



Co-funded by the
Erasmus+ Programme
of the European Union



AGENCIJA ZA
MOBILNOST I
PROGRAME EU



Summer course: STEM Ambassadors Program

**Teaching mathematics in STEM context
for STEM students**

**Project Number:
2019-1-HR01-KA203-060804**



Co-funded by the
Erasmus+ Programme
of the European Union



AGENCIJA ZA
MOBILNOST I
PROGRAME EU



Modular arithmetic and applications to Cryptography
Summer course: STEM Ambassadors
Program

Monday, 09.05. 2022

09:30-11:00	What is a modular arithmetic? <ul style="list-style-type: none">• Clock arithmetic• Equivalence class
11:00 – 11:30	Refreshment break
11:30– 13:00	Operations on modular arithmetic <ul style="list-style-type: none">• Modular operations• Modular exponentiation
13:00– 15:00	Lunch break
15:00 – 16:30	Applications: Cryptography <ul style="list-style-type: none">• Caesar Cipher• Frequency analysis• Vigenere Cipher
16:30– 17:00	Refreshment break
17:00-18:00	Questions and Exit ticket (40 minutes test on the gained knowledge)
18:00	Dinner



Co-funded by the
Erasmus+ Programme
of the European Union



AGENCIJA ZA
MOBILNOST I
PROGRAME EU



Definite Integration – Applications to Business and Economy
Summer course: STEM Ambassadors
Programme

Tuesday, 10.05. 2022

09:30-11:00	Indefinite Integral <ul style="list-style-type: none">• Review – Derivatives• Antiderivatives and Area under a curve• The Definite Integral
11:00 – 11:30	Refreshment break
11:30– 13:00	The Fundamental Theorem of Calculus <ul style="list-style-type: none">• Integration rules• Example 1: Net excess profit – Wolfram Alpha / Wolfram Mathematica
13:00– 15:00	Lunch break
15:00 – 16:30	Interactive exercises in Wolfram Mathematica <ul style="list-style-type: none">• Curve plot, Area under a curve, Subtraction of areas• Examples: Net excess profit and Net earnings
16:30– 17:00	Refreshment break
17:00-18:00	Exit ticket (one hour test on the gained knowledge)
18:00	Dinner



Co-funded by the
Erasmus+ Programme
of the European Union



AGENCIJA ZA
MOBILNOST I
PROGRAME EU



Statistical reasoning in R
Summer course: STEM Ambassadors
Programme

Wednesday, 11.05. 2022

09:30-11:00	Descriptive statistics <ul style="list-style-type: none">• Visual data representation• Relative frequencies and numerical data values• Mean and dispersion measures
11:00 – 11:30	Refreshment break
11:30– 13:00	Theoretical models and simulations of random events <ul style="list-style-type: none">• Probabilities• Simulations in R
13:00– 15:00	Lunch break
15:00 – 16:30	Statistical reasoning <ul style="list-style-type: none">• Hypotheses formulation• Statistical tests
16:30– 17:00	Refreshment break
17:00-18:00	Exit ticket (one hour test on the gained knowledge)
18:00	Dinner



Co-funded by the
Erasmus+ Programme
of the European Union



AGENCIJA ZA
MOBILNOST I
PROGRAME EU



Vectors and Vectors application
Summer course: STEM Ambassadors
Programme

Thursday, 12.05. 2022

09:30-11:00	Vector algebra <ul style="list-style-type: none">• Vectors and scalars• Vectors in space• Operation with vectors
11:00 – 11:30	Refreshment break
11:30– 13:00	Products of two vectors <ul style="list-style-type: none">• The scalar product of two vectors• The vector product of two vectors
13:00– 15:00	Lunch break
15:00 – 16:30	Vector application <ul style="list-style-type: none">• Lines in space• Planes
16:30– 17:00	Refreshment break
17:00-18:00	Exit ticket (one hour test on the gained knowledge)
18:00	Dinner



Co-funded by the
Erasmus+ Programme
of the European Union



AGENCIJA ZA
MOBILNOST I
PROGRAME EU



Programme

Friday, 13.05. 2022

09:30-11:00	Introduction to Orthogonal axonometry method
11:00 – 11:30	Refreshment break
11:30– 13:00	Orthogonal axonometry of various objects
13:00– 15:00	Lunch break
15:00 – 16:30	Dynamic Programming
16:30– 17:00	Refreshment break
17:00-18:00	Exit ticket (one hour test on the gained knowledge)
18:00	Dinner

MODULAR ARITHMETIC AND CRYPTOGRAPHY

CONTENTS

1. Introduction.....	3
2. Modular Arithmetic	
2.1. Integers, divisors and common divisors.....	6
2.2. Clock arithmetic.....	7
2.3. Congruent modula.....	10
3. Operations on modular Arithmetic	
3.1. Addition, subtraction and multiplication.....	14
3.2. Modular exponentiation.....	15
4. Applications of modular arithmetic: Cryptography	
4.1. General principles of Cryptography.....	18
4.2. Caesar ciphers.....	22
4.3. Cryptanalysis.....	27
4.4. Vigenere Cipher.....	29
5. More questions	31

Chapter 1

Introduction

Have you ever wondered why you can place your credit card number on Amazon's web page to pay online and no eavesdropper could exploit it for his Christmas shopping?

Do you want to know how the British cracked the fantastic ENIGMA machine of the Germans in World War II ?



How can we keep our secrets on a computer for ourselves although people may try very hard to find out?

Since ancient times, people desiring to transmit messages privately have devised methods of encoding messages, so that no person but the intended recipient could read the message. The ability to successfully encode and decode messages has played a central role in the development of financial markets and in history- altering military

turnarounds. We use cryptography to refer to the study of how information can be made secretive enough so that bad people can't read it. Cryptography is a very exciting and developing area of contemporary mathematics, with connections to number theory.

Let us consider a person Alice who would like to send a secret message to another person Bob. Perhaps Alice and Bob are childhood friends and are planning a surprise birthday party for a mutual friend. Or perhaps Alice and Bob have never met, but Alice would like to send Bob her credit card information so she can pay for something Bob is selling. In both cases, Alice and Bob would like to guarantee several things:

- (a) Alice would like to ascertain that Bob has received her message;
- (b) Both Alice and Bob would like to know that no one else has seen the secret message;
- (c) Bob would like to ascertain that the message he believes to have come from Alice has indeed come from Alice. It is not immediately clear how we can guarantee each of these except in the case where Alice and Bob actually meet up and Alice whispers the message into Bob's ear. What should they do, however, if they are far apart?

If they send a message through the postal service, there is a small chance that an eavesdropper might intercept the message before it reaches Bob. Even if they use the telephone, or an email, or a text, there is a chance that the intended message and information will make its way to the wrong hands. These kinds of questions motivate the need to develop methods of encoding and decoding information so that messages can be communicated securely.

Simple ways of encoding messages were known since antiquity. Sometimes letters were switched for other letters, or for numbers, and so an eavesdropper quickly looking at an encoded message would only see gibberish. However, this approach has many limitations. For starters, how would Alice communicate to Bob the scheme which she used to encode the message and which he, consequently, will need to decode it? If he can determine this by himself, perhaps through some guesswork, then what would stop someone else from doing the same? Many somewhat sophisticated methods have been developed over the centuries for encoding and decoding secret messages, though in Section 4 we will focus on one that is built on what is called modular arithmetic, a system of arithmetic that in some sense only has a finite number of numbers.

Chapter 2

Modular Arithmetic

Every reader is familiar with arithmetic from the time they are three or four years old. It is the study of numbers and various ways in which we can combine them, such as through addition and subtraction, multiplication and division. Since even before they were in grade school, every reader knew that adding 2 and 2 together gives us 4, and can make that calculation now without almost any thinking.

The reader is also likely familiar with another kind of arithmetic, even if we don't always think of it as such. If it is 4 o'clock now, what will the time be in 25 hours? If we didn't know from watches and clocks, we would probably have answered 29 o'clock. But we are familiar with watches, clocks, and the standard conventions of time-keeping, and so every reader would probably have answered the answer with 5 o'clock. How can we add 25 to 4 and end up with 5? The reason is that in this system 25 o'clock is the same as 1 o'clock, 26 is the same as 2, and so forth. In many time-keeping systems, we don't even use numbers larger than 12, and instead use a.m. and p.m. to denote the earlier and latter halves of a 24-hour period. Such systems, that "wrap around" after hitting some limit, are called modular arithmetic systems, and play an important role both in theoretical and applied mathematics.

Modular arithmetic motivates many questions that don't arise when studying classic arithmetic. For example, in classic arithmetic, adding a positive number a to another number b always produces a number larger than b . In modular arithmetic this is not always so.

For example, if it is now 4 o'clock and we "add" 23 hours, the time will then be 3 o'clock, which doesn't appear to be larger than 4 o'clock. In fact, it is no longer clear whether it makes sense at all to discuss "larger" and "smaller" in such systems.

This particular example should motivate us think, if even momentarily, about modular arithmetic systems and the ways in which they are similar to and different from the classical arithmetic with which we are familiar. The next several sections will investigate these systems which have a finite number of numbers, and in which numbers "wrap around" after going too high.

2.1. The Integers, Divisors, Common Divisors

Recall that the set of integers Z consists of the natural numbers N (1, 2, 3, 4, . . .), along with their negatives $(-1, -2, -3, -4, \dots)$ and zero (0). The set of integers comes attached with a number of natural binary operations: addition $+$, subtraction $-$, and multiplication \cdot , which should surely be very familiar.

It is not always possible to divide one integer by another and obtain an integer result: there is no integer n such that $1/2 = n$, for example. Rather than trying to define the division operation we will instead focus on the idea of "divisibility".

Division. If $a \neq 0$, we say that a divides b , written $a \mid b$, if there is an integer k with $b = ka$.

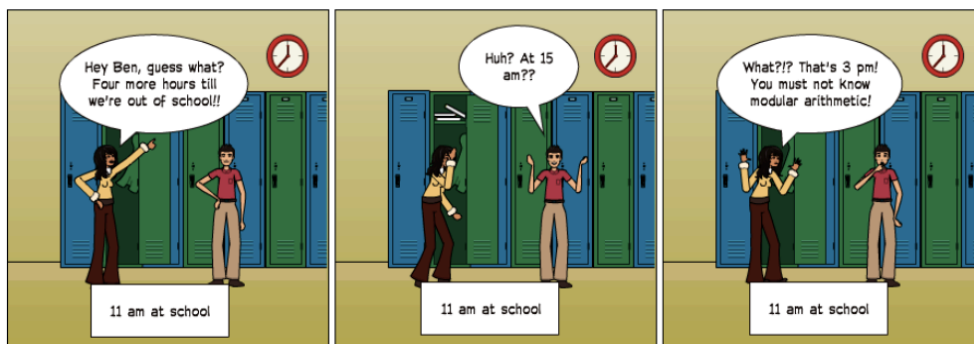
Example 2.1.1. $2 \mid 4$, $(-7) \mid 7$, and $6 \nmid 0$.

Proposition 2.1.2. There are a number of basic properties of divisibility that follow immediately from the definition and properties of arithmetic:

1. If $a \mid b$, then $a \mid bc$ for any c .
2. If $a \mid b$ and $b \mid c$, then $a \mid c$.
3. If $a \mid b$ and $a \mid c$, then $a \mid (xb+yc)$ for any x and y .
4. If $a \mid b$ and $b \mid a$, then $a = \pm b$.
5. If $a \mid b$, and $a, b > 0$, then $a \leq b$.
6. For any $m \neq 0$, $a \mid b$ is equivalent to $(ma) \mid (mb)$.

