "block". The block above it of $2k-1$ triangles is called the "above block". We wish to cover all the triangles in the last row of the big triangle T with diamonds. There can be horizontal diamonds, but there also can be vertical ones, and for them we will need to borrow upward triangles from the above row. When we finish, it is easy to conclude there will be $n-1$ holes in the remaining triangle

T' of side $n-1$. We just need to pick the diamonds in such a way that the induction hypothesis is respected. Call a triangle which contains a part of the base a "base" triangle. Call a triangle which contains as many holes as its sidelength "full", and one which contains more than its sidelength "deadly". We must not create "deadly" triangles after cutting the last row. It is clear a deadly triangle may appear only as a "base" triangle for T'.

We claim that if a block from the base contained r holes, at most r holes are added to the block above it after removing the last row. We do the following procedure: the space between two consecutive holes in the last row can be filed in with diamonds except one downward triangle which has to be connected with a triangle above (clearly at least one of the downward triangles can be matched otherwise there would be too many holes breaking the condition). It is is easy to verify the claim after such a conclusion. Moreover, it is clear that any arrangement is of the above type, as otherwise we would create too many holes in T' (between any two consecutive holes we must create a hole in T', and if we somewhere create more it would be over). Now we claim a "deadly" triangle arises from a full triangle. Indeed, if we have a base deadly triangle with sidelength j an $j+l$ holes we can extend it one unit downward. This triangle has sidelength $j+1$ and had originally also at least $j+l$ holes according to our claim. So $l=1$ and this triangle was full at the beginning. It only remains to manage the full base triangles, thus.

If two full triangles of sidelength k, l intersect in a triangle of sidelength m, we deduce this small triangle is also full and the big triangle of length $k+l-m$ is too. So we may do such unions unless $k+l-m$ is not n. When we finish, we have two possible cases:

a) we have some non-intersecting full triangles. Then every full base triangle is contained in one of them otherwise we could still unite it with one. Applying the induction hypothesis for them we can tile the bottom rows of them with diamonds without not making them full. Completing the tiling we achieve the goal.

b) We have intersecting triangles, but uniting them would form the big triangle T for which we can not apply the induction step. It is easy to see this is possible only when there are two such triangles of sidelength k, l which intersect in a triangle of sidelength m and $k+l-m=n$. Then we have at most m holes in the common part. So at least $k + l - m = n$ holes in the union of these two triangles. It means the remaining parallelogram at the top contains no hole and then it can be tiled with diamonds. The two triangles can also be tiled. Moreover any tiling tiles also the common part of sidelength m completely. So we may tile one triangle, and tile the remaining part of the second. This is the desired tiling of T.

Problem 7.46. (**IMO 2005 shortlist**) Let $n \geq 3$ be a given positive integer. We wish to label each side and each diagonal of a regular n-gon $P_1 \ldots P_n$ with a positive integer less than or equal to r so that:

i) every integer between 1 and r occurs as a label;

ii) in each triangle $P_i P_j P_k$, two of the labels are equal and greater than the third.

Given these conditions:

a) Determine the largest positive integer r for which this can be done.

b) For that value of r, how many such labellings are there?

Solution. Let $[XY]$ denote the label of the segment XY, where X and Y are vertices of the polygon. Consider any segment MN with the maximum label $[MN] = r$. By the condition (ii), for any $P_i \neq M, N$, exactly one of $P_i M$ and $P_i N$ is labelled by r. Thus the set of all vertices of the n-gon splits into two complementary groups: $\mathcal{A} = \{P_i : [P_i M] = r\}$ and $\mathcal{B} = \{P_i : [P_i N] = r\}$.

We claim that a segment XY is labelled by r if and only if it joins two points with one in $\mathcal{A}$ and the other in $\mathcal{B}$. Assume without loss of generality that $X \in \mathcal{A}$. If $Y \in \mathcal{A}$, then $[XM] = [YM] = r$, so $[XY] < r$. If $Y \in \mathcal{B}$, then $[XM] = r$ and $[YM] < r$, so $[XY] = r$, by (ii), as claimed.

We conclude that a labelling satisfying (ii) is uniquely determined by groups $\mathcal{A}$ and $\mathcal{B}$ and labellings satisfying (ii) within $\mathcal{A}$ and $\mathcal{B}$.

(a) We prove by induction on n that the greatest possible value of r is $n - 1$. The degenerate cases $n = 1$ and $n = 2$ are trivial. If $n \geq 3$, the number of different labels of segments joining vertices in $\mathcal{A}$ (respectively $\mathcal{B}$) does not exceed $|\mathcal{A}|-1$ (respectively $|\mathcal{B}|-1$), while all segments joining a vertex in $\mathcal{A}$ and a vertex in $\mathcal{B}$ are labelled by r. Therefore

$$r \le (|\mathcal{A}|-1)+(|\mathcal{B}|-1)+1 = n-1.$$

The equality is attained if all the mentioned labels are different.

(b) Let a_n be the number of labellings attained with $r=n-1$. We prove by induction that $a_n = \frac{n!(n-1)!}{2^{n-1}}$. This is trivial for $n=1$, so let $n\ge 2$. If $|\mathcal{A}|=k$ is fixed, the groups $\mathcal{A}$ and $\mathcal{B}$ can be chosen in $\binom{n}{k}$ ways. The set of labels used within $\mathcal{A}$ can be selected among $1,2,\dots,n-2$ in $\binom{n-2}{k-1}$ ways. Now the segments within groups $\mathcal{A}$ and $\mathcal{B}$ can be labelled so as to satisfy (ii) in a_k and a_{n-k} ways, respectively. This way every labelling has been counted twice, since choosing $\mathcal{A}$ is equivalent to choosing $\mathcal{B}$. It follows that

$$\begin{aligned} a_n &= \frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}\binom{n-2}{k-1}a_k a_{n-k} \\ &= \frac{n!(n-1)!}{2(n-1)}\sum_{k=1}^{n-1}\frac{a_k}{k!(k-1)!}\cdot\frac{a_{n-k}}{(n-k)!(n-k-1)!} \\ &= \frac{n!(n-1)!}{2(n-1)}\sum_{k=1}^{n-1}\frac{1}{2^{k-1}}\cdot\frac{1}{2^{n-k-1}} = \frac{n!(n-1)!}{2^{n-1}}. \end{aligned}$$

Problem 7.47. In a convex n-gon ($n \ge 403$), there are $200n$ diagonals drawn. Prove that one of them intersects at least 10000 others.

Solution. More generally, we prove that if $n > 150$ and if no diagonal intersects more than 10000 others, we have drawn at most $200(n-51)$ diagonals. We do this by induction on n.

If $n < 343$, we have $\frac{n(n-3)}{2} \le 200(n-51)$ and it is clear.

For the induction step, assume that we have proved the result for all integers less than n and $n > 343$. Suppose that we have drawn more than $200(n-51)$ diagonals. Then we have a vertex from which exit at least 343 diagonals. Take the 170th diagonal. This divides the polygon into two smaller polygons. Let the polygons have sides k, l, with $k+l=n+2$. We note that $k, l > 150$. Then by the induction hypothesis, the number of diagonals in the first polygon is at most $200(k-51)$ and in second at most $200(l-51)$. Summing, since we have at least $200(n-51)$ diagonals, we get at least 10000 diagonals intersecting the 170-th diagonal, which is contradiction.

Problem 7.48. Consider a convex n-gon such that no three of its diagonals are concurrent. In how many parts do the diagonals divide the n-gon?

Solution. Let us prove that diagonals divide our n-gon into

$$\frac{(n-1)(n-2)(n^2-3n+12)}{24}$$

parts. The base case $n=3$ is clear. Assume that the statement is true for all n-gons and let us prove it for $n+1$-gons. Consider an $n+1$-gon $A_1A_2\dots A_{n+1}$ that is divided by the diagonal A_1A_n into n-gon $A_1A_2\dots A_n$ and triangle $A_1A_nA_{n+1}$. The diagonals in the n-gon $A_1A_2\dots A_n$ divide it into $F(n)$ regions. When we add back the triangle $A_1A_nA_{n+1}$ we gain 1 more region. Now consider adding the diagonals A_kA_{n+1} for $2\le k\le n-1$ one at a time. If we add a diagonal which is cut by m other diagonals, then that means it splits $m+1$ regions. Since A_kA_{n+1} meets every diagonal A_iA_j with $1\le i<k$ and $k<j\le n$, we see that A_kA_{n+1} cuts $(k-1)(n-k)$ other diagonals and hence adds $(k-1)(n-k)+1$ regions. For the triangle $A_1A_nA_{n+1}$ and these $n-2$ diagonal we get a total of $n-1$ contributions of $+1$, so we obtain the following recursive formula

$$F(n+1)=F(n)+(n-1)+1\cdot(n-2)+\cdot 2(n-3)+\ldots+(n-3)\cdot 2+(n-2)\cdot 1.$$

This can be reduced to

$$F(n+1)=F(n)+\frac{n^3}{6}-\frac{n^2}{2}+\frac{4n}{3}-1,$$

which gives us the desired result after direct calculation.

Problem 7.49. We are given some unit squares which are translations of each other in the plane, such that from any $n+1$, at least two intersect. Prove that we can place at most $2n-1$ needles in the plane, such that every square is stabbed by at least one needle.

Solution. Introduce a system of coordinates with axes parallel to the sides of the squares. The proof is by induction on n. For $n = 1$, any two squares intersect. To see this, consider the leftmost of the squares (pick one if there is more than one leftmost squares), then every other square must intersect it. We conclude that every other square must meet the line through its right hand side. Intersecting each square with this line, we get a collection of unit intervals such that any two intersect. Take an uppermost one. Then any of our unit intervals contains the lowest point in this interval. Thus a needle placed at this point meets every square.

Now let us prove the result for $n = k + 1$ using the result for $n = k$, where $k \geq 1$. Let $ABCD$ be the leftmost square, such that A and B are the lower-left and upper-left vertices of the square. It is clear that all squares intersecting $ABCD$ must contain either C or D, so we can place two needles near them such that all squares intersecting $ABCD$ are stabbed (including $ABCD$). Next, the squares not intersecting $ABCD$ must satisfy the conditions in the hypothesis: if we take $k + 1$ of them and $ABCD$, then two squares must intersect. Since $ABCD$ does not intersect the other squares, we conclude that two of the $k + 1$ squares intersect. So applying the induction hypothesis, we can stab all the squares not intersecting $ABCD$ by $2k - 1$ needles. Together with the two previously placed needles, we get $2k+1 = 2(k+1)-1$ needles. This completes our proof.

8 Games

Problem 8.1. Three players play the following game. There are 54 candies on the table. Players play by turns (one after another). In each turn it is allowed to take either one, three, or five candies, such that the same number of candies cannot be taken in two consecutive turns. The winner is the person who takes the last set of candies. Given perfect play of all the players, who will win?

Solution. Let us prove, by mathematical induction, that if the number of candies is of the form $9n$, where $n \in \mathbb{N}^*$, then the third player wins.

If $n = 1$, then the third player can take all the candies left and win.

Assume now that the statement holds true for some $n \geq 1$ and we prove that it holds true for $n + 1$. We have $9(n + 1)$ candies; if the first player takes a candies and the second player takes b candies, then the third player takes $9 - a - b$ candies (that is, whichever of $1, 3, 5$ is neither a nor b). We are now left with $9n$ candies and from the induction hypothesis we obtain that the third player wins. As $9 \mid 54$, the third player will win the game we are given in the problem.

Problem 8.2. Let a_0, a_1, a_2,. . ., a_{2016} be distinct positive numbers. Consider the following two-player game: players take turns to write the numbers a_0, a_0, a_1, a_2,. . ., a_{2016} instead of $*$ in the polynomial $*x^{2016} + *x^{2015} + ... + *$ (one number in each turn). If the polynomial obtained in the end has an integer root, then the second player wins; otherwise, the first player wins. Given perfect play by both sides, who will win?

Solution. We shall prove that the first player wins. We begin by proving the following result:

Lemma. If $\max(|b_1|, |b_2|, ..., |b_n|) = |b_n|$ and $|k| \geq 2$, $n \geq 2$, then

$$|b_n k^{n-1}| - |b_{n-1} k^{n-2}| - ... - |b_1| \geq |b_n|.$$

Proof. We proceed by induction on n. The base case is $n = 2$, for which we have:

$$|b_2 k| - |b_1| \geq 2|b_2| - |b_1| \geq |b_2|.$$

Assume that the statement holds true for some $n \geq 2$ and we prove that it holds true for $n+1$.

We have that

$$\begin{aligned}|b_{m+1}k^m| - |b_m k^{m-1}| - \ldots - |b_1| &\geq 2|b_{m+1}||k^{m-1}| - |b_m||k^{m-1}| - \ldots - |b_1| \\ &\geq |b_{m+1}||k^{m-1}| - |b_{m-1}||k^{m-2}| - \ldots - |b_1| \\ &\geq |b_{m+1}|.\end{aligned}$$

It follows that

$$|b_{m+1}||k^m| - |b_m||k^{m-1}| - \ldots - |b_1| \geq |b_{m+1}|.$$

This completes the proof of our lemma.

Back to our original problem, without loss of generality, we can assume that $\max(a_0, a_1, \ldots, a_{2016}) = a_{2016}$.

If the first player in his first turn writes a_{2016} as the coefficient of x^{2016}, then the obtained polynomial does not have a root which is less than or equal to -2. Indeed, if $|k| \geq 2$ and

$$a_{2016}k^{2016} + a_{i_1}k^{2015} + \ldots + a_{i_{2015}} = 0,$$

then, according to the above lemma, we have

$$|a_{i_{2015}}| = |a_{2016}k^{2015} + \ldots + a_{i_{2014}}||k| \geq 2(|a_{2016}k^{2015}| - \ldots - |a_{i_{2014}}|) \geq 2|a_{2016}|.$$

Therefore, $a_{2016} \geq a_{i_{2015}} \geq 2a_{2016}$, which leads to a contradiction.

Notice that since the numbers $a_0, a_1, \ldots a_{2016}$ are positive integers, the polynomial we obtain at the end of the game cannot have a root which is greater than or equal to 0.

By the work we have done so far, we know that the first player can with his first move ensure that only -1 could be a root of our polynomial. Split the remaining coefficients into two groups those which are coefficients of x^k for k even and those which have k odd. We will call this choice of groups the parity of the coefficient (since at $x = -1$ the even group coefficients will get multiplied by $+1$ and the odd group by -1). For his next 1006 moves, the first player always writes a coefficient of the opposite parity to the previous move by the second player. Thus when we get to the last 4 turns there will be two coefficients of each parity left.

