Introduction to Combinatorics

University of Toronto Scarborough

Lecture Notes

Stefanos Aretakis

January 29, 2018

Contents

1 Lecture 1

Introduction

Combinatotics is about...

... counting *without really counting all possible cases one by one.*

More broadly:

Combinatorics is about...

... derivining properties of structures satisfying given conditions *without analyzing each and every possible case separately.*

Analyzing, deriving and counting common properties of structures satisfying given conditions can in principle be quite challenging and require a non trivial amount of focus and concentration. This is the challenging part of our course.

The fun part of our course is that the structures we will be considering are very *elementary* (no involved defitions, symbols etc.). Hence, for most of the time we will not be obstracted by general and abstract nonsense.

In other words, we should view our course as a fun and challenging way to learn how to learn. (Question: what is the Greek word for "learning how to learn": Answer: Mathematics!).

Outline of our course

- Chapter 1: Basic Combinatorics Principles
  - The Pigeonhole Principle
  - The Principle of Extremals
  - The Principle of Invariants
  - Applications in Game Theory
- Chapter 2: Counting
  - The Bijection Principle
  - The Additive and Multiplicative Principles
  - Permutations
  - Combinations
- Chapter 4: Generating Functions

The Prisonners' Problem

Let's consider the so-called "Prisonners' problem" as a way to see a few Combinatorial principles in action:

We consider an island full of male prisonners such that the following conditions hold:

1. there are 100 prisonners,

2. all have green eyes,

3. they can all see that all other prisonners have green eyes,
4. they, however, do not know that they themselves have green eyes,
5. no communication is allowed between the prisonners,
6. right of the prisonners: any prisonner can go at any midnight to the guards and claim that they have green eyes. If the prisonner has indeed green eyes, then the prisonner is freed otherwise the prisonner is killed.

The prisonners are aware of their right but unfortunately they do not know if they have green eyes and also there is no way of communicating with any other prisonner. Hence, as is, there is no way the prisonners can find out that they have green eyes and hence escape by making use of their only right (condition 6 above).

However, the situation changes if we add one more condition. Specifically, the prisonners are all gathered together and are told the following statement:

- (additional condition) at least one of you have has green eyes...

Even though at first glance the above statement seems useless, since everyone can see that all the remaining 99 prisonners have green eyes[1], the prisonners manage to realize that they themselves have green eyes and hence make use of their right and escape.

How is this possible? How did the additional condition above shake the dynamics of our problem?

Solution to the Prisonners' Problem

For simplicity, let's assume that there are less than 100 prisonners to make the problem more manageable (REDUCTION PRINCIPLE). Let's assume that we have only 2 prisonners: P_1, P_2.

Let's consider the point of view of one of them, for example P_1:

- P_1's reasoning on the first day: If I have green eyes, then I can make use of my right and escape. hence, I would like somehow to prove to myself that I do have green eyes. How can I do this? What if I assume that I do not have green eyes and reach a conclusion that is obviously wrong (CONTRADICTION ASSUMPTION)? Then my assumption that I do not have green eyes must have been wrong, and hence I will be able to safely conclude that I have green eyes and make use of my right.

  If I do not have green eyes then P_2 must be able to see that I do not have green eyes. However, since we were told that at least one of us has green eyes (additional condition) then P_2 must be able to conclude that it is him who has green eyes and hence make use of his right the first midnight.

However,P_2 does not make use of his right to escape since he cannot conclude if he has green eyes, since he see that P_1 has green eyes and therefore cannot conclude, using the additional condition, that he is the one who has green eyes. Therefore, at the beginning of the second day, P_1 makes the following realization:

[1] and everyone can see that everyone else sees at least someone–in fact at least 98 prisonners– with green eyes

- P_1's reasoning on the second day: Since P_2 is still here it must mean that he is not certain if he has green eyes or not which means that my assumption that I do not have green eyes (which would mean that P_2 would realize that he has green eyes) is wrong. Therefore, I must have green eyes and therefore P_2 must be watching me.

Therefore, P_1 realizes on the second day that he has green eyes and makes use of his right the second midnight and escapes.

Exactly, the same reasoning applies to P_2 (symmetrically) and so P_2 also escapes the second night. Therefore, they both realize that they have green eyes on the second day and make use of their right the same night.

This concludes the solution to the prisonners' problem in the case of $n = 2$ prisonners P_1, P_2.

Let's consider next the case of $n = 3$ prisonners.

Let's consider the point of view of one of them, for example P_1:

- P_1's reasoning on the first day: If I do not have green eyes then P_2, P_3 are not looking at me and and are only looking at each other. Hence, the problem should be reduced to the case of two prisonners (INDUCTION PRINCIPLE) and thefore P_2, P_3 should both escape on the second day.

However,P_2, P_3 do not make use of their right to escape since they cannot conclude if they have green eyes, since they are also observing P_1 since they can see that he has green eyes.

- P_1's reasoning on the third day: Since P_2, P_3 are still here it must mean that they are not certain if they have green eyes or not which means that my assumption that I do not have green eyes (which would mean that P_2, P_3 would realize that they have green eyes on the second day) is wrong. Therefore, I must have green eyes.

Therefore, P_1 realizes on the third day that he has green eyes and makes use of his right the second midnight and escapes.

Exactly, the same reasoning applies to P_2, P_3 (symmetrically) and so P_2 and P_3 also escape the third night. Therefore, they all realize that they have green eyes on the third day and make use of their right the same night.

This concludes the solution to the prisonners' problem in the case of $n = 3$ prisonners P_1, P_2, P_3.

The general problem is solved similarly, or more precisely inductively. Each prisonners assumes that he does not have green eyes and therefore the problem is reduced to the case of 99 prisonners with by induction (INDUCTION PRINCIPLE) should terminate on the 99th day. But this does not happen, and hence every prisonner realizes on the 100th day that they have green eyes and all escape on the same night.

- Question: How can we change the additional condition, slightly, so that all 100 prisonners escape on the second day, instead of the 100th day?

Please also see the following link for an animated explanation of the problem:
https://www.youtube.com/watch?v=98TQv5IAtY8

Combinatorial Principles: Contradiction, Reduction and Induction

The following three principles were used and were of fundamental importance in solving the prisonners' problem:

- Contradiction: Assume that what we want to show does not hold and use this extra assumption to reach an obviously wrong conclusion (such as $0 = 1$). This is one of the most powerful techniques in mathematics, precisely because it allows for the introduction of one additional assumption (the contradiction assumption) which can be freely used. A nice application of contradiction is the standard proof for the fact that there are infinitely many prime numbers.
- Reduction Principle: Address the problem by first lowering its complexity, for example by lowering the number of variables in the problem.
- Induction Principle: Solve the general problem using the solution for the reduced problem.

2 Lecture 2

In this lecture we will study the very simple but surprisingly important and with far reaching applications Pigeonhole Principle.

- If $n + 1$ pigeons are put in n pigeonholes then there is a pigeonhole with at least 2 pigeons.

Consider for example the case where $n = 4$ with 5 pigeons in 4 pigeonholes. Clearly, there are many different ways to put 5 pigeons in 4 pigeonhole and we will certainly not consider each of these cases separately. We need to show that all of these cases have a common property, namely that there is always a pigeonhole that contains at least two pigeons. In order to prove this, without going through all possible cases separately, we need to argue by contradiction. Specifically we assume that every pigeonhole contains at most one pigeon. Since we have 4 pigeonholes and each of them contains at most one pigeon, we must have at most 4 pigeons in total. This is however contradiction since we have 5 pigeons in total. The proof for the general case with n pigeonholes is done similarly (exercise).

If we increase the number of pigeons while keeping the number of pigeonholes the same, then we can guarantee the existence of a pigeonhole with even more pigeons:

- If $m \cdot n + 1$ pigeons are put in n pigeonholes then there is a pigeonhole with at least $m + 1$ pigeons.

Proof: Exercise.

Let's consider a few applications of pigeonhole's principle.

Problem 1. Among 13 people there are always 2 who were born on the same month. Among 37 people there are always 4 that were born on the same month.

Proof. Indeed, $13 = 12 + 1$ and $37 = 12 \cdot 3 + 1$. □

The above problem has 13 variables (people) and we want to show that a small subset of them (consisting of 2 variables) has much greater structure (namely that they were born on the same month) compared to the orginal set of 13 variables (for which we can in general make no specific statement about their date of birth).

Systematic approach to using the pigeonhole principle

Let v be the number of variables in our problem (e.g. people). If we want to show that at least l of them has a property P then we set:

- p = the number of all possible cases that the variables might satisfy (e.g. $p = 12$ in Problem 1 above since the DOB can be in any of the 12 months),
- $v = m \cdot p + 1$,
- $l = m + 1$.

Then, by pigeonhole principle, we have that, if the initial v variables (pigeons) satisfy any of the p conditions (pigeonholes), then there are at least l variables (pigeons) which satisfy one condition in common (ie. in one pigeonhole). Note that

$$p = \frac{v-1}{l-1}.$$

Hence, we need to make sure that there are in total $p = \frac{v-1}{l-1}$ different cases (properties) for the v variables.

Problem 2. Consider the set

$$A = \{1, 2, 3, \cdots, 18\}$$

and randomly choose four mutually distinct numbers $x_1, x_2, x_3, x_4 \in A$. Prove that there are always two x_i, x_j such that

$$|x_i - x_j| \leq 5.$$

Proof. Of course, in general the difference between two numbers in A is at most 18-1=17. We need to show that if we have four numbers in the set A then the difference of two of them is at most 5 (which is greatly smaller than the general bound 17).

Following our systematic approach, we have $v = 4$ variables x_1, x_2, x_3, x_4 and we want to show that at least $l = 2$ of them satisfy a more refined property. Hence we need to show that our variables satisfy one of $p = \frac{4-1}{2-1} = 3$ properties. This motivates us to decompose the set A in three sets as follows

$$A = \{1, 2, 3, 4, 5, 6\} \cup \{7, 8, 9, 10, 11, 12\} \cup \{13, 14, 15, 16, 17, 18\}.$$

Since we have 4 variables in A, each of our variables must lie in one of the above three sets. Therefore, by pigeonhole principle, at least 2 of our variables lie in the same set! But the difference of any two numbers belonging to the same set is at most 5. This finishes the proof. □

Remarks:

1. Informally speaking, the pigeonhole principle is used when we want to show that in a given set of variables there is a smaller subset of variables which satisfy a more refined property and hence has more structure than the original set of all variables. This is useful when one wants to partition a given the set of variables into smaller sets such that at least one of them has more structure (divide and conquer principle).

2. The pigeonhole principle is used in order to show only the existence of variables satisfying a given (refined) property. There is no explicit construction of these variables. No algorithm is provided on how to get these variables.

Problem 3. Among n consecutive natural numbers there is always one of them which is divisible by n.

Proof. The special feature of this problem is that given our n variables we want to show that there is at least one (and not two or more as is typical in pigeonhole principle) which satisfies an extra refined property.

The trick is to consider the remainders of these n numbers when divided by n. We want to show that one of these remainders is zero. We argue by contradiction. Assume that they all give remainders different from 0. Then there are $n-1$ possibilities for the remainders to be: $1, 2, \cdots, n-1$. But we have n numbers, and so we have n remainders, each of which can only take one of $n-1$ values. Hence, by the pigeonhole principle, there are two numbers which leave the same remainder when divided by n. But then this implies that their difference is divisible by n (exercise: why?). This however is impossible, since the difference of any two number in a any set of n consecutive numbers is always between 1 and $n-1$ and hence not divisible by n. This is contradiction and it shows that the desired result indeed holds.

□

Problem 4. We are given 33 people with the following property *among any 9 of them there are always 2 with the same height.* Show that there are at least 5 people in the original set of 33 people with exactly the same height.

Proof. Following our systematic approach, we have $v = 33$ variables and we want to show that at least $l = 5$ of them satisfy a refined property. Then we need to show that the 33 people satisfy at most

$$p = \frac{33-1}{5-1} = 8$$

conditions.

In other words, in order to show that there are at least 5 people with the same height, we need to show that there are 8 possible heights for the original 33 people. This is however true since in view of the assumption of the problem, if we had 9 possible heights then picking one person with each of these heights would yield a group of 9 people with no two people sharing the same height.

□

Remark: Why is the following reasoning wrong: Consider 9 people out of these 33 people. Two of them have the same height. We are left with 33-9=24 people. Choose 9 people out of these 33 people. By assumption, two of them have the same height. We have two pairs so far and we are left with 24-9=15 people. Choose 9 people out of these 15 people. Two of them have the same height. Then we have 3 pairs of people with the same height. This gives us

6 people with the same height, which in particular gives us 5 people with the same height. What is wrong?

3 Lecture 3

In this lecture we will study more problems and applications of the Pigeonhole Principle.

Recall that according to the pigeonhole principle:

- If $m \cdot n + 1$ pigeons are put in n pigeonholes then there is a pigeonhole with at least $m + 1$ pigeons.

Recall also our systematic approach to using the pigeonhole principle:

Let v be the number of variables in our problem (e.g. people). If we want to show that at least l of them has a property P then we set:

- $p =$ the number of all possible cases that the variables might satisfy (e.g. $p = 12$ in Problem 1 above since the DOB can be in any of the 12 months),
- $v = m \cdot p + 1$,
- $l = m + 1$.

Then, by pigeonhole principle, we have that, if the initial v variables (pigeons) satisfy any of the p conditions (pigeonholes), then there are at least l variables (pigeons) which satisfy one condition in common (ie. in one pigeonhole). Note that

$$p = \frac{v-1}{l-1}.$$

Hence, we need to make sure that there are in total $p = \frac{v-1}{l-1}$ different cases (properties) for the v variables.

Problem 1. Consider the real numbers $a_1, a_2, \cdots, a_n$ such that

$$0 \leq a_i \leq n$$

for all $i = 1, 2, \cdots, n$. Show that there are two distinct a_i, a_j such that

$$|a_i - a_j| \leq \frac{n}{n-1}.$$

Proof. We divide the interval $[0, n]$ in $n - 1$ equal sub-intervals with end points:

$$0, \frac{n}{n-1}, \frac{2n}{n-1}, \frac{3n}{n-1}, ..., \frac{(n-2)n}{n-1}, n.$$

The size of each of these intervals is exactly $\frac{n}{n-1}$. By pigeonhole principle, since we have n numbers in $n - 1$ sub-intervals, there must be two numbers a_i, a_j which lie in the same sub-interval. The difference of these numbers must be bounded by the size of the sub-interval which is equal to $\frac{n}{n-1}$. Hence this two numbers satisfy the desired property. $\square$

Problem 2. Consider six mutually distinct numbers $x_i \in \{1,2,3,4,5,6,7,8,9\}$, $i = 1,2,3,4,5,6$. Show that there are four $x_{i_1}, x_{i_2}, x_{i_3}, x_{i_4}$ of these numbers such that

$$x_{i_1} + x_{i_2} = x_{i_3} + x_{i_4}.$$

Proof. We want to show a refined property for the sums of all pairs, namely that there are two such sums which coincide in value. We therefore want to apply pigeonhole principle for the sums of all pairs.