Proposition 2.1.3. (Quotient With Remainder): If a and b are positive integers, then there exist unique integers q and r such that $a = qb + r$ with $0 \leq r < b$. Furthermore, $r = 0$ if and only if $b \mid a$.

2.2. Clock Arithmetic

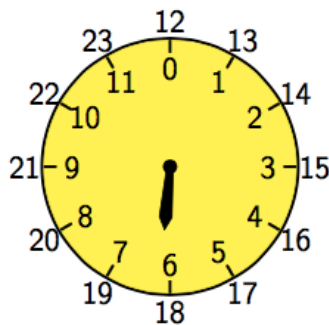


Warm-up question: If my birthday was on a Tuesday last year, and this year is not a leap year, what day of the week will my birthday be on this year?

How can we begin to understand modular arithmetic? One way is to think about hours on a clock. Imagine a clock face with numbers from zero to eleven, one hand to show the hour (there is no minute hand). We can add numbers on the clock by moving the hand in a clockwise direction. We can also subtract by moving the hand in anti-clockwise direction as follows.

The basic principles of modular arithmetic are often discussed at primary school level, except that it is likely to be called clock arithmetic. The idea is quite simple: 3 hours after 10 o'clock is 1 o'clock (because $10 + 3 = 13$ and we subtract 12). Similarly, 17 hours after 9 o'clock will be 2 o'clock, because $17 + 9 = 26$ and we must subtract $2 \times 12 = 24$ to find where the hand is pointing on the clock face. And 9 hours before 4 o'clock is 7 o'clock, since $4 - 9 = -5$, and now we add 12 to get the answer 7, the required number on the clock face. Clock arithmetic thus deals only with the numbers from 1 to 12 and whenever a calculation takes you outside that range, you add or subtract a multiple of 12 to get back onto the clock face.

Clock Arithmetic or a Circle as a Number Line One way to turn a circle into a number line is to divide it into twelve equal parts. In this case, one step is usually called one hour.



Notice that 0 coincides with 12, and as the hour hand moves to the right, 1 coincides with 0 coincides with 12. The hour hand moves from 0 to 1, from 13, 2 with 14, and so on. The hour hand rotates clockwise which corresponds with numbers 1 to 2, ... from 11 to 12 just as it would have on the straight increasing when moving to the right on a number line. However, 12 is equivalent to 0 on this number line. However, 12 equals 0 on this circle, so there it goes circle, which can be written as follows:

$$12 \equiv 0 \pmod{12}.$$

This can be read as 12 is congruent to 0 modulo 12. The usual “=” sign is reserved for the straight number line; we use “ \equiv ” on the circle instead. The symbol “mod 12” tells us that the circle is divided into 12 equal parts, so that 12 coincides with 0, 13 with 1, etc. In the new notation we have:

$$12 \equiv 0 \pmod{12}, 13 \equiv 1 \pmod{12}, \dots 23 \equiv 11 \pmod{12}$$

Example 2.2.1. The hour hand of a clock is pointing to the “10”. After 1000 hours have gone by, where is the hour hand pointing to?



The hour hand goes back to “10” every 12 hours (12 hour cycles). So cast out the multiples of 12 hours by using mod12 arithmetic.

$$1000 \bmod 12 = 4 \text{ hours remainder}$$

Answer: The hour hand is pointing to 4 hours past its starting point, or 4 hours past the “10.” In other words, the hour hand is pointing to the “2.”

When we divide two integers we will have an equation that looks like the following:
 $A/B = Q \text{ remainder } R$

A is the dividend

B is the divisor

Q is the quotient

R is the remainder

Sometimes, we are only interested in what the **remainder** is when we divide A by B. For these cases there is an operator called the modulo operator (abbreviated as mod). Using the same A, B, Q, and R as above, we would have: $A \bmod B = R$. We would say this as A modulo B is equal to R. Where B is referred to as the **modulus**.

Example 2.2.2. $13/5 = 2 \text{ remainder } 3$ that means $13 \bmod 5 = 3$

Now we observe what happens when we increment numbers by one and then divide them by 3.

$0/3 = 0 \text{ remainder } 0$
 $1/3 = 0 \text{ remainder } 1$
 $2/3 = 0 \text{ remainder } 2$
 $3/3 = 1 \text{ remainder } 0$
 $4/3 = 1 \text{ remainder } 1$
 $5/3 = 2 \text{ remainder } 2$
 $6/3 = 2 \text{ remainder } 0$

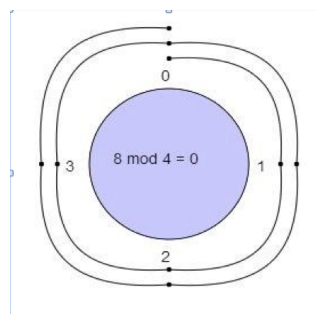
The remainders start at 0 and increases by 1 each time, until the number reaches one less than the number we are dividing by. After that, the sequence repeats.

By noticing this, we can visualize the modulo operator by using circles.

Example 2.2.3. $8 \bmod 4 = ?$

With a modulus of 4 we make a clock with numbers 0, 1, 2, 3.

We start at 0 and go through 8 numbers in a clockwise sequence 1, 2, 3, 0, 1, 2, 3, 0.

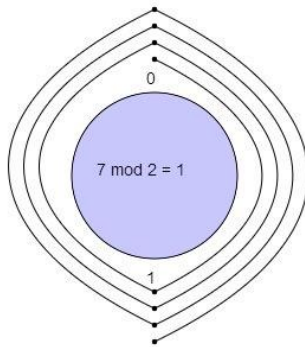


We ended up at 0 so $8 \bmod 4 = 0$

Example 2.2.4. $7 \bmod 2 = ?$

With a modulus of 2 we make a clock with numbers 0, 1.

We start at 0 and go through 7 numbers in a clockwise sequence 1, 0, 1, 0, 1, 0, 1.

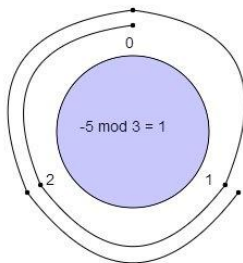


We ended up at **1** so $7 \bmod 2 = 1$

Example 2.2.5. $-5 \bmod 3 = ?$

With a modulus of 3 we make a clock with numbers 0, 1, 2.

We start at 0 and go through 5 numbers in counter-clockwise sequence (5 is **negative**)
2, 1, 0, 2, 1.



We ended up at **1** so $-5 \bmod 3 = 1$

2.3. Congruence modula

In regular arithmetic, if two numbers a and b have the same value then we write

$$a = b$$

If two numbers have the same remainder when divided by n then we write

$$a \equiv b \pmod{n}$$

and say “ a is congruent to b (when modding by n)” or we say: “ a and b are the same (in mod n arithmetic)”.

For example, $65 \bmod 7 = 2$

$$16 \bmod 7 = 2$$

$$\text{so } 65 \equiv 16 \pmod{7}$$

And we say “65 is congruent to 16 (when modding by 7)”.

In other words, “65 is congruent to 16 (in mod7 arithmetic)”. Another way of thinking of this is that even though 65 and 16 are not the same in regular arithmetic, they are the same in mod 7 arithmetic.

Definition 2.3.1 If m is a positive integer and m divides $b - a$, we say that a and b are congruent modulo m (or equivalent modulo m), and write “ $a \equiv b \pmod{m}$ ”.

Although this definition looks somewhat technical, the idea is very simple. For some fixed integer m , two numbers are roughly the same if they differ by multiples of m .

- Observe that if $m \mid (b - a)$, then $(-m) \mid (b - a)$ as well, so we do not lose anything by assuming that the modulus m is positive.
- In general, the statement $a \equiv b \pmod{m}$ can be thought of as saying “ a and b are equal, up to a multiple of m ”.
- Notation: As shorthand we often also write “ $a \equiv b \pmod{m}$ ”, or even just “ $a \equiv b$ ” (when the modulus m is clear from the context).

Example 2.3.2. $3 \equiv 9 \pmod{6}$, since 6 divides $9 - 3 = 6$.

Example 2.3.3. $-2 \equiv 28 \pmod{5}$, since 5 divides $28 - (-2) = 30$.

Example 2.3.4. $0 \equiv -666 \pmod{3}$, since 3 divides $-666 - 0 = -666$.

If m does not divide $b - a$, we say a and b are not congruent mod m , and write $a \not\equiv b \pmod{m}$. For example: $2 \not\equiv 7 \pmod{3}$, because 3 does not divide $7 - 2 = 5$.

- Modular congruences behave quite similarly to equalities:

1. For any a , $a \equiv a \pmod{m}$.
2. For any a and b , $a \equiv b \pmod{m}$ if and only if $b \equiv a \pmod{m}$.
3. For any a , b and c if $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$ then $a \equiv c \pmod{m}$.
4. For any a , b , c and d if $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$ then $a + c \equiv b + d \pmod{m}$.

m).

Example 2.3.5. Find the possible values of m that satisfy each congruence:

$$(a) \ 13 \equiv 3 \pmod{m} \qquad (b) \ 15 \equiv 4 \pmod{m}$$

Solution: (a) By definition of congruence, $13 - 3 = 10$ must be divisible by m . So the possible values of m are the positive divisors of 10 : $m \in \{2, 5, 10\}$.

Remember by definition $m \neq 1$

(b) $15 - 4 = 11$ must be divisible by m . So m can only be 11.

Example 2.3.6. If today is Wednesday, what day of the week will it be in 100 days' time?

Solution:

Sun	Mon	Tue	Wed	Thurs	Frđ	Sat
0	1	2	3	4	5	6

Since we have numbers 0, 1, 2, 3, 4, 5 and 6, we will do our arithmetic in modulo 7. Today is Wednesday (day 3) hence to find 100 days later, we can write

$$3 + 100 = 103 \equiv 5 \pmod{7}$$

So the answer is the 5th day which is Friday.

MOD TIPS

1. When you calculate $a \bmod n$, the only possibilities for a result (remainder) are 0 to $n - 1$.

- (a) $14 \bmod 3$ would be either 0, 1 or 2.
- (b) $3 \bmod 9$ would be 0, 1, 2, 3, 4, 5, 6, 7 or 8.
- (c) $10523 \bmod 115$ would be only one number from 0, 1, 2, 3, ..., 114.