Suppose at the fourth turn from the end, the second player assigns a coefficient with parity p. Then the first player is left with three numbers, one of which will be assigned parity p and the other two the opposite parity. The first thing to note is that whichever value gets assigned to the sole remaining coefficient with parity p will determine the value of the polynomial at -1. The second thing to note is that the different assignments give different values for the polynomial at -1. Since at most one assignment can make the value 0, the first player can make this assignment in a way that guarantees that -1 is not a root of the polynomial. This completes our proof.

Problem 8.3. A regular pack of 52 cards (with 26 red cards and 26 black cards) is shuffled and dealt out to you one card at a time. At any moment, based on what you have seen so far, you can say "I predict that the next card will be red". You can only make this prediction once. Which strategy gives you the greatest chance of being right?

Solution. Let us introduce some terminology first: by strategy, we will understand the following: for every possible shuffling of the deck, we construct a corresponding word on 26 R (red) and 26 B (black) letters and denote by W the set of all such words. By a strategy, we mean a set S of words on R and B of length between 0 and 51 such that every word in W has exactly one element in S as its 'head', i.e. it begins with that word. Moreover, every element in S has to be the head of some element in W. This will determine, based on the cards that you have seen (which is basically a word, w) if you choose to predict ($w \in S$) or to wait ($w \notin S$) at each step. We also naturally assume that all possible configurations of cards have the same probability. With this understanding of the problem, we can prove that every strategy gives us a winning chance of 50%.

Let us use induction in the more general context of x red and y black cards in order to show that every strategy gives the probability $\frac{x}{x+y}$. We run our induction on $x+y$. For $x+y=2$, assume $x=y=1$, as the other cases are trivially checked. Then we can predict red from the beginning, or alternatively we have to shout red after we see the first revealed card. Clearly, all possibilities are $'RB'$ or $'BR'$. The first strategy prevails only for the first case, while the second strategy only for the second. Hence, both strategies have a 50% probability.

Now assume that the probability we achieve for $x+y$ cards is $\frac{x}{x+y}$. Let us prove that for X red cards and Y black cards with $X+Y=x+y+1$, the corresponding probability is $\frac{X}{X+Y}$. If a strategy S contains the empty word and hence no other, it has winning probability $\frac{X}{X+Y}$. Otherwise, the player would see the first card. In this case, partition S into two subsets: the set SR of words that start with R and the set SB of words that start with B. Conditioning on the event that the first card in the deck is red, SR becomes a strategy in its own right for a deck with $X-1$ red and Y black cards-provided we ignore the first letter of every word in SR. By induction, SR has winning probability $\frac{X-1}{X+Y-1}$. Similarly, SB has winning probability $\frac{X}{X+Y-1}$. Therefore, the overall winning probability of S is

$$\begin{aligned}\frac{XY}{(X+Y)(X+Y-1)}+\frac{X(X-1)}{(X+Y)(X+Y-1)} &= \frac{X(X+Y-1)}{(X+Y)(X+Y-1)}\\ &= \frac{X}{X+Y},\end{aligned}$$

as required. This completes the proof in the case of deterministic strategies, in which the player makes a prediction at any step before the last card.

Problem 8.4. On an infinite chessboard consisting of unit squares (x, y) with $x, y \geq 0$ two players play the following game: initially a king is positioned somewhere on the board, but not on $(0,0)$, and they alternatively move it either down or left or down-left. The player who *loses* is the one who moves the king into the $(0,0)$ square. Find the initial position of the king for which the first player wins.

Solution. We will prove by induction on $x+y \geq 1$ that the first player has a winning strategy if and only if either $x=0$ and y is even or $y=0$ and x is even or $x, y \geq 1$ and not both x and y are even.

Clearly for $x+y=1$, the first player is forced to move into $(0,0)$ and lose. Since the pairs $(0,1)$ and $(1,0)$ are not of the form above, this proves the base case.

For $x+y>1$, if one of x and y is 0, then without loss of generality $x=0$. The first player has to move into $(0,y-1)$ so, by induction, he wins if and only if y is even.

If $x=1$ or $y=1$, then without loss of generality $x=1$, so the first player can move into one of $(0,y)$ and $(0,y-1)$ which gives him a winning strategy.

Suppose $x,y\geq 2$. If x and y are both even, the first player has to move into one of $(x-1,y)$,$(x-1,y-1)$ or $(x,y-1)$, all of which have positive coordinates that are not both even, so, by induction, he loses. Else, one of $(x-1,y)$,$(x-1,y-1)$ and $(x,y-1)$ has even, non-negative coordinates, so he can move into that one, winning by induction.

Problem 8.5. (**TOT 2003**) In a game, Boris has 1000 cards numbered $2,4,\ldots,2000$, while Anna has 1001 cards numbered $1,3,\ldots,2001$. The game lasts 1000 rounds. In an odd-numbered round, Boris plays any card of his. Anna sees it and plays a card of hers. The player whose card has the larger number wins the round, and both cards are discarded. An even-numbered round is played in the same manner, except that Anna plays first. At the end of the game, Anna discards her unused card. What is the maximal number of rounds each player can guarantee to win, regardless of how the opponent plays?

Solution. We will show by induction on n that if there are $4n+1$ cards, then with optimal play Anna will win $n+1$ rounds and Boris will win $n-1$.

We will use two observations repeatedly. First, the game does not depend on the exact values of the cards, only that the cards from highest to lowest alternate between Anna and Boris with Anna having the highest and lowest. Second, suppose the player going second in a round has the option of playing two cards both of which result in the same winner for that round. Then playing the smaller of those two cards is always at least as good for that player as playing the larger. This is clear since the result of that round will be the same and at later rounds he will have a better card in his hand.

Now we first prove by induction on n that Anna can guarantee to win at least $n+1$ rounds. For the base case suppose $n=1$, that is, there are 5 cards. Whatever card k Boris plays first, Anna responds with card $k+1$ winning the first round. She then plays her highest card (which is the highest unplayed card) and wins the second round.

For the induction step, suppose Boris plays card k first. Then Anna plays card $k+1$ and wins the round. She then plays card 1. Since Boris will definitely win this round, by the second observation above he should play his lowest card (either card 2, if $k \neq 2$, or card 4 is $k = 2$). After these plays we are left with $4(n-1)+1$ cards which alternate between Anna and Boris. Hence by the first observation above and the induction hypothesis, Anna can win n of the remaining rounds and $n+1$ rounds in total.

Next, we prove that Boris can guarantee to win at least $n-1$ rounds. Boris will play card 2 first. By the second observation above, Anna should either play card 1 (losing) or 3 (winning with the smallest possible card). Since either way she is left with the lowest unplayed card, we may assume she plays card 3 and wins. Thus after this first round, we are left with $4n-1$ cards alternating between Anna and Boris, but now Anna goes first.

We will prove by induction on n that in this slightly modified game, Boris can guarantee to win $n-1$ rounds. The base case $n = 1$ is trivial since of course Boris can guarantee to win at least 0 rounds.

For the induction step, first assume Anna plays a card $k < 4n-1$, then Boris responds with card $k+1$ winning this round and returns card 2. As above, we may assume Anna plays card 3 and wins. This reduces the game to one equivalent to the $4(n-1)-1$ card game and hence by induction Boris can guarantee to win $n-2$ of the remaining rounds, hence $n-1$ total.

Now assume Anna plays her top card $4n-1$. Then Boris plays card 2 losing that round. He then plays card $4n-2$ in the next round. Since Anna will lose this round, we may assume she plays her lowest card 1. Thus again Boris has won one round and we are reduced to a game equivalent to the $4(n-1)-1$ card game. Hence by induction Boris can guarantee to win $n-2$ of the remaining rounds, hence $n-1$ total.

Since Anna can guarantee to win $n+1$ rounds and Boris can guarantee to win $n-1$ rounds (and this adds to the total number $2n$ of rounds), this must be optimal play for both players.

Problem 8.6. There are $n > 1$ balls in a box. Now two players A and B are going to play a game. At first, A can take out $1 \leq k < n$ ball(s). When one player takes out m ball(s), then the next player can take out $1 \leq \ell \leq 2m$ ball(s). The person who takes out the last ball wins. Find all positive integers

n such that B has a winning strategy.

Solution. Let $F_0 = 0, F_1 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$. The answer is that the second player wins if and only if n is of the form F_k for $k \geq 3$.

Let (n, m) denote the position where there are n balls left and the last move took m of them. Let $n = F_{i_1} + F_{i_2} + \cdots + F_{i_r}$ be the Zeckendorf representation of n, with $i_1 > i_2 > \cdots > i_r \geq 2$ and no two i_j are consecutive. We will prove by strong induction on n that (n, m) is losing for the first player if and only $m < F_{i_r}/2$.

The base case $n = 1$ is obvious. Suppose it is true for all values up to n, and consider (n, m). Write $n = F_{i_1} + F_{i_2} + \cdots + F_{i_r}$ as above. If $m \geq F_{i_r}/2$, then we take F_{i_r} balls away. Because $i_{r-1} > i_r + 1$, we have $F_{i_{r-1}} > 2F_{i_r}$, so this leaves our opponent in a losing position by the inductive hypothesis. Therefore the original position is a winning one.

It remains to show we cannot leave out opponent in a losing position if $m < F_{i_r}/2$. Suppose we move taking x balls. Let

$$n - x = F_{i_1} + F_{i_2} + \cdots + F_{i_{r-1}} + F_{j_1} + F_{j_2} + \cdots + F_{j_s}.$$

The Zeckendorf representation guarantees we can write all numbers $1, 2, \ldots F_{j_s - 1} - 1$ using only terms from $F_2, F_3, \ldots F_{j_s - 2}$. So $n - x + 1, n - x + 2, \ldots, n - x + F_{j_s - 1} - 1$ all have a representation starting with the terms that $n - x$ has. This is not true of n, so therefore $n > n - x + F_{j_s - 1} - 1$, which is equivalent to $x \geq F_{j_s - 1}$. But this is greater than $F_{j_s}/2$, so this is a winning position by the inductive hypothesis and the induction is complete.

To finish the problem, note that the starting position is as though we have $m = \frac{n-1}{2}$. It is straightforward to see that $\frac{n-1}{2} < F_{i_r}/2$ means that $n = F_{i_r}$, giving us the answer we claimed in the first paragraph.

Problem 8.7. (**Russia 2002**) We are given one red and $k > 1$ blue cells, and a pack of $2n$ cards, numbered from 1 to $2n$. Initially, the pack is situated on the red cell and arranged in an arbitrary order. In each move, we are allowed to take the top card from one of the cells and place it either onto the top of another cell on which the number on the top card is greater by 1, or onto an empty cell. Given k, what is the maximal n for which it is always possible to move all the cards onto a single blue cell?

Solution. We shall prove that the answer is $n = k - 1$. First, let us give a counterexample when $n = k$:

Assume that initially on the red cell, the cards are in the order (from top to bottom) $1, 3, \ldots, 2k-1, 2k$ and then the rest of the cards are in arbitrary order. Following the rules of the game, none of the odd numbers from 1 to $2k - 1$ can be placed on top of each other. So it must be that after k moves, each of the k blue cells contain one of the odd numbers from 1 to $2k - 1$. But then we are left with the card with number $2k$ which cannot be placed in any of the cells.

Now we show that when $n = k - 1$, we can always achieve our goal. We will do this by induction on k.

First, we call a position of the process *good* if the cards from 1 to $k - 1$ are located in once cell and the cards from k to $2k - 2$ are in another cell. Once we have a good position, we move the cards $1, 2, \ldots, k-2$ successively in the $k - 2$ remaining empty cell, so that each occupies one of these cells. Then in the cell with the cards from k to $2k - 2$, we place $k - 1$ on top, then $k - 2$, and so on, until we obtain we have all the cards in one stack, which completes the game.

Therefore, it suffices to prove by induction that we can obtain a good position for any k and $n = k - 1$. The base case is $k = 1$, when we only have two cards, and we put them in separate cells, no matter what their order is. Then we place 1 on top of 2 and we are done.

Assume now that the result holds for some $k \geq 1$ and we prove it for $k + 1$. We prove that we can obtain a position where the cards $2k - 1$ and $2k$ are in one cell, without disturbing the process which leads to a good position for $n = k - 1$ for the remaining k cells. To do this, we simply do the process from the induction hypothesis to get a good position for $n = k - 1$ and whenever we meet the cards $2k - 1$ and $2k$ we put them separately in the extra cell available. The only situation when we cannot do this is when the card $2k - 1$ comes before the card $2k$, and by the time we reached the card $2k$, all the k blue cells are occupied. If this is the case, we proceed as follows:

When we reach the card with number $2k$, we look at the k cells (as we exclude the one containing $2k - 1$) which are now full and consider the top card on each of them. There are $2k - 2$ possible card numbers and k cells, so by the Pigeonhole Principle, two of these card numbers must be consecutive. Say our consecutive card numbers are l and $l+1$. Since l could only have gone on top of $l+1$ or on an empty cell, it must be that l is the only number in its cell. So we place l on top of $l+1$, then we place $2k$ in the cell where l was, put $2k-1$ on top of it, and now we put l in the cell where $2k-1$ was. With this sequence of moves, we obtained a cell with $2k-1$ and $2k$ and the configuration for the rest of the cards is the same as when we reached the card $2k$, so we have not disturbed the process for the remaining $2k-2$ cards.

This shows that we can achieve the position where the cards $2k-1$ and $2k$ are in one cell, cards from 1 to $k-1$ in another, and the cards from k to $2k-2$ in a third separate cell. To finish the induction step, we must prove that this configuration leads to a good position. We place the cards from k to $2k-3$ successively in the remaining $k-2$ empty cells. Now $2k-2$ is on its own in a cell, and we place it on top of the cell with the cards $2k$ and $2k-1$. Then we successively place the cards from $2k-3$ to $k+1$ in the same pile. Now we have a pile with cards from $k+1$ to $2k$, one with cards from 1 to $k-1$, a cell containing k and the rest of the cells empty. We now successively place each of the cards from 1 to $k-2$ in the $k-2$ empty cells. Finally, we successively place the cards from $k-1$ to 1 on the pile containing the card with number k. In this manner, we obtained a good position, completing our induction.

Problem 8.8. (**Mathematical Reflections**) A and B play the following game on a $2n+1 \times 2m+1$ board: A has a pawn in the bottom left corner (square $(1,1)$) and wants to get to the top right corner (square $(2n+1, 2m+1)$). At each turn, A moves the pawn in an adjacent square (having a common edge) and B either does nothing or blocks a square for the rest of the game, but in such a way that A can still get to the top right corner. Prove that B can force A to do at least $(2n+1)(2m+1)-1$ moves before reaching the top right corner.