We have in total 15 possible sums that we can form. For example

$$x_1 + x_2,$$

$$x_1 + x_3,$$

$$x_1 + x_4,$$

$$x_1 + x_5,$$

$$x_1 + x_6,$$

$$x_2 + x_3,$$

are 6 of these 15 possible sums. The values that these 15 sums can take are between 3 and 17 (why?). There are exactly 15 numbers between 3 and 17. Hence we have 15 sums taking 15 possible values. We, therefore, cannot yet apply the pigeonhole principle. For this reason, we consider the following two cases:

Case I: The 15 sums take in fact at most 14 values. Then by pigeonhole principle two of the sum must take the same value and hence must be equal.

Case II: The 15 sums take exaclty 15 values. Then our sums must take all values between 3 and 17. Hence, there is a pair (a, b) whose sum $a + b = 3$ and there is a pair (c, d) whose sum is $c + d = 17$. But then we must have

$$a = 1, b = 2, c = 8, d = 9$$

and hence, obviously,

$$a + c = b + d$$

which shows the desired result. □

Problem 3. Among any six people there are always three who are mutually friends or three who are mutually unknown. Can we replace six by five?

Proof. Please think about this, or look at notes of Lecture 5. □

4 Lecture 4

Recall the pigeonhole principle:

- If $m \cdot n + 1$ pigeons are put in n pigeonholes then there is a pigeonhole with at least $m + 1$ pigeons.

Methodology: If we are given n classes of objects such that

- objects belonging to the same class are different, but,
- objects belonging to different classes might coincide,

and we want to show that there is the same object in at least l different classes then we

1. we count all objects from all classes, and say we compute that this number is v, and
2. we count all possible values that each object can take, and say that this number is m.

If

$$v = m \cdot (l - 1) + 1,$$

then we have $m \cdot (l - 1) + 1$ objects (pigeons) taking at most m values (pigeonholes) and hence there must be at least l objects in all our classes which take the same value (i.e. they coincide). Hence, one object will belong to at least l classes simultaneously.

Problems

[please attempt to solve them before looking at the solution at the following pages]

Problem 1. Show that in any collection of $n+1$ numbers from the set $\{1, 2, 3, \cdots, 2n\}$, where n is a natural number, there are two which are consecutive.

Problem 2. Show that in any collection of $n+1$ numbers from the set $\{1, 2, 3, \cdots, 2n\}$, where n is a natural number, there are two such that one is a multiple of the other.

Problem 3. Consider n natural numbers: $x_1, \cdots, x_n$. Show that there is always a sequence of them such that their sum

$$x_i + x_{i+1} + \cdots + x_j$$

is divisible by n.

Problem 4. A student studied for a period of 37 days according to the following rules

1. every day he studied for at least 1 hour,
2. every day he studied for an integer number of hours without exceeding 12 hours,
3. he had to study for 60 hours in total.

Show that there was a period of a few consecutive days when he studied exactly for 13 hours.

Problem 5. Let $A_1, \cdots, A_{2000}$ be subsets of a set M such that each set A_i contains at least two thirds of the elements of M. Show that there is an element of M which belongs to at least 1334 of the 2000 subsets A_i.

Problem 6. Let p be a prime number different from 2 and 5. Show that among the numbers

$$1, 11, 111, \ldots, 11\cdots1 (p\ 1's)$$

there is always one divisible by p.

Solutions

Problem 1. We are given $n+1$ numbers in the set

$$\{1, 2, 3, \cdots, 2n\}$$

or in other words we are given $n+1$ numbers in the n sets

$$\{1, 2\} \cup \{3, 4\} \cup \{5, 6\} \cup \cdots \cup \{2n-3, 2n-2\} \cup \{2n-1, 2n\}.$$

By pigeonhole principle, two of the given numbers must belong to the same set and hence must be consecutive. $\square$

Problem 2. Every natural number m can be written as follows:

$$m = 2^k \cdot o$$

where o is an odd number known as the greatest odd divisor of m.

Now for each of the given $n+1$ numbers of our exercise we consider the associated greatest odd divisors. Since we have $n+1$ numbers we also must have $n+1$ odd divisors. But in the set $\{1, \cdots, 2n\}$ we have exactly n odd numbers. Hence, by pigeonhole principle, there must exist two numbers whose greatest odd divisors must be equal. These numbers must have the property that one if a multiple of the other (why?).

$\square$

Problem 3. Consider the n sums

$$\begin{aligned}
&S_1 = x_1, \\
&S_2 = x_1 + x_2, \\
&S_3 = x_1 + x_2 + x_3, \\
&\cdots \\
&S_n = x_1 + x_2 + x_3 + \cdots + x_n.
\end{aligned}$$

If one of the above sums is a multiple of n then we are done. If no sum above is a multiple of n then when divided by n they must all leave remainders between 1 and $n-1$. Since we have n sums and $n-1$ possible remainders, by pigeonhole principle, there must be two sums which leave the same remainder. It is easy to see (exercise) that the difference of these two sums is now a multiple of n. But this difference is a sum of the form of the statement of the problem and hence we have proved that such a sum (of consecutive terms) is a multiple of n.

$\square$

Problem 4. Let A_i denote the number of hours that the student studied during the first i days. Then we have the following

1. $A_{i+1} \geq A_i + 1$ (why?),
2. $A_{i+1} \leq A_i + 12$ (why?),
3. $A_i \neq A_j$ for $i \neq j$ (why?),
4. $A_{37} = 60$.

We want to show that there are i, j with $i \geq j+2$ such that

$$A_i = A_j + 13.$$

Following our methodology, we realize that we have the following two classes

$$\text{Class 1:} \quad \{A_1, A_2, \cdots, A_{37}\}$$

and

$$\text{Class 2:} \quad \{A_1 + 13, A_2 + 13, \cdots, A_{37} + 13\}.$$

We want to show that there are two objects from the two classes above which coincide. Clearly we have $37 + 37 = 74$ objects in total. But what are the values that these 74 numbers can take? Clearly we have

$$1 \leq A_i \leq 60 < 73$$

and

$$1 \leq A)i + 13 \leq 73.$$

Hence the given 74 numbers can only take at most 73 values. Hence, by pigeonhole principle, there must be two numbers which take the same value (i.e. they are equal). Hence, for some i and j , with $i \geq j+2$ (why?), we must have

$$A_i = A_j + 13$$

which shows the required result.

$\square$

Problem 5. We have 2000 subsets, that is 2000 classes of objects and we want to show that at least 1334 of their elements coincide.

Following our methodology, we compute that the total number of all objects in all the 2000 subsets is at least

$$2000 \cdot \frac{2}{3} \cdot |M|,$$

where $|M|$ is the number of elements of the set M. But all the elements in these subsets are simply elements of the set M, and hence can be at most $|M|$ elements.

Therefore, we have $2000 \cdot \frac{2}{3} \cdot |M|$ elements (pigeons) taking at most $|M|$ values (pigeonholes), and hence by the pigeonhole principle, there must be at least $2000 \cdot \frac{2}{3}$ elements taking the same value. That is at least $2000 \cdot \frac{2}{3}$ are the same. That is, there is an element belonging to at least $2000 \cdot \frac{2}{3} = 1333.33333$ sets. Hence, we obtain the desired result.

$\square$

Problem 6. We are given p numbers and we want to show that at least one of them is a multiple of p. We argue by contradiction. Assume that none of them is a multiple of p. Then arguing as above, by pigeonhole principle, there must exist two numbers out of the p numbers whose difference is a multiple of p. But the difference of these numbers is always of the form

$$(11 \cdots 1) \cdot 10^k.$$

Hence, p must divide $(11 \cdots 1) \cdot 10^k$. But p does not divide 10^k and hence p must divide $11 \cdots 1$, which contradicts our assumption that p does not divide any number with unit digits! Hence, p must divide of one these numbers.

$\square$

5 Lecture 5

Ramsey Theory

In this lecture we will introduce various problems that are part of a very important area in Combinatorics known as Ramsey theory. The problems we will presently consider are applications of the pigeonhole principle.

Clearly, if we say consider all people on earth (about 7 billion people) then there must three of them who are mutually friends or three of them who are unknown to each other. We would like to consider a smaller group of people such that they will always have this property (namely that there will always be three of them who are mutually friends or three of them who are unknown to each other). The smallest number of (random) people that we need to consider so that there will always be three of them who are friends or three of them who are unknown to each other is in fact only six. We therefore need to prove the following problem:

Problem of six people: Among any six people there are always three who are mutually friends or three who are mutually unknown.

Proof. Consider one person (call him P_0) of the six people. Then this person is either friend or unknown to each of the remaining five people. By pigeonhole principle, there must be at least three of the remaining five people, call them P_1, P_2, P_3, such that P_0 is either a friend of P_1, P_2, P_3 or P_0 is unknown to P_1, P_2, P_3.

Case I. Assume that P_0 is a friend of P_1, P_2, P_3.

Sub-case 1. If there is a pair of friends, call them P_i, P_j (with $i, j \in \{1, 2, 3\}$), among these three people P_1, P_2, P_3 then P_0, P_i, P_j are mutally friends, and so we are done (since we found three people out of the six people who are mutually friends).

Sub-case 2. If there is NO a pair of friends among these three people P_1, P_2, P_3 then P_1, P_2, P_3 are mutally unknown to each other, and so we are done (since we found three people out of the six people who are mutually unknown to each other).

Case II. Assume that P_0 is a unknown to P_1, P_2, P_3.

Sub-case 1. If there is a pair of unknown people, call them P_i, P_j (with $i, j \in \{1, 2, 3\}$), among these three people P_1, P_2, P_3 then P_0, P_i, P_j are mutally unknown, and so we are done (since we found three people out of the six people who are mutually unknown to each other).

Sub-case 2. If there is NO a pair of unknown people among these three people P_1, P_2, P_3 then P_1, P_2, P_3 are mutally friends with each other, and so we are done (since we found three people out of the six people who are mutually friends with each other).

So, no matter how the initial six people are related to each other, there will always be a group of three people which consists of friends or a group of three people which consists of people unknown to each other.

□

It is convienient to represent people by points on the plane and relations of people by colored edges connecting the points.

For example, let the points A, B represent two people P_1, P_2. Then we join A, B with a

- blue edge if $P_1.P_2$ are friends or,
- red edge if $P_1.P_2$ are unknown to each other.

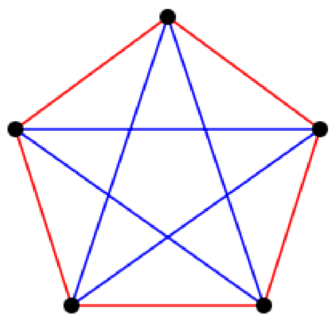
Then the above problem can be restated as follows:

Version 2 of Problem of six people: Consider any six points on the plane and color all edges which connect these six points either red or blue. Show that there must always be a red triangle or a blue triangle.

Proof. Consider one of these points, call it A. Then A is connected to five more points and so there are five edges which terminate at A. Each of these five edges is either red or blue. We have five edges (pigeons) and two colors (pigeonhole), and so by pigeonhole principle, we have at there must be three edges which terminate at A which are either red or blue. Assume that there are three edges which are red (similarly we deal with the case if they are blue). Let B, C, D be the other endpoints of these three red edges. In other words, we have the red segments AB, AC, AD. If either of the segments BC, BD, CD is red, then we have a red triangle (why?). If none of the edges BC, BD, CD is red then they are all blue, and hence the triangle BCD is blue. Hence, in all cases, there must always exist a red or a blue triangle. □

Note that the graph of six points contains exactly 15 colored edges. Since we use only two colors, we have in total $2^15 = 32$k different possibilities for the colorings of the graph. Hence, the above theorem, says that every graph out of the 32k possible graphs must contain a red or a blue triangle. It is remarkable that we were able to show a result for 32k different graphs WITHOUT considering each one of them, but instead by using the pigeonhole principle.

What if we consider five points, instead of six. Then can we always find a monochromatic triangle? The answer is no. See the following picture for a case where there is no monochromatic triangle in a doubly colored graph of five points.



Hence, six is the smallest number that has the property that we can always find blue or a red triangle always.

As we see, the above problem is an example where we seek regularity (e.g. a monochromatic triangle) amid disorder (e.g. a randomly colored graph with two colors).

This is part of a combinatorial theory called Ramsey theory, that, in general, seeks regularity amid disorder. We will not spend more time on Ramsey theory in this course. However, it is interesting to include a few more definitions and provide a few known results (without any proofs).

Definition of Ramsey numbers $R(m, n)$. The natural $R(m, n)$ is defined as the smallest natural number that has the following property:

- Consider $R(m,n)$ points on the plane, and all edges connecting all pairs of these $R(m,n)$ points. Color each of these edges either red or blue (two possible colors). Then there is always a blue m-gon, such that all sides and diagonals are blue, or there is a red n-gon such that all sides and diagonals are red.

The theorem we showed above gives us that

$$R(3,3) = 6.$$

Results that are known:

$$\begin{aligned} R(2,n) &= n, \\ R(3,4) &= 9, \\ R(3,5) &= 14, \\ R(3,6) &= 18, \\ R(3,7) &= 23, \\ R(3,8) &= 28, \\ 42 \leq & R(5,5) \leq 55, \\ 102 \leq & R(6,6) \leq 169. \end{aligned}$$

Note that the numbers $R(5,5), R(6,6)$ are not known! Note that computationally it would almost take infinite time for a computer to determine the exact value, of say $R(5,5)$, between 42 and 55 (why?).

Exercise: Define $R(m,n,k)$ and show that $R(3,3,3) = 17$.

6 Lecture 6

We next consider two fundamental applications of pigeonhole principle. The first one concerns maximally ordered numbers in a random list of numbers (with applications to computer science) and the second concerns powers of 2 (with applications to number theory).

Theorem (Erdös–Szekeres theorem) Consider n^2+1 real numbers $a_1, a_2, \cdots, a_{n^2+1}$. Show that there is always an increasing or decreasing sub-sequence consisting of $n+1$ numbers.

Exercise: Does the theorem hold if we replace n^2+1 with n^2.

Proof. Let's first consider a few special cases:

- $n = 1$: Then we are given $2 = 1^2 + 1$ numbers a_1, a_2. Clearly we will have $a_1 \leq a_2$ or $a_2 \leq a_1$. So there will always be an increasing or decreasing sub-sequence with $2 = 1+1$ numbers.

- $n = 2$: Then we are given $5 = 2^2+1$ numbers a_1, a_2, a_3, a_4, a_5. We want to find $3 = 2+1$ of them which form an increasing or a decreasing chain. Consider as a special case, just to gain intuition, the following 5 (randomly selected) numbers:

$$2, 3, 1, 1.5, 4$$

Note that we can consider the following increasing sub-list: $2, 3, 4$. If we replace the final 4 by something less than 3 (so that we break the above increasing list) then we would get this list of numbers:

$$2, 3, 1, 1.5, 2.5$$

in which case we still have the increasing list $1, 1.5, 2.5$ (which simply does not start from 2, but from 1. If we replace 2.5 with 0, so that we break again this increasing list, then we would have the numbers

$$2, 3, 1, 1.5, 0$$

in which case there is no increasing sequence with 3 numbers! But then we can find a decreasing sequence of three numbers: 2,3,1 or 3,1, 0.