2. If a is less than n , and a is NOT negative, then $a \bmod n = a$

- $2 \bmod 18 = 2$
- $23 \bmod 24 = 23$
- $114 \bmod 1000 = 114$
- $0 \bmod 15 = 0$

3. If a is a multiple of n then $a \bmod n = 0$.

- $63 \bmod 9 = 0$

- $48 \bmod 12 = 0$
- $60000 \bmod 6 = 0$
- $992 \bmod 8 = 0$
- $-77 \bmod 7 = 0$

4. Since $0 \bmod n = 0$ for any positive integer n , then 0 and any multiple of n are congruent. In other words $0 \equiv a \pmod{n}$ where a is any multiple of n . This means that 0 is the same thing as any multiple of n in $\bmod n$ arithmetic. Therefore, 0 can be replaced by any multiple of n in $\bmod n$ arithmetic. We can use this fact to our advantage when dealing with negative numbers in modular arithmetic: just add on a positive multiple of n that is “bigger” than your negative number.

$$\begin{aligned} \text{(a) } -12 \bmod 5 \\ &\equiv 15 - 12 \\ &\equiv 3 \pmod{5} \end{aligned}$$

$$\begin{aligned} \text{(b) } -7 \bmod 13 \\ &\equiv 13 - 7 \\ &\equiv 6 \pmod{13} \end{aligned}$$

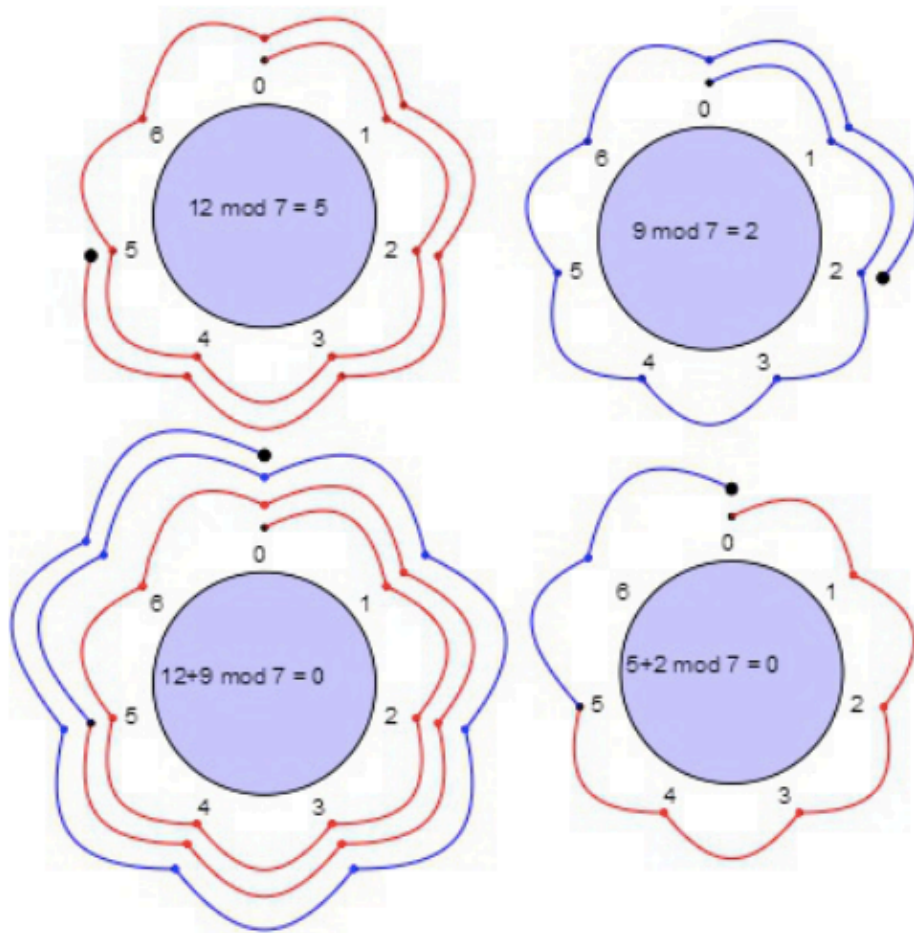
$$\begin{aligned} \text{(c) } -8765 \bmod 12 \\ &\equiv 12000 - 8765 \\ &\equiv 3235 \\ &\equiv 7 \pmod{12} \end{aligned}$$

Chapter 3

Operations on Modular Arithmetic

After considering the basic definition of modular arithmetic, we next consider some of its basic properties. It turns out that modular arithmetic follows many of the same rules of classical arithmetic, thus making it very easy to work with. In order to highlight what is going on, we try to compare and contrast modular arithmetic to classical arithmetic.

Intuition Behind Modular Addition: Observe the figure below. If we want to calculate $12+9 \bmod 7$ we can easily go around the modular circle for a sequence of $12+9$ steps clockwise (as shown in the bottom left circle).



We can take a shortcut by observing that every 7 steps we end up in the same position on the modular circle. These complete loops around the modular circle don't contribute to our final position. We ignore these complete loops around the circle by calculating each number mod 7 (as shown in the two upper modular circles). This will give us the number of clockwise steps, relative to 0, that contributed to each of their final positions around the modular circle.

Now, we only have to go around the circle clockwise the total of the number of steps that contributed to each of numbers final position (as shown in the bottom right modular circle). This method applies, in general, to any two integers and any modular circle.

3.1. Addition, subtraction and multiplication

Suppose we have the following two congruence relations:

$$a \equiv b \pmod{m}$$

$$c \equiv d \pmod{m}.$$

Are we able to combine these to obtain

$$a+c \equiv b+d \pmod{m},$$

$$a-c \equiv b-d \pmod{m},$$

$$a \times c \equiv b \times d \pmod{m}$$

Modular arithmetic obeys the same laws as conventional arithmetic, such as the commutative laws $a + b \equiv b + a$ and $ab \equiv ba$, the associative laws $(a + b) + c \equiv a + (b + c)$ and $(ab)c \equiv a(bc)$ and the distributive laws (laws of brackets) $a(b + c) \equiv ab + ac$ and $(a + b)c \equiv ac + bc$, all of these modulo n .

Example 3.1.1. How can we simplify 20×21 in arithmetic modulo 19? We first note that $20 \equiv 1 \pmod{19}$ and also that $21 \equiv 2 \pmod{19}$. We can combine these equations to obtain $20 \times 21 \equiv 1 \times 2 \equiv 2 \pmod{19}$.

Example 3.1.2. Can we simplify 17^{753} in arithmetic modulo 9? We first note that $17 \equiv -1 \pmod{9}$, because 17 and -1 differ by a multiple of 9. Combine this congruence relation as many times as we would like. In particular, by combining 753 copies, we obtain $17^{753} \equiv (-1)^{753} \pmod{9}$. Since $(-1)^n = -1$ for any odd integer n , we have $17^{753} \equiv -1 \pmod{9}$. Finally, if we would like to have a simple, positive answer, then we can add 9 to obtain a final answer of 8.

We have by now seen that in arithmetic modulo m , there is no difference between writing 1, $1 + m$, $1 + 2m$, and so forth, at least as far as addition, subtraction, and multiplication are concerned. For this reason, writing $4+11 \equiv 15 \pmod{13}$ is “just as correct” as writing $4 + 11 \equiv 2 \pmod{13}$, and “just as correct” as writing $4 + 11 \equiv 11 \pmod{13}$. As far as arithmetic modulo 13 is concerned, 2, 15, and -11 are exactly the same number. However, in some applications it is convenient to agree upon a standard way to represent numbers.

Example 3.1.3. Suppose we want to know the remainder of 17×18 when it is divided by 19. We can do this in two different ways. First, we can multiply the two numbers directly and obtain 306; some calculation will show that 306 is congruent to 2 modulo 19. Alternatively, we know that $17 \equiv 2 \pmod{19}$ and $18 \equiv 1 \pmod{19}$. Multiplying both sides we see that $17 \times 18 \equiv (2) \times (1) \equiv 2 \pmod{19}$.

3.2. Modular Exponentiation

Most technological applications of modular arithmetic involve exponentials with very large numbers. For example, a typical problem related to encryption might involve solving one of the following two equations:

$$67930^{32319} \equiv a \pmod{103969}$$

$$67930^b \equiv 48560 \pmod{103969}.$$

It turns out that $a = 6582$ and $b = 32320$ solve these equations, but those answers are not obvious at all from looking at the equations. More importantly, it is not even clear how we would go about determining a and b . In this section we will look at some problems involving modular exponentiation and some techniques we can use to solve such problems.

Now let us consider the remainders of 10, 100, 1000, and so forth when we divide them by 3. The first thing we notice is that the remainder of 10 after dividing it by 3 is 1. In the language of modular arithmetic we can write:

$$10^1 \equiv 1 \pmod{3}$$

If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then

$a \times c \equiv b \times d \pmod{m}$. In our particular case, we know that

$$10^1 \times 10^1 \equiv 1 \times 1 \pmod{3}, 10^2 \equiv 1 \pmod{3}.$$

We can then use the same technique, through induction, to show that all integer powers of 10 are congruent to 1 mod 3, since we can continue multiplying our resulting equation by the initial equation $10^1 \equiv 1 \pmod{3}$. In other words, all positive integer powers of 10, when divided by 3, give us a remainder of 1!

Example 3.2.1. Consider the very large number $7^{1383921}$ and how we might determine its remainder after dividing it by 4. Of course we know that the only possible remainder are 0, 1, 2, and 3, but it is not clear how to determine which of those it is. Simple calculations show the following pattern:

$$7^1 \equiv 3 \pmod{4},$$

$$7^2 \equiv 1 \pmod{4},$$

$$7^3 \equiv 3 \pmod{4},$$

$$7^4 \equiv 1 \pmod{4}, \dots$$

It seems that if n is odd, then $7^n \equiv 3 \pmod{4}$, and if n is even, then $7^n \equiv 1 \pmod{4}$. We can prove that this pattern will repeat as n increases by noticing that $7^2 \equiv 1 \pmod{4}$.

Thus $7^n \equiv 3 \pmod{4}$ then $7^{n+2} \equiv 3 \pmod{4}$, and likewise if $7^n \equiv 1 \pmod{4}$ then

$7^{n+2} \equiv 1 \pmod{4}$. Therefore, the pattern repeats with a period of 2. Determining the

remainder of $7^{1383921}$ when dividing by 4 is then straightforward, since the exponent $n = 1383921$ is odd, the remainder must be 3.

Example 3.2.2. Suppose we want to determine the standard form of 17^2 in mod 19 arithmetic. One way in which we can do this is by considering the square of 17, which is 289, divide that by 19 and then take the remainder. However, since we know that $17 \equiv 2 \pmod{19}$, we can multiply this congruence equation by itself to obtain $17^2 \equiv 2^2 \equiv 4 \pmod{19}$. We can easily verify that the remainder of 289, when divided by 19, is indeed 4.

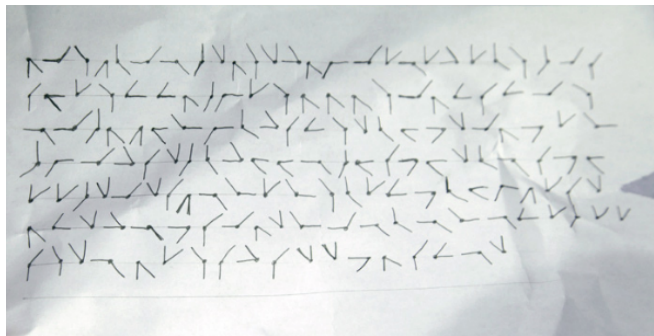
Example 3.2.3. Suppose we want to determine the standard form of $18^{489391312}$ in mod 19 arithmetic. We should first notice that in mod 19 arithmetic, 18 is congruent to -1, and so $18^{489391312} \equiv (-1)^{489391312} \pmod{19}$. It is relatively easy to see that if n is odd then $(-1)^n = -1$, and if n is even then $(-1)^n = 1$. Since 489391312 is even, $18^{489391312} \equiv 1 \pmod{19}$.

Chapter 4

Applications of modular arithmetic: Cryptography

4.1 General principles of Cryptography

- Cryptography is the name given to encoding and transmitting information in a way that makes it difficult for someone else to intercept and use.
 - Many of the earliest uses of cryptography were to send secure military information and orders that could not be decoded by enemy forces.
 - In the modern setting, secure cryptography is at the heart of internet commerce: for example, it allows merchants and credit card companies to exchange purchasing information without anyone else being able to eavesdrop.