Solution. We shall prove this by induction on $m+n$. For $m=0$ or $n=0$ the result is clear.

Now assume $m \geq 1$ and $n \geq 1$.

B can apply the following strategy: block (in order) the following squares:

$$(2,2), (3,2), \ldots, (2n,2), (2n,3), \ldots, (2n,2m).$$

Note that it would require $k+1$ moves for player A to get to the k-th of these squares. Therefore player B can block these squares without interfering with A. Also not that this gives the first $2n+2m-3$ moves for B. Once B blocks these squares A can only win by passing through one of the two squares $(2n+1, 2m-1)$ or $(2n-1, 2m+1)$. It will take A at least $2n+2m-2$ moves for A to reach one of these squares, so he cannot be there yet.

We then have two cases depending on which of these two squares A reaches first:

Case 1. A first gets to $(2n+1, 2m-1)$ after at least $2n+2m-2$ moves. B can block the square $(2n+1, 2m)$ and then A will have to take the way back up to $(1,1)$ in at least $2n+2m-2$ moves, then to $(1,3)$ in at least 2 moves. Then, the game has essentially reduced to the analogous game on a $(2n-1)\times(2m-1)$ board. By induction B can make A go from $(1,3)$ to $(2n-1, 2m+1)$ in at least $(2n-1)(2m-1)-1$ moves, and then in another 2 moves to $(2n+1, 2m+1)$. Thus, the total number of moves is at least

$$(2n+2m-2)+(2n+2m-2)+2+(2n-1)(2m-1)-1+2 = (2n+1)(2m+1)-1.$$

Case 2. A first gets to $(2n-1, 2m+1)$ after at least $2n+2m-2$ moves. B can block the square $(2n, 2m+1)$. Now to get to the top right corner, A must first return to square $(1,3)$. This is again the same game on a $(2n-1)\times(2m-1)$ board. By induction, B can make A do at least $(2n-1)(2m-1)-1$ moves to get to $(1,3)$. Then A will do at least 2 moves to get to $(1,1)$ and then at least $2n+2m$ to get to $(2n+1, 2m+1)$. Thus we have a total of at least

$$(2n+2m-2)+(2n-1)(2m-1)-1+2+(2n+2m) = (2n+1)(2m+1)-1 \quad \text{moves.}$$

This completes our induction and thus the proof of the result.

Problem 8.9. (44th tournament of Ural, 4th tour). Consider the following two-player game: there are given two piles of stones, one consisting of 1914 stones and the other of 2014. In his turn, a player is allowed to remove either two stones from the bigger pile or one stone from the smaller pile; if at some stage the piles are equal, then he is allowed to remove from one of the piles either one or two stones. A player loses when cannot play his turn anymore. Given perfect play by both sides, who will win?

Solution. Let us prove that the second player wins. If the first player takes one stone, then the second player takes two stones. Otherwise, if the first player takes two stones, the second player takes one stone. In this manner, after each series of two turns, the larger pile decreases by 2 and the smaller pile decreases by 1. Thus after every pair of turns the total number of stones will always be 1 modulo 3. Also, after 99 pairs of turns, in one of the piles we will have 1816 stones and in the other one 1815 stones.

Now let us prove by mathematical induction that from this point onward after every turn of the second player, the difference between the number of stones in the piles is less than or equal to 3. We use downwards induction on $a+b$, where a and b represent the number of stones in the two piles.

Note that the statement holds true for $(1816, 1815)$, which is the base case.

For the induction step, assume the result for $a+b \equiv 1 \pmod 3$, $a+b \leq 1816+1815$ and we prove it for $a+b-3$. Note that from the configuration (n, n) we obtain $(n-1, n-2)$. From $(n, n+1)$ we obtain either $(n-1, n-1)$ or $(n, n-2)$. From $(n, n+2)$ we obtain $(n-1, n)$. Finally, from $(n, n+3)$ we obtain $(n-1, n+1)$. By symmetry, this covers all the possible cases.

At some point, the number of stones in the smaller pile will be equal to one of 0, 1 or 2. Moreover, we will end up with one of the following cases: $(0,1)$, $(1,3)$, $(2,2)$ or $(2,5)$. For the case $(2,5)$, after one pair of turns we can obtain $(1,3)$. In all of the remaining situations it is easy to see that the second player wins.

9 Miscellaneous Topics

9.1 Geometry

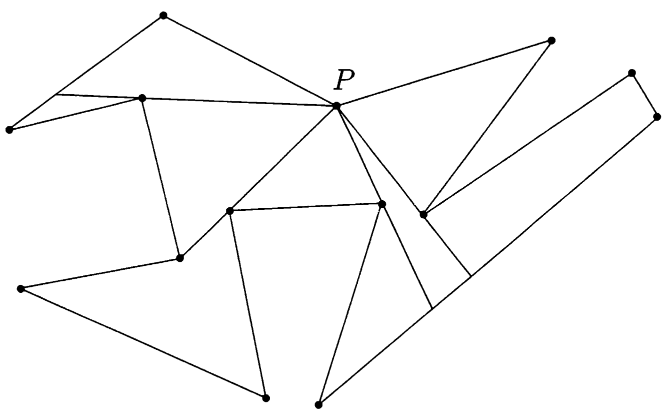
Problem 9.1.1. a) Prove that any n-gon can be cut into triangles by non-intersecting diagonals.

b) Prove that the sum of the inner angles of any n-gon is equal to $(n-2)180°$. Hence prove that the number of triangles into which an n-gon is cut by non-intersecting diagonals is equal to $n-2$.

Solution. We will first prove the following:

Lemma. Any n-gon $(n \geq 4)$ has at least one diagonal that completely lies inside it.

Proof. If the polygon is convex, there is nothing to prove. Otherwise, the exterior angle of the polygon at some vertex P is greater than $180°$. The visible part of any side subtends an angle smaller than $180°$ from a vertex at point P, therefore, parts of at least two sides subtend an angle with vertex at P. Consequently, there exist rays exiting point P such that on these rays, the change of (parts of) sides visible from P occurs (see the figure below). Each of such rays determines a diagonal that lies entirely inside the polygon.



a) We prove the statement by induction on n. For $n = 3$, things are clear.

Suppose now that we proved the statement for all k-gons, where $k < n$ (some $n \geq 4$) and let us prove it for an n-gon. By the lemma that we just proved, any n-gon can be divided by a diagonal into two polygons and the number of vertices of every of the smaller polygons is strictly less than n. By the induction hypothesis, both smaller polygons can be divided into triangles and hence we are done.

b) We proceed by induction on n. For $n = 3$, the result is immediate.

Suppose we proved the statement for all k-gons, where $k < n$ (some $n \geq 4$) and let us prove it for an n-gon. Again, by the lemma, any n-gon can be divided by a diagonal into two polygons. If the number of sides of one of the smaller polygons is equal to $k+1$, then the number of sides of the other one is equal to $n-k+1$ and both numbers are smaller than n. Therefore, the sum of the angles of these polygons are equal to $(k-1)\cdot 180°$ and $(n-k-1)\cdot 180°$, respectively. It is also clear that the sum of the angles of the n-gon is equal to the sum of the angles of these polygons, i.e. it is equal to $(k-1+n-k-1)\cdot 180° = (n-2)\cdot 180°$.

Problem 9.1.2. For any positive integer $n > 1$, prove that there exist 2^n points in the plane, no three collinear, such that no $2n$ of them form a convex polygon.

Solution. We use mathematical induction. For $n = 2$ we can take a triangle and a point inside the triangle. These 4 points do not form a convex quadrilateral.

For the induction step, suppose we found such a set S_n for some $n \geq 2$. Draw all segments among points in S_n and now choose a vector v with a sufficiently small magnitude such that the line joining any point of S_n with its translate by v does not intersect any of the segments drawn. Denote this translation by T. We will prove that the set S_{n+1} consisting of points of S_n and their translates (S'_n) is an example for $n + 1$. Indeed, suppose that we have found a $(2n + 2)$-sided convex polygon. We may replace each vertex that is in S'_n by the corresponding vertex in S_n, preserving the convexity of the polygon. At most two points from S'_n can occur among the vertices of the polygon together with their pre-images from S_n, since a convex polygon cannot have three parallel sides. Hence the new convex polygon has at least $2n+2-2 = 2n$ vertices among S_n, which contradicts the induction hypothesis.

Problem 9.1.3. Let $A_1 \ldots A_n$ be a convex polygon inscribed in a circle such that among its vertices, there are no two which form a diameter. Prove that if among the triangles $A_pA_qA_r$, as p, q, r range over $1, \ldots, n$, there is at least one acute triangle, then there are at least $n-2$ such acute triangles.

Solution. We prove the result by induction on n. For $n = 3$, the statement is clear. Now let $n \geq 4$. Fix one acute triangle $A_pA_qA_r$ and remove from the n-gon $A_1 \ldots A_n$ a vertex A_k distinct from A_p, A_q and A_r. The induction hypothesis applies to the obtained $(n-1)$-gon. Moreover, if for example the point A_k lies on the arc A_pA_q and $\angle A_kA_pA_r \leq \angle A_kA_qA_r$, then the triangle $A_kA_pA_r$ is acute, since $\angle A_pA_kA_r = \angle A_pA_qA_r$, $\angle A_pA_rA_k < \angle A_pA_rA_q$ and $\angle A_kA_pA_r \leq 90°$ (and thus $\angle A_kA_pA_r < 90°$).

Problem 9.1.4. a) Prove that the projection of a point P of the circumscribed circle of a cyclic quadrilateral $ABCD$ onto the Simson lines of this point with respect to triangles BCD, CDA, DAB and BAC lie on one line. (*This line is called the Simson line of P with respect to the inscribed quadrilateral*).

b) Prove by induction that we can similarly define the Simson line of a point P with respect to an inscribed n-gon as the line that contains the projections of point P on the Simson lines of all $(n-1)$-gons obtained by deleting one of the vertices of the n-gon.

Solution. Notice that part a) of the question will be the base case for the proof by induction that we will do for b). So we start by proving this statement. Let B_1, C_1, D_1 be the projections of point P to lines AB, AC and AD, respectively. Points B_1, C_1, D_1 lie on the circle with diameter AP. Lines B_1C_1, C_1D_1 and D_1B_1 are the Simson lines of point P with respect to triangles ABC, ACD and ADB, respectively. Therefore, the projections of the point P to the Simson lines of these triangles lie on one line, which is the Simson line of the triangle $\triangle B_1C_1D_1$. Arguing symmetrically, we see that the projections of P onto any three of the four Simson lines are collinear. Hence all four are collinear.

b) Let P be a point of the circumscribed circle of the n-gon $A_1 \ldots A_n$. Let $B_2, B_3, \ldots, B_n$ be the projections of P to the lines $A_1A_2, \ldots, A_1A_n$, respectively. Notice that the points $B_2, \ldots, B_n$ lie on the circle with diameter A_1P.

Let us prove by induction that the Simson line of P with respect to the n-gon $A_1 \dots A_n$ coincides with the Simson line of point P with respect to $(n-1)$-gon $B_2 \dots B_n$. The base case $n = 4$ was proved in part a).

By the induction hypothesis, the Simson line of the $(n-1)$-gon $A_1 A_3 \dots A_n$ coincides with the Simson line of the $(n-2)$-gon $B_3 \dots B_n$. Hence, the projections of P to the Simson line of the $(n-1)$-gons whose vertices are obtained by consecutively deleting the points $A_2, \dots A_n$, from the collection $A_1, \dots, A_n$ lie on the Simson line of the $(n-1)$-gon $B_2 \dots B_n$. The projection of the point P to the Simson line of the $(n-1)$-gon $A_2 \dots A_n$ lies on the same line, because our arguments show that any $n-1$ of the considered n points of projections lie on one line.

Problem 9.1.5. On the circle of radius 1 with center O there are given $2n+1$ points $P_1, \dots P_{2n+1}$ which lie on one side of a diameter. Prove that

$$|\overrightarrow{OP_1} + \ldots + \overrightarrow{OP_{2n+1}}| \geq 1.$$

Solution. We prove the statement by induction on n. For $n = 0$, the result is clear.

Assume now that the statement is proved for $2n+1$ vectors. In a system of $2n+3$ vectors, consider two vectors for which the angle between them is maximal. Without loss of generality, suppose that these vectors are $\overrightarrow{OP_1}$ and $\overrightarrow{OP_{2n+3}}$. By the induction hypothesis, the length of

$$\overrightarrow{OR} = \overrightarrow{OP_2} + \cdots + \overrightarrow{OP_{2n+2}}$$

is not shorter than 1. The vector $\overrightarrow{OR}$ belongs to the interior of the angle $\angle P_1 O P_{2n+3}$ and therefore it forms an acute angle with the vector

$$\overrightarrow{OS} = \overrightarrow{OP_1} + \overrightarrow{OP_{2n+3}}$$

which bisects $\angle P_1 O P_{2n+3}$. Hence

$$|\overrightarrow{OS} + \overrightarrow{OR}| \geq |\overrightarrow{OR}| \geq 1.$$

Problem 9.1.6. Let l_1 and l_2 be two parallel lines and $A, B \in l_1$, $A \neq B$. Using only a ruler, divide the segment AB into n equal parts, where $n \geq 2$.

Solution. We prove the result by induction on n. Notice that in order to obtain the required configuration, it suffices to find a point P_n on the segment AB such that $AP_n = \frac{1}{n}AB$. Then applying the same algorithm to the segment P_nB we find the next division point $P_n^{(2)}$, then the next point by working with the segment $P_n^{(2)}B$, etc.

For $n = 2$, let S be a an arbitrary point which lies in the half plane determined by l_2 and not containing l_1. Let $\{C\} = AS \cap l_2$, $\{D\} = BS \cap l_2$. Also, let $AD \cap BC = \{T_2\}$ and $ST_2 \cap l_1 = \{P_2\}$. We claim that P_2 is the midpoint of AB:

Let $ST_2 \cap l_2 = \{Q_2\}$. Then $\triangle T_2P_2B \backsim \triangle T_2Q_2C$, $\triangle ABT_2 \backsim \triangle DCT_2$, $\triangle SAP_2 \backsim \triangle SCQ_2$ and $\triangle SAB \backsim \triangle SCD$, hence

$$\frac{P_2B}{Q_2C} = \frac{T_2B}{T_2C} = \frac{AB}{CD} \quad \text{and} \quad \frac{P_2A}{Q_2C} = \frac{SA}{SC} = \frac{AB}{CD}.$$

This gives that

$$\frac{P_2B}{Q_2C} = \frac{P_2A}{Q_2C},$$

so $P_2A = P_2B$, as required.