Hence, we see that in all of the above cases, there is always an increasing or decreasing sub-list of 3 numbers. How can we show this property for all possible lists of 5 numbers?

Let's consider the (general) numbers a_1, a_2, a_3, a_4, a_5 and let's assume that there is no list of three of them which is increasing, as was the case for the numbers $2, 3, 1, 1.5, 0$. We will then show that there are three of them which form a decreasing list.

- Consider now the longest decreasing lists starting with the number a_1. If this lists has (at least) three numbers then we are done. Assume then that this list has one or two numbers.
- Consider now the longest decreasing lists starting with the number a_2. If this lists has (at least) three numbers then we are done. Assume then that this list has one or two numbers.
- Consider now the longest decreasing lists starting with the number a_3. If this lists has (at least) three numbers then we are done. Assume then that this list has one or two numbers.
- Consider now the longest decreasing lists starting with the number a_4. This list has at most two numbers a_4, a_5.
- Consider now the longest decreasing lists starting with the number a_5. This list has exaclty only number: a_5.

Hence, we have five numbers and for each of them we consider the length of the longest decreasing sequence starting with them. The lenghts are either 1, 2. Hence we have five numbers and for each of them we have a list of either 1 or 2 numbers. By pigeonhole principle, there must be three numbers with the property that three, call them x_1, x_2, x_3, of the five numbers must have maximal associated decreasing sequences of the same length (either 1 or 2). Say that this length is 2.

We claim that

$$x_1 \leq x_2 \leq x_3.$$

Let's show this. Let's assume that we have $x_1 > x_2$. Then by considering the longest decreasing list of 2 numbers starting with x_2 and joining to this list the number x_1 we would obtain a decreasing list of 3 numbers which starts from x_1. But then the maximal decreasing sequence starting from x_1 would contain 3 numbers and not 2, as is assumed above. Hence, contradiction. Hence, we must not have $x_1 > x_2$ and hence we must have $x_1 \leq x_2$. Similarly we show that $x_2 \leq x_3$. Hence, we have

$$x_1 \leq x_2 \leq x_3.$$

But this gives us a list of three numbers which is increasing.

So there is always a list of three numbers which is either increasing or decreasing.

General case $n \in \mathbb{N}$: Consider for each number a_i the longest decreasing sequence starting from that number a_i (here $i = 1, 2, \cdots, n^2 + 1$). This list will contain say l_i numbers.

- If $l_i \geq n + 1$, for some i, then we are done!
- If $l_i < n + 1$, for all i, then we have

  $$1 \leq l_i \leq n.$$

  Hence, we have $n^2 + 1$ numbers l_i which take at most n values. By pigeonhole principle, there must be $n + 1$ numbers l_i which take the same value, that is which are equal. Let these numbers be:

  $$l_{k_1}, l_{k_2}, \cdots, l_{k_{n+1}}.$$

  Consider the associated numbers from our list

  $$a_{k_1}, a_{k_2}, \cdots, a_{k_{n+1}}.$$

  Then all these numbers have the property that the longest decreasing list starting with each of them has the same number of numbers for all of them!
  This implies that these numbers must form an increasing sequence

  $$a_{k_1} \leq a_{k_2} \leq \cdots \leq a_{k_{n+1}}.$$

  (why?–argue as above by contradiction assuming that for instance $a_{k_1} > a_{k_2}$) Hence we have a list of $n + 1$ numbers which is increasing.

$\square$

The above result shows that if we are given sufficiently many numbers (that is $n^2 + 1$ numbers) then we can find a subset (with $n + 1$ numbers) which has much more structure, that is, it is either increasing or decreasing.

We next consider applications of pigeonhole principle in number theory. We start with the following

Problem (Diophantine approximations). Let a be a positive irrational number and let $\epsilon > 0$ be any positive number. Show that there are natural numbers n, m such that

$$|n \cdot a - m| \leq \epsilon. \tag{1}$$

As a corollary, there are natural numbers n, m such that

$$\left| a - \frac{m}{n} \right| | \leq \frac{\epsilon}{n} \leq \epsilon.$$

Proof. First of all, it is easy to find infinitely many m, n such that

$$\left| a - \frac{m}{n} \right| \leq \frac{1}{n}.$$

Indeed, the above inequality is equivalent to

$$|n \cdot a - m| \leq 1.$$

Consider now any natural number $n > \frac{1}{a}$ and consider the number $n \cdot a$. Then clearly this number must be between two consecutive natural numbers $m, m+1$. Then for this m we have

$$m \leq n \cdot a \leq m + 1$$

which implies that

$$|n \cdot a - m| \leq 1.$$

If $\epsilon > 1$ then we are done (why?). If $\epsilon < 1$, then we consider the multiples of ϵ:

$$0 < \epsilon < 2\epsilon < 3\epsilon < \cdots < (k-1)\epsilon < 1 < k\epsilon.$$

Hence, the multiples of ϵ divide the interval $[0, 1]$ in k intervals:

$$[0, \epsilon],$$
$$[\epsilon, 2\epsilon],$$
$$[2\epsilon, 3\epsilon],$$
$$[3\epsilon, 4\epsilon],$$
$$\ldots$$
$$[(k-1)\epsilon, k\epsilon].$$

Let's now consider $k+1$ different numbers of the form $n \cdot a - m$ which are less than 1 (by the above). Since we have $k+1$ numbers in a set of k intervals (the ones with endpoints the multiples of ϵ) two of them must be in the same set! But then their difference is bounded by ϵ, since ϵ is the size of the interval. The difference of numbers of the form $n \cdot a - m$ is of the same form. Hence, we proved that there exists numbers n, m such that

$$|n \cdot a - m| \leq \epsilon.$$

Hence,

$$\left| a - \frac{m}{n} \right| \leq \frac{\epsilon}{n} \leq \epsilon.$$

□

Problem (Application in number theory). Show that there is a power of 2 which starts with 2017, that is there is $n \in \mathbb{N}$ such that

$$2^n = 2017 \cdots .$$

Proof. We want to show that there are natural numbers n, m such that

$$2017 \cdot 10^m \leq 2^n \leq 2018 \cdot 10^m$$

or equivalently, if $a = \log 2$ then

$$\log 2017 + m \leq n \log 2 \leq \log 2018 + m$$

or

$$\log 2017 \leq n \log 2 - m \leq \log 2018.$$

If we let $a = \log 2$, then by above problem, see equation (1), we have that there are natural numbers n', m' such that

$$|n' \log 2 - m'| \leq (\log 2018 - \log 2017)/2.$$

Then an appropriate multiple of the number $n' \log 2 - m'$ will lie between the numbers $\log 2018$ and $\log 2017$ (why?). This multiple gives us the desired result. $\square$

7 Lecture 7

The principle of extremals

We next consider a class of combinatorial problems the solution of which makes use of the so-called *principle of extremals.* According to this principle:

- if we want to show that a certain construction exists, then we consider the largest or the smallest among a specific class of structures. The largest one, or the smallest, might be the one that satisfies the given properties.
- if we want to show that a certain construction cannot possible exist, then, arguing by contradiction, we assume that at leaste one such construction exists. We next, consider the largest (or the smallest) such construction and using it and the assumptions of the problem we deduce that there is an even larger (or smaller respectively) construction. But this contradicts the fact that we started with the largest (or smallest respectively) construction. Hence, no such construction can exist.

As we shall see, the principle of extremals is an extremely strong principle and can provide very elegant answers to surprisingly complex problems.

We next present a few applications of the principle of extremals.

Problem 1. Consider 100 numbers $a_1, a_2, \cdots, a_{100}$, each number being on the vertex of a regular 100-gon. We assume that any number is the average of its two neighbors, that is

$$a_i = \frac{a_{i-1} + a_{i+1}}{2}.$$

Show that all these numbers are equal, that is $a_1 = a_2 = \cdots = a_{100}$.

Proof. Consider the largest number, say a_j, of these 100 numbers. By assumption a_j is the average of its two neighbors a_{j-1}, a_{j+1}. However since a_j is the largest number, we necessarily have

$$a_j \geq a_{j+1},$$
$$a_j \geq a_{j-1}.$$

This is possible only if

$$a_{j-1} = a_j = a_{j+1}.$$

Therefore, a_{j-1}, a_{j+1} are also the greatest numbers. So by the above reasoning they are also equal to their neighbors. We proceed inductively.

□

Problem 2. Consider n points on the plane. Color each of these points either red or blue. We assume that on any edge with endpoints of the same color there is a point of different color (that is, there is always a red point between two blue points, and a blue point between two red points). Show that all the points must lie on the same line.

Proof. Assume that not all points like on the same line. Then we can form triangles using our points. We consider that triangle with the smallest area. Let is be ABC. Two of its vertices, say B, C, will have the same color. Hence, by assumption, there will be another point, call it D, of different color on the side BC. But then clearly the triangle ABD has smaller area, contradiction. Hence there cannot be any triangles and hence all points lie on the same line.

□

Problem 3. Show that $\sqrt{2}$ is irrational.

Proof. Assume that $\sqrt{2}$ is rational. There is a multiple of $\sqrt{2}$ which is a natural number. Consider the following set

$$A = \left\{ n \in \mathbb{N} : n\sqrt{2} \text{ is a natural number } \right\}$$

By assumption the set A is a non-empty subset of the natural numbers. Hence, it must have a smallest element. Let this smallest number in A be n_0. Consider the number

$$n_1 = n_0\sqrt{2} - n_0.$$

Clearly by our assumptions n_1 is a natural number (why?). Furthermore

$$n_1\sqrt{2} = 2n_0 - \sqrt{2}n_0$$

is also a natural number (why?). Hence, n_1 must be a member of the set A. But $n_1 < n_0$ (why?) which contradicts the fact that n_0 is the smallest element of A. Hence there cannot be a smallest element in A. Hence A must be empty. Which means that $\sqrt{2}$ is indeed irrational.

□

Problem (exercise). Show that given $n + 2$ natural numbers there are always two whose sum or difference is divisible by $2n + 1$.

8 Lecture 8

Recall that according to the principle of extremal:

- if we want to show that a certain construction exists, then we consider the largest or the smallest among a specific class of structures. The largest one, or the smallest, might be the one that satisfies the given properties.
- if we want to show that a certain construction cannot possible exist, then, arguing by contradiction, we assume that at leaste one such construction exists. We next, consider the largest (or the smallest) such construction and using it and the assumptions of the problem we deduce that there is an even larger (or smaller respectively) construction. But this contradicts the fact that we started with the largest (or smallest respectively) construction. Hence, no such construction can exist.

We next present a few applications of the principle of extremals.

Problem 1. Consider a solar system with 2017 planets such that all planets have mutually distinct distances from each other. Consider an astronomer on each of these planets (so 2017 astronomers in total) who observe the closest planet to them. Show that there must be a planet which is not observed by any astronomer.

Proof. Consider first for simplicity (reduction principle) the case where we have three planets P_1, P_2, P_3. Arguing using the principle of extremals, we consider the two planets with the smallest distance. Let's assume that these planets are P_1, P_2. Clearly, the astronomer on P_1 observes the planet P_2 and the astronomer on P_2 observes P_1. The observer on P_3 observes either P_1 or P_2 and hence nobody observes P_3. So we proved the desired result for the case of three planets.

What about the case of 2017 planets $P_1, P_2, \cdots, P_{2017}$?

First note that it suffices to show that there are at least two astronomers who observe the same planet. Indeed, there are exactly 2017 planets and exactly 2017 astronomers each of which observes one and only one planet. So, if two astronomers observe the same planet, there must be a planet which is not observed.

Arguing using the principle of extremals, we consider the two planets, say P_1, P_2, with the smallest distance. As before, the astronomers on these planets observe each other. We next consider two cases:

- Case I: If there is a third astronomer that observes either P_1 or P_2 then either P_1 or P_2 is observed by at least two astronomers. Hence we are done.
- Case II: No other astronomer observes P_1 or P_2. Then, the reduced planetary system $P_3, P_4, ...P_{2017}$ satisfies the same property as the bigger one. We proceed by induction. We argue for the reduced planetary system $P_3, P_4, ...P_{2017}$ as above. We consider the pair of planets with the smallest distance and consider cases I, II as above. If case I holds then we are done. If case II holds then we are able to remove two more planets and end up with 2013 planets. We keep doing the same thing until we reach three planets. The case of three planets was however addressed at the beginning.

$\square$

Problem 2. Consider a collection C of points on the plane such every point of C is the midpoint of two other points of C. Show that C must contain infinitely many points.

Proof. Assume that we have finitely many points. Consider the longest segment AB defined by these (finitely many) points. By assumption there is segment CD whose midpoint is B. But then AB is the median of the triangle ACD and hence (by a well-known result in plane geometry)

$$AB < (AC + AD)/2$$

which implies that either AC or AD is longer than AB, which contradicts the maximality assumption for the length of AB. Hence, no such longest segment can exist and so there must infinitely many points (because for any collection of finitely many points there are always two with the longest distance).

□

Problem 3. (Sylvester–Gallai theorem) Consider a collection D of n points on the plane with the property that for any two points in D there is a third point in D which lies on the line defined by the two points. Show that all points in D must lie on the same line (that is, they must be all co-linear).

Proof. Assume that not all points lie on the same line. So there must be triangles that are formed with vertices being points in the collection D.

Since we assume that such triangles exist, we can consider that triangle CAB with the smallest height AL (where L is the projection of A on BC). By assumption, there is a point K of the collection D on the line AB. Let's assume that C lies between the points K and B.

Then the height from B of the triangle BCK is smaller than the height from A of the triangle CAB (why?). But we have assumed that the triangle CAB has the smallest height. This is contradiction. Hence, there cannot be any triangle with the smallest height, hence there cannot be any triangles to begin with. Hence all points must lie on the same line (in which case they indeed do not form any triangle).

□

Problem 4. Consider n points on the plane such that every point is connected via edges with at least three other points. Show that there must exist a closed path with an even number of edges. (equivalently, using terms from graph theory, the problem says that for any finite graph with the property that every vertex has degree at least three, there is always an even cycle).

Proof. Consider the longest path in our finite graph:

$$A_1 \to A_2 \to \cdots A_{k-1} \to A_k.$$

By assumption, the point A_1 must be connected with at least three points. One of them is A_2. The other two points must necessarily be points of that path (otherwise we would be able to obtain an even larger path contradicting the fact that we considered the longest path). Let's assume that A_1 is connected with A_i and A_j with $i < j$. Then we have three cycles:

$$\text{cycle 1. } A_1 \to A_2 \to \cdots A_{k-1} \to A_k.$$

$$\text{cycle 2. } A_1 \to A_2 \to \cdots A_i \to A_1.$$

$$\text{cycle 3. } A_1 \to A_i \to \cdots A_j \to A_1.$$

If i is even then cycle 1 is even. If j is even, then cycle is even. If i, j are both odd, then cycle 2 is even, since its length is $j - i$. Hence, in all cases, there is always an even cycle.