- In analysis of cryptography, it is useful to have a standard list of placeholder names:
 - Alice and Bob refer to two parties attempting to exchange information. (Generally, Alice wants to send a message to Bob, though the communication can be two-directional.)
 - Eve refers to a non-malicious eavesdropper, who can listen in to the communications between Alice and Bob, but will not alter them.
 - Mallory refers to a malicious eavesdropper, who can listen to Alice and Bob's communications and may also attempt to impersonate them or alter their messages.
- In general, a cryptographic system works as follows:
 - Alice wishes to send a secure message to Bob.
 - Alice takes her unencrypted message, her plaintext, and encrypts it

- somehow to obtain a ciphertext.
 - Alice then sends the ciphertext to Bob, who then decodes it to recover Alice's original message.
- We will generally write plaintexts in **bold lowercase** and ciphertexts in **BOLD UPPERCASE**.
 - Note: For the ease of readability, when it is reasonable we will include spaces (because it is hard to read a lengthy text with no spaces) when rendering plaintexts and ciphertexts, but when we encode messages we will not use the spaces.
- Historically, most cryptography relied on making the message appear nonsensical and unreadable, or by hiding it in some other more innocuous location (e.g., by encoding the message in the first letter of each word in a document).
 - This latter procedure is sometimes called *steganography*, the hiding of secret information in plain sight. It is also interesting, but is not really the purpose of cryptography.
 - One of the most classical cryptosystems is the Caesar shift algorithm (so named because it was used by Julius Caesar): simply shift each letter of the plaintext forward a fixed number of letters in the alphabet (wrapping around from Z to A, as needed).
- We will also mention a few different types of attacks on cryptosystems:
 - Ciphertext-only attack: Eve only has a copy of the ciphertext and wants to decode it.
 - Known-plaintext attack: Eve has a copy of the ciphertext and the associated plaintext. In this case Eve's goal is to break the encryption system so she can read future ciphertexts that are encoded using the same system.
 - Chosen-plaintext attack: Eve is able to choose a plaintext and see how it encodes to a ciphertext. (For example, if the encryption algorithm is implemented as software on a computer, Eve would have access to the part of the program that encodes messages.) Again, Eve's goal is to try to break the encryption algorithm so she can read future ciphertexts.
 - Chosen-ciphertext attack: Eve is able to choose a ciphertext and see how it decodes to a plaintext. (For example, if the encryption algorithm is implemented as software on a computer, Eve would have access to the part of the program that decodes messages.)

In the following (see <https://www.khanacademy.org/computing/computer-science/cryptography/crypt/v/caesar-cipher>) there is an excellent video on ciphers.

What are the eras of cryptography? Cryptography has been through numerous phases of evolution. Early ciphers in cryptography were designed to allow encryption and decryption to take place by hand, while those which are developed and used today are only possible due to the high computational performance of modern machines (i.e the computer you are using right now). The major eras which have shaped cryptography are listed below.

Classical

The classical algorithms are those invented pre-computer up until around the 1950's. The list below is roughly ordered by complexity, least complex at the top.

- Atbash Cipher
- ROT13 Cipher
- Caesar Cipher
- Affine Cipher
- Rail-fence Cipher
- Baconian Cipher
- Polybius Square Cipher
- Simple Substitution Cipher
- Codes and Nomenclators Cipher
- Autokey Cipher
- Beaufort Cipher
- Porta Cipher
- Vigenere Cipher
- Playfair Cipher
- ADFGVX Cipher

Mechanical

Mechanical Ciphers are those that were developed around the second World War, which rely on sophisticated gearing mechanisms to encipher text.

- Enigma Cipher
- Lorenz Cipher
- Jefferson disk

Cryptographic machines: Before the advent of the modern computer, machines existed that simplified the use of encryption and made more complex encryption schemes feasible. Initially, such devices were simple mechanical machines, but as technology progressed, we began to see the inclusion of electronics and considerably more complex systems. The Jefferson Disk, invented by Thomas Jefferson in 1795, is a purely mechanical cryptographic machine. It is composed of a series of disks, each marked with the letters a to z around its edge, as shown in the following Figure. On each disk, the letters are arranged in a different order; each disk is also marked with a

unique designator to facilitate arranging them in a particular order. The device built by Jefferson contained 36 disks, with each disk representing one character in the message.



Jefferson disks

Enigma Cipher

The Enigma cipher was a field cipher used by the Germans during World War II. The Enigma is one of the better known historical encryption machines, and it actually refers to a range of similar cipher machines. The first Enigma machine was invented by a German engineer named Arthur Scherbius at the end of the first world war. It was used commercially from the early 1920s on, and was also adopted by the military and governmental services of a number of nations — most famously by Nazi Germany before and during World War II. A variety of different models of Enigma were produced, but the German military model, the Wehrmacht Enigma, is the version most commonly discussed.



Modern

Modern algorithms are those that are used in current technology e.g. block ciphers, public key cryptosystems etc. These algorithms are very secure (otherwise they would not be used), but in many cases we can practice on weakened versions of the algorithms.

4.2 Caesar Cipher

The Caesar cipher is a classic example of ancient cryptography and is said to have been used by Julius Caesar. The Caesar cipher is based on transposition and involves shifting each letter of the plaintext message by a certain number of letters, historically three, as shown in the following figure. The ciphertext can be decrypted by applying the same number of shifts in the opposite direction. This type of encryption is known as a substitution cipher, due to the substitution of one letter for another in a consistent fashion.

S	E	C	R	E	T	M	E	S	S	A	G	E
V	II	F	U	H	W	P	II	V	V	D	J	II

A more recent variation of the Caesar cipher can be found in the ROT13 cipher. ROT13 uses the same mechanism as the Caesar cipher but moves each letter 13 places forward. The convenience of moving 13 places lies in the fact that applying another round of encryption with ROT13 also functions as decryption, as two rotations will return us to the original starting place in the alphabet. Utilities for performing ROT13 can be found in the basic set of tools that ship with many Linux and UNIX operating systems. There are a number of simple systems that are built around simple transposition.

Example 4.2.1. To pass an encrypted message from one person to another, it is first necessary that both parties have the 'key' for the cipher, so that the sender may encrypt it and the receiver may decrypt it. For the Caesar cipher, the key is the number of characters to shift the cipher alphabet.

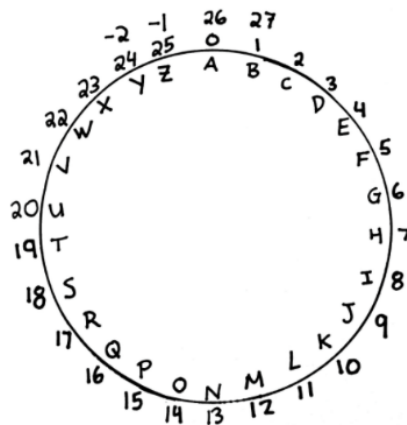
Here is a quick example of the encryption and decryption steps involved with the Caesar cipher. The text we will encrypt is 'defend the east wall of the castle', with a shift (key) of 1.

```
plaintext: defend the east wall of the castle
ciphertext: efgfoe uif fbtu xbmm pg uif dbtumf
```

It is easy to see how each character in the plaintext is shifted up the alphabet. Decryption is just as easy, by using an offset of -1.

```
plain:  abcdefghijklmnopqrstuvwxyz
cipher: bcdefghijklmnopqrstuvwxyza
```

How do Modular Arithmetic and Caesar Ciphers relate? Since there are 26 letters in the English alphabet, let's relate the letters a-z by numbers 0-25 as shown by the following figure .



Using the Caesar Cipher: Line up the wheels so that the “a” lines up with “D”. Notice going from “a” to “D” was a shift of 3 letters over. Thus we can encrypt the word “TOP SECRET ” by relating “t” with 19 on the wheel, adding 3 to get 22, and then we turn this back into a letter, which gives us “w”. Similarly

“o” → 14 → 17 → r.

Plaintext: top secret

“p” → 15 → 18 → s

Ciphertext: wrsvhfuhw

“s” → 18 → 21 → v.

“e” → 4 → 7 → h.

“c” → 2 → 5 → f.

“r” → 17 → 20 → u.

“e” → 4 → 7 → h

“t” → 19 → 22 → w.

- Mathematically we can describe the Caesar shift as follows:

- First, we choose a key k .
 - We encrypt the message by applying the function $f(x) = x + k \pmod{26}$ to the numbers making up the message.
 - To decrypt, we apply the inverse function $f^{-1}(x) = x - k \pmod{26}$ to the encrypted message.
- We can generalize the shift cipher by using a more complicated encoding function. Let us instead consider the class of linear encoding functions of the form $f(x) = ax + b \pmod{26}$ for some choices of a and b . These functions are affine functions, so the associated cipher is called an affine cipher.
 - We can make a table for the encryption of each letter under the affine cipher to save time when encoding long messages.

For example, here is the encoding table for the affine cipher $f(x) = 7x + 3 \pmod{26}$:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
3	10	17	24	5	12	19	0	7	14	21	2	9	16	23	4	11	18	25	6	13	20	1	8	15	22
d	k	r	y	f	m	t	a	h	o	v	c	j	q	x	e	l	s	z	g	n	u	b	i	p	w

The encoding of the plaintext secret message is **ZFRSFGJFZZDT**

Example 4.2.1 We agree with our friend to use the Shift Cipher with key $k=19$ for our message.

We encrypt the message "**KHAN**", as follows:

ENCRYPTION

K	H	A	N
10	7	0	13
+	19	19	19
<hr/>			
(29	26	19
	32)	mod 26
	3	0	19
	6		
<hr/>			
D	A	T	G

So, after applying the Shift Cipher with key $K=19$ our message text "**KHAN**" gave us **cipher text "DATG"**.

We give the message "DATG" to our friend.

How to decrypt: For every letter in the cipher text **C** :

1. Convert the letter into the number that matches its order in the alphabet starting from 0, and call this number **Y**.

(A=0, B=1, C=2, ..., Y=24, Z=25)

2. Calculate: **X = (Y - K) mod 26**

3. Convert the number **X** into a letter that matches its order in the alphabet starting from 0.

(A=0, B=1, C=2, ..., Y=24, Z=25)

Our friend now decodes the message using our agreed upon **key K=19**. As follows:

DECRYPTION

	D	A	T	G
	3	0	19	6
-	19	19	19	19
<hr/>				
(-16	-19	0	-13
) mod 26				
	10	7	0	13
<hr/>				
	K	H	A	N

So, after decrypting the Shift Cipher with key **K=19** our friend deciphers the cipher text "**DATG**" into the message text "**KHAN**".

Why is the Shift Cipher insecure? A cipher should prevent an attacker, who has a copy of the cipher text but does not know the key, from discovering the contents of the message. Since **we only have 26 choices for the key**, someone can easily try all of the 26 keys, one by one, until they recover the message.

How to Crack the Caesar Cipher:

As we've discovered, there are only 25 different shifts we can use to encrypt a message with a Caesar cipher. Because of this, the Caesar cipher is considered to be a very weak type of cryptography. We call the act of testing all 25 options until finding the key, the method of brute force.