Assume now that for some $n \geq 2$ we have constructed a point P_n on the segment AB such that $AP_n = \frac{1}{n}AB$. Let S, C, D be as above and let $\{T_n\} = SP_n \cap AD$, $\{Q_n\} = SP_n \cap l_2$. Let $T_{n+1} = AD \cap CP_n$. Also, let $\{Q_{n+1}\} = ST_{n+1} \cap l_2$, $\{P_{n+1}\} = ST_{n+1} \cap l_1$. We claim that $AP_{n+1} = \frac{1}{n+1}AB$: We have $\triangle CQ_{n+1}T_{n+1} \backsim \triangle P_nP_{n+1}T_{n+1}$, $\triangle CT_{n+1}D \backsim \triangle P_nT_{n+1}A$, so

$$\frac{P_{n+1}P_n}{CQ_{n+1}} = \frac{P_nT_{n+1}}{CT_{n+1}} = \frac{AP_n}{CD}.$$

We also have $\triangle SAP_{n+1} \backsim \triangle SCQ_{n+1}$ and $\triangle SAB \backsim \triangle SCD$, which gives

$$\frac{AP_{n+1}}{CQ_{n+1}} = \frac{SA}{SC} = \frac{AB}{CD}.$$

Combining the two derived equalities, we get

$$\frac{P_{n+1}P_n}{AP_{n+1}} = \frac{AP_n}{AB}.$$

Now, $P_{n+1}P_n = AP_n - AP_{n+1}$ and $AP_n = \frac{1}{n}AB$, which gives $AP_{n+1} = \frac{1}{n+1}AB$, as required.

Problem 9.1.7. Given $2n+1$ points in the plane such that no three of them are collinear, construct a $(2n+1)$-gon (self-intersecting polygons are allowed) for which the given points serve as the midpoints of its sides.

Solution. We prove the result by induction on n.

When $n = 1$, we need to construct a triangle, knowing the midpoints of its sides. We accomplish this by drawing through each of the three points a parallel to the line joining the other two points.

For the induction step, assume that we have established the result for $2n-1$ points and we shall prove it for $2n+1$ points. Let the $2n+1$ points be $M_1, \dots, M_{2n+1}$. Let us call the $(2n+1)$-gon that we want to construct $A_1A_2 \dots A_{2n+1}$.

Consider the quadrilateral $A_1A_{2n-1}A_{2n}A_{2n+1}$. We must have that M_{2n-1}, M_{2n} and M_{2n+1} are the midpoints of the sides $A_{2n-1}A_{2n}$, $A_{2n}A_{2n+1}$ and $A_{2n+1}A_1$, respectively. Let M be the midpoint of A_1A_{2n-1}. Then $M_{2n-1}M_{2n}M_{2n+1}M$ is a parallelogram. Since M_{2n-1}, M_{2n} and M_{2n+1} are given, M is uniquely determined by this condition.

Now $M_1, M_2, \dots, M_{2n-2}, M$ are the midpoints of the $(2n-1)$-gon $A_1A_2 \dots A_{2n-1}$, which can be constructed from the induction hypothesis. To finish the induction step, we construct the segments A_1A_{2n+1} and $A_{2n-1}A_{2n}$ (where A_1 and A_{2n-1} were already determined) such that M_{2n+1} and M_{2n-1} are their corresponding midpoints. By the way we constructed M, the remaining point M_{2n} will be the midpoint of $A_{2n}A_{2n+1}$, as required.

Problem 9.1.8. We are given a convex polygon in the plane. Show that we can pick a triangle formed by three of its vertices such that the sum of squares of its side lengths is at least the sum of squares of the side lengths of the polygon.

Solution. We prove the result by induction on the number $n \geq 3$ of sides. If the polygon is a triangle, we have nothing to prove.

If $n > 3$, the polygon has n angles with sum $(n-2)\pi \geq n\frac{\pi}{2}$, so it contains an angle which is obtuse. Without loss of generality, let the polygon be $A_1A_2\ldots A_n$ and let $\angle A_{n-1}A_nA_1$ be obtuse. Therefore, $(A_{n-1}A_1)^2 \geq (A_1A_n)^2 + (A_nA_{n-1})^2$, so the polygon $A_1A_2\ldots A_{n-1}$ has the sum of squares of its side lengths at least as big as the sum of squares of the side lengths of $A_1A_2\ldots A_n$. We are now done using the induction hypothesis for $A_1A_2\ldots A_{n-1}$.

Problem 9.1.9. (**TOT 2001**) Prove that there exist 2001 convex polyhedra such that any three of them do not have any common points, but any two of them touch each other (i.e. they have at least one common boundary point but no common inner points).

Solution. We shall describe the more general situation of constructing n convex polyhedra satisfying the hypotheses of the problem (in our case $n = 2001$).

We begin by considering a system of 3-dimensional coordinates, with the standard Ox, Oy and Oz axes. In this system we consider an infinite straight circular cone K with the vertex at the origin and the axis directed along Oz. Let $C(t)$ be a circle with the center $O(t) = (0,0,t)$ obtained by intersecting K with the plane $z = t$. Now consider a regular n-gon inscribed in $C(1)$ and label its vertices by $A_1, \ldots A_n$. Let B_i be the middle points of the arcs $\widehat{A_iA_{i+1}}$, where $i = 1, \ldots, n$ and we consider the indices modulo n. Further, let $A_i^t, B_i^t \in C(t)$ be the points of the intersection of the lines OA_i, OB_i with the plane $z = t$, where $t > 0$ and $i = 1, 2, \ldots, n$.

To proceed further with our construction we need a little lemma:

Lemma. For any $t_0 > 0$ and any i, $1 \leq i \leq n$, there exists $T > t_0$ such that for all $t \geq T$, the parallel translation of $C(t_0)$ by the vector $\overrightarrow{B_i^{t_0}B_i^t}$ lies inside the segment the region $S_i(t) = A_i^tB_i^tA_{i+1}^t$ bounded by an arc and a straight segment.

Proof. This follows immediately from the fact that the distance from B_i^t to $A_i^tA_{i+1}^t$ is proportional to t.

We now proceed to construct the polyhedra satisfying the conditions of the problem by induction. We start with any $t_1 > 0$. We choose any convex polygon M_1 inside a circle $C(t_1)$ and form an infinite "up prism" P_1 with base M_1 and the lateral edges parallel to OB_1.

Assume now that $n \geq 1$ and that we have already defined the numbers $0 < t_1 < \ldots < t_n$, we constructed the convex polygons $M_1, \ldots M_n$ contained in circles $C(t_1), \ldots, C(t_n)$ and formed the infinite prisms $P_1, \ldots P_n$ with bases $M_1, \ldots, M_n$ and lateral edges parallel to $OB_1, \ldots, OB_n$, satisfying the conditions of the problem.

Using the Lemma that we stated above, there exists $t_{n+1} > t_n$ such that $M_n(t_{n+1})$ lies inside the region $S_n(t_{n+1})$ and all the previous polygons remain within their areas. To define M_{n+1}, we proceed as follows:

Since it has to touch each of the previous prisms, in order to find the points of tangency, we make a parallel translation of every segment $A_i^{t_{n+1}} A_{i+1}^{t_{n+1}}$ until it touches the polygon $M_i(t_{n+1})$, where $1 \leq i \leq n$. If we connect the tangent points (we choose just one, if there are several for one tangency), we obtain a convex $M_{n+1}(t_{n+1})$.

To make the forthcoming arguments easier, we introduce the translated lines $l_i(t_{n+1})$, which separate the points M_i and M_{n+1}.

Now we can form an infinite "up prism" P_{n+1} with base $M_{n+1}(t_{n+1})$ and lateral edges parallel to OB_n. Now we cut the prism by the plane $z = T > t_{n+1}$.

We have to show that the obtained prisms satisfy the hypotheses of the question. From the inductive hypothesis, it suffices to check that P_{n+1} intersects P_i, $1 \leq i \leq n$ only at the plane $z = t_{n+1}$. Assume by contradiction that this is not the case. This means that there exists a common point $C \in z = t$ plane for all these prisms, where $t > t_{n+1}$. Consider the straight lines parallel to OB_{n+1} and OB_i passing through C. Denote the points of intersection of these lines with the plane $z = t_{n+1}$ by $C_{n+1}^{t_{n+1}}$ and $C_i^{t_{n+1}}$, respectively. Notice that $C_{n+1}^{t_{n+1}} \in M_{n+1}(t_{n+1})$ and $C_i^{t_{n+1}} \in M_i(t_{n+1})$. Moreover, the vectors $\overrightarrow{C_i^{t_{n+1}} C_{n+1}^{t_{n+1}}}$ and $\overrightarrow{B_i^{t_{n+1}} B_{n+1}^{t_{n+1}}}$ have opposite directions and are not 0. But this cannot happen, as $C_i^{t_{n+1}}$ and $B_i^{t_{n+1}}$ lie on one side of $l_i(t_{n+1})$, while $C_i^{t_{n+1}}$ and $B_{n+1}^{t_{n+1}}$ lie on the other side. This gives our desired contradiction and finishes the induction and the proof.

Problem 9.1.10. **(IMO 1992)** Let S be a finite set of points in three-dimensional space. Let S_x, S_y, S_z be the sets consisting of the orthogonal projections of the points of S onto the yz-plane, zx-plane and xy-plane, respectively. Prove that

$$|S|^2 \leq |S_x| \cdot |S_y| \cdot |S_z|,$$

where $|A|$ denotes the number of elements in the finite set A.
(Note: the orthogonal projection of a point onto a plane is the foot of the perpendicular from that point to the plane).

Solution. We shall prove the statement by induction on $|S|$. Let $|S_x| = a$, $|S_y| = b$, $|S_y| = c$. For a set S consisting of just one point, the statement is true.

Suppose that the statement is true for all sets with fewer than n points. Consider now a set S with $|S| = n$. Choose a plane parallel to one of the coordinate planes, which contains no points of S and divides S into two non-empty subsets S_1 and S_2. Then $n = |S_1| + |S_2|$, where $|S_1| < n$ and $|S_2| < n$. By the induction hypothesis,

$$|S_1| < a_1 b_1 c_1, \quad |S_2|^2 < a_2 b_2 c_2,$$

where a_i, b_i, c_i are the number of elements in the projections of S_i, $i = 1, 2$ onto the coordinate planes yz, zx and xy, respectively. Without loss of generality, we may assume that the dividing plane is parallel to the coordinate plane xy. Then

$$a_1 + a_2 = a, \quad b_1 + b_2 = b, \quad c_1 < c, \quad c_2 < c,$$

and by the Cauchy-Schwarz inequality

$$\begin{aligned} |S|^2 &= (|S_1| + |S_2|)^2 \leq \left(\sqrt{a_1 b_1 c_1} + \sqrt{a_2 b_2 c_2}\right)^2 \\ &\leq \left(\sqrt{a_1 b_1}\sqrt{c} + \sqrt{a_2 b_2}\sqrt{c}\right)^2 \leq c(a_1 + a_2)(b_1 + b_2) \\ &= abc. \end{aligned}$$

Problem 9.1.11. Prove that any convex n-gon which is not a parallelogram can be enclosed by a triangle whose sides lie along three sides of the given n-gon.

Solution. We first prove the following weaker statement:

Lemma. Any convex n-gon can be enclosed by a triangle or by a parallelogram whose sides lie along three (or four sides, respectively) of the given n-gon.

Proof. We shall prove the result by induction on $n \geq 3$. For $n = 3$, there is nothing to prove. When $n = 4$, either the quadrilateral is a parallelogram, or it has two non-parallel opposite sides. In the latter case, these two sides, together with the side farther from their intersection of the remaining two form a triangle whose interior contains the quadrilateral.

Assume now that $n \geq 5$ and we have proved the result for all k-gons, with $k < n$. Let $\mathcal{M}$ be a convex n-gon and let AB be an arbitrary side of $\mathcal{M}$. Since $n \geq 5$, there exists another side which is neither adjacent nor parallel to AB. Call this side CD. Now we extend both AB and CD until they intersect at a point O. Without loss of generality, assume that BC lies in the interior of triangle AOD (the other case being the same). Since $n \geq 5$, the polygonal line between B and C in $\mathcal{M}$ has length at least 2. So if we replace this polygonal line by the polygonal line BOC we obtain another polygon $\mathcal{M}_1$ having fewer than n sides and containing the polygon $\mathcal{M}$ inside it, al the sides of $\mathcal{M}_1$ being simultaneously the sides of $\mathcal{M}$.

From the induction hypothesis, there exists a triangle or a parallelogram formed by the sides of $\mathcal{M}_1$ which contains $\mathcal{M}_1$ within its interior. Since $\mathcal{M}_1$ contains the $\mathcal{M}$ and the sides of $\mathcal{M}_1$ are simultaneously the sides of $\mathcal{M}$, this shows that the assertion holds for $\mathcal{M}$ as well.

This completes the induction step and hence the proof of our lemma.

Having proved the above lemma, we need to consider two situations. If the polygon which contains $\mathcal{M}$ inside itself is a triangle, there is nothing to prove. So let us assume that the polygon is a parallelogram $ABCD$ and that $\mathcal{M}$ is not a parallelogram itself.

Without loss of generality, let A be the vertex of the parallelogram which is not a vertex of $\mathcal{M}$. Let P be the vertex of $\mathcal{M}$ nearest to A lying on the side AB of the parallelogram. From our construction, there is a side PQ of $\mathcal{M}$ which lies inside the parallelogram. $\mathcal{M}$ is convex, so $\mathcal{M}$ lies completely in the half plane bounded by PQ and containing the points B, C, D. So it must be that $\mathcal{M}$ lies completely inside the triangle determined by the lines PQ, BC and CD. This completes the proof of our statement.

Problem 9.1.12. Given a circle and n points in the plane, construct an n-gon (self-intersecting polygons are allowed) which is inscribed in the given circle and such that the lines determined by its sides pass through the given points.

Solution. We start with the following, more general problem:

Given a circle in the plane, one can construct an n-gon inscribed in the circle such that the k lines determined by k consecutive sides of it $(1 \leq k \leq n)$ pass through k given points in the plane and the remaining $n-k$ sides are parallel to some given lines.

Notice that this result implies our question when $k = n$.

We prove the above result by induction on k. We shall adjust the statement of the above problem as we go along, imposing some restrictions on given $n-k$ lines so that our induction works (Notice that even with further restrictions on the $n-k$ given lines, the statement of our problem will still hold, as we are interested in the case $k = n$). When $k = 1$, we need to construct an n-gon $A_1A_2 \dots A_n$ inscribed in a given circle, such that the line A_1A_n passes through a given point P and the $n-1$ remaining sides $A_1A_2, A_2A_3, \dots A_{n-1}A_n$ are parallel to the given lines $l_1, l_2, \dots l_{n-1}$.