□

9 Lecture 9–11

The principle of invariants

The principle of invariants is yet another very powerful principle that is applicable to problems which involve a process and the goal is to see what the possible final states of that process are. A general statement of this principle is the following:

- Assume that we have a process which takes place in discrete steps $S_1, S_2, S_3, ..., S_k$ where S_k is the final step of the process.
- Assume also that the variables of the problem take different values at different steps (that is, the system changes from step to step).
- According to the principle of invariants, for each step S_i we need to obtain an appropriate quantity Q_i that depends on the values variables at the S_i step. The quantity Q_i should be such that
$$Q_1 = Q_2 = \cdots = Q_k.$$
In other words, we would like the quantity Q_i not to change from step to step meaning that it is indeed a genuine invariant of the process.
- Using the initial state of the system we determine the value of the invariant.
- By the invariant property, the value we found above should also coincide with the value at the final state.
- Finally, knowing the value of the invariant at the final state should allow us to obtain significant information about the possible final states of the process.

The most challenging part in applying this principle is finding what the correct invariant quantity is. There is no general rule as to how to obtain an invariant quantity. It depends on the specific problem under consideration. We next present a few applications of the principle of invariants.

Problem 1. The numbers $1, 2, 3, \cdot, 200$ are written on the blackboard. At each step

- We select two numbers, say x, y, that are written on the blackboard
- We delete the numbers x, y from the blackboard
- We add the number $x + y$ on the blackboard.

What are the possible final numbers on the blackboard.

Proof. Let Q be the sum of all numbers written on the blackboard at each step. Clearly Q is an invariant, that is it does not change from step to step (why?). Initially Q is equal to 19900. Hence, the final number on the blackboard must be 19900.

□

Problem 2. We write the numbers 1,0,1,0,0,0 (in this order) on a circle. At each step

- We select two consecutive such numbers on the circle.
- We add 1 to both of the numbers we selected.

Is it possible to perform this process in such a way so that we obtain equal numbers after a finite number of steps?

Proof. Let $x_1, x_2, x_3, x_4, x_5, x_6$ be the numbers on the circle at each step. We consider the quantity

$$Q = x_1 -, x_2 + x_3 - x_4 + x_5 - x_6.$$

Note that this quantity is initially equal to 1. This quantity is an invariant of our process (why?). Hence, we can never reach a state where wll the numbers are equal since such a state would have $Q = 0$. But we showed that $Q = 1$ always.

□

Problem 3. Can we put the numbers 1,2,3,..., 2019 in a row to form one number which is a perfect square?

Proof. What all the numbers that we can form have in common? All the possible numbers we can form have equal sum of digits

$$1 + 2 + 3 + \cdots + 2019 = \frac{1}{2} 2019 * 2020 = 2039190.$$

This number is a multiple of 3, but not a multiple for 9. However any perfect square which is a multiple for three is also a multiple of 9. So we cannot form a perfect square.

□

The next problem is a very impressive application of the principle of invariants.

Problem 4. The numbers 1,2,3,..., 10 are written on the blackboard. At each step:

- We select two numbers, say x, y that are written on the blackboard
- We delete the numbers x, y from the blackboard
- We add the number $x + y + xy$ on the blackboard.

What are the possible final numbers on the blackboard.

Proof. In the first problem, the invariant quantity was simply the sum of all numbers written on the blackboard. However, the sume is not an invariant for this problem and hence is not useful here. On the other hand, we observe that

$$x + y + xy + 1 = (x + 1)(y + 1).$$

This means that if we add 1 to all numbers that are written on the blackboard and then take the product of the resulting numbers, then this product is an invariant (why?)!! Initially, this product is equal to 11!. Hence, the final number must be $11! - 1$ (why?). $\square$

For the next few problems we will need a generalized version of an invariant. That is, we will need to construct quantities Q for each step, which although its value changes from step to step, its parity (for example) does not change. That is Q will always be an even number, or will always be an odd number, or will always be a number of the form $3n+1$ etc. It is the latter property of Q that will serve as our invariant.

Problem 1. Let n be an odd natural number. The numbers $1, 2, 3, \cdots, 2n$ are written on the blackboard. We follow the following process. At each step:

- We select two, say $x > y$, of the numbers that are written on the blackboard
- We delete them from the blackboard
- We add the difference number $x - y|$ on the blackboard.

is it possible to perform the process such that 2 is the final number written on the blackboard?

Proof. Let Q be the sum of the numbers written on the blackboard at each step. How does Q change from step to step? Since we replace x, y by $|x-y|$ we have that

$$Q(i+1\text{-step}) = Q(i\text{-step}) - x - y + (x - y) = Q(i\text{-step}) - 2y.$$

Hence Q changes by an even number. Since Q is initially odd (why?) it must always be odd. Hence, the final number cannot be even!

□

Problem 2. There are 100 0's and 100 1's written on the blackboard. We follow the following process:

- At each step we choose two of the numbers written on the blackboard.
- If the numbers we chose are equal, then we delete them from the blackboard and add a 0 on the blackboard.
- If the numbers we chose are not equal, then we delete them from the blackboard and add a 1 on the blackboard.

Is it possible to perform the steps appropriately such that the final (unique) number that is written on the blackboard is 1?

Proof. Let Q be the sum of all numbers on the blackboard. Clearly, Q is initially even. At each step, Q either is unchanged or is decreased by 2. Hence Q remains even always. Therefore, the final number must be even, that is 0. Hence it cannot be 1.

□

Problem 3. We have three tubes containing 8, 9, 10 coins respectively. We follow the following process:

- At each step we select two of the three tubes.
- We remove one coin from each of the two tubes we selected.

- We add these two coins to the third tube.

Can we perform this process such that all coins are transferred to the same tube?

Proof. At each step, the difference of the number of coins in any two tubes is either unchanged or is changed by 3 or -3. Suppose we are able to put all coins in one of the tubes. Then the other two tubes will have zero coins (and hence the difference will also be zero). But this difference is initially 1. So it is impossible to transfer all coins to the same tube. $\square$

Semi-invariants

For a type of problems we will not be able to find an invariant quantity; however, we might be able to find what is know as a semi-invariant quantity. A semi-invariant quantity is a quantity that is always increasing or always decreasing from step to step. Using this monotonicity property of the semi-invariant quantity might yield important properties of the final states of the process under consideration.

A spectacular application of the use of semi-invariants is the so-called Conway's army.

Problem (Conway's army). Please see the following link:
`https://en.wikipedia.org/wiki/Conway's_Soldiers`

10 Lecture 12–15

We spent the first half of the course in analyzing structures and finding ways to obtain properties of them. We will spend the second half of the course in counting structures. We will need to find ways to count all possible configurations that satisfy given properties without considering each and every possible configuration. In other words, we need to find ways to count without really counting.

We start we some basic combinatorial principles.

- **Additive principle: If there are n ways to perform process A and m ways to perform process B then there are $n+m$ ways to perform process A <u>or</u> the process B.**

It is very important to emphasize that $n+m$ refers to the number of all possible ways of performing either of the two processes and not both of them.

A trivial, but instructive, example is the following: A school has two classes. The first class has 20 students and the second class has 30 students. The total number of students of the school is computed using the additive principle since any student belongs to either classes but not in both of them. Hence the total number is 20+30=50.

Multiplicative principle:

The multiplicative principle concerns processes happening simultaneously, supplementing therefore the additive principle.

- **If there are n ways to perform process A and m ways to perform process B then there are $n \cdot m$ ways to perform process A <u>and</u> process B.**

For example, in the case of the school above, there are $20 \cdot 30$ ways to choose a student from the first class and a student from the second class (forming thus a pair of students). Indeed, choosing a student from the first class can be done in 20 ways. Choosing a student from the second class can be done in 30 ways. Hence, according to the multiplicative principle, choose a pair from both classes can be done in $20 \cdot 30 = 600$ ways.

The multiplicative principle (also known as the rule of product) is of fundamental importance in combinatorics and for this reason is known as the fundamental principle of counting.

Principle of bijections

The principle of bijections allows to show that two different processes can be performed in the same number of ways. We first need to recall a few basic definitions.

Let A, B be two sets. A function $f : A \to B$ is called

- injective if for all $x_1 \neq x_2$ in A we have $f(x_1) \neq f(x_2)$ in B,
- surjective if for all $y \in B$ there is an $x \in A$ such that $f(x) = y$,
- bijective if it is both injective and surjective.

Principle of bijections: Two (finite) sets A, B have the same number of elements if there is a bijection function from A to B.

The case of infinite sets is more involved. For example, there is a bijection between all even natural numbers and all natural numbers, but there is no bijection between all natural numbers and all real numbers.

Permutations

Consider n objects. For example, consider the first n natural numbers

$$T_n = \{1, 2, 3, \cdots, n\}$$

A permutation of T_n is a re-arrangment of T_n, that is it is a new ordered list where all the natural numbers from 1 to n appear exactly once in some specific order.

For example, the following are permutations of T_3:

$$(1,2,3),\ (1,3,2),\ (2,1,3),\ (2,3,1),\ (3,1,2),\ (3,2,1).$$

In fact these 6 permutations are all permutations of T_3.

More generally, there are

$$n! = 1 \cdot 1 \cdot 2 \cdot 3 \cdots n$$

(n factorial) permutations of T_n and hence $n!$ permutations of any n distinct objects. This follows immediately by the multiplicative principle. Indeed, any permutation can be determined by first determining its first entry (there are n possible ways for this), then determining its second entry (there are $n-1$ ways for this) etc... until we reach the final entry (which can only be determined in 1 way, since all the other $n-1$ entries have been already selected). Hence, by the multiplicative principle the entire process can be completed in $1 \cdot 1 \cdot 2 \cdot 3 \cdots n = n!$ ways.

Ordered subsets

An ordered subset with k elements of T_n is a list which consists of k elements taken from T_n such that each element in the list appears not more than once. For example, the following lists

$$(1,2), (1,3), (2,1), (2,3), (3,1), (3,2)$$

are ordered subsets of T_3 with $k = 2$ elements.

By the multiplicative principle, and arguing as above, there are

$$n \cdot (n-1) \cdot (n-2) \cdots (n-(k-1)) = \frac{n!}{(n-k)!}$$

ordered subsets of T_n with k elements. Indeed, there are n ways to select the first entry, $n-1$ ways to select the second entry, ..., $n-(k-1)$ ways to select the k^{th} entry.

Combinations

A combination of T_n with k elements is a subset of T_n with k elements for which every element appears at most once and for which order does not matter. In other words, the order of the elements does not play a role in combinations and, moreover, no repetition of elements is allowed.

For example, the following

$$\{1,2,3\}, \{3,2,1\}, \{1,3,2\}$$

represent the **same** combination of T_4 with 3 elements. On the other hand, the following

$$\{1,2,2\}, \{3,3,1\}, \{1,1,2\}$$

are **not** combinations since we have elements which are appear more than once.

Question: How many combinations of T_n with k elements are there?

Answer: The only difference between combinations with k elements and ordered-subsets with k elements is that order does not play a role in combinations whereas order plays a role in ordered subsets. Since we have k elements, there are in total $k!$ different re-orderings (permutations) of any ordered subset with k elements. On the other hand, there are $\frac{n!}{(n-k)!}$ ordered subsets with k elements. However, since $k!$ ordered subsets represent the same combination, we have that the total number of all combinations with k elements is

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

The number $\binom{n}{k}$ is called the binomial coefficient because of the following property:

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^k = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n. \tag{2}$$

This identity follows easily by observing the following: for each k the term x^k is formed by choosing k of the n factors of the product:

$$(1+x)^n = (1+x) \cdot (1+x) \cdot (1+x) \cdots (1+x).$$

Clearly there are $\binom{n}{k}$ to choose k factors and hence there are $\binom{n}{k}$ terms of the form x^k in the expanded sum.

Identities of the binomial coefficients

The following identities hold

1. Special cases: $\binom{n}{0}$=1, $\binom{n}{1}$=n, $\binom{n}{2} = n(n-1)/2$, $\binom{n}{n} = 1$.
2. Symmetry identity: $\binom{n}{k} = \binom{n}{n-k}$.
3. Pascal identity: $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$.

These identities are easily proved algebraically using the formula for the binomial coefficients. However, we can also provide combinatorial proofs which require no computations.

Combinatorial proofs of identities: There are two generals methods to prove identities in a combinatorial way:

1. The first method makes use of the bijection principle. That is, we need to find a bijection between the elements counted on the left hand side of the identity and the elements counted on the right hand side of the identity.

2. The second method computes the elements of the same set in two different ways. Since the same number is computed in two different ways, the final results from these two ways must coincide leading to the desired identity.

Combinatorial proof of the symmetry identity:

The symmetry identity easily follows from the bijection principle. Indeed, there is a bijection between combinations with k elements and combinations with $n-k$ elements. Indeed, if we choose a combination with k elements then we have left out $n-k$ elements which clearly form an $n-k$ combination. Hence, by the bijection principle, the number $\binom{n}{k}$ of all combinations with k elements are equal to the number $\binom{n}{n-k}$ of all combinations with $n-k$ elements.

Combinatorial proof of the Pascal identity:

Pascal's identity follows from counting the same set in two different ways. The LHS gives the number of all k combinations. We will prove that the RHS gives the numbers of all k combinations as well. Consider the element $\{n\}$ of the set T_n. A k combination either contains $\{n\}$ or does not contain $\{n\}$. There are in total $\binom{n-1}{k-1}$ combinations with k elements which contain $\{n\}$, since any such k combination is determined by the choice of the first $k-1$ elements from the set T_{n-1} (since the k^{th} element is $\{n\}$). On the other hand, there are $\binom{n-1}{k}$ combinations with k elements which do not contain $\{n\}$ since each such combination is completely determined by the choice of k elements from the set T_{n-1} (since the element $\{n\}$ is now excluded). The identity follows immediately from the additive principle.