Our ciphertext is the following:

YMJHFJXFWHNUMJWNXTSJTKYMJJFWQNJXYPSTBSFSIXNRUQJXYHNUMJWX

To find out what the original was, we try decrypting it with each of the 25 possible keys, calculating the fitness for each trial decryption:

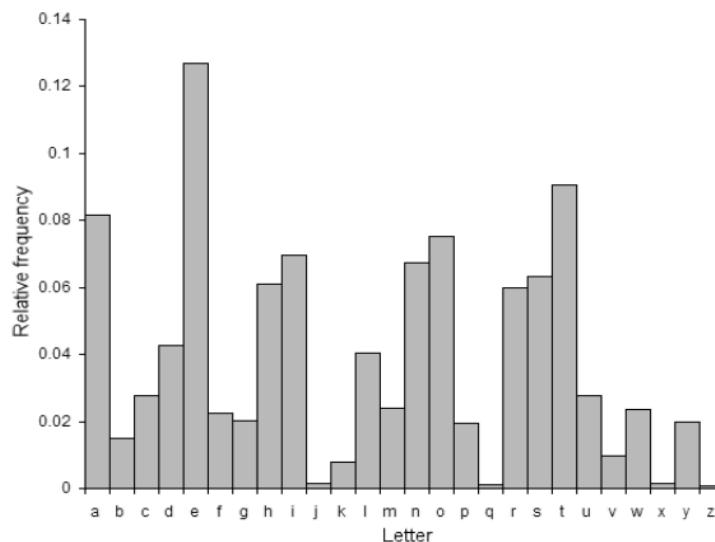
key	plaintext	fitness

1 :	XLIGEIWEVGMTLIVMWSRISJXLIIEVPM...	-442.22
2 :	WKHFDHVDUFLSKHULVRQHRIWKHHUOL...	-495.20
3 :	VJGECGUCTEKRJGTUQPGQHVGJGCTNK...	-484.13
4 :	UIFDBFTBSDJQIFSJTPOFFGUIFFBSMJ...	-490.73
5 :	THECAESARCIPHERISONEOFTHEEARLI...	-246.02
6 :	SGDBZDRZQBHOGDQHRNMDNESGDDZQKH...	-485.69
7 :	RFCAYCQYPAGNFCPGQMLCMDRFCCYPJG...	-481.17
8 :	QEBZXBPXOZFMEBOFPLKBLCQEBBXOIF...	-478.19
9 :	PDAYWAOWNYELDANEOKJAKBPDAAWNHE...	-415.66
10 :	OCZ XVZNV MXDKCZMDNJIZJAO CZZVMGD...	-488.75
11 :	NBYWUYMULWCJBYLCMIHYIZNBYYULFC...	-490.46
12 :	MAXVTXLTKVBIAXKBLHGXYHMAXXTKEB...	-490.82
13 :	LZWUSWKSJUAHZWJAKGFWGXLZWW SJDA...	-483.63
14 :	KYVTRVJRITZGYVIZJFEVFWKYVVRICZ...	-475.01
15 :	JXUSQUIQHSYFXUHYIEDUEVJXUQHBX...	-466.90
16 :	IWTRPHTHPRXEW TGXHDCTDUIWTPGAX...	-458.49
17 :	HVSQOSGOFQWDV SFWGCBSC THVSSOFZW...	-474.67
18 :	GURPNRFNEPVCUREVFBARBSGURNEYV...	-460.86
19 :	FTQOMQEMDOUBTQDUEAZQARFTQQMDXU...	-467.13
20 :	ESPNLPDLCNTAS PCTDZYPZQESPPLCWT...	-454.29
21 :	DROMKOCKBMSZROBSCYXOYPDROOKBVS...	-461.91
22 :	CQNLJNBALRYQNARBXWNXOCQNNJAUR...	-479.58
23 :	BPMKIMAI ZKQXPMZQAWVMWNBPMI ZTQ...	-473.52
24 :	AOLJHLZHYJPWOLYPZVULVMAOLLHYSP...	-474.57
25 :	ZNKIGKYGXIOVNKKOYUTKULZNKKGXRO...	-494.13

Cryptanalysis

Cryptanalysis is the art of breaking codes and ciphers. The Caesar cipher is probably the easiest of all ciphers to break. Since the shift has to be a number between 1 and 25, (0 or 26 would result in an unchanged plaintext) we can simply try each possibility and see which one results in a piece of readable text. If you happen to know what a piece of the ciphertext is, or you can guess a piece, then this will allow you to immediately find the key.

If this is not possible, a more systematic approach is to calculate the frequency distribution of the letters in the cipher text. This consists of counting how many times each letter appears. Natural English text has a very distinct distribution that can be used help crack codes. This distribution is as follows:



English Letter Frequencies

This means that the letter **e** is the most common, and appears almost 13% of the time, whereas **z** appears far less than 1 percent of time. Application of the Caesar cipher does not change these letter frequencies, it merely shifts them along a bit (for a shift of 1, the most frequent ciphertext letter becomes **f**). A cryptanalyst just has to find the shift that causes the ciphertext frequencies to match up closely with the natural English frequencies, then decrypt the text using that shift. This method can be used to easily break Caesar ciphers by hand.

4.3. Vigenere Cipher

A major weakness of the Caesar cipher is that there are not many ways to encrypt a message. Also long messages encrypted with the Caesar cipher are easily cracked using “frequency analysis”. A stronger cipher is the Vigenere cipher. Here’s how it works!

- Here is the procedure for the Vigenère cipher:
- First, we choose a keyword, which (numerically) is a vector of some length n .
- We then break the message into letter blocks of length n , and then encrypt each block of letters by adding the keyword vector to it. We then put all of the blocks together in the appropriate order.
- To decrypt, we simply do the inverse: break the ciphertext into blocks of length n and subtract the keyword vector.

Example 4.3.1: Encode the message *twentysix* using the Vigenère cipher with keyword *one*.

Here is a table of the encryption procedure:

Plaintext	t	w	e	n	t	y	s	i	x
#	19	22	4	13	19	24	18	8	23
Key letter	o	n	e	o	n	e	o	n	e
Key #	14	13	4	14	13	4	14	13	4
Encoding	7	9	8	1	6	2	6	21	1
Ciphertext	h	j	i	b	g	c	g	v	b

Thus we obtain the ciphertext . **HJIBGCGVB**

Note that the letter t appears twice in the plaintext, but is represented in the ciphertext by two different characters: the first time by H and the second by G. Inversely, the letter G appears twice in the ciphertext, but represents different letters from the plaintext.

Example 4.3.2: Encode the message I love math using the Vigenère cipher with keyword “car”.

c	a	r	c	a	r	c	a	r	
+2	0	+17	+2	0	+17	+2	0	+17	
I	l	o	v	e	m	a	t	h	
8	11	14	21	4	12	0	19	7	
10	11	31=5	23	4	29=3	2	19	24	
K	L	F	X	E	D	C	T	Y	

Thus we obtain the ciphertext : **KLFXEDCTY**

Now by using the similar system we can decode Vigenere ciphers

Example 4.3.3 : Suppose we arranged our secret codeword to be “dog” and I sent you the secret message below. Try to decode it.

Secret Message: ZVE GCKV BUECJB HGOY ZR AK?

D	o	g	d	o	g	d	o	g	d	o	g	d	o	g	d	o	g	d	o	g
-3	-14	-6	-3	-14	-6	-3	-14	-6	-3	-14	-6	-3	-14	-6	-3	-14	-6	-3	-14	-6
Z	V	E	G	C	K	V	B	U	E	C	J	B	H	G	O	Y	Z	R	A	K
25	21	4	6	2	10	21	1	20	4	2	9	1	7	6	14	24	25	17	0	10
22	7	24	3	14	4	18	13	14	1	14	3	24	19	0	11	10	19	14	12	4
W	H	Y	D	O	E	S	N	O	B	O	D	Y	T	A	L	K	T	O	M	E

MORE QUESTIONES

1... Suppose that Alice and Bob are trying to send secret letters to each other in the mail. In order for the letters to stay a secret, they want to think of a way to send the messages in a “secret code” so that anybody who tries to intercept the message wouldn’t be able to read it even if they managed to intercept it. One way they can do this is using a Caesar Cipher. In a Caesar cipher, the alphabet is shifted a certain number of places and each letter is replaced by the corresponding letter.

For example, say Alice and Bob agree that they want to shift the letters by three: TO ENCRYPT

(1) Using the cipher key, they would first convert the letters in their message to their corresponding numbers to get a numerical message:

Letter Message	M	E	E	T		M	E		A	T		T	H	E		Z	O	O
Numerical Message	12																	

(2) Then they would then shift all the numbers in their message up by three:

Unshifted Numerical Message	12																	
Shifted Numerical Message	15																	

(a) Can we have numbers greater than 25 in our shifted numerical message? Why or why not?

(b) What do we do if a numbers in the numerical message is greater than 25?

(3). Then, using the cipher key, they would convert the resulting numerical message into a letter message.

Shifted Numerical Message	15																	
Encrypted Letter Message	P																	

TO DECRYPT

(1) Using the cipher key, they would first convert the letters in their message to their corresponding numbers to get a numerical message:

Encrypted Letter Message	R	N		V	H	H		B	R	X		W	K	H	Q
Numerical Message	17														

(2) Then they would then shift all the numbers in their message down by three:

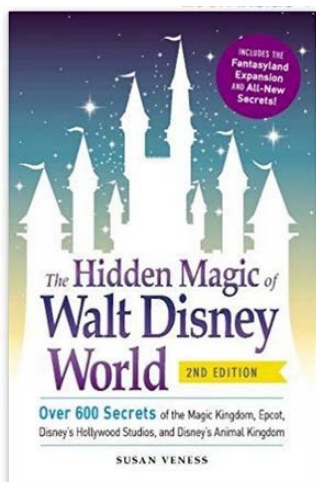
Unshifted Numerical Message	17														
Shifted Numerical Message	14														

(3) Then, using the cipher key, they would convert the resulting numerical message into a letter message.

Shifted Numerical Message	14														
Decrypted Letter Message	O														

2.... You win the lottery and decide to take a 1000-day trip. You leave on a Monday. On what day of the week do you return?

3..... *The Hidden Magic of Walt Disney World* (2nd edition) has over 600 secrets of the Magic Kingdom, Epcot, Disney's Hollywood Studios, and Disney's Animal Kingdom. Can you find the secret digit # that is missing from its ISBN14#0587809



The following is the 10-Digit ISBN Check Digit Formula. If the number is $a_9a_8a_7a_6a_5a_4a_3a_2a_1a_0$ where a_0 is the check digit, then the formula is

$$a_0 = (a_9 + 2a_8 + 3a_7 + 4a_6 + 5a_5 + 6a_4 + 7a_3 + 8a_2 + 9a_1) \bmod 11$$

4.....

The *Little Debbie Fudge Rounds (Big Pack)* - 12 ct. (pack of 2) has a UPC with an illegible digit #. If the main part of the UPC is 70#53820676 with a check digit of 8, then what must the digit # be?



The following is the UPC number Check Digit Formula. If the number is $a_{11}a_{10}a_9a_8a_7a_6a_5a_4a_3a_2a_1a_0$ where a_0 is the check digit, then the formula is

$$a_0 = -(3a_{11} + a_{10} + 3a_9 + a_8 + 3a_7 + a_6 + 3a_5 + a_4 + 3a_3 + a_2 + 3a_1) \bmod 10$$

5..... Find the remainder when 19^{53} is divided by 8.