Let B_1 be an arbitrary point on the circle. Construct an inscribed polygon $B_1B_2 \dots B_n$ whose sides $B_1B_2, B_2B_3, \dots, B_{n-1}B_n$ are parallel to the lines $l_1, l_2, \dots, l_{n-1}$. Basic geometry shows that that the polygon $A_1A_2 \dots A_n$ we need to construct should then satisfy the following: The arcs $\overset{\frown}{A_1B_1}, \overset{\frown}{A_2B_2}, \dots, \overset{\frown}{A_nB_n}$ are all equal and moreover, the pairs of arcs $\overset{\frown}{A_1B_1}$ and $\overset{\frown}{A_2B_2}$, $\overset{\frown}{A_2B_2}$ and $\overset{\frown}{A_3B_3}, \dots$ have opposite directions (e.g. if $\overset{\frown}{A_1B_1}$ is directed clockwise, then $\overset{\frown}{A_2B_2}$ is directed anti-clockwise).

This implies that when n is even, $\overset{\frown}{A_1B_1}$ and $\overset{\frown}{A_nB_n}$ should be oppositely directed and $A_1B_1B_nA_n$ is an isosceles trapezoid having bases A_1A_n and B_1B_n. Hence A_1A_n should be parallel to B_1B_n. Moreover, it should pass through the point P, which determines A_1 and A_n uniquely (For this construction to be possible, we would need that P lies inside the strip determined by the two tangents to the circle which are parallel to B_1B_n The same argument we used above to show that A_1A_n is parallel to B_1B_n, shows that B_1B_n will be parallel to the same fixed direction, regardless of the choice of B_1. So to make our argument work, we need to choose a configuration for the lines $l_1, \dots, l_{n-1}$ so that P lies in the required strip). Once we have determined A_1 and A_n, the rest of the points $A_2, \dots, A_{n-1}$ are uniquely determined by the conditions we specified.

When n is odd, $\overset{\frown}{A_1B_1}$ and $\overset{\frown}{A_nB_n}$ have the same direction, and the quadrilateral $A_1B_1A_nB_n$ is an isosceles trapezoid with bases A_1B_n and B_1A_n. Since the diagonals A_1A_n and B_1B_n are equal, we need to draw through the point P a straight line on which the given circle cuts off a chord A_1A_n equal to the known chord B_1B_n. This is a a straight line which is tangent to the circle which has the same center as the given circle and is tangent to B_1B_n. (Again, we impose the appropriate conditions on the lines $l_1, \ldots, l_{n-1}$ so that this construction is possible).

Assume now that we have solved the problem for some $1 \leq k < n$. We want to inscribe in a given circle an n-gon $A_1 \ldots A_n$, whose $k+1$ consecutive sides $A_1A_2, A_2A_3, \ldots, A_{k+1}A_{k+2}$ pass through $k+1$ given points $P_1, P_2, \ldots, P_{k+1}$ and the remaining $n-k-1$ sides are parallel to some $n-k-1$ given lines. Again, we shall assume that the $n-k-1$ lines are such that all the constructions we are making are possible.

Similarly as in the case $k = 1$, we seek for some conditions which would determine the points $A_1, \ldots, A_n$ uniquely. If $A_1 \ldots A_n$ is the required polygon, consider the following:

We draw a line through A_1 which is parallel to P_1P_2 and we denote its intersection with the circle by A_2'. We also denote the intersection of the lines $A_2'A_3$ and P_1P_2 by P_2'. Then $\angle A_2P_1P_2 = \angle A_2A_1A_2' = \angle A_2A_3P_2'$ and $\angle A_2P_2P_1 = \angle P_2'P_2A_3$. So the triangles $P_1A_2P_2$ and $P_2'P_2A_3$ are similar and we have

$$\frac{P_1P_2}{A_3P_2} = \frac{A_2P_2}{P_2'P_2} \Rightarrow P_2'P_2 = \frac{A_3P_2 \cdot A_2P_2}{P_1P_2}.$$

Now, the product $A_3P_2 \cdot A_2P_2$ depends only on P_2 and the circle (and not on A_2 and A_3, by the power of the point property), so it can be explicitly determined. So the length of $P_2'P_2$ can be explicitly found and so P_2' can be constructed. Now the k consecutive sides $A_2'A_3, A_3A_4, \ldots, A_{k+1}A_{k+2}$ of the n-gon $A_1A_2'A_3 \ldots A_n$ pass through the k points $P_2', P_3, \ldots, P_{k+1}$, while the remaining $n-k$ sides are parallel to the known lines. From the induction hypothesis, we can construct the polygon $A_1A_2'A_3 \ldots A_n$ satisfying these properties and then the above conditions determine the point A_2 uniquely, establishing the induction step.

9.2 Induction in Calculus

Problem 9.2.1. Let $f : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$, $f(x) = \arctan(x)$. Show that if n is a positive integer, we have that

$$f^{(n)}(0) = \begin{cases} 0 & \text{if } n = 2k; \\ (-1)^k (2k)! & \text{if } n = 2k+1, \end{cases}$$

where $f^{(n)}$ denotes the n-th derivative of f.

Solution. We have that

$$f'(x) = \frac{1}{1+x^2} \quad \text{and} \quad f''(x) = \frac{-2x}{(1+x^2)^2},$$

so the result holds for $n = 1$ and $n = 2$, as $f'(0) = 1$ and $f''(0) = 0$.

We prove the result of the question by induction of step 2. The base cases $n = 1$ and $n = 2$ were checked above. To perform the induction step, we look for a recursion which relates $f^{(n+1)}(0)$ to the lower order derivatives. To derive such a recursion, notice from the above that we have

$$f'(x)(1+x^2) = 1.$$

We differentiate this equality n times and we use Leibniz's formula to get

$$\binom{n}{0} f^{(n+1)}(x)(1+x^2) + \binom{n}{1} f^{(n)}(x) \cdot 2x + \binom{n}{2} f^{(n-1)}(x) \cdot 2 = 0,$$

which gives $f^{(n+1)}(0) = -n(n-1) f^{(n-1)}(0)$.

Assuming now that $f^{(n)}(0) = 0$ for some even $n \geq 2$, we have that

$$f^{(n+2)}(0) = -(n+1) n f^{(n)}(0) = 0.$$

If we assume now that $f^{(2k-1)}(0) = (-1)^{k-1} \cdot (2k-2)!$, for some $k \geq 1$, from the above recurrence we obtain

$$f^{(2k+1)}(0) = -(2k)(2k-1) f^{(2k-1)}(0) = (-1)^k \cdot (2k)!.$$

Therefore, the induction step is proved in both cases and we are done.

Problem 9.2.2. Let $f : [-1,1] \to (-\frac{\pi}{2}, \frac{\pi}{2})$, $f(x) = \arcsin(x)$. Prove that

$$f^{(n)}(0) = \begin{cases} 0 & \text{if } n = 2k; \\ (1 \cdot 3 \cdot 5 \ldots \cdot (2k-1))^2 & \text{if } n = 2k+1, \end{cases}$$

where $f^{(n)}$ denotes the n-th derivative of f.

Solution. We have that

$$f'(x) = \frac{1}{\sqrt{1-x^2}} \quad \text{and} \quad f''(x) = \frac{x}{(1-x^2)\sqrt{1-x^2}} = \frac{x}{1-x^2} f'(x),$$

so the result is verified for $n = 1$ and $n = 2$, as $f'(0) = 1$ and $f''(0) = 0$.

Moreover, by differentiating the relation $f''(x)(1-x^2) = f'(x) \cdot x$ n times and using Leibniz's formula we have

$$\begin{aligned} \binom{n}{0} f^{(n+2)}(x)(1-x^2) + \binom{n}{1} f^{(n+1)}(x)(-2x) + \binom{n}{2} f^{(n)}(x)(-2) \\ = \binom{n}{0} f^{(n+1)}(x) \cdot x + \binom{n}{1} f^{(n)}(x). \end{aligned}$$

This shows that $f^{(n+2)}(0) = n^2 f^{(n)}(0)$. Using this recursion relation and the base cases $f'(0) = 1$ and $f''(0) = 0$, we immediately obtain by induction that

$$f^{(n)}(0) = \begin{cases} 0 & \text{if } n = 2k; \\ (1 \cdot 3 \cdot 5 \ldots \cdot (2k-1))^2 & \text{if } n = 2k+1, \end{cases}$$

as required.

Problem 9.2.3. Let $f : [0, \infty) \to \mathbb{R}$ be given by

$$f(x) = \int_0^1 e^{-t} t^{x-1} \, dt.$$

Prove that for any non-negative integer n we have

$$f(n+1) = n! - \frac{1}{e} \sum_{k=0}^{n} \frac{n!}{(n-k)!}.$$

Solution. We prove the result by induction on n. For $n = 0$ we have

$$f(1) = \int_0^1 e^{-t}\, dt = e^{-t}\Big|_0^1 = 1 - \frac{1}{e},$$

so the identity holds for $n = 0$.

Assume now that the relation holds for some $n \geq 0$. To perform the induction step, we need to find a suitable relation between $f(x+1)$ and $f(x)$:

$$\begin{aligned}
f(x+1) &= \int_0^1 e^{-t} t^x\, dt = \int_0^1 \left(-e^{-t}\right)' t^x\, dt \\
&= -e^{-t} t^x\Big|_0^1 + x \cdot \int_0^1 e^{-t} t^{x-1}\, dt \quad \text{using integration by parts} \\
&= -\frac{1}{e} + x \cdot f(x).
\end{aligned}$$

This gives $f(n+1) = -\frac{1}{e} + n \cdot f(n)$.
Assume now that

$$f(n+1) = n! - \frac{1}{e} \sum_{k=0}^{n} \left(\frac{n!}{(n-k)!} \right).$$

Then

$$\begin{aligned}
f(n+2) &= -\frac{1}{e} + (n+1) \cdot f(n+1) \\
&= -\frac{1}{e} + (n+1) \cdot \left(n! - \frac{1}{e} \sum_{k=0}^{n} \left(\frac{n!}{(n-k)!} \right) \right) \\
&= -\frac{1}{e} + (n+1)! - \frac{1}{e} \sum_{k=0}^{n} (n+1) \frac{n!}{(n-k)!} \\
&= (n+1)! - \frac{1}{e} \left((n+1)! + \sum_{k=0}^{n} \frac{(n+1)!}{(n-k)!} \right) \\
&= (n+1)! - \frac{1}{e} \sum_{k=0}^{n+1} \frac{(n+1)!}{(n+1-k)!},
\end{aligned}$$

which proves the induction step.

Problem 9.2.4. Prove that for any $n \in \mathbb{Z}$, $n \geq 1$, one has

$$\lim_{x\to 0} \frac{n!x^n - \sin(x)\sin(2x)\dots\sin(nx)}{x^{n+2}} = \frac{n(2n+1)}{36} \cdot n!.$$

Solution. We shall prove that the limit is well defined for each $n \geq 1$ while we are finding the suitable recurrence relation.

Let L_n be the corresponding limit to $n \geq 1$. For $n = 1$, we have by repeated application of L'Hôpital's rule that

$$L_1 = \lim_{x\to 0} \frac{x - \sin(x)}{x^3} = \frac{1}{6}.$$

Also, one has

$$\begin{aligned}
L_n - nL_{n-1} &= \lim_{x\to 0} \frac{nx\sin(x)\dots\sin(x(n-1)) - \sin(x)\sin(2x)\dots\sin(nx)}{x^{n+2}} \\
&= \lim_{x\to 0} \frac{\sin(x)}{x} \cdot \frac{\sin(2x)}{x} \cdots \frac{\sin((n-1)x)}{x} \lim_{x\to 0} \frac{nx - \sin(nx)}{x^3} \\
&= (n-1)! \lim_{x\to 0} \frac{n - n\cdot\cos(nx)}{3n^2} \\
&= n! \lim_{x\to 0} \frac{n\sin(nx)}{6n} \\
&= n! \cdot \frac{n^2}{6}.
\end{aligned}$$

Therefore we have $L_n = nL_{n-1} + n! \cdot \frac{n^2}{6}$.
We now prove by induction the statement

$$P(n): \quad L_n = \frac{n(2n+1)}{36} \cdot (n+1)!.$$

For $n = 1$ we proved the result above, since $\frac{1\cdot 3\cdot 2!}{36} = \frac{1}{6}$.
Assume now that $P(n-1)$ is true. From the above recurrence we have

$$\begin{aligned}
L_n &= n \cdot \frac{(n-1)(2n-1)\cdot n!}{36} + \frac{n^2\cdot n!}{6} = \frac{n!\cdot n}{36}[(n-1)(2n-1) + 6n] \\
&= \frac{n!\cdot n(n+1)(2n+1)}{36} = \frac{n(2n+1)}{36}(n+1)!.
\end{aligned}$$

This establishes the induction step and completes our proof.

Problem 9.2.5. Prove that if m, n are non-negative integers with $m > n$, then

$$\int_0^{\pi} \cos^n(x)\cos(mx)\ dx = 0.$$

Using this, deduce the value of

$$J_n = \int_0^{\pi} \cos^n(x)\cos(nx)\ dx, \quad \text{for all non-negative integers} \quad n.$$

Solution. Let $I_n = \int_0^{\pi} \cos^n(x)\cos(mx)\ dx$ and let $P(n)$ be the mathematical statement

$$P(n): \quad I_n = 0 \quad \text{for any} \quad m > n \geq 0.$$

For $n = 0$ we have

$$I_0 = \int_0^{\pi} \cos(mx)\ dx = \left.\frac{\sin(mx)}{m}\right|_0^{\pi} = 0.$$

Assume now that $P(n)$ holds for some $n \geq 0$. Then

$$\begin{aligned}
I_{n+1} &= \int_0^{\pi} \cos^{n+1}(x)\cos(mx)\ dx \\
&= \frac{1}{2}\int_0^{\pi} \cos^n(x)\left(\cos\left((m+1)x\right) + \cos\left((m-1)x\right)\right)\ dx \\
&= \frac{1}{2}\int_0^{\pi} \cos^n(x)\cos\left((m+1)x\right)\ dx + \frac{1}{2}\int_0^{\pi} \cos^n(x)\cos\left((m-1)x\right)\ dx \\
&= 0 \quad \text{using the induction hypothesis, as} \quad m > n+1 \Rightarrow m-1 > n.
\end{aligned}$$

So $P(n+1)$ is true, and by the principle of induction, $P(n)$ holds for all non-negative integers n, which is what we wanted.
For the second part of the question, notice that

$$J_0 = \int_0^{\pi} dx = \pi,$$

and

$$J_1 = \int_0^{\pi} \cos^2(x)\ dx = \int_0^{\pi} \frac{1+\cos(2x)}{2}\ dx = \frac{1}{2}\left(x + \sin(2x)\right)|_0^{\pi} = \frac{\pi}{2}.$$

We prove by induction on n that $J_n = \frac{\pi}{2^n}$, for all $n \geq 0$. We have just verified the base cases and assuming the result for some $n \geq 0$, we have

$$J_{n+1} - \frac{1}{2}J_n = \int_0^\pi \cos^n(x)\left(\cos\left((n+1)x\right)\cos(x) - \frac{1}{2}\cos(nx)\right)\, dx$$
$$= \frac{1}{2}\int_0^\pi \cos^n(x)\cos\left((n+2)x\right)\, dx.$$

From the first part of the question we know that

$$\int_0^\pi \cos^n(x)\cos\left((n+2)x\right)\, dx = 0,$$

so we have that $J_{n+1} = \frac{1}{2}J_n$, which proves the induction step.