More identities of the binomial coefficients
The following identities hold:

$$
\begin{aligned}
\sum_{k=0}^{n} \binom{n}{k} &= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n, \\
\sum_{k=0}^{n} k\binom{n}{k} &= 0 \cdot \binom{n}{0} + 1 \cdot \binom{n}{1} + 2 \cdot \binom{n}{2} + \cdots + n \cdot \binom{n}{n} = n2^{n-1}, \\
\sum_{k=0}^{n} \binom{n}{k}^2 &= \binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n},
\end{aligned}
$$

Proof. The first identity can be proved *algebraically* by setting $x = 1$ in the the binomial identity (2). However, we can also provide a combinatorial proof of this identity. The left hand side, by the additive principle, counts all possible combinations that we can form. The right hand side counts the same thing. Indeed, for each of the n elements there are 2 possibilities: either the elements is a member of the combination or not. Since we have n elements, and for each of them we have 2 possibilities, by the multiplicative principle, there are in total 2^n possibilities and hence 2^n combinations in total. Note that at the last step we had to use the multiplicative principle and not the additive principle, since each combinations is determined by knowing if every element is part of it or not. That is we have to know the state of all elements simultaneously and hence

The second identity can be proved *analytically* by first differentiating the binomial identity (2) and then setting $x = 1$ in the resulting identity. However, we can again provide a combinatorial proof of this identity. The term $k\binom{n}{k}$ counts all possible teams (combinations) with k players (elements) with an assigned leader. Indeed, there are $\binom{n}{k}$ teams with k players and for each of these teamas there k ways to choose the leader. Hence, by the additive principle, the LHS counts the number of all teams with an assigned leader. We will show that the RHS counts the same thing. Indeed, since we have n players to begin with, there are n possibilities for the assigned leader. For each of the remaining $n-1$ players we have 2 possibilities: either they are a member of the team or they are not a member of the team. Hence, by the multiplicative principle we have in total $n2^{n-1}$ different ways to make a team with a leader.

The third identity can also be proved combinatorially. In fact, using the symmetry identity, it suffices to prove that

$$\sum_{k=0}^{n} \binom{n}{k} \cdot \binom{n}{n-k} = \binom{2n}{n}.$$

The RHS counts all combinations with n elements from $2n$ objects. We can divide these $2n$ objects in two sets S_1, S_2 each of which contains n sets. Hence, the RHS counts all combinations with n elements that we can form using the elements in the sets S_1, S_2. We will show that the LHS counts the same thing. Indeed, any n combination of the $2n$ elements of the set $S_1 \cup S_2$ will contain k elements from S_1 and $n-k$ elements from S_2, for some k between 0 and n. There are $\binom{n}{k}$ ways to choose the k elements from S_1 and $\binom{n}{n-k}$ ways to choose the $n-k$ elements from S_2. By the multiplicative principle, there are $\binom{n}{k} \cdot \binom{n}{n-k}$ such combinations. The result follows from the additive principle where we sum all possibilities for k.

□

Applications

1. The path problem

Consider all points (x, y) on the plane with integer coordinates. How many different paths are there from $(0, 0)$ to (n, n), if we can only move towards the right, that is

$$(x, y) \to (x+1, y)$$

and towards the top of the plane, that is

$$(x, y) \to (x, y+1)?$$

The answer is $\binom{2n}{n}$. Indeed, any such path is completely determined by the sequence of the $2n$ individual moves, n of which are towards the right (each such move is denoted by R) and the other n moves are towards the top of the plane (each of which is denoted by T). How many different lists of length $2n$ are there if n of the elements of the lists are R and the other n elements are T? We clearly only need to specify the n locations where have the symbol R in our lists. There are $2n$ positions and we need to specify (that is, to choose) n of them. This can be done in $\binom{2n}{n}$ ways.

2. Solutions to linear equations

Question 1: How many solutions $(x_1, x_2, \cdots, x_n)$ to the equation

$$x_1 + x_2 + \cdots + x_n = k$$

are there if

$$x_i \in \{0, 1\}?$$

Answer: Clearly, k of the unknowns have to be equal to 1 and the remaining $n-k$ of the unknowns have to equal to 0. If we choose the k variables which will be equal to 1 then we automatically choose the remaining $n-k$ variables that will be set equal to 0. Hence it suffices to choose k of then n variables. This can be done in $\binom{n}{k}$ ways.

Question 2: How many solutions $(x_1, x_2, \cdots, x_n)$ to the equation

$$x_1 + x_2 + \cdots + x_n = k$$

are there if

$$x_i \in \{0, 1, 2, 3, ..., k\}?$$

Answer: We have k units to distribute among the n unknowns. However, in contrast to the above example, in this case we can distribute more than one unit to each variable. In other words, we can choose each variable more than once. In fact, each variable can be chosen either

- 0 times (in which case is given the value 0) or
- 1 times (in which case is given the value 1) or
- 2 times (in which case is given the value 2) or
- $\cdots$
- k times (in which case is given the value k).

This means that we need to choose k of the given n variables but where we allow repetition, that is we allow each variable to be chosen more than once (and up to k times).

The above leads us to consider **combinations with repetition**. According to assignment 3 (or see also lecture notes 17) there are in total $\binom{n+k-1}{k}$ combinations with k elements with repetition. Hence, there are in total $\binom{n+k-1}{k}$ different solutions to the above equation.

11 Lecture 16

We next present two problems that appeared in mathematical olympiads. These problems are beautiful applications of the additive and the multiplicative principles.

Problem 1. Consider 3 lines on the plane passing through the point O creating 6 different sectors. We put 5 different points in each of these sectors (so 30 points in total). Show that there are at least 1000 triangles with vertices 3 of these 30 points and such that the point O is either in the interior or on the boundary of these triangles.

Proof. Let's enumerate the sectors by 1,2,3,4,5,6. If the vertices of a triangle belong in the sectors (1,3,5) or (2,4,6) then they must contain O. How many such triangles are there? There are 5 points in sector 1, 5 points in sector 3 and 5 points in sector 5. Hence, by the multiplicative principle, there are $5 \cdot 5 \cdot 5$ triangle with vertices in the sectors (1,3,5). Similarly, there are $5 \cdot 5 \cdot 5$ triangles with vertices in the sectors (2,4,6). This gives us 250 triangles with the desired property. We need to find 750 more such triangles.

Consider an pair of points from opposite sectors, that is from sectors (1,4), (2,5), (3,6). Consider the case where the points are taken from the sectors (1,4). Consider the line that connects these points. Then the point O must lie in one of the two parts of the plane that are created by this line. Either the sectors (2,3) or the sectors (5,6) also line in the same part. Suppose that the sectors (2,3) line on that part of the plane. Then any triangle with vertices from the sectors (1,4,2) or (1,4,3) contain the point O. How many such triangles are there? By the multiplicative principle there are $5 \cdot 5 \cdot 5$ triangles with vertices on the sectors (1,4,2) and $5 \cdot 5 \cdot 5$ triangles with vertices on the sectors (1,4,3) giving in total 250 triangles. However we have in total 3 pairs of opposite sectors to begin with, hence by the additive principle, there are 750 more triangles that contain O giving a total of 1000 such triangles.

□

Problem 2: Compute the number of all squares on the plane with vertices of the form (x, y) with $x, y \in \mathbb{N}$ such that $1 \leq x \leq n, 1 \leq y \leq n$.

Answer: We will first compute the number of all squares with sides parallel to the x and the y axes and of type $k \times k$. The bottom left vertex (a, b) of any such square must satisfy (why?)

$$1 \leq a \leq n-k, 1 \leq b \leq n-k.$$

By the multiplicative principle, there are $(n-k)^2$ such points and hence $(n-k)^2$ such triangles.

There are squares whose sides are not parallel to the horizontal and the vertical axes. Every such square can be inscribed, however, in a square with horizontal and vertical sides (why?). Hence, in order to compute the number of all squares, it suffices to compute all parallel squares (that is with horizontal and vertical sides) and all the squares inscribed in them. Given any $k \times k$ parallel square, we can find $k-1$ inscribed squares (why??). Hence, any $k \times k$ parallel squares gives rise to k squares in total, 1 is the parallel square itself and $k-1$ are the inscribed squares. Hence, by the multiplicative principle, there are $k \cdot (n-k)^2$ squares that are created by the $k \times k$ paralles squares. By the additive principle we have in total

$$S(n) = \sum_{k=1}^{n} k \cdot (n-k)^2$$

squares. The number $S(n)$ can be computed easily using

$$\sum_{k+1}^{n} k = n(n+1)/2,$$

$$\sum_{k+1}^{n} k^2 = n(n+1)(2n+1)/2,$$

$$\sum_{k+1}^{n} k^3 = n^2(n+1)^2/4.$$

It follows (why?)

$$S(n) = n^2(n+1)(n-1)/12.$$

12 Lecture 17

Combinations with repetition

A combination with repetition with k elements of n objects is an un-ordered collection of k elements from n given objects where each element might appear repeatedly up to k times.

For example, there are 3 combinations of 2 elements from 3 objects, namely

$$\{1,2\}\,,\{1,3\}\,,\{2,3\}$$

but there are 6 combinations with repetition of 2 elements from 3 objects, namely

$$\{1,2\}\,,\{1,3\}\,,\{2,3\}\,,\{1,1\}\,,\{2,2\}\,,\{3,3\}\,.$$

We will next show that there are

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

combinations with repetition with k elements from n objects.

We observe that any combination with repetition with k elements from the set $\{1,2,3,\cdots,n\}$ gives rise to a sequence of $(n-1)+k$ objects such that k of them are the k objects of the combination with repetition and $(n-1)$ are arrows $\to$ that are used to declare that we done with using one number in our combination and we move to the next number. For example, the combination

$$\{1,2\}$$

gives rise to the sequence with $n-1+k=2+2=4$ entries:

$$(1,\to,2,\to).$$

The above list basically provides a way to exhaust the set $\{1,2,3\}$ in the following way:

1. we write 1 as the first entry, since 1 is a member of the combination
2. we are done with 1, so we use the arrow $\to$ to declare that we need to move to 2
3. we write 2 as the next entry, since 2 is a member of the combination
4. we are done with 2, so we use the arrow $\to$ to declare that we need to move to 3 where the process terminates since we cannot use 3 anymore because we have already chosen 2 members.

The above imply that every combination with repetition with 2 members gives rise to an exhaustion of the set $\{1,2,3\}$ that consists of 4 in total steps, as above. An another example, consider the case of the combination with repetition

$$(2,\to,2,\to)$$

of $\{1,2,3\}$. The above provides a way to exhaust the set $\{1,2,3\}$ in the following way:

1. 1 is not a member of the combination, so we need to use $\to$ to move to 2
2. we write 2 as the next entry, since 2 is a member of the combination

3. we write 2 as the next entry, since 2 is again member of the combination
4. we are done with 2, so we use the arrow $\rightarrow$ to declare that we need to move to 3 where the process terminates since we cannot use 3 anymore because we have already chosen 2 members.

This produces the following sequence:

$$(\rightarrow, 2, 2, \rightarrow).$$

More generally, every combination with repetition with k elements produces a unique sequence with $n-1+k$ entries ($n-1$ arrows and k objects). The main observation is that every such sequence is uniquely determined by the location of the $n-1$ arrows. Hence, it suffices to know in how many ways we can place $n-1$ arrows in $n-1$ of the $n-1+k$ entries of the sequence. This can clearly be done in $\binom{n-1+l}{n-1}$ ways.

Permutations with repetition

Consider the triplet $(1, 1, 2)$. How many different re-arrangments (permutations) of this triplet are there? Clearly we have the following 3 permutations

$$(1, 1, 2), (1, 2, 1), (2, 1, 1).$$

Note that we only obtained 3 permutations instead of 6, because it does not make any sense to flip 1 and 1. In other words, we have 6! permutations modulo the ones that flip 1 and 1 we are 2! and hence we obtain $6!/2! = 3$ permutations where two of the three objects are repeated.

More generally, if we have n such that k_1 are of the same type (that is the same repeated object), k_2 are of the same type, ..., k_l are of the same type (hence l different types of objects), such that

$$n = k_1 + k_2 + \cdots + k_l$$

then there are

$$\frac{n!}{k_1! \cdot k_2! \cdots k_l!}$$

permutations with repetition.

The path problem revised

Recall the path problem where we want to find the number of all possible paths on the plane that pass through points with natural coordinates, start from $(0, 0)$, finish at (n, m), and consist of the following horizontal moves:

$$(x, y) \rightarrow (x + 1, y)$$

and vertical moves

$$(x, y) \rightarrow (x, y + 1).$$

Each such path is determined by a sequence of $n+m$ moves, n of which are horizontal and m are vertical. Hence, the total number of paths is equal to all permutations with repitition of $n+m$ objects where n of them are identical (horizontal moves) and m of them are also identical (vertical moves). According to the formula above, this is equal to

$$\frac{(n+m)!}{n! \cdot m!} = \binom{n+m}{n}.$$

Restricted paths

Now suppose we want to consider all paths from $(0,0)$ to (n,n) which however satisfy one additional restriction, namely that they do not intersect the line $y = x+1$. This line will be called the *forbidden* line. The paths that intersect the forbidden line are called *bad* paths, and the paths that do not intersect the forbidden line are called *good* paths.

We will compute the number of bad paths first. Then the number of good paths is simply equal to the number of all paths (given by the formula above) minus the number of all bad paths.

Consider the point $(n-1, n+1)$. This is the reflection of (n,n) across the forbidden line $y = x+1$. We will show that there is a bijection between all bad paths from $(0,0)$ to (n,n) and all regular paths from $(0,0)$ to $(n-1, n+1)$. Indeed, any path from $(0,0)$ to $(n-1, n+1)$ must intersect the forbidden line (why?). By refleting across the forbidden line the portion of the path after its intersection with the forbidden line we obtain a bad path from $(0,0)$ to (n,n). This process yields a bijection (why?).

Since there are $\binom{2n}{n-1}$ paths from $(0,0)$ to $(n-1, n+1)$, there must be $\binom{2n}{n-1}$ bad paths from $(0,0)$ to (n,n).

Hence, there are in total

$$c_n = \binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n}$$

good paths. The natural numbers c_n are known as the *Catalan* numbers.

Inclusion–Exclusion principle

Let $|A|$ denote the cardinality of a set A. The inclusion-exclusion principle says that

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

That is the number of the elements of the union$A \cup B$ is equal to the number of the elements in A plus the number of the elements in B minus the number of the elements that were already counted twice (there are elements in the intersection of the sets). Similarly, we obtain

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

Hence, the inclusion-exclusion principle allows us to compute the cardinality of the union of sets in terms of the cardinality of the intersection of sets. This is very convenient, since in general it is easier to compute the cardinality of the intersection.

The following in an application in number theory.

How many natural numbers between 1 and 600 are there which are multiples of 2 or 3?

There are $600/2 = 300$ multiples of 2. There are $600/3 = 200$ multiplies of 3. And there are $600/6 = 100$ multiples of 2 and 3. Hence, by the inclusion-exclusion principle there are $300 + 200 - 100 = 400$ numbers which are multiples of 2 or 3. And clearly, there are $600 - 400 = 200$ numbers which are neither a multiple of 2 nor of 3.

The above example is important because it is related to a fundamental function in number theory, namely Euler's ϕ function. For any natural number, $\phi(n)$ is equal to the number of natural numbers $m \leq n$ which have no common prime divisor with n. We can produce a formula for $\phi(n)$ using the inclusion-exclusion principle.