6..... Which day of the week will it be

- (a) 56 days after a Saturday?
- (b) 264 days after a Friday?
- (c) 312 days after a Wednesday?

7..... A man left ANKARA for IZMİR at 6.00 am and had a flat tire on the way. The journey took 16 hours. When did he arrive in Ibadan?

8..... Find all possible values of m if

(a) $20 \equiv 8 \pmod{m}$

(b) $-7 \equiv -2 \pmod{m}$

9..... Find the remainder when each expression is divided by 3.

(a) 8^{92}

(b) $14^{25} + 274^{12}$

10..... Solve the following equations

(a) $x + 3 \equiv 4 \pmod{8}$

(b) $2x \equiv 1 \pmod{7}$

(c) $3x + 2 \equiv 4 \pmod{8}$

(d) $2x + 5 \equiv 3 \pmod{4}$

(e) $2x^2 + 3x \equiv 14 \pmod{5}$

SEEU MathSTEM teaching methodology, Ohrid, May 2022

Lesson plan: **Definite Integration: Applications to Business and Economics**

Lecturers: Vladimir Radevski, Murat Sadiku, Halil Snopce

1. Lesson plan: DESCRIPTION OF THE LESSON

Study program: Undergraduate studies in Computer Sciences

Subject: Calculus / Business Mathematics

Topic: Definite Integration: Applications to Business and Economics

Duration: 2h + 1h (Lecture and practical auditorial + Practical Wolfram Mathematics Examples)

Description: Introduction to Definite Integration – Area as the Limit of a Sum
The Definite Integral – Area Under a Curve – The Fundamental Theorem of Calculus
Examples paper and pencil
Wolfram Mathematics for Definite Integration – Interactive experimentation
Example 1: Increasing of the *total manufacturing cost* by change of *the level of production* based on *marginal cost function*.
Example 2: *Rate of profitability comparison* for two *input investments*.

Goals:

1. To introduce the students to the concept of Definite Integration and the calculation of the area under a curve of a function.
2. To show and examine application of the Definite Integration to Business and Economics.
3. To demonstrate the usage of Wolfram Mathematics in examination of a real-world problem solving by organized interactive activity with its graphical presentation potential. To introduce the concept of mathSTEM methodology in teaching mathematics for STEM students.

Objectives:

1. Students will understand the concept of Definite Integration and its application for calculating the Area under a Curve.
2. Students will be introduced the Fundamental Theorem of Calculus supported by examples of application.
3. Students will be introduced to the mathSTEM methodology developed for this topic by introducing the experimentation environment in Wolfram Mathematics.
4. Students will experiment interactively with the mathSTEM application for Definite Integration on the Wolfram Mathematics platform.

Materials: A detailed lesson plan for the lecture

A power point presentation supporting the formal definitions, theorem statements and exercises to be solved during the lecture.

Wolfram mathematics application running on presenter's computer and allowing interactions visible to the class.

Summary worksheet with exercises for further practice.

Moodle entry in the mathSTEM platform for further reference.

Assessment: The students' will be assessed at the end of the class. A feedback form will be offered to the students at the end of the class on mathSTEM methodology.

2. Lesson plan: LESSON CONTENT

Definite Integration – Applications to Business and Economy

PART ONE: AREA AS THE LIMIT OF A SUM, AREA UNDER A CURVE

1. Review of Indefinite integral:

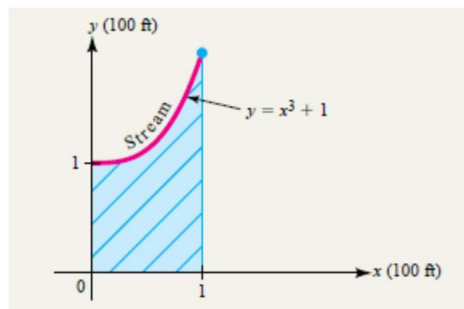
$$\int f(x)dx = F(x) + C \qquad F'(x) = f(x)$$

Diagram illustrating the components of the indefinite integral equation $\int 3x^2 dx = x^3 + C$:

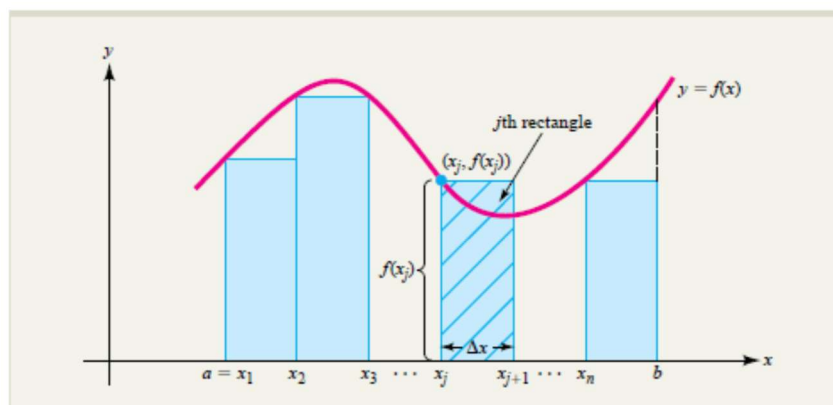
- integrand:** $3x^2$
- constant of integration:** C
- integral symbol:** \int
- variable of integration:** x

2. Antiderivatives and Area under a curve:

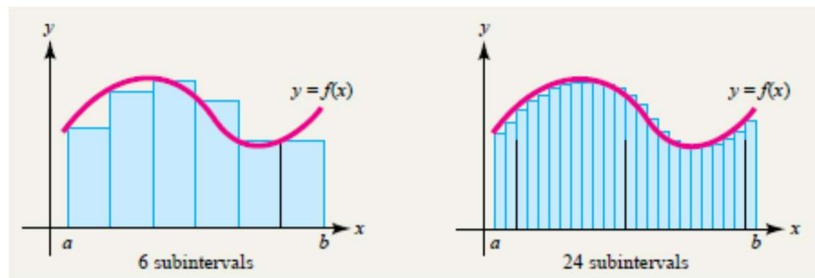
A surprising connection between antiderivatives and sums of areas under a given curve:



Determining a land value by finding the area under the curve.



An approximation of area under a curve by rectangles.



The approximation improves as number of rectangles increase.

Area Under a Curve ■ Let $f(x)$ be continuous and satisfy $f(x) \geq 0$ on the interval $a \leq x \leq b$. Then the region under the curve $y = f(x)$ over the interval $a \leq x \leq b$ has area

$$A = \lim_{n \rightarrow +\infty} [f(x_1) + f(x_2) + \cdots + f(x_n)] \Delta x$$

where x_j is the left endpoint of the j th subinterval if the interval $a \leq x \leq b$ is divided into n equal parts, each of length $\Delta x = \frac{b-a}{n}$.

3. The Definite integral:

The Definite Integral ■ Let $f(x)$ be a function that is continuous on the interval $a \leq x \leq b$. Subdivide the interval $a \leq x \leq b$ into n equal parts, each of width $\Delta x = \frac{b-a}{n}$, and choose a number x_k from the k th subinterval for $k = 1, 2, \dots, n$. Form the sum

$$[f(x_1) + f(x_2) + \cdots + f(x_n)] \Delta x$$

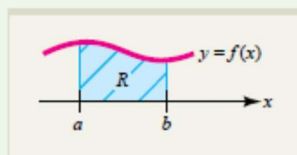
called a **Riemann sum**.

Then the **definite integral** of f on the interval $a \leq x \leq b$, denoted by $\int_a^b f(x) dx$, is the limit of the Riemann sum as $n \rightarrow +\infty$; that is,

$$\int_a^b f(x) dx = \lim_{n \rightarrow +\infty} [f(x_1) + f(x_2) + \cdots + f(x_n)] \Delta x$$

The function $f(x)$ is called the **integrand**, and the numbers a and b are called the **lower and upper limits of integration**, respectively. The process of finding a definite integral is called **definite integration**.

Area as a Definite Integral ■ If $f(x)$ is continuous and $f(x) \geq 0$ on the interval $a \leq x \leq b$, then the region R under the curve $y = f(x)$ over the interval $a \leq x \leq b$ has area A given by the definite integral $A = \int_a^b f(x) dx$.



4. The Fundamental Theorem of Calculus:

The Fundamental Theorem of Calculus ■ If the function $f(x)$ is continuous on the interval $a \leq x \leq b$, then

$$\int_a^b f(x) dx = F(b) - F(a)$$

where $F(x)$ is any antiderivative of $f(x)$ on $a \leq x \leq b$.

$$F(x) \Big|_a^b = F(b) - F(a)$$

Thus,

$$\int_a^b f(x) dx = F(x) \Big|_a^b = F(b) - F(a)$$

NOTE You may wonder how the fundamental theorem of calculus can promise that if $F(x)$ is *any* antiderivative of $f(x)$, then

$$\int_a^b f(x) dx = F(b) - F(a)$$

To see why this is true, suppose $G(x)$ is another such antiderivative. Then $G(x) = F(x) + C$ for some constant C , so $F(x) = G(x) - C$ and

$$\begin{aligned} \int_a^b f(x) dx &= F(b) - F(a) \\ &= [G(b) - C] - [G(a) - C] \\ &= G(b) - G(a) \end{aligned}$$

since the C 's cancel. Thus, the valuation is the same regardless of which antiderivative is used. ■

Example:

Find the area of the parcel of land described in the introduction to this section; that is, the area under the curve $y = x^3 + 1$ over the interval $0 \leq x \leq 1$, where x and y are in hundreds of feet. If the land in the parcel is appraised at \$12 per square foot, what is the total value of the parcel?

Solution

The area of the parcel is given by the definite integral

$$A = \int_0^1 (x^3 + 1) dx$$

Since an antiderivative of $f(x) = x^3 + 1$ is $F(x) = \frac{1}{4}x^4 + x$, the fundamental theorem of calculus tells us that

$$\begin{aligned} A &= \int_0^1 (x^3 + 1) dx = \left. \frac{1}{4}x^4 + x \right|_0^1 \\ &= \left[\frac{1}{4}(1)^4 + 1 \right] - \left[\frac{1}{4}(0)^4 + 0 \right] = \frac{5}{4} \end{aligned}$$

Because x and y are measured in hundreds of feet, the total area is

$$\frac{5}{4} \times 100 \times 100 = 12,500 \text{ ft}^2$$

and since the land in the parcel is worth \$12 per square foot, the total value of the parcel is

$$V = (\$12/\text{ft}^2)(12,500 \text{ ft}^2) = \$150,000$$

5. Integration rules:

Rules for Definite Integrals

Let f and g be any functions continuous on $a \leq x \leq b$. Then,

1. **Constant multiple rule:** $\int_a^b k f(x) dx = k \int_a^b f(x) dx$ for constant k

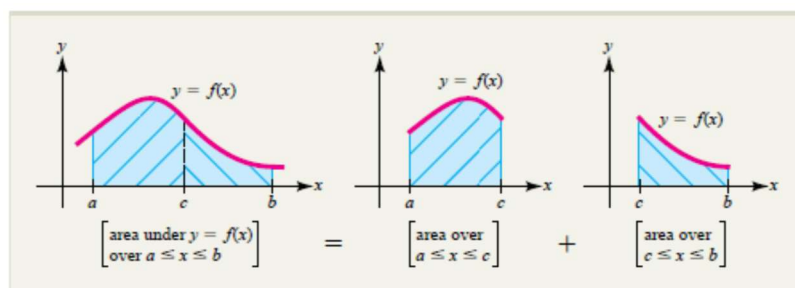
2. **Sum rule:** $\int_a^b [f(x) + g(x)] dx = \int_a^b f(x) dx + \int_a^b g(x) dx$

3. **Difference rule:** $\int_a^b [f(x) - g(x)] dx = \int_a^b f(x) dx - \int_a^b g(x) dx$

4. $\int_a^a f(x) dx = 0$

5. $\int_b^a f(x) dx = -\int_a^b f(x) dx$

6. **Subdivision rule:** $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$



The subdivision rule for definite integrals.