Problem 9.2.6. Let $a_1, a_2, \ldots, a_{2001}$ be non-zero real numbers. Prove that there exists a real number x, such that

$$\sin(a_1 x) + \sin(a_2 x) + \ldots + \sin(a_{2001} x) < 0.$$

Solution. We begin by proving the following general result:

Lemma. Let $f_1(x), \ldots, f_n(x)$ be periodic functions defined on $\mathbb{R}$, such that $f(x) = f_1(x) + \ldots + f_n(x)$ has a limit when $x \to +\infty$. Then $f(x)$ is a constant function.

Proof. We prove the lemma by induction on n. The base case is $n = 1$. Assume that T_1 is a period for the function $f_1(x)$. If the function $f(x) = f_1(x)$ is not a constant function, then there are x_1 and x_2, such that $f(x_1) \neq f(x_2)$. Hence, it follows that $f(x_i + nT_1) = f(x_i) \to f(x_i)$, $i = 1, 2$, for $n \in \mathbb{N}$, $n \to +\infty$. Therefore, if $x \to +\infty$, the function $f(x)$ has no limit, which leads to a contradiction.

Assume now that the statement holds for $n = k \geq 1$, $k \in \mathbb{N}$, and we show that it also holds for $n = k + 1$.

Let $f(x) = f_1(x) + \ldots + f_{k+1}(x)$ and let T_i be a period for $f_i(x)$, $i = 1, \ldots, k+1$. Consider the following function

$$f(x + T_{k+1}) - f(x) = (f_1(x + T_{k+1}) - f_1(x)) + \ldots + (f_k(x + T_{k+1}) - f_k(x)).$$

Note that T_i is a period for the function $f_i(x+T_{k+1})-f_i(x)$, $i=1,\ldots,k$. On the other hand, $f(x+T_{k+1})-f(x)\to 0$, when $x\to+\infty$. From the induction hypothesis, $f(x+T_{k+1})-f(x)$ will be a constant function. Obviously, if the limit of that constant function is equal to 0, then the function is equal to 0. Therefore, we have obtained that the function $f(x)$ is a periodic function. Thus, according to the statement for the case $n=1$, we have that $f(x)$ is a constant function. This completes the proof of the lemma.

Back to the original question, we assume by contradiction that for any $x>0$ we have

$$\sin(a_1x)+\sin(a_2x)+..+\sin(a_{2001}x)\geq 0.$$

Note that this implies that the function

$$f(x)=-\frac{1}{a_1}\cos(a_1x)-\frac{1}{a_2}\cos(a_2x)-\ldots-\frac{1}{a_{2001}}\cos(a_{2001}x)$$

is non-decreasing. Indeed, we have that

$$f'(x)=\sin(a_1x)+\sin(a_2x)+\ldots+\sin(a_{2001}x)\geq 0.$$

On the other hand, the function $f(x)$ is bounded. As $f(x)$ is bounded and non-decreasing, it has a limit when $x\to+\infty$. So $f(x)$ satisfies the hypotheses of the above lemma. Thus, $f(x)$ is a constant function.
We deduce that

$$f''(x)=a_1\cos(a_1x)+a_2\cos(a_2x)+\ldots+a_{2001}\cos(a_{2001}x)\equiv 0.$$

Therefore,

$$a_1+a_2+\ldots+a_{2001}=f''(0)=0.$$

In a similar way, we obtain that

$$a_1^n+a_2^n+\ldots+a_{2001}^n=(-1)^{\frac{n-1}{2}}\cdot f^{n+1}(0)=0,$$

where n is any odd positive integer. Without loss of generality, one can assume that

$$|a_1|\leq|a_2|\leq\ldots\leq|a_{2001}|.$$

Hence, we deduce that, when $n \to +\infty$, (and n is odd)

$$0 = \left(\frac{a_1}{a_{2001}}\right)^n + \left(\frac{a_2}{a_{2001}}\right)^n + \ldots + \left(\frac{a_{2000}}{a_{2001}}\right)^n + 1 \to 0.$$

Now, if $x \in (0,1)$, then $x^n \to 0$ when $n \to +\infty$. So if all of $a_1, \ldots, a_{2000}$ were strictly less than a_{2001}, then we would have

$$\left(\frac{a_1}{a_{2001}}\right)^n + \left(\frac{a_2}{a_{2001}}\right)^n + \ldots + \left(\frac{a_{2000}}{a_{2001}}\right)^n \to 0,$$

contradicting the previous relation. As n is odd, this shows that one of the numbers $a_1, a_2, \ldots, a_{2000}$ will be equal to $-a_{2001}$. Deleting a_{2001} and the number having the opposite sign, we obtain in the same manner that among the rest of the numbers there are also numbers with opposite signs. Continuing these steps, we deduce that one of the numbers is equal to 0 (since we started with 2001 numbers, and 2001 is odd). This gives a contradiction and establishes the result that we wanted.

Problem 9.2.7. Let $f : \mathbb{R} \to \mathbb{R}$ be a function defined by

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{if } x \neq 0; \\ 0 & \text{if } x = 0. \end{cases}$$

Prove that f is infinitely differentiable at 0 and $f^{(n)}(0) = 0$, where $f^{(n)}$ represents the n-th derivative of f.

Solution. First notice that f is infinitely differentiable on $(-\infty, 0)$ and $(0, \infty)$, respectively. For $x \neq 0$ we have

$$f'(x) = \frac{2}{x^3} e^{-\frac{1}{x^2}}, \quad f''(x) = \frac{6x^2 + 4}{x^6} e^{-\frac{1}{x^2}}.$$

We first prove by induction that for $x \neq 0$,

$$f^{(n)}(x) = \frac{P_n(x) e^{-\frac{1}{x^2}}}{x^{3n}},$$

where $P_n(x)$ is a polynomial with real coefficients of degree $2n-2$. The base cases for $n=1$ and $n=2$ were derived above.
Assume now that the result holds for some $n \geq 1$. Then

$$\begin{aligned} f^{(n+1)}(x) &= P_n'(x)x^{-3n}e^{-\frac{1}{x^2}} - 3n \cdot P_n(x)x^{-3n-1}e^{-\frac{1}{x^2}} + P_n(x)x^{-3n} \cdot \frac{2}{x^3} \cdot e^{-\frac{1}{x^2}} \\ &= \frac{P_{n+1}(x)e^{-\frac{1}{x^2}}}{x^{3n+3}}, \end{aligned}$$

where

$$P_{n+1}(x) = P_n'(x) \cdot x^3 - 3n \cdot P_n(x) \cdot x^2 + 2P_n(x).$$

Since $P_n(x)$ has degree $2n-2$ from the induction hypothesis, P_{n+1} has degree $2n$, establishing the induction step.
We now prove that for any positive integer m we have

$$\lim_{x\to 0} \frac{e^{-\frac{1}{x^2}}}{x^m} = 0.$$

When $m = 2k$, by substituting $y = x^{-2}$ we have

$$\begin{aligned} \lim_{x\to 0} \frac{e^{-\frac{1}{x^2}}}{x^m} = \lim_{y\to\infty} \frac{y^k}{e^y} &= \lim_{y\to\infty} \frac{ky^{k-1}}{e^y} \quad \text{using L'Hôpital's rule} \\ &= \lim_{y\to\infty} \frac{k(k-1)y^{k-2}}{e^y} = \dots \\ &= \lim_{y\to\infty} \frac{k!}{e^y} = 0. \end{aligned}$$

For $m = 2k-1$ we have

$$\lim_{x\to 0} \frac{e^{-\frac{1}{x^2}}}{x^m} = \lim_{x\to 0} x \cdot \lim_{x\to 0} \frac{e^{-\frac{1}{x^2}}}{x^{2k}} = 0.$$

Now we can compute that $f^{(n)}(0) = 0$ by induction on n. For the base case we have

$$f'(0) = \lim_{h\to 0} \frac{f(h)-f(0)}{h} = \lim_{h\to 0} \frac{e^{-\frac{1}{h^2}}}{h} = 0,$$

by the $m = 1$ case of the previous limit. For the induction step, we compute that

$$f^{(n+1)}(0) = \lim_{h\to 0} \frac{f^{(n)}(h) - f^{(n)}(0)}{h} = \lim_{h\to 0} \frac{f^{(n)}(h)}{h} = \lim_{h\to 0} \frac{P_n(h)e^{-\frac{1}{h^2}}}{h^{3n+1}} = 0,$$

where the limit is a linear combination of the limits computed above.

Problem 9.2.8. Let $n > 1$ be a positive integer, $0 < a_1 < \ldots < a_n$ and let $c_1,\ldots,c_n$ be non-zero real numbers. Prove that the number of roots of the equation

$$c_1 \cdot a_1^x + \ldots + c_n \cdot a_n^x = 0$$

is not larger than the number of negative elements of the sequence c_1c_2, c_2c_3, $\ldots$, $c_{n-1}c_n$.

Solution. We prove the statement by induction on n.

We begin by treating the base case $n = 2$:

If $c_1c_2 < 0$, then the equation $c_1 \cdot a_1^x + c_2 \cdot a_2^x = 0$ has only one root, namely $x = \log_{\frac{a_2}{a_1}} \left(-\frac{c_1}{c_2}\right)$.

If $c_1c_2 > 0$, then the equation $c_1 \cdot a_1^x + c_2 \cdot a_2^x = 0$ does not have any roots.

We now show that if the statement holds for $n = k$, $k \in \mathbb{Z}$, $k > 1$, then it also holds for $n = k + 1$. We do this via a proof by contradiction:

Assume that the number of roots of $c_1 \cdot a_1^x + \ldots + c_{k+1} \cdot a_{k+1}^x = 0$ is larger than the number of negative elements of the sequence c_1c_2, c_2c_3, $\ldots$, c_kc_{k+1}.

Consider the function

$$f(x) = c_1 \cdot \left(\frac{a_1}{a_{k+1}}\right)^x + \ldots + c_k \cdot \left(\frac{a_k}{a_{k+1}}\right)^x + c_{k+1}.$$

Note that the number of zeroes of the function

$$f'(x) = c_1 \ln \frac{a_1}{a_{k+1}} \cdot \left(\frac{a_1}{a_{k+1}}\right)^x + \ldots + c_k \ln \frac{a_k}{a_{k+1}} \cdot \left(\frac{a_k}{a_{k+1}}\right)^x$$

is bigger than the number of negative elements of the sequence c_1c_2, c_2c_3, $\ldots$, $c_{k-1}c_k$ (Here we used the fact that between any two zeroes of the function there is at least one zero of the derivative, so if f has m zeroes, then f' has at least $m-1$ zeroes). On the other hand, from the induction hypothesis, the number of roots of $f'(x) = 0$ is not larger than the number of negative elements of the sequence

$$\left(c_1 \ln \frac{a_1}{a_{k+1}}\right) \cdot \left(c_2 \ln \frac{a_2}{a_{k+1}}\right), \ldots, \left(c_{k-1} \ln \frac{a_{k-1}}{a_{k+1}}\right) \cdot \left(c_k \ln \frac{a_k}{a_{k+1}}\right),$$

which is equal to the number of negative elements of the sequence $c_1c_2, c_2c_3, \ldots, c_{k-1}c_k$, as $\ln \dfrac{a_i}{a_{k+1}} < 0$, $i = 1, \ldots, k$. This leads to a contradiction and completes the proof of our induction step, establishing the result of the question.

Problem 9.2.9. Prove that for any $|x| < 1$ and any positive integer k one has

$$\sum_{n=0}^{\infty} x^n \binom{n+1}{k} = \frac{x^{k-1}}{(1-x)^{k+1}}.$$

Solution. We will prove by induction on k that for $k \geq 0$

$$\sum_{n=k-1}^{\infty} \binom{n+1}{k} x^{n+1-k} = \frac{1}{(1-x)^{k+1}},$$

which implies the desired identity by a simple rearrangement.

The base case $k = 0$ is just the infinite geometric series

$$\sum_{r=0}^{\infty} x^r = \frac{1}{1-x},$$

where we have substituted $r = n+1$.

For the inductive step, we differentiate the formula for k to get

$$\sum_{n=k-1}^{\infty} (n+1-k) \binom{n+1}{k} x^{n+1-k-1} = \frac{k+1}{(1-x)^{k+2}}.$$

Since the $n = k-1$ term in the sum vanishes we may drop it. Hence we get

$$\sum_{n=k}^{\infty} \binom{n+1}{k+1} x^{n+1-k-1} = \sum_{n=k}^{\infty} \frac{n+1-k}{k+1} \binom{n+1}{k} x^{n+1-k-1} = \frac{1}{(1-x)^{k+2}}.$$

This completes the inductive step and the proof.

9.3 Induction in Algebra

Problem 9.3.1. (**Iran 1985**) Let α be an angle such that $\cos(\alpha) = \frac{p}{q}$, where p and q are two integers. Prove that the number $q^n \cos(n\alpha)$ is an integer, for any $n \in \mathbb{N}^*$.

Solution. From the addition formula for the cosine we have

$$\cos((n \pm 1)\alpha) = \cos(\alpha)\cos(n\alpha) \mp \sin(\alpha)\sin(n\alpha).$$

Hence adding these two formulas

$$\cos((n+1)\alpha) = 2\cos(\alpha)\cos(n\alpha) - \cos((n-1)\alpha).$$

From this formula, the problem follows by an easy induction. For the base cases $n = 1$ and $n = 2$ we have $q\cos(\alpha) = p \in \mathbb{Z}$ and

$$q^2\cos(2\alpha) = q^2(2\cos^2(\alpha) - 1) = 2p^2 - q^2 \in \mathbb{Z}.$$

For the inductive step, we have

$$q^{n+1}\cos((n+1)\alpha) = 2 \cdot q\cos(\alpha) \cdot q^n\cos(n\alpha) - q^2 \cdot q^{n-1}\cos((n-1)\alpha) \in \mathbb{Z}.$$

Problem 9.3.2. Prove that there is a monic polynomial $P \in \mathbb{Z}[X]$ of degree n, such that $P(2\cos x) = 2\cos nx$ and $P\left(x + \dfrac{1}{x}\right) = x^n + \dfrac{1}{x^n}$.