If we assume that
$$n = p_1^{k_1} \cdot p_2^{k_2}$$
where p_1, p_2 are the prime divisors of n then we immediately obtain that there are n/p_1 multiples of p_1, there are n/p_1p_2 multiples of p_2 and there are n/p_1p_2 multiples of both p_1 and p_2. Hence there are
$$\frac{n}{p_1} + \frac{n}{p_2} - \frac{n}{p_1p_2}$$
numbers which are divisible by either p_1 or p_2. Hence,
$$\phi(n) = n - \frac{n}{p_1} - \frac{n}{p_2} + \frac{n}{p_1p_2} = n \cdot \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right).$$
In a similar fashion we can produce a formula for general natural numbers $n = p_1^{k_1} p_2^{k_2} \cdots p_l^{k_l}$.

13 Lecture 18

Recurrence relations

Recurrence relations is a very powerful method in combinatorics to compute all the states of a system that depends on n variables.

Let a_n denote the number of all different states that a system with n variables can take. Instead of computing a_n directly, we assume that a_{n-1} (and possibly a_{n-2}) are known and we compute a_n in terms of a_{n-1} (and a_{n-2} if needed).

This technique is somehow the analogue of the inductive principle in number theory where we assuming that a statement holds for $k = n - 1$ (and needed for $k = n - 2$) we show that it holds for $k = n$.

Hence, we want to obtain a relation of the following form:
$$a_n = f(a_{n-1}) \tag{3}$$
or
$$a_n = f(a_{n-1}, a_{n-2}). \tag{4}$$
Solving (3) requires knowing a_1 and solving (4) requires knowing a_1, a_2.

There is a huge theory developed for solving recurrence equations of the form (3) and (4). In this course, we will only consider basic equations that can be solved with elementary methods. For example, it is important to recall the formulas for the geometric progressions and the arithmetic progressions:

Arithmetic progression
$$a_n = a_{n-1} + d \Rightarrow a_n = (n - 1)d + a_1.$$

The method of **telescopic sums** can be used to derive the above formula. Indeed, if we use the equation

$$a_k - a_{k-1} = d$$

for $k = 2, 3, ..., n$ and we sum them then all but the first and last terms on the LHS will cancel (telescopic sum) yielding

$$a_n - a_1 = (n-1)d.$$

Geometric progression

$$a_n = r \cdot a_{n-1} \Rightarrow a_n = r^{n-1} \cdot a_1.$$

The method of **telescopic product** can be used to derive the above formula. Indeed, if we use the equation

$$\frac{a_k}{a_{k-1}} = r$$

for $k = 2, 3, ..., n$ and we multiply them then all but the first and last terms on the LHS will cancel (telescopic product) yielding

$$\frac{a_n}{a_1} = r^{n-1}.$$

Problems

Problem 1. Compute the number of all subsets of the set

$$I_n = \{1, 2, 3, ..., n\}.$$

Answer.

Let a_n be the number of all subsets of I_n. Then clearly $a_1 = 2$ since we have two subsets in I_1, namely the empty set and the subset $\{1\}$. Similarly, we see that $a_2 = 4$ and $a_3 = 8$. We will show that $a_n = 2^n$ for all n using the method of recurrent relations.

All subsets in I_n can be split in two categories:

- Category I consists of all subsets of I_n which contain n.
- Category II consists of all subsets of I_n which do not contain n.

How many subsets are there in Category I and how many in Category II?

Clearly, any subset in Category II is also a subset of I_{n-1} (why?). Hence, there are a_{n-1} subsets in Category II. Also, any subset in Category I is in fact the union of $\{n\}$ and a subset in Category II. This yields a bijection between the subsets in Category I and the subsets in Category II. Hence, by the bijection principle, there are a_{n-1} subsets in Category I. Hence, by the additive principle we finally get

$$a_n = a_{n-1}$$

which with $a_1 = 2$ yields that

$$a_n = 2^n.$$

Problem 2. Consider an equilateral triangle OAB and add n points on each of the sides OA, OB, AB so in total we have $3n$ new points on the three sides. Specifically, on OA we consider the points $P_1, P_2, ..., P_n$ such that

$$OP_1 = P_1P_2 = ... = P_nA.$$

Similarly, on OB we consider the points $Q_1, Q_2, ..., Q_n$ such that

$$OQ_1 = Q_1Q_2 = ... = Q_nB$$

and on AB we consider the points $T_1, T_2, ..., T_n$ such that

$$AT_1 = T_1T_2 = ... = T_nB.$$

We consider all the segments that connect the points P_i, Q_i, T_i and we consider all the intersection points of these segments. Consecutive intersection points are connected by what we call quantum segments.

How many paths (which cannot revisit the same vertex more than once) are there from O to A if we can only move along quantum segments such that we can only move in the following directions:

- parallel to the vector $\vec{AB}$,
- parallel to the vector $\vec{BA}$,
- parallel to the vector $\vec{OA}$,
- parallel to the vector $\vec{OB}$.

and if $n = 6$?

Answer.

Let a_n be the number we want to compute.

First observe that once a path has reached a point on the line AB then there is only one way to continue to the path so it terminates at A. This implies that the number of paths from O to A is the same as the number of paths from O to any other point on the line AB (why?). This is a very crucial observation for this problem.

Next consider the line that connects the points P_nQ_n. This is the last line L parallel to AB. Any path from O to A must intersect the line L at one of its $n+1$ points. Once a path has reached its last point on the line L then there are only two ways to continue the path to the point A (why?). Hence, it suffices to compute the number of all paths from O to any of the $n+1$ points of the line P_nQ_n. Clearly there are

$$(n + 1) \cdot a_{n-1}$$

such paths (why??). Hence,

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

There desired number a_6 can be computed using the method of telescopic products (exercise).

14 Lecture 19

We start with a few more applications the method of recurrent relations.

Problem. How many different sets of n pairs can be formed from $2n$ people?

Answer.

Let a_n be the number we want to compute.

Consider any person. Then there are $2n-1$ pairs that can be formed with this one person. After forming any of these pairs, we need to form $n-1$ pairs from the remaining $2n-2$ people. This can be done in a_{n-1} ways. Hence,

$$a_n = (2n-1)a_{n-1}.$$

Clearly $a_1 = 1$. Hence, by the method of telescopic products we obtain that a_n is equal to the product of all odd numbers less than $2n$. Indeed, by multiplying the following equations

$$\begin{aligned}
\frac{a_n}{a_{n-1}} =&(2n-1)\\
\frac{a_{n-1}}{a_{n-2}} =&(2n-3)\\
\frac{a_{n-2}}{a_{n-3}} =&(2n-5)\\
&\dots\\
\frac{a_3}{a_2} =&5\\
\frac{a_2}{a_1} =&3
\end{aligned}$$

we obtain

$$\frac{a_n}{a_{n-1}} \cdot \frac{a_{n-1}}{a_{n-2}} \cdot \frac{a_{n-2}}{a_{n-3}} \cdots \frac{a_3}{a_2} \cdot \frac{a_2}{a_1} = (2n-1)\cdot(2n-3)\cdot(2n-5)\cdots 5\cdot 3\cdot 1$$

and hence (since $a_1 = 1$) we obtain

$$a_n = \prod_{k=1}^{n}(2k-1).$$

Now, a_n can be rewritten as follows

$$\begin{aligned}
a_n =&(2n-1)\cdot(2n-3)\cdot(2n-5)\cdots 5\cdot 3\cdot 1\\
=&\frac{(2n)\cdot(2n-1)\cdot(2n-2)\cdot(2n-3)\cdot(2n-4)\cdot(2n-5)\cdots 5\cdot 4\cdot 3\cdot 2\cdot 1}{(2n)\cdot(2n-2)\cdot(2n-4)\cdots 4\cdot 2}\\
=&\frac{(2n)\cdot(2n-1)\cdot(2n-2)\cdot(2n-3)\cdot(2n-4)\cdot(2n-5)\cdots 5\cdot 4\cdot 3\cdot 2\cdot 1}{2^n\cdot n\cdot(n-1)\cdot(n-2)\cdots 2\cdot 1}\\
=&\frac{(2n)!}{2^n\cdot n!}.
\end{aligned}$$

That is,

$$a_n = \frac{(2n)!}{n!\cdot 2^n}.$$

Problem (Hanoi Towers). We are given three pegs A, B, C and n disks of graduated size with holes in their centers. Initially the discs are at peg A such that any disc is always on top of a bigger disk. We want to transfer all discs to another peg by moving one disk at a time and without placing a larger disk on top of a smaller disk. What is the minimal number of moves required to transfer the n disks?

Answer.

Let a_n be the number we want to compute. It takes a_{n-1} moves to transfer the first $(n-1)$ disks from peg A to peg B. Then it takes one move to transfer the biggest disk from peg A to peg C. And it takes again a_{n-1} moves to transfer the $n-1$ disks from peg B to peg C. This is really the minimal number of moves (why?). Hence,

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

Clearly, $a_1 = 1$. To solve this recurrence relation, we add 1 to both sides and we obtain

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

Hence, if we denote

$$x_n = a_n + 1$$

then

$$x_n = 2x_{n-1}$$

with $x_1 = a_1 + 1 = 2$. Hence $x_n = 2^n$ and hence

$$a_n = 2^n - 1.$$

Problem. Consider an equilateral triangle OAB and add n points on each of the sides OA, OB, AB so in total we have $3n$ new points on the three sides. Specifically, on OA we consider the points $P_1, P_2, ..., P_n$ such that

$$OP_1 = P_1P_2 = ... = P_nA.$$

Similarly, on OB we consider the points $Q_1, Q_2, ..., Q_n$ such that

$$OQ_1 = Q_1Q_2 = ... = Q_nB$$

and on AB we consider the points $T_1, T_2, ..., T_n$ such that

$$AT_1 = T_1T_2 = ... = T_nB.$$

How many parallelograms are formed by all segments that connect the points O, A, B, P_i, Q_i, T_i?

Answer.

We call nodes the intersections of the segments. Clearly there are

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

nodes.

Let a_n be the number we want to compute. All parallelograms can be divided in two classes:

- Type 1 consists of all parallelograms which do not have any vectices on AB,
- Type 2 consists of all parallelograms which have at least one vertex on AB.

Clearly, there are a_{n-1} parallelograms of type 1. We will next compute the number of all parallelograms of type 2. For we will establish a bijection between all parallelograms of type 2 and specific pairs (u, v) of vertices such that v is one of the $n+2$ nodes on the line AB. The bijection is the following: Any parallelogram has two opposite angles equal to 60 degrees and two opposite angles equal to 120 degrees. We associate any parallelogram to the pair of vertices of the angles which are equal to 60 degrees. This association is a bijection (why??).

Any pair of vertices comes that from a parallelogram in the above way is called admissible. Clearly, if (u, v) is an admissible pair of vertices then u lies in one of the three planar sections (angles) that are created by the three lines that pass through v and each of which has angle 60 degrees. This immediately implies that u cannot line on any of the lines that pass through v.

We will next compute all admissible pairs (u, v) such that $v \in AB$. We have $n+2$ nodes on AB. For each of these $n+2$ nodes (call it v), we need to compute the total number of admissible nodes, that is the total number of nodes which of nodes which do not lie on any of the three lines that pass through the node v. Clearly, one of these three lines is the line AB which in total contains $n+2$ nodes. The number of the nodes on the other two lines passing through v, and excluding v since it has already been counted in the nodes of AB, is equal to $n+1$ (why??).

Therefore, there are

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

admissible pairs for the node v. There are $n+2$ modes v on AB hence in total we have

$$\frac{n(n+1)(n+2)}{2} = 3\binom{n+2}{3}$$

admissible pairs, and hence parallelograms of type 2.

We have

$$a_n = a_{n-1} + 3\binom{n+2}{3}.$$

Also, clearly $a_0 = 0, a_1 = 3$. By the telescopic sums method we obtain

$$a_n - a_0 = \sum_{k=0}^{n} \binom{k+2}{3}$$

and hence, by the bonus problem of assignment 3, we obtain

$$a_n = 3\binom{n+3}{4}$$

.

Remark: For an alternative, more direct, proof, see: http://www.laurentlessard.com/bookproofs/countin parallelograms/

15 Lecture 20–21

Generating Functions

The method of generating functions is a very powerful tool in combinatorics developed by Euler in 1748. The main idea is to turn combinatorial considerations into algebraic manipulations. The latter algebrain operations can in principle be done by a computer system.

In most problems in combinatorics, given any natural number n, we want to compute the number a_n of all possible ways to perform a given process that depends on n. The method of generating functions provides a new way to compute a_n by first computing the so-called associated generating function of a_n:

$$g_{a_n}(x) = a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n + \ldots = \sum_{i=0}^{\infty} a_i x^i.$$

Clearly we have

$$g_{a_n}(x) = g_{b_n}(x) \text{ if and only if } a_n = b_n \text{ for all } n \in \mathbb{N}.$$

We will not consider issues of convergence here. We will assume that x takes values such that the powerseries (i.e. the generating functions) are finite without worrying about the exact convergence interval.

The main idea of the method is to compute a_n by first computing the associated generating function $g_{a_n}(x)$. Once the generating function has been found, then the sequence is trivially computed since it is simply given by the coefficients of the monomials x^n for all n.

Algebraic operations of generating functions

1. Sum of generating functions:

$$g_{a_n}(x) + g_{b_n}(x) = g_{c_n}(x),$$

where

$$c_n = a_n + b_n.$$

2. Product of generating functions:

$$g_{a_n}(x) \cdot g_{b_n}(x) = g_{c_n}(x),$$

where

$$c_n = a_0 \cdot b_n + a_1 \cdot b_{n-1} + a_2 \cdot b_{n-2} + \cdots + a_n \cdot b_0.$$

The sequence c_n is known as the Cauchy product or the convolution of the sequences a_n, b_n.

Special examples of generating functions

1. The most important example of a generating function is the one associated to the constant sequence

$$a_n = 1$$

that is

$$g_{a_n}(x) = 1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}. \tag{5}$$

Note that this formula follows from the following identity

$$1 + x + x^2 + \cdots x^n = \frac{1 - x^{n+1}}{1-x}$$

and using that $x^n \to 0$ as $x \to \infty$. Furthermore, if

$$a_n = b$$

then

$$g_{a_n}(x) = \frac{b}{1-x}.$$

On the other hand, if

$$a_n = b^n$$

then

$$g_{a_n}(x) = \frac{1}{1-bx}.$$

2. If $a_k = 1$ for $k = n$ and $a_k = 0$ for all $k \neq n$ (for some fixed n) then

$$g_{a_n}(x) = x^n.$$

3. If

$$a_k = \binom{n}{k} \text{ for } k \leq n, \text{ and } a_k = 0 \text{ for } k > n$$

then

$$g_{a_n}(x) = (1+x)^n.$$

The above follows immediately by the binomial identity.