6. Interactive Exercises (Net Change):

Net Change ■ If $Q'(x)$ is continuous on the interval $a \leq x \leq b$, then the net change in $Q(x)$ as x varies from $x = a$ to $x = b$ is given by

$$Q(b) - Q(a) = \int_a^b Q'(x) dx$$

Example:

At a certain factory, the marginal cost is $3(q - 4)^2$ dollars per unit when the level of production is q units. By how much will the total manufacturing cost increase if the level of production is raised from 6 units to 10 units?

Solution

Let $C(q)$ denote the total cost of producing q units. Then the marginal cost is the derivative $\frac{dC}{dq} = 3(q - 4)^2$, and the increase in cost if production is raised from 6 units to 10 units is given by the definite integral

$$\begin{aligned} C(10) - C(6) &= \int_6^{10} \frac{dC}{dq} dq \\ &= \int_6^{10} 3(q - 4)^2 dq = (q - 4)^3 \Big|_6^{10} \\ &= (10 - 4)^3 - (6 - 4)^3 \\ &= \$208 \end{aligned}$$

PART TWO: EXAMPLES IN BUSINESS AND ECONOMY USING WOLFRAM MATHEMATICS

EXAMPLE 1: Manufacturing cost increase.

Total manufacturing cost increase: Given Marginal cost find by how much will the total manufacturing cost increase for a certain raise of level of production.

At a certain factory, the marginal cost is $3(q - 4)^2$ dollars per unit when the level of production is q units.

By how much will the total manufacturing cost increase if the level of production is raised from 6 units to 10 units?

SOLUTION:

(1) $C(q)$ total cost of producing q units.

(2) $\frac{dC}{dq} = 3(q - 4)^2$ the derivative specifying the marginal cost

The increase in cost if production is raised from 6 units to 10 units is given by the definite integral:

$$(3) \quad C(10) - C(6) = \int_6^{10} 3(q - 4)^2 dq = (q - 4)^3 \Big|_6^{10} = \$208.$$

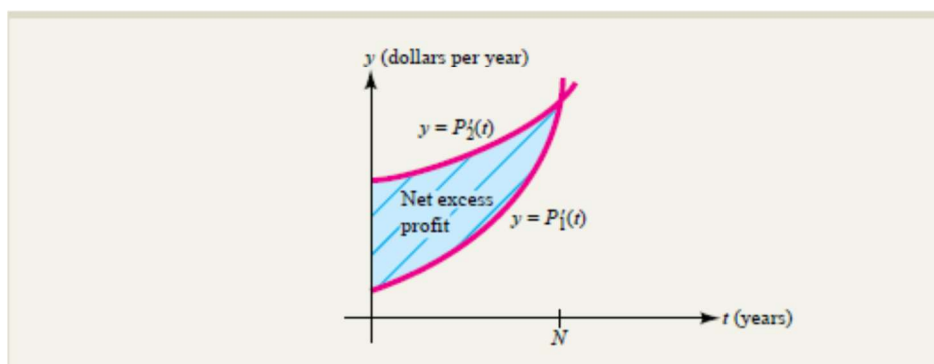
End of solution.

EXAMPLE 1: Net excess profit.

Net Excess Profit The area between curves can sometimes be used as a way of measuring the amount of a quantity that has been accumulated during a particular procedure. For instance, suppose that t years from now, two investment plans will be generating profit $P_1(t)$ and $P_2(t)$, respectively, and that their respective rates of profitability, $P_1'(t)$ and $P_2'(t)$, are expected to satisfy $P_2'(t) \geq P_1'(t)$ for the next N years; that is, over the time interval $0 \leq t \leq N$. Then $E(t) = P_2(t) - P_1(t)$ represents the **excess profit** of plan 2 over plan 1 at time t , and the **net excess profit** $NE = E(N) - E(0)$ over the time interval $0 \leq t \leq N$ is given by the definite integral

$$\begin{aligned} NE &= E(N) - E(0) = \int_0^N E'(t) \, dt \\ &= \int_0^N [P_2'(t) - P_1'(t)] \, dt \end{aligned} \quad \begin{array}{l} \text{since } E'(t) = [P_2(t) - P_1(t)]' \\ \quad = P_2'(t) - P_1'(t) \end{array}$$

This integral can be interpreted geometrically as the area between the rate of profitability curves $y = P_1'(t)$ and $y = P_2'(t)$ as shown in Figure 5.13. Example 5.4.3 illustrates the computation of net excess profit.



Net excess profit as the area between rate of profitability curves.

Suppose that t years from now, one investment will be generating profit at the rate of $P'_1(t) = 50 + t^2$ hundred dollars per year, while a second investment will be generating profit at the rate of $P'_2(t) = 200 + 5t$ hundred dollars per year.

- For how many years does the rate of profitability of the second investment exceed that of the first?
- Compute the net excess profit for the time period determined in part (a). Interpret the net excess profit as an area.

Solution

- The rate of profitability of the second investment exceeds that of the first until

$$\begin{aligned}
 P'_1(t) &= P'_2(t) \\
 50 + t^2 &= 200 + 5t \\
 t^2 - 5t - 150 &= 0 && \text{subtract } 200 + 5t \text{ from both sides} \\
 (t - 15)(t + 10) &= 0 && \text{factor} \\
 t = 15, -10 && \text{since } uv = 0 \text{ if and only if } u = 0 \text{ or } v = 0 \\
 t = 15 \text{ years} && \text{reject the negative time } t = -10
 \end{aligned}$$

- The excess profit of plan 2 over plan 1 is $E(t) = P_2(t) - P_1(t)$, and the net excess profit NE over the time period $0 \leq t \leq 15$ determined in part (a) is given by the definite integral

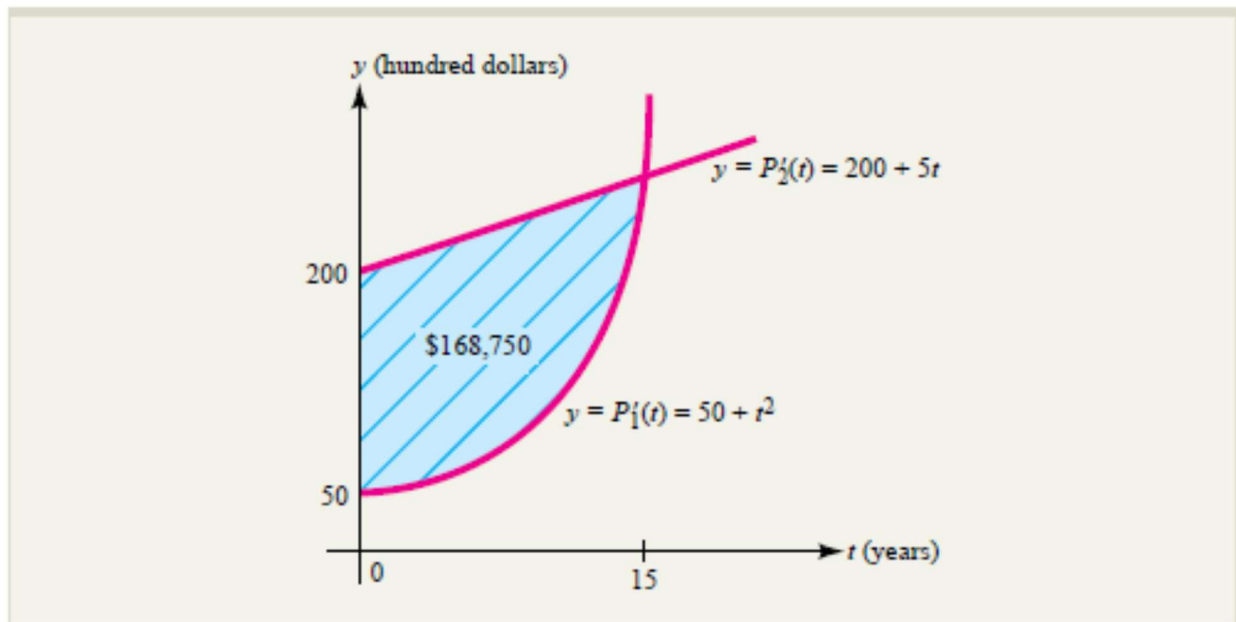
$$\begin{aligned}
 \text{NE} &= E(15) - E(0) = \int_0^{15} E'(t) \, dt && \text{fundamental theorem of calculus} \\
 &= \int_0^{15} [P'_2(t) - P'_1(t)] \, dt && \text{since } E(t) = P_2(t) - P_1(t) \\
 &= \int_0^{15} [(200 + 5t) - (50 + t^2)] \, dt \\
 &= \int_0^{15} [150 + 5t - t^2] \, dt && \text{combine terms} \\
 &= \left[150t + 5\left(\frac{1}{2}t^2\right) - \left(\frac{1}{3}t^3\right) \right]_0^{15} \\
 &= \left[150(15) + \frac{5}{2}(15)^2 - \frac{1}{3}(15)^3 \right] - \left[150(0) + \frac{5}{2}(0)^2 - \frac{1}{3}(0)^3 \right] \\
 &= 1,687.50 \text{ hundred dollars}
 \end{aligned}$$

Thus, the net excess profit is \$168,750.

The graphs of the rate of profitability functions $P'_1(t)$ and $P'_2(t)$ are shown in Figure 5.14. The net excess profit

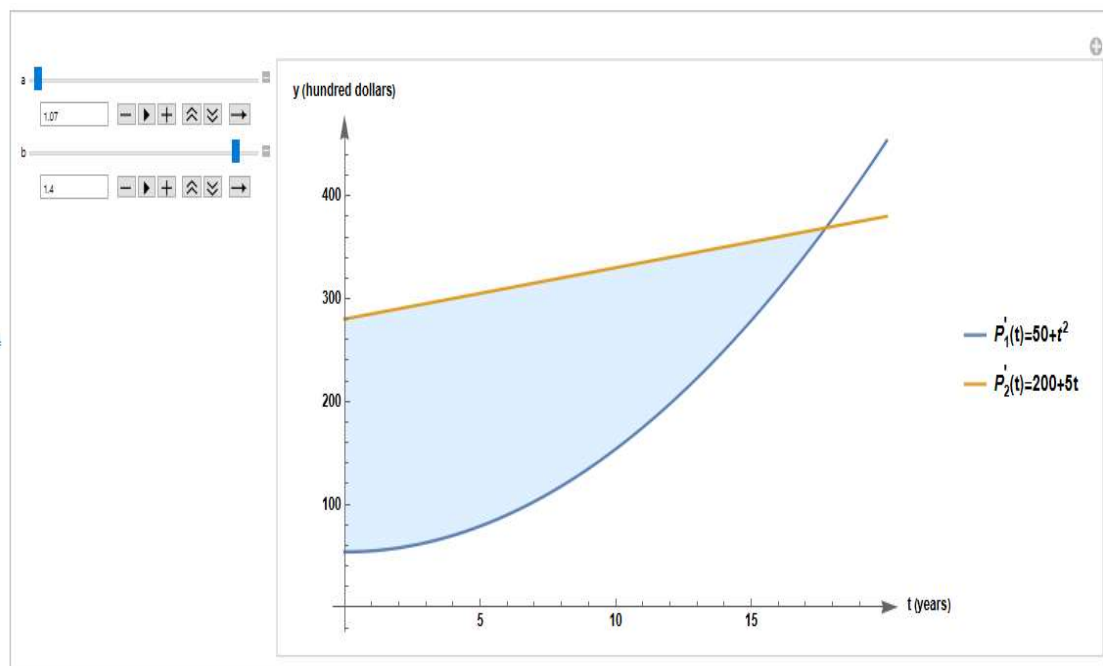
$$\text{NE} = \int_0^{15} [P'_2(t) - P'_1(t)] \, dt$$

can be interpreted as the area of the (shaded) region between the rate of profitability curves over the interval $0 \leq t \leq 15$.