Solution. We use the identity

$$2\cos nx \cdot 2\cos x = 2\cos(n+1)x + 2\cos(n-1)x$$

derived from $\cos(x+y) + \cos(x-y) = 2\cos x\cos y$ and also the identity

$$\left(x^n + \frac{1}{x^n}\right)\left(x + \frac{1}{x}\right) = \left(x^{n+1} + \frac{1}{x^{n+1}}\right) + \left(x^{n-1} + \frac{1}{x^{n-1}}\right).$$

Note that the identities are analogous, since setting $x = e^{it}$ gives $x + \frac{1}{x} = 2\cos t$.

Therefore, we can define

$$p_0(x) = 1, p_1(x) = x \quad \text{and} \quad p_{n+1}(x) = xp_n(x) - p_{n-1}(x).$$

An immediate induction on n shows that the relations $p_n(2\cos x) = 2\cos nx$ and $p_n(x + \frac{1}{x}) = x^n + \frac{1}{x^n}$ hold. Another induction on n shows that p_n are monic polynomials of degree n.

Problem 9.3.3. Prove that there is no polynomial $P \in \mathbb{R}[X]$ of degree $n \geq 1$ such that $P(x) \in \mathbb{Q}$ for all $x \in \mathbb{R} \backslash \mathbb{Q}$.

Solution. We will prove the statement of problem using induction on the degree $n \geq 1$.

The base case is $n = 1$. Suppose that $P(x) = ax + b$, where $a, b \in \mathbb{R}$ and $a \neq 0$, such that $P(x) \in \mathbb{Q}$, for all $x \in \mathbb{R} \setminus \mathbb{Q}$. For $x \in \mathbb{R} \setminus \mathbb{Q}$ we also have that $x + 1, \frac{x}{2} \in \mathbb{R} \setminus \mathbb{Q}$, and so $P(x)$, $P(x+1), P\left(\frac{x}{2}\right) \in \mathbb{Q}$. Then

$$a = P(x+1) - P(x) \in \mathbb{Q} \quad \text{and} \quad b = 2P\left(\frac{x}{2}\right) - P(x) \in \mathbb{Q}.$$

Hence, $x = \dfrac{P(x) - b}{a} \in \mathbb{Q}$ and that contradicts that $x \in \mathbb{R} \setminus \mathbb{Q}$.

Let $n \geq 2$. Suppose that the statement of problem holds for polynomials of degree $m \in \{1, 2, ..., n-1\}$ and we prove that there is no polynomial $P \in \mathbb{R}[X]$ of degree n such that $P(x) \in \mathbb{Q}$ for all $x \in \mathbb{R} \setminus \mathbb{Q}$. Assume the contrary, so $P(x) = a_0 x^n + a_1 x^{n-1} + \cdot + a_{n-1} x + a_n$, where $a_0 \neq 0$, and $P(x) \in \mathbb{Q}$ for all $x \in \mathbb{R} \setminus \mathbb{Q}$.

Since $x+1 \in \mathbb{R} \backslash \mathbb{Q}$, then $P(x+1) \in \mathbb{Q}$. Setting $P_1(x) := P(x+1) - P(x)$ we have $\deg P_1(x) = n - 1 < n$ and $P_1(x) \in \mathbb{Q}$, for any $x \in \mathbb{R} \setminus \mathbb{Q}$. Thus we get a contradiction to the induction hypothesis. This completes our proof.

Problem 9.3.4. Let $F(X)$ and $G(X)$ be two polynomials with real coefficients such that the points

$$(F(1), G(1)), (F(2), G(2)), \ldots, (F(2011), G(2011))$$

are vertices of a regular 2011-gon.
Prove that $\deg(F) \geq 2010$ or $\deg(G) \geq 2010$.

Solution. Without loss of generality, we can assume that the center of the 2011-gon is located at the origin, $(0, 0)$. Therefore, we have

$$F(k) = \cos\left(\alpha + \frac{2k\pi}{2011}\right), \quad G(k) = \sin\left(\alpha + \frac{2k\pi}{2011}\right) \tag{1}$$

for all $1 \leq k \leq 2011$ and some fixed $\alpha \in \mathbb{R}$. We now prove by induction on $n \geq 1$ that if $F(X)$ and $G(X)$ satisfy (1) for $k = 1, 2, \ldots, n$, then at least one of them has degree at least $n-1$ (so taking $n = 2011$ will prove our statement). The base case $n = 1$ is clear.

Assume now that the result holds for all integers less than some $n \geq 2$ (with $n \leq 2011$) and $F(X)$ and $G(X)$ satisfy (1) for all $k = 1, 2, \ldots, n$. Define

$$F^*(X) = \frac{F(X+1) - F(X)}{-2\sin\frac{\pi}{2011}}$$

and

$$G^*(X) = \frac{G(X+1) - G(X)}{2\sin\frac{\pi}{2011}}.$$

Now for all $k = 1, 2, \ldots, n-1$ we have

$$F^*(k) = \sin\left(\alpha + \frac{(2k+1)\pi}{2011}\right), \quad G^*(k) = \cos\left(\alpha + \frac{(2k+1)\pi}{2011}\right).$$

Notice now that $\deg(F^*) \leq \deg(F) - 1$ and $\deg(G^*) \leq \deg(G) - 1$ and F^*, G^* satisfy the hypotheses of our question. By the induction hypothesis, one of them has degree at least $n-2$, which shows that at least one of F or G has degree at least $n-1$.

Problem 9.3.5. Is $\cos 1^\circ$ rational?

Solution. To answer this question, we combine two strategies: proof by contradiction and induction. Namely, we will prove that if $\cos(1^\circ)$ is rational, then so is $\cos(n^\circ)$ for $n \in \mathbb{N}$. As $\cos(30^\circ) = \frac{\sqrt{3}}{2}$ is irrational, this gives a contradiction.

The base cases are $n = 1$, which is true from our assumption and $n = 2$, for which we have

$$\cos 2^\circ = 2\cos^2 1^\circ - 1 \in \mathbb{Q}.$$

Assume now that we have proved the result for some $n \geq 2$. We use the following identity:

$$\cos((k+1)^\circ) + \cos((k-1)^\circ) = 2\cos(k^\circ)\cos(1^\circ).$$

This shows that if $\cos(1°), \cos((n-1)°), \cos(k°)$ are rational, then $\cos((k+1)°)$ is also rational.
By induction, $\cos(30°)$ would be rational, which we know is false.

Problem 9.3.6. (**Poland 2000**) Let P be a polynomial of odd degree satisfying the identity

$$P(x^2 - 1) = P(x)^2 - 1.$$

Prove that $P(x) = x$ for all real x.

Solution. Setting $x = y$ and $x = -y$ in the given equation, we find that $P(y)^2 = P(-y)^2$ for all y. Thus, one of the polynomials $P(x) - P(-x)$ or $P(x) + P(-x)$ vanishes for infinitely many x, and hence for all x. Because P has odd degree, the latter must be the case, which shows that P is an odd polynomial. In particular, $P(0) = 0$, which in turn implies that

$$P(-1) = P(0^2 - 1) = P(0)^2 - 1 = -1, \quad \text{and hence} \quad P(1) = 1.$$

Set $a_0 = 1$ and let $a_n = \sqrt{a_{n-1} + 1}$ for all $n \geq 1$. Note that $a_n > 1$ when $n \geq 1$. We prove by induction on n that $P(a_n) = a_n$ for all $n \geq 0$. The base case $n = 0$ was done above.
Assuming the result for some $n \geq 0$, we have

$$P(a_{n+1})^2 = P(a_{n+1}^2 - 1) + 1 = P(a_n) + 1 = a_n + 1,$$

which implies that $P(a_{n+1}) = \pm a_{n+1}$. If $P(a_{n+1}) = -a_{n+1}$, then

$$P(a_{n+2})^2 = P(a_{n+1} + 1) = 1 - a_{n+1} < 0,$$

giving a contradiction. Therefore, $P(a_{n+1}) = a_{n+1}$ and the induction is complete.

To finish the question, notice that for $1 < x < \frac{1+\sqrt{5}}{2}$, we have $x < \sqrt{x+1} < \frac{1+\sqrt{5}}{2}$. Hence an easy induction shows that $a_0 = 1 < a_1 < a_2 < \cdots$. This proves that all the a_n's are distinct, so $P(x) = x$ for infinitely many values of x, hence $P(x) = x$ for all x.

Problem 9.3.7. (**Bulgaria**) Prove that there exists a quadratic polynomial $f(X)$ such that $f(f(X))$ has 4 non-positive real roots and $f^n(X)$ has 2^n real roots, where f^n denotes the composite of f with itself n times.

Solution. Replacing $f(X)$ with $-f(X)$ if necessary, we can assume that $f(X)$ has positive leading coefficient (notice that this does not affect the location of the roots). For $f(f(X))$ to have 4 real roots, it is clear that $f(X)$ must itself have two real roots, say $x_1 < x_2$. If $x_2 > 0$, then one can show there exists a positive real s such that $f(s) = x_2$ (as the sign of f between x_1 and x_2 is negative). But then we would get $f(f(s)) = f(x_2) = 0$, so $f(f(X))$ would have a positive real root, which is a contradiction. Hence we must have $x_1 < x_2 \leq 0$.

Now notice that the roots of $f(f(X))$ are the solutions of the equations $f(x) = x_1$ and $f(x) = x_2$. Both equations must have 2 real roots. Let m be the minimal value which f attains on $\mathbb{R}$ (which exists, as f has positive leading term). Then we must have $m < x_1 < x_2 \leq 0$. Notice that this implies that the roots of $f(f(X))$ lie in the interval (x_1, x_2). Examples of an f satisfying the criteria we mentioned so far are easy to construct, e.g. take $f(X) = (X+1)(X+9)$.

We now prove that a quadratic polynomial constructed as above satisfies all the conditions of the problem, namely we show that $f^n(X)$ has all its roots in the interval (x_1, x_2). We do this by induction on n. The case $n = 2$ was treated above. For the induction step, assume that the result holds for some $n \geq 1$ and let $y_1, \ldots, y_{2^n}$ be the roots of $f^n(X)$. Now notice that the roots of $f^{n+1}(X)$ are the solutions to $f(x) = y_k$, where $1 \leq k \leq 2^n$. Since $m < x_1$, all equations of the form $f(x) = y_k$ will have two real roots which lie in (x_1, x_2). Thus $f^{n+1}(X)$ has 2^{n+1} real roots in (x_1, x_2). This completes our proof.

Problem 9.3.8. Prove that if a rational function which is not a polynomial takes rational values at all positive integers, then it is the quotient of two coprime polynomials, both having integer coefficients.

Solution. Let the rational function be $R(X) = \frac{P(X)}{Q(X)}$, where P and Q are relatively prime polynomials. We will prove a slightly stronger statement, assuming only that R outputs rational values for all sufficiently large positive integers. Let $r = \deg(P) + \deg(Q)$. We prove the result by induction on r. The base case $r = 0$ is immediate.

For the induction step, consider if necessary $\frac{1}{R(X)}$ instead of $R(X)$, so that we ensure that $\deg(P) \geq \deg(Q)$. Also, let a be a positive integer for which

$Q(a) \neq 0$. Then $\frac{P(a)}{Q(a)}$ is rational and the rational function

$$\frac{1}{X-a}\left(R(X) - \frac{P(a)}{Q(a)}\right) = \frac{P_1(X)}{Q(X)}$$

outputs rational values for all positive integers $x > a$. Notice also that

$$P_1(X) = \frac{P(X)Q(a) - Q(X)P(a)}{Q(a)(X-a)}$$

is a polynomial whose degree is less than that of $P(X)$. So

$$\deg(P_1) + \deg(Q) < \deg(P) + \deg(Q),$$

and by induction, we are done.

Problem 9.3.9. (**IMO 2007 shortlist**) Let $n > 1$ be an integer. In the space, consider the set

$$S = \{(x, y, z) \mid x, y, z \in \{0, 1, \ldots, n\}, x + y + z > 0\}.$$

Find the smallest number of planes that jointly contain all $(n+1)^3 - 1$ points of S but none of them passes through the origin.

Solution. The answer to the question is $3n$ planes. To find $3n$ planes satisfying the requirements, one can take the planes $x = i$, $y = i$ or $z = i$ $(i = 1, 2, \ldots, n)$ which cover the set S but none of them contains the origin. Another such collection consists of all planes $x+y+z = k$ for $k = 1, 2, \ldots, 3n$.

We show that $3n$ is the smallest possible number.

Lemma. Consider a non-zero polynomial $P(x_1, \ldots, x_k)$ in k variables. Suppose that P vanishes at all points $(x_1, \ldots, x_k)$ such that $x_1, \ldots, x_k \in \{0, 1, \ldots, n\}$ and $x_1 + \ldots + x_k > 0$, while $P(0, 0, \ldots, 0) \neq 0$. Then $\deg P \geq kn$.

Proof. We use induction on k. The base case $k = 0$ is clear, since $P \neq 0$. Denote for clarity $y = x_k$.

Let $R(x_1, \ldots, x_{k-1}, y)$ be the residue of P modulo

$$Q(y) = y(y-1)\ldots(y-n).$$

Polynomial $Q(y)$ vanishes at each $y = 0, 1, \ldots, n$, hence

$$P(x_1, \ldots, x_{k-1}, y) = R(x_1, \ldots, x_{k-1}, y)$$

for all $x_1, \ldots, x_{k-1}, y \in \{0, 1, \ldots, n\}$. Therefore, R also satisfies the condition of the Lemma. Moreover, as a polynomial in y, R has degree at most n. Clearly, $\deg R \le \deg P$, so it suffices to prove that $\deg R \ge nk$.

Now, expand the polynomial y in the powers of y:

$$\begin{aligned} R(x_1, \ldots, x_{k-1}y) = {} & R_n(x_1, \ldots, x_{k-1})y^n + R_{n-1}(x_1, \ldots, x_{k-1})y^{n-1} + \ldots \\ & + R_0(x_1, \ldots, x_{k-1}). \end{aligned}$$

We show that the polynomial $R_n(x_1, \ldots, x_{k-1})$ satisfies the condition of the inductive hypothesis:

Consider the polynomial $T(y) = R(0, \ldots, 0, y)$ of degree at most n. This polynomial has n roots: $y = 1, \ldots, n$. On the other hand $T(y) \not\equiv 0$ since $T(0) \neq 0$. Hence $\deg T = n$, and its leading coefficient is $R_n(0, \ldots, 0) \neq 0$. In particular, in the case $k = 1$ we obtain that the coefficient of R_n is non-zero.