Differentiation of generating functions

Recall from calculus the following identity

$$\frac{d}{dx}\left(\frac{1}{1-x}\right) = \frac{1}{(1-x)^2}$$

On the other hand, by differentiating (9) we obtain

$$\frac{1}{(1-x)^2} = \frac{d}{dx}\left(\frac{1}{1-x}\right) = 1 + 2x + 3x^2 + \cdots + (n+1)x^n + \cdots. \tag{6}$$

Hence, $\frac{1}{(1-x)^2}$ is the generating function for the sequence $a_n = n+1$. Moreover,

$$\frac{d}{dx}\left(\frac{1}{(1-x)^2}\right) = \frac{2}{(1-x)^3}.$$

Moreover, by differentiating (10) we obtain

$$\frac{1}{(1-x)^3} = \frac{1}{2}\frac{d}{dx}\left(\frac{1}{(1-x)^2}\right) = 1 + 3x + \cdots + \frac{(n+1)(n+2)}{2}x^n + \cdots$$

Hence, $\frac{1}{(1-x)^3}$ is the generating function for the sequence $a_n = \frac{(n+1)(n+2)}{2}$.

Exercise: Compute the generating functions of the sequences $a_n = n$ and $b_n = n^2$.

Combinatorics and generating functions

In this section we will consider problems where we need to select a certain number of elements (usually we want to select n elements) from a certain set of elements such that each of the elements satisfies given restrictions.

A brief general description of the method is as follows (note that the method will become more clear by reading the examples that follow):

Step 1.

We introduce a variable x which indicates if an element has been chosen. Specifically,

- x^0 indicates that an element has not been selected,
- x^1 indicates that an element has been selected 1 time,
- x^2 indicates that an element has been selected 2 times,
- ....
- x^n indicates that an element has been selected n times.

Step 2.

For each of the elements of the set we construct an associated powerseries, that is, sums of those powers of x meet the restrictions of the element.

Step 3.

We consider the product of all the powerseries of Step 2. This product gives the generating function associated to the problem. The coefficient of x^n in the generating function gives the desired sequence a_n.

The above will become more clear in the examples that follow:

Example 1.

For example, consider the set

$$S = \{a, b, c\}$$

such that

- a can be selected only an even number of times,
- b can be selected only an odd number of times,
- c can be selected at most 10 times.

Find the number of ways that we can select n elements by repeatedly selecting elements from S so that the above restrictions are met.

Answer: Let a_n be the number of ways of the desired selection. We will first compute the generating function $g_{a_n}(x)$.

Step 1:

We introduce the variable x which indicates if an elements has been selected.

Step 2:

For the element a we have the following considerations:

- x^0 indicates that a has not been selected (i.e. selected 0 times, and 0 is even),
- x^2 indicates that a has been selected 2 times,
- x^4 indicates that a has been selected 4 times,

- ....
- x^{2n} indicates that a has been selected $2n$ times,
- ....

Hence, the associated powerseries for the element a is

$$1 + x^2 + x^4 + \cdots + x^{2n} + \cdots$$

For the element b we have the following considerations:

- x^1 indicates that b has been selected 1 time,
- x^3 indicates that b has been selected 3 times,
- x^5 indicates that b has been selected 5 times,
- ....
- x^{2n+1} indicates that b has been selected $2n+1$ times,
- ....

Hence, the associated powerseries for the element b is

$$1 + x^3 + x^5 + \cdots + x^{2n+1} + \cdots$$

For the element c we have the following considerations:

- x^0 indicates that c has not been selected,
- x^1 indicates that c has been selected 1 time,
- x^2 indicates that c has been selected 2 times,
- ....
- x^{10} indicates that c has been selected 10 times.

Hence, the associated powerseries for the element c is

$$1 + x + x^2 + x^3 + \cdots + x^{10}.$$

Step 3.

Finally we obtain the generating function by multiplying the powerseries associated to each of the variables a, b, c:

$$g_{a_n}(x) = (1+x^2+x^4+\cdots+x^{2n}+\cdots)\cdot(1+x^3+x^5+\cdots+x^{2n+1}+\cdots)\cdot(1+x+x^2+x^3+\cdots+x^{10}).$$

The coefficient of x^n in the above expression gives the desired number a_n. Note that if n is large, say $n = 1000$ then it is difficult to find the value a_n by hand, however it is always easy for a computer to quickly determine the value of a_n.

Example 2.

In how many ways can we select 4 elements from the set

$$S = \{a, b, c, d\}$$

such that

- a can be selected at most 2 times,
- b can be selected at most once,
- c can be selected at most 2 times,
- d can be selected at most once.

Answer: We will first compute the generating function $g_{a_n}(x)$. Then the answer will be equal to the coefficient of x^4.

Step 1:

We introduce the variable x which indicates if an elements has been selected.

Step 2:

For the element a we have the following considerations:

- x^0 indicates that a has not been selected (i.e. selected 0 times, and 0 is even),
- x^1 indicates that a has been selected once,
- x^2 indicates that a has been selected 2 times,

Hence, the associated powerseries for the element a is

$$1 + x + x^2$$

For the element b we have the following considerations:

- $1 = x^0$ indicates that b has been selected 0 times,
- x^1 indicates that b has been selected 1 time.

Hence, the associated powerseries for the element b is

$$1 + x$$

For the element c we have the following considerations:

- x^0 indicates that c has not been selected,
- x^1 indicates that c has been selected 1 time,
- x^2 indicates that c has been selected 2 times,

Hence, the associated powerseries for the element c is

$$1 + x + x^2$$

For the element d we have the following considerations:

- $1 = x^0$ indicates that d has been selected 0 times,
- x^1 indicates that d has been selected 1 time.

Hence, the associated powerseries for the element d is

$$1 + x$$

Step 3.

Finally we obtain the generating function by multiplying the powerseries associated to each of the variables a, b, c, d:

$$g_{a_n}(x) = (1 + x + x^2) \cdot (1 + x) \cdot (1 + x + x^2) \cdot (1 + x)$$

The coefficient of x^4 in the above expression gives the desired number.

Example 3.

Solutions to linear equations with restrictions. Find the number of all natural solutions (x_1, x_2, x_3, x_4) to the linear equation

$$x_1 + x_2 + 2x_3 + 3x_4 = n$$

if

- $x_1 \geq 2$,
- x_2 takes only even values,
- $x_3 \geq 0$,
- $x_4 \geq 3$.

Answer: We will first compute the generating function $g_{a_n}(x)$. Then the answer will be equal to the coefficient of x^n.

Step 1:

We introduce the variable x which represent a unit. In total we need to select (i.e. distribute) n units.

Step 2:

For the element x_1 we have the following considerations:

- x^2 indicates that x_1 has been given 2 units (that is $x_1 = 2$),
- x^3 indicates that $x_1 = 3$,
- ...

Hence, the associated powerseries for the element x_1 is

$$\begin{aligned} x^2 + x^3 + x^4 + \cdots =& x^2 \cdot (1 + x + x^2 + \cdots) \\ =& x^2 \cdot \frac{1}{1 - x}. \end{aligned}$$

For the element x_2 we have the following considerations:

- $1 = x^0$ indicates that $x_2 = 0$,
- x^2 indicates that $x_2 = 2$,

- ...
- x^{2n} indicates that $x_2 = 2n$ (even),
- ...

Hence, the associated powerseries for the element x_2 is

$$\begin{aligned} 1 + x^2 + x^4 + \cdots + x^{2n} + \cdots =& 1 + (x^2) + (x^2)^2 + \cdots + (x^2)^n + \cdots \\ =& \frac{1}{1 - x^2}. \end{aligned}$$

For the element x_3 we have the following considerations: It is the value $2x_3$ that we need to consider. Since $x_3 \geq 0$ we have that $2x_3$ takes any even value greater or equal to 0. Hence,

- $1 = x^0$ indicates that $2x_3 = 0$,
- x^2 indicates that $2x_3 = 2$,
- ...
- x^{2n} indicates that $2x_3 = 2n$ (even),
- ...

Hence, the associated powerseries for the element $2x_3$ is

$$\begin{aligned} 1 + x^2 + x^4 + \cdots + x^{2n} + \cdots =& 1 + (x^2) + (x^2)^2 + \cdots + (x^2)^n + \cdots \\ =& \frac{1}{1 - x^2}. \end{aligned}$$

For the element x_4 we have the following considerations:It is the value $3x_4$ that we need to consider. Since $x_4 \geq 3$ we have that $3x_4$ takes any value greater or equal to 9 which is a multiple of 3. Hence,

- x^9 indicates that $3x_4 = 9$,
- x^{12} indicates that $3x_4 = 12$,
- ...
- x^{3n} indicates that $3x_4 = 3n$ (multiple of 3),
- ...

Hence, the associated powerseries for the element d is

$$\begin{aligned} x^9 + x^{12} + \cdots + x^{3n} + \cdots =& x^9 \cdot (1 + x^3 + x^6 + \cdots + x^{3n} + \cdots \\ =& x^9 \cdot \frac{1}{1 - x^3}. \end{aligned}$$

Step 3.

Finally we obtain the generating function by multiplying the powerseries associated to each of the variables $x_1, x_2, 2x_3, 3x_4$:

$$g_{a_n}(x) = x^2 \cdot \frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^2} \cdot x^9 \cdot \frac{1}{1-x^3} = x^{11} \cdot \frac{1}{1-x} \cdot \frac{1}{(1-x^2)^2} \cdot \frac{1}{1-x^3}.$$

The coefficient of x^n in the above expression gives the desired number.

Example 4.

In how many ways can we toss a die three times and get a total of 14?

Answer: We toss the die three times and hence we have three results to consider (the sum of which has to be 14). Each individual result is a natural number between 1 and 6. Hence if x indicates a unit, then x^k indicates that the result of a toss if exactly k. In total we need to have 14 units so we need to compute the coefficient of x^{14} in the associated generating function.

The powerseries corresponding to each die toss is

$$x + x^2 + x^3 + x^4 + x^5 + x^6$$

and since we we toss the die three time the final generating function is

$$(x + x^2 + x^3 + x^4 + x^5 + x^6)^3.$$

The coefficient of x^{14} in the above function is the desired number (it can be easily computed by a computer).

16 Lecture 22

We present two problems that can be solved using the method of generating functions.

Problem 1.

Compute the generating function of a_n if a_n is the number of ways to select n elements from the set

$$S = \{e_1, e_2, \cdots, e_m\}$$

such that the element e_i can be selected at most k_i times.

Answer: For any $i = 1, 2, ..., m$, the element e_i can be selected at most k_i times and hence its associated powerseries is

$$1 + x + x^2 + \cdots + x^{k_i}.$$

The final generating function is simply the product of all the powerseries of all elements e_i:

$$g_{a_n}(x) = \prod_{i=1}^{m}(1 + x + x^2 + \cdots + x^{k_i}).$$

The coefficient of x^n gives the desired answer.

Problem 2.

Find the number of ways to select $2n$ balls from n identical red balls, n identical blue balls and n identical white balls.

Answer: We will use the method of generating functions. The associated powerseries of each of the three types of balls (ie. red, blue and white balls) is

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Hence, the generating function is

$$g_{a_n}(x) = \frac{(1-x^n)^3}{(1-x)^3}.$$

Note that

$$(1-x^{n+1})^3 = 1 - 3x^{n+1} + 3(x^{n+1})^2 - (x^{n+1})^3 = 1 - 3x^{n+1} + 3x^{2n+2} - x^{3n+3}$$

and recall from Lecture notes 20–21 that

$$\frac{1}{(1-x)^3} = 1 + 3x + \cdots + \frac{(n+1)(n+2)}{2}x^n + \cdots$$

Hence, we want to compute the coefficient of x^{2n} in the product

$$g = (1 - 3x^{n+1} + 3x^{2n+2} - x^{3n+3}) \cdot \left(1 + 3x + \cdots + \frac{(n+1)(n+2)}{2}x^n + \cdots\right)$$

The answer is given by

$$\frac{(2n+1)(2n+2)}{2} - 3\frac{n(n+1)}{2} = \frac{(n+1)(n+2)}{2}.$$

Remark: The problem is equivalent to finding the number of all natural solutions (x_1, x_2, x_3) of the linear equation

$$x_1 + x_2 + x_3 = 2n$$

such that $0 \le x_i \le n$ for $i = 1, 2, 3$.

17 Lecture 23

Applications of Generating Functions

Consider first the following problem which can be solved using the method of generating functions.

Problem 1.

Consider 5 boxes in a row.

- Let a_n denote the number of ways that we can distribute n identical objects into these 5 boxes such that the first, the third and the fifth boxes are always non-emptry (that is, they all contain at each one objects).
- Let b_n denote the number of ways that we can distribute n identical objects into these 5 boxes such that the second and the fourth boxes contain at least two objects.

Prove that $a_n = b_{n+1}$ for all n.

Proof. We will use the method of generating functions (it is left as an exercise to the students to find alternative solutions which do not make use of generating functions).

We will compute the generating functions g_{a_n} and g_{b_n}.

Let x_i denote the number of objects we put in the i^{th} box. We need to find the number of solutions $(x_1, x_2, x_3, x_4, x_5)$ of the linear equation

$$x_1 + x_2 + x_3 + x_4 + x_5 = n$$

with additional restrictions on x_i.

We start with g_{a_n}. By assumption, we need to have

$$x_1 \geq 1, x_3 \geq 1, x_5 \geq 1.$$

Hence the powerseries associated to each of the variables x_1, x_3, x_5 is

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

The only restrictions for the variables x_2, x_4 are simply

$$x_2 \geq 0, x_4 \geq 0.$$

Hence the powerseries associated to each of the variables x_2, x_4 is

$$1 + x + x^2 + \cdots x^n + \cdots = \frac{1}{1-x}.$$

Hence, the generating function g_{a_n} of a_n is

$$g_{a_n}(x) = \frac{x}{1-x} \cdot \frac{1}{1-x} \cdot \frac{x}{1-x} \cdot \frac{1}{1-x} \cdot \frac{x}{1-x} = \frac{x^3}{(1-x)^5}. \tag{7}$$

We next compute g_{b_n}. By assumption, we need to have

$$x_2 \geq 2, x_4 \geq 2.$$

Hence the powerseries associated to each of the variables x_2, x_4 is

$$x^2 + x^3 + \cdots + x^n + \cdots = x^2 \cdot (1 + x + x^2 + \cdots x^n + \cdots) = \frac{x^2}{1-x}.$$

The only restrictions for the variables x_1, x_3, x_5 are simply

$$x_1 \geq 0, x_3 \geq 3, x_5 \geq 0.$$

Hence the powerseries associated to each of the variables x_1, x_3, x_5 is

$$1 + x + x^2 + \cdots x^n + \cdots = \frac{1}{1-x}.$$

Hence, the generating function g_{b_n} of b_n is

$$g_{b_n}(x) = \frac{1}{1-x} \cdot \frac{x^2}{1-x} \cdot \frac{1}{1-x} \cdot \frac{x^2}{1-x} \cdot \frac{1}{1-x} = \frac{x^4}{(1-x)^5}. \tag{8}$$

Hence, from (7) and (8) it immediately follows that

$$g_{b_n}(x) = x \cdot g_{a_n}(x)$$

We will show that the above relation of the two generating functions implies the desired identity $a_n = b_{n+1}$. We have

$$\begin{aligned} g_{b_n}(x) =& x \cdot g_{a_n}(x) = x \cdot (a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n \cdots) \\ =& a_0 x + a_1 x^2 + a_2 x^3 + \cdots + a_n x^{n+1} \cdots \end{aligned} \tag{9}$$

On the other hand, we also have (by the definition of generating functions)

$$g_{b_n}(x) = b_0 x + b_1 x + b_2 x^2 + \cdots + b_n x^n + b_{n+1} x^{n+1} + \cdots \tag{10}$$

Therefore, from (9) we have the coefficient of x^{n+1} in $g_{b_n}(x)$ is a_n and from (10) we have the coefficient of x^{n+1} in $g_{b_n}(x)$ is b_{n+1}. Hence, we must have

$$a_n = b_{n+1}$$

for all n. □

Note that a_n is the left shift of b_n. Indeed, if we shift every entry of b_n towards the left then the new n^{th} entry is the old $(n+1)^{th}$ entry which was shifted to the left. Hence, indeed, $a_n = b_{n+1}$ is the left shift of b_n.