Net excess profit for one investment plan over another.

Interactive activity with Wolfram Mathematics - Example 1



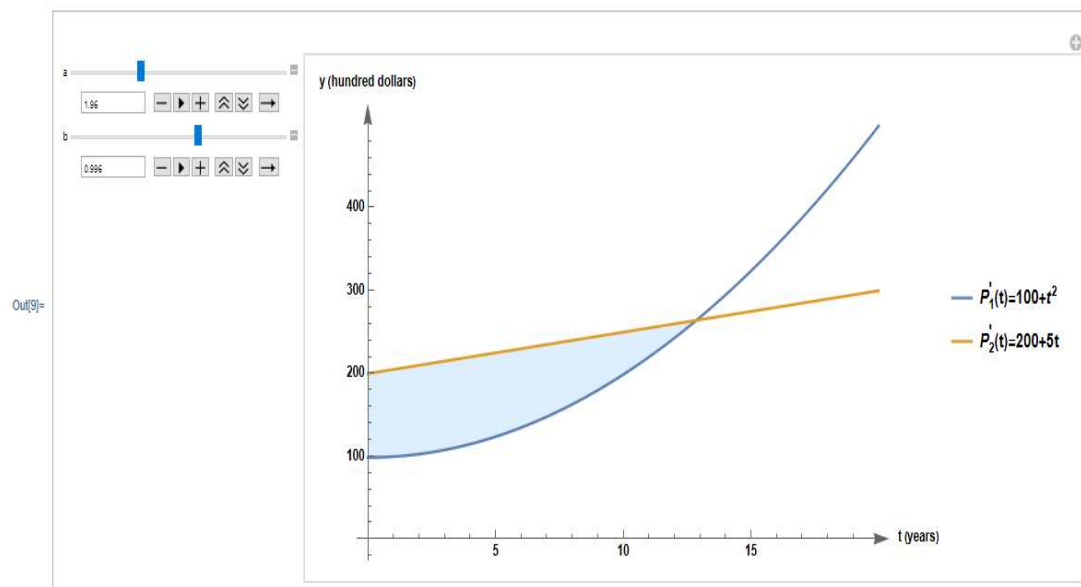
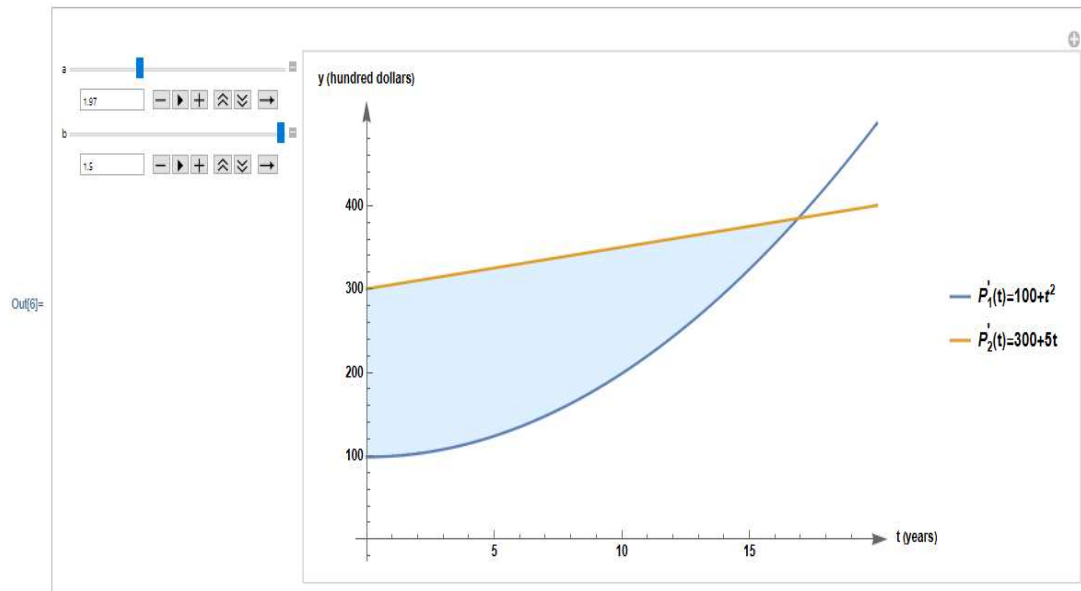
Activity 1: Use the provided Wolfram Mathematics application. Plot the profit generator functions $P'_1(t) = 50 + t^2$ and $P'_2(t) = 200 + 5t$

1. Note the one function is linear and another a parabola.
2. Consider the domain $x \geq 0$. Discuss the ranges of the functions.
3. Discuss the expectations for intersections.

4. Discuss the area between the curves, the y-axis and the line $x = x_0$, where x_0 is the intersection.
5. Conclude on resulting area being the subtraction of two areas.
6. Involve definite integral in calculations.

Activity 2: Use the provided application in Wolfram Mathematics. Examine the change of the area between the curves according to the respective changes in the functions.

1. Try several linear functions with the same slope.
2. Try several parabolas with different y-axis intercepts.
3. Conclude of the resulting areas between the two curves.
4. Examine few more features of the powerful WM package.



EXAMPLE 2: Net earnings from an industrial machine.

Suppose that when it is t years old, a particular industrial machine generates revenue at the rate $R'(t) = 5000 - 20t^2$ dollars per year and that the operating and servicing costs related to the machine accumulate at rate $C'(t) = 2000 + 10t^2$ dollars per year.

(A) How many years pass before the profitability of the machine begins to decline?

(B) Compute the net earnings generated by the machine over the time period determined by (A).

Solution:

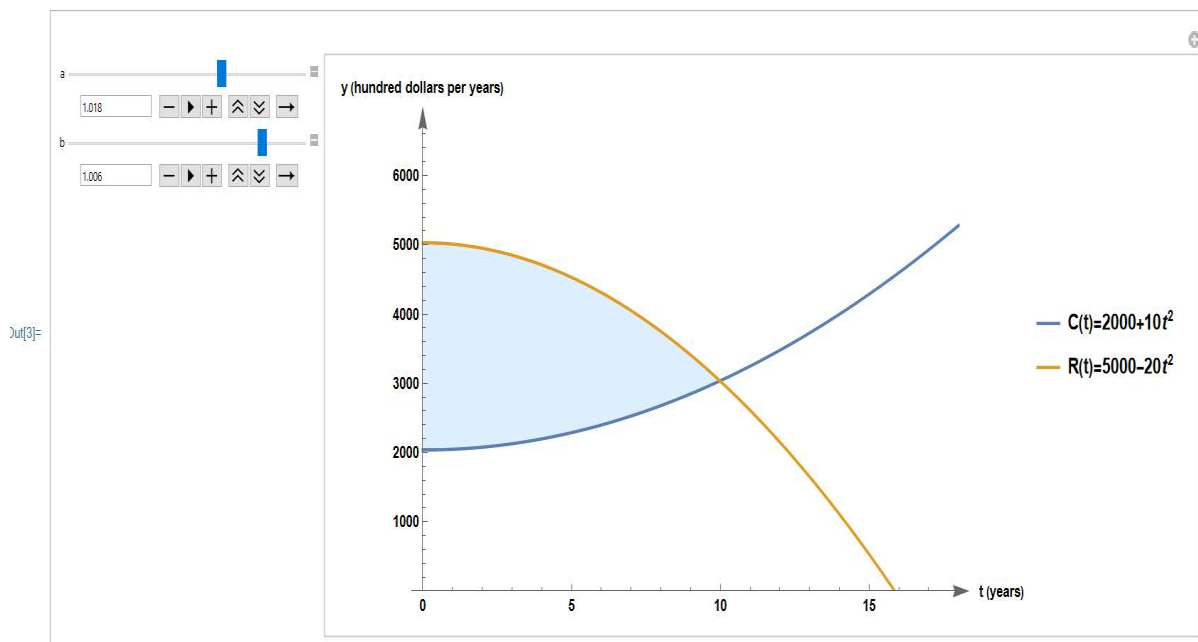
$$\begin{aligned} \text{(A)} \quad P(t) &= R'(t) - C'(t) = 3000 - 30t^2 && \text{the profit associated with the machine after } t\text{-years} \\ P'(t) &= 0 && \text{profitability begins to decline} \\ t &= 10 \text{ years} \end{aligned}$$

(B) The net earnings NE over the time period 0 to 10 are given by the difference

$$NE = P(10) - P(0) = \int_0^{10} P'(t) dt = \int_0^{10} (3000 - 30t^2) dt = (3000t - 10t^3) \Big|_0^{10} = \$20,000$$

End of solution

Interactive activity with Wolfram Mathematics Example 2



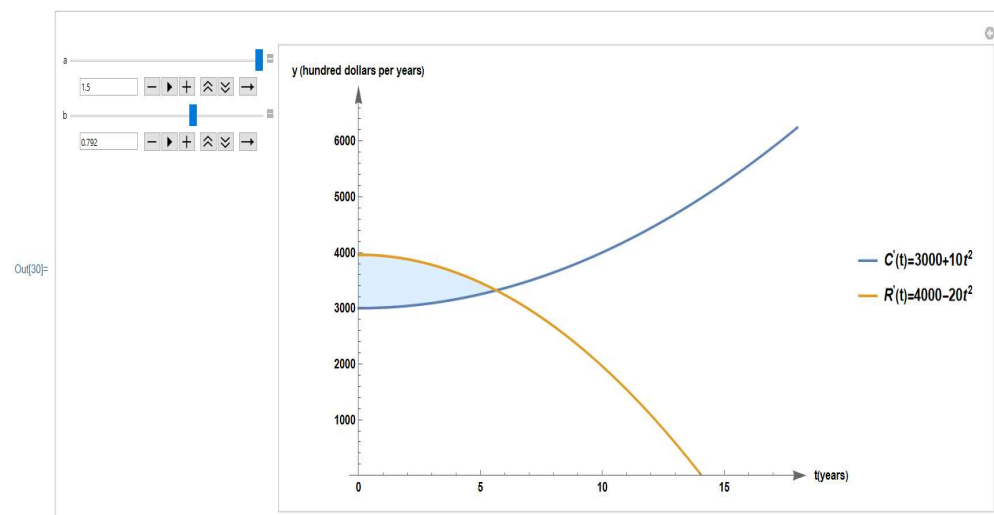