Similarly, take any numbers $a_1, \ldots, a_{k-1} \in \{0, 1, \ldots, n\}$ with $a_1 + \ldots + a_{k-1} > 0$. Substituting $x_i = a_i$ into $R(x_1, \ldots, x_{k-1}, y)$, we get a polynomial in y which vanishes at all points $y = 0, \ldots, n$ and has degree at most n. Therefore, this polynomial is null, hence $R_i(a_1, \ldots, a_{k-1}) = 0$ for all $i = 0, 1, \ldots, n$. In particular, $R_n(a_1, \ldots, a_{k-1}) = 0$.

Thus, the polynomial $R_n(x_1, \ldots, x_{k-1})$ satisfies the condition of the induction hypothesis. So we have that $\deg R_n \ge (k-1)n$ and $\deg P \ge \deg R \ge \deg R_n + n \ge kn$. This establishes the lemma.

Now we can finish the solution of the problem. Suppose that there are N planes covering all the points of S but not covering the origin. Let their equations be $a_i x + b_i y + c_i z + d_i = 0$. Consider the polynomial

$$P(x, y, z) = \prod_{i=1}^{N} (a_i x + b_i y + c_i z + d_i).$$

This polynomial has total degree N and has the property that $P(x_0, y_0, z_0) = 0$ for any $(x_0, y_0, z_0) \in S$, while $P(0, 0, 0) \neq 0$. Hence by the above lemma we get that $N = \deg P \ge 3n$, as desired.

Problem 9.3.10. Prove that for every positive integer n there exists a polynomial $p(x)$ with integer coefficients such that $p(1), p(2), \dots, p(n)$ are distinct powers of 2.

Solution. We prove the result by induction on n. For $n = 1$ we can take for example $p(x) = x$.

Now assume we have found a polynomial p such that $p(1)$, $p(2)$, ..., $p(n)$ are distinct powers of 2. We now look for a polynomial $Q(x)$ of the form

$$Q(x) = 2^a p^j(x) + A(x-1)(x-2)\dots(x-n),$$

where a is such that 2^a does not divide $(n+1)!$. Notice that $Q(1) = p(1)$, $Q(2) = p(2)$, ..., $Q(n) = p(n)$.

So to have Q satisfy the induction step, we want $Q(n+1)$ to be a power of 2 distinct from $p(1)$, $p(2)$, ..., $p(n)$. Let $B = p(n+1)$. We must have that $2^a B^j + An!$ is a power of 2 different from $p(1)$, $p(2)$, ..., $p(n)$. Let 2^s be the largest power dividing B and 2^t be the largest power dividing $n!$, so that $n! = 2^t K$ with K odd. Now $(K, B) = 1$ because if a prime number q divides both K and B, then $q \le n$ and $q \mid (p(n+1) - p(n+1-q))$, hence $q \mid p(n+1-q)$, which is impossible since $p(n+1-q)$ is a power of 2. So there exists j such that $B^j \equiv 1 \pmod{K}$.

We know that there exist infinitely many numbers m such that $2^m \equiv 1 \pmod{K}$. Thus $2^{m+a} - 2^a B^j$ is divisible by K and also by 2^a hence by $n!$. So we may take

$$A = \frac{2^{m+a} - 2^a B^j}{n!}.$$

Then the polynomial

$$Q(x) = 2^a p^j(x) + A(x-1)(x-2)\dots(x-n)$$

satisfies

$$Q(1) = 2^a p(1)^j,\ Q(2) = 2^a p(2)^j, \dots,\ Q(n) = 2^a p(n)^j,\ Q(n+1) = 2^{m+a}.$$

Since there are infinitely many m to satisfy our condition, we can take m such that 2^{m+a} is different from all $2^a p(1)^j, 2^a p(2)^j, \dots, 2^a p(n)^j$. This $Q(x)$ is the required polynomial for our induction step.

Problem 9.3.11. (**Moscow 2013**) Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $f(x) \in \mathbb{Z}$ for all $x \in \mathbb{Z}$. For every prime p, there exists a polynomial $Q_p(X)$ of degree less than 2013, with integer coefficients such that $f(n) - Q_p(n)$ is divisible by p for all positive integers n. Prove that there exists a polynomial $g(x) \in \mathbb{R}[X]$ such that for all positive integers we have $f(n) = g(n)$.

Solution. We replace the 2013 by k and we prove by induction on k that if $\deg(Q_p(X)) \leq k$, then the polynomial g exists. For $k = 0$, $Q_p(X)$ must be a constant c and so $f(n) - Q_p(n) = f(n) - c$. so $p \mid f(n) - c$, which shows that for all $m, n \in \mathbb{N}^*$, $f(m) - f(n)$ is divisible by p. This is true for all primes p, so by fixing m, n and taking p large enough, we conclude that f must be constant integer. This establishes the base case.

Before we perform the induction step, we need to establish some preliminary results. Firstly, notice that if $h(X)$ is a polynomial of degree $d \geq 1$, then $\Delta h = h(X+1) - h(X)$ is a polynomial of degree $d - 1$. We now prove the following:

Lemma. If for all positive integer n, we have $\Delta h(n) = P(n)$, for some polynomial P of degree less than or equal to $d-1$, then for all positive integers x we have $h(x) = h_1(x)$, for some polynomial h_1 of degree less than or equal to d.

Proof. We prove the lemma by induction on d. For $d = 1$, we know that $h(x) = h(0) + cx$, for all positive integers x. Now assume that

$$\Delta h(x) = x^d + \dots,$$

and define

$$h_0(x) = h(x) - \frac{a}{d+1}x(x-1)\dots(x-d).$$

Then

$$\Delta h_0 = \Delta h(x) - ax(x-1)\dots(x-d+2).$$

Then $\deg(\Delta h_0) \leq d-1$, so by the induction hypothesis, for all positive integers x we have $h_0(x) = R(x)$, for some polynomial R of degree less than or equal to d. Now from $h_0(x) = h(x) - \frac{a}{d+1}x(x-1)\dots(x-d)$, we have that $\deg(h) \leq d+1$, whence for all positive integers x, $h(x) = R_1(x)$, for some polynomial R_1 of degree at most $d+1$. This completes the proof of the lemma.

Going back to our original problem, for the induction step $k-1 \to k$, notice that Δf satisfies the hypotheses of our question (as $\Delta f(x) - \Delta Q_p(x)$ is divisible by p whenever x is a positive integer). Moreover, $\Delta Q_p(x)$ has degree at most $k-1$, so by the induction hypothesis, $\Delta f(x)$ takes polynomial values at all positive integers. Applying the above lemma, we obtain the conclusion.

Notations and Abbreviations

Notations

We assume familiarity with the standard mathematical terminology. Most of the notations used throughout the book were introduced in the Theory section of the relevant chapters. We list below some of the essential notations.

- $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, stand for the set of non-negative integers, integers, rationals, reals and complex numbers, respectively. For $A \in \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, A^* denotes the set $A \setminus \{0\}$. Similarly, $A_{<0}$ or A_- denote the negative elements of A, $A_{>0}$ or A_+ stand for the positive elements of A, $A_{\geq 0}$ represents the set of non-negative elements of A, and $A_{\leq 0}$ is the set of non-positive elements of A.

- For S a finite set, $|S|$ denotes the number of elements in S. If S is infinite, $|S|$ may also stand for the cardinality of S (as defined in Chapter 1).

- For a real number x, both $\lfloor x \rfloor$ and $[x]$ denote the largest integer which is less than or equal to x, while $\lceil x \rceil$ represents the smallest integer greater than or equal to x. $\{x\}$ is the fractional part of x, defined as $\{x\} = x - \lfloor x \rfloor$.

- e denotes the Euler's constant $e = \lim\limits_{n \to \infty} \left(n + \frac{1}{n}\right)^n \approx 2.71828$.

- For $F = \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ or $\mathbb{Z}/n\mathbb{Z}$, we denote by $F[X]$ the set of polynomials in the variable X with coefficients in F, and $F(X)$ the set of rational functions in X with coefficients in F. A rational function is the ratio of two polynomials.

- For A, B points in the plane or space, AB denotes the line through A and B, the segment through A and B, or the distance between A and B, depending on the context.

- For S a plane figure, $[S]$ denotes the area of S. If S is a solid, $[S]$ denotes the volume of S.

Abbreviations

We indicated the sources of the problems where possible. The meaning of the abbreviations is explained below.

- AMM - American Mathematical Monthly
- AoPS - Art of Problem Solving
- APMO - Asia Pacific Mathematics Olympiad
- BMO - Balkan Mathematical Olympiad
- ELMO - An annual Mathematical Olympiad which takes place at MOP
- GMA - Graduate Mathematics Association
- GMB - Gazeta Matematică, Seria B
- IMC - International Mathematics Competition
- IMO - International Mathematical Olympiad
- INMO - Indian National Mathematical Olympiad
- JBMO - Junior Balkan Mathematical Olympiad
- OMM - Mexican Mathematical Olympiad
- MO - Mathematical Olympiad
- MOSP - Mathematical Olympiad Summer Program
- MR - Mathematical Reflections
- RMM - Romanian Master of Mathematics
- TOT - Tournament of Towns
- TST - Team Selection Test
- USAMO - United States of America Mathematical Olympiad
- USSR - Union of Soviet Socialist Republics

Bibliography

[1] N. Agahanov, O. Podlipsky, *Olimpiade Matematice Ruseşti-Moscova 1993-2002*, Editura GIL, Zalău, 2004.

[2] T. Andreescu, Z. Feng, G. Lee Jr., *Mathematical Olympiads 2000-2001: Problems and Solutions from Around the World*, The Mathematical Association of America, 2003.

[3] T. Andreescu, I. Boreico, O. Mushkarov, N. Nikolov, *Topics in Functional Equations*, XYZ Press, 2012.

[4] T. Andreescu, D. Andrica *360 Problems for Mathematical Contests*, XYZ Press, 2015.

[5] A. Bjorndahl, *Puzzles and Paradoxes in Mathematical Induction.*

[6] D. Djukic, V. Jankovic, I. Matic, N. Petrovic, *The IMO Compendium*, Springer, 2006.

[7] *Gazeta Matematică, Seria B*, electronic edition.

[8] L.I. Golovina, I.M. Yaglom, *Induction in Geometry*, Little Mathematics Library, 1979.

[9] A. Engel, *Probleme de matematică-strategii de rezolvare*, Editura GIL, 2006.

[10] D. Leites, *The 60-odd years of the Moscow Mathematical Olympiads*, preprint version, 1997.

[11] L. Panaitopol, M.E. Panaitopol, M. Lascu, *Inducţia Matematică*, Editura GIL, Zalău, 2001.

[12] V. Prasolov, *Problems in plane and solid Geometry v.1 Plane Geometry*, ebook.

[13] D.O. Shklarsky, N.N. Chentzov, I.M. Yaglom, *The USSR Olympiad Problem Book*, Dover Publications, 1993.

[14] A.M. Slinko, *USSR Mathematical Olympiads 1989-1992*, Australian Mathematics Trust, 1997.

[15] P. Soberon, *Problem-Solving Methods in Combinatorics. An approach to olympiad problems*, Birkäuser, 2013.

[16] www.artofproblemsolving.com

[17] www.awesomemath.org/mathematical-reflections

[18] www.kvant.mccme.ru

[19] www.maa.org/press/periodicals/american-mathematical-monthly

Other Books from XYZ Press

1. Andreescu, T., Elliott, S., *114 Exponent and Logarithm Problems from the AwesomeMath Summer Program*, 2017.

2. Andreescu, A., Andreescu, T., Mushkarov, O., *113 Geometric Inequalities from the AwesomeMath Summer Program*, 2017.

3. Bosch, R., *Cuban Mathematical Olympiads (2001-2016)*, 2017.

4. Matei, V., Reiland, E., *112 Combinatorial Problems from the AwesomeMath Summer Program*, 2016.

5. Andreescu, T., Mortici, C., Tetiva, M., *Pristine Landscapes in Elementary Mathematics*, 2016.

6. Andreescu, A., Vale, V., *111 Problems in Algebra and Number Theory*, 2016.

7. Andreescu, T., *Mathematical Reflections - two special years*, 2016.

8. Andreescu, T., Pohoata, C., Korsky, S., *Lemmas in Olympiad Geometry*, 2016.

9. Mihalescu, C., *The Geometry of Remarkable Elements. Points, lines, and circles*, 2016.

10. Andreescu, T., Pohoata, C., *110 Geometry Problems for the International Mathematical Olympiad*, 2015.

11. Andreescu, T., Boreico, I., Mushkarov, O., Nikolov, N., *Topics in Functional Equations*, 2nd edition, 2015.

12. Andreescu, T., Ganesh, A., *109 Inequalities from the AwesomeMath Summer Program*, 2015.

13. Andreescu, T., Pohoata, C., *Mathematical Reflections - two great years*, 2014.

14. Andreescu, T., Ganesh, A., *108 Algebra Problems from the AwesomeMath Year-Round Program*, 2014.

15. Andreescu, T., Kisačanin, B., *Math Leads for Mathletes – a rich resource for young math enthusiasts, parents, teachers, and mentors*, Book 1, 2014.

16. Becheanu, M., Enescu, B., *Balkan Mathematical Olympiads – the first 30 years*, 2014.

17. Andreescu, T., *Mathematical Reflections – two more years,* 2013.

18. Andreescu, T., Rolinek, M., Tkadlec, J., *107 Geometry Problems from the AwesomeMath Year-Round Program*, 2013.

19. Andreescu, T., Rolinek, M., Tkadlec, J., *106 Geometry Problems from the AwesomeMath Summer Program*, 2013.

20. Andreescu, T., *105 Algebra Problems from the AwesomeMath Summer Program*, 2013.

21. Andreescu, T., Kane, J., *Purple Comet Math Meet! - the first ten years*, 2013.

22. Andreescu, T., *Mathematical Reflections - the next two years*, 2012.

23. Andreescu, T., Dospinescu, G., *Straight from the Book*, 2012.

24. Andreescu, T., *Mathematical Reflections – the first two years*, 2011.

25. Andreescu, T., Dospinescu, G., *Problems from the Book*, 2nd edition, 2010.

Printed by "Combinatul Poligrafic". Com. nr. 70140