The proof above shows that shifting the sequence to left amounts to multiplying the generating function by $\frac{1}{x}$[2] and similarly shifting the sequence to the right amounts to multiplying the generating function by x. These observations are very important since they can be used to solve recurrence relations which provide a relation between a sequence a_n and its shifted sequences a_{n+1} or a_{n-1}.

Generating functions and recurrence solutions

Let's first show how to using generating functions to solve the following general recurrent relation:

$$a_n = r \cdot a_{n-1} + f(n) \tag{11}$$

for all $n \geq 1$, where $r \in \mathbb{R}$ and $f(n)$ is a given function of n (that is a given sequence). We multiply (11) by x^n and we add over all $n = 1, 2, 3, ...$ to obtain

$$\sum_{n\geq 1} a_n x^n = r \cdot \sum_{n\geq 1} a_{n-1} x^n + \sum_{n\geq 1} f(n) x^n$$

Therefore,

$$\sum_{n\geq 0} a_n x^n - a_0 = rx \cdot \sum_{n-1\geq 0} a_{n-1} x^{n-1} + \sum_{n\geq 1} f(n) x^n$$

and so

$$\sum_{n\geq 0} a_n x^n - a_0 = rx \cdot \sum_{n\geq 0} a_n x^n + \sum_{n\geq 1} f(n) x^n$$

[2] Note that we assume that $a_{-1} = b_0 = 0$.

Hence

$$g_{a_n}(x) - a_0 = rx \cdot g_{a_n}(x) + \sum_{n \geq 1} f(n)x^n \tag{12}$$

We can easily solve (12) with respect to g_{a_n} to obtain

$$g_{a_n}(x) = \frac{a_0 + \sum_{n \geq 1} f(n)x^n}{1 - rx}. \tag{13}$$

Hence, (13) provides the generating function of a_n. Therefore, the sequence a_n can be computed. Note that we need to know a priori the value of a_0. The following result is very useful in computing the actual a_n from (13):

$$\frac{1}{1 - rx} = 1 + rx + r^2x^2 + \cdots r^n x^n + \cdots = g_{b_n = r^n}(x)$$

Concluding, a_n **is the convolution of the sequences** $b_n = r^n$ **and the sequence whose generating function is** $x \cdot g_{f(n)}(x) + a_0$ **that is of the sequence**

$$(a_0, f(1), f(2), f(3), ...)$$

Example 1.

Solve

$$a_n = 2a_{n-1} + 2^n$$

for all $n \geq 1$ with $a_0 = 1$.

Answer: We multiply both sides with x^n and we add over all $n \geq 1$ to obtain

$$\sum_{n \geq 1} a_n x^n = 2 \sum_{n \geq 1} a_{n-1} x^n + \sum_{n \geq 1} 2^n x^n$$

Hence

$$\sum_{n \geq 0} a_n x^n - a_0 = 2x \sum_{n \geq 1} a_{n-1} x^{n-1} + \sum_{n \geq 1} 2^n x^n$$

Therefore,

$$g_{a_n} - a_0 = 2x g_{a_n} + \sum_{n \geq 1} 2^n x^n$$

Hence we reached to

$$g_{a_n}(x) = \frac{a_0 + \sum_{n \geq 1} 2^n x^n}{1 - 2x} = \frac{1 + \sum_{n \geq 1} 2^n x^n}{1 - 2x} = \frac{1}{(1 - 2x)^2}$$

Hence, we obtain that (why?)

$$a_n = (n + 1) \cdot 2^n.$$

Example 2. (Fibonacci sequence)

Solve the Fibonacci relation

$$F_n = F_{n-1} + F_{n-2}$$

for all $n \geq 2$ with $F_0 = 0, F_1 = 1$.

Answer:

We multiply the Fibonacci relation with x^n and add over all $n \geq 2$ to obtain:

$$\sum_{n\geq 2} F_n x^n = \sum_{n\geq 2} F_{n-1} x^n + \sum_{n\geq 2} F_{n-2} x^n$$

Hence,

$$\sum_{n\geq 0} F_n x^n - F_0 - F_1 x = x \sum_{n\geq 2} F_{n-1} x^{n-1} + x^2 \sum_{n\geq 2} F_{n-2} x^{n-2}$$

and so

$$\sum_{n\geq 0} F_n x^n - F_0 - F_1 x = x \sum_{n-1\geq 1} F_{n-1} x^{n-1} + x^2 \sum_{n-2\geq 0} F_{n-2} x^{n-2}$$

which yields

$$\sum_{n\geq 0} F_n x^n - F_0 - F_1 x = x \sum_{k\geq 1} F_k x^k + x^2 \sum_{k\geq 0} F_k x^k$$

Thus

$$g_{F_n} - F_0 - F_1 x = x \cdot \Big(g_{F_n} - F_0 \Big) + x^2 g_{F_n}$$

Therefore,

$$g_{F_n} = \frac{F_0 + F_1 x - F_0 x}{1 - x - x^2} = \frac{x}{1 - x - x^2}.$$

It is easily verified that

$$g_{F_n} = \frac{A}{1 - ax} + \frac{B}{1 - bx}$$

where

$$A = \frac{1}{\sqrt{5}}, B = -\frac{1}{\sqrt{5}}, a = \frac{1+\sqrt{5}}{2}, b = \frac{1-\sqrt{5}}{2}.$$

This implies that (why?)

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

Partitions of natural numbers

We will conclude our discussion on generating functions by a very nice application of generating functions in Number Theory. In particular, we will consider the problem of counting all *partitions* of natural numbers.

A partition of a positive integer n is a collection of positive integers with sum equal to n. For example,

$$3 = 2 + 1$$

is a partition of 3. The precise order of the positive integers does not matter. Hence,

$$3 = 2 + 1$$

and

$$3 = 1 + 2$$

represent the same partition of 3. Hence, since order does not matter, it is useful to write the summands in increasing order. For example, we have the following partitions:

$$\begin{aligned}
1 &= 1 \\
2 &= 1+1 \\
2 &= 2 \\
3 &= 1+1+1 \\
3 &= 1+2 \\
3 &= 3 \\
4 &= 1+1+1+1 \\
4 &= 1+1+2 \\
4 &= 1+3 \\
4 &= 2+2 \\
4 &= 4
\end{aligned}$$

Hence, we see that 1 has 1 partition, 2 has 2 partitions, 3 has three partitions, 4 has 5 partitions.

In general we denote by p_n the number of different partitions of a natural number n. Any partition is completely determined by the number of 1's present, the of 2's present, the number of 3's present, ..., the number of n's present. For example for the following partition of 35

$$35 = 1+2+3+3+5+7+7+7$$

we have that

- 1 is present 1 time
- 2 is present 1 time
- 3 is present 2 times,
- 4 is present 0 times,
- 5 is present 1 time,
- 6 is present 0 times,
- 7 is present 3 times,
- 8 is present 0 times,
- all natural numbers greater of equal to 9 are present 0 times.

Therefore, we can rewrite this partition of 35 as follows

$$35 = 1\cdot 1 + 2\cdot 1 + 3\cdot 2 + 4\cdot 0 + 5\cdot 1 + 6\cdot 0 + 7\cdot 3 + 8\cdot 0 + ... + k\cdot 0 + ...$$

More generally, any partition of n corresponds to a solution $(x_1, x_2, ..., x_n)$ of the equation

$$n = x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 + 6x_6 + 7x_7 + 8x_8 + \cdots + ix_i + \cdots$$

where x_i denotes the number of i's present in the partition, and hence $x_i \geq 0$.

Problem 1.

Compute the generating function $g_{p_n}(x)$ of the number p_n of all partitions of n.

Answer: It suffices to compute the generating function of the number of all solutions $(x_1, x_2, ..., x_n)$ of the equation

$$n = x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 + 6x_6 + 7x_7 + 8x_8 + \cdots + ix_i + \cdots$$

Let x correspond to a unit. The variable x_i contributes ix_i units in the partition of n since x_i denotes the number of i's present in the partition of n. Hence, x_i contributes either 0 units, or i units, or $2i$ units, etc. Therefore, the associated powerseries to the variable x_i is

$$1 + x^i + (x^i)^2 + (x^i)^3 + ...$$

Hence the generating function of p_n is equal to the product of all the associated powerseries and hence equal to

$$g_{p_n}(x) = \big(1+x+x^2+\cdots\big)\cdot\big(1+x^2+(x^2)^2+\cdots\big)\cdot\big(1+x^3+(x^3)^2+\cdots\big)\cdot ...\cdot\big(1+x^i+(x^i)^2+\cdots\big)\cdots$$

The i^{th} factor provides all the possibilities for the contribution of the number i in the partitions of n. Clearly we have

$$g_{p_n}(x) = \frac{1}{1-x}\cdot\frac{1}{1-x^2}\cdot\frac{1}{1-x^3}\cdot\dots\cdot\frac{1}{1-x^i}\cdots$$

therefore

$$g_{p_n}(x) = \prod_{k=1}^{\infty}\frac{1}{1-x^k}.$$

Partitions with restrictions

The theory of generating functions provides neat solutions to partition problems with additional restrictions.

Problem 1.

Derive the generating function of the number of partitions of n which consist of the number of 1,2 and 3 only.

Answer: We need to find the number of all solutions (x_1, x_2, x_3) to the equation

$$x_1 + 2x_2 + 3x_3 = n$$

with $x_i \geq 0$. The powerseries associated to x_1 is

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

The powerseries associated to x_2 is

$$1 + x^2 + (x^2)^2 + \cdots = \frac{1}{1-x^2}.$$

The powerseries associated to x_3 is

$$1 + x^3 + (x^3)^2 + \cdots = \frac{1}{1 - x^3}.$$

Hence, the desired generating function is

$$\frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^3}.$$

Problem 2.

Derive the generating function of the number of partitions of n which consist of *distinct* numbers (that it, each number cannot appear more than once).

Answer: We need to find the number of all solutions $(x_1, x_2, x_3, ..., x_i, ...)$ to the equation

$$x_1 + 2x_2 + 3x_3 + ...ix_i + ... = n$$

with $0 \leq x_i \leq 1$ since each number can appear at most once. The powerseries associated to x_1 is

$$1 + x.$$

The powerseries associated to x_2 is

$$1 + x^2.$$

More generally, the powerseries associated to x_i is

$$1 + x^i.$$

Hence, the desired generating function is

$$(1 + x) \cdot (1 + x^2) \cdot (1 + x^3) \cdots = \prod_{k=1}^{\infty} (1 + x^k).$$

Problem 3.

Derive the generating function of the number of partitions of n which consist of *odd* numbers.

Answer: We need to find the number of all solutions $(x_1, x_3, x_5, ..., x_{2k+1}, ...)$ to the equation

$$x_1 + 2x_2 + 3x_3 + ... + (2k + 1)x_{2k+1} + ... = n$$

with $x_{2k+1} \geq 0$. The powerseries associated to x_1 is

$$1 + x + x^2 + \cdots = \frac{1}{1 - x}$$

The powerseries associated to x_3 is

$$1 + x^3 + (x^3)^2 \cdots = \frac{1}{1 - x^3}$$

More generally, the powerseries associated to x_{2k+1} is

$$1 + x^{2k+1} + (x^{2k+1})^2 + \cdots = \frac{1}{1 - x^{2k+1}}$$

Hence, the desired generating function is

$$\frac{1}{1-x}\cdot\frac{1}{1-x^3}\cdot\frac{1}{1-x^5}\cdots=\prod_{k=0}^{\infty}\frac{1}{(1-x^{2k+1})}.$$

Let's consider the partitions of 6 in problems 2 and 3.

The partitions of 6 with distinct numbers are the following

$$\begin{aligned}6&=1+5\\6&=2+4\\6&=1+2+3\\6&=6\end{aligned}$$

Hence, there are 4 partitions of 6 with distinct numbers.

Let's consider now the partitions of 6 with odd numbers. We have the following partitions

$$\begin{aligned}6&=1+1+1+1+1+1\\6&=1+1+1+3\\6&=3+3\\6&=1+5\end{aligned}$$

Hence, there are 4 partitions of 6 with odd numbers.

The fact that we have 4 partitions of 6 with distinct number and 4 partitions of 6 with odd number is not a coincidence. Once of the most important results in partition theory is due to Euler and states the following

Theorem 1 (Euler). *The number of partitions of any natural number n with distinct numbers is equal to the number of partitions of n with odd numbers.*

Proof. Let d_n denote the number of partitions of n with distinct numbers and let o_n partitions of n with odd numbers.

It suffices to prove that the generating functions g_{d_n}, g_{o_n} of the corresponding sequences are equal. By problems 2 and 3 above we have

$$g_{d_n}(x)=(1+x)\cdot(1+x^2)\cdot(1+x^3)\cdots=\prod_{k=1}^{\infty}(1+x^k)$$

and

$$g_{o_n}(x)=\frac{1}{1-x}\cdot\frac{1}{1-x^3}\cdot\frac{1}{1-x^5}\cdots=\prod_{k=0}^{\infty}\frac{1}{(1-x^{2k+1})}.$$

Using the identity

$$1+a=\frac{1-a^2}{1-a}$$

we obtain

$$\begin{aligned}g_{d_n}(x)=&(1+x)\cdot(1+x^2)\cdot(1+x^3)\cdot(1+x^4)\cdot(1+x^5)\cdots\\=&\frac{1-x^2}{1-x}\cdot\frac{1-x^4}{1-x^2}\cdot\frac{1-x^6}{1-x^3}\cdot\frac{1-x^8}{1-x^4}\cdot\frac{1-x^{10}}{1-x^5}\cdots\\=&\frac{1}{1-x}\cdot\frac{1}{1-x^3}\cdot\frac{1}{1-x^5}\cdots\\=&g_{o_n}(x),\end{aligned}$$

where we observed that the numerator and the denumerator contain all terms of the form $(1 - x^{2k})$ and hence all these terms cancel out.

Since the generating functions are equal we obtain $d_n = o_n$ for all n. $\square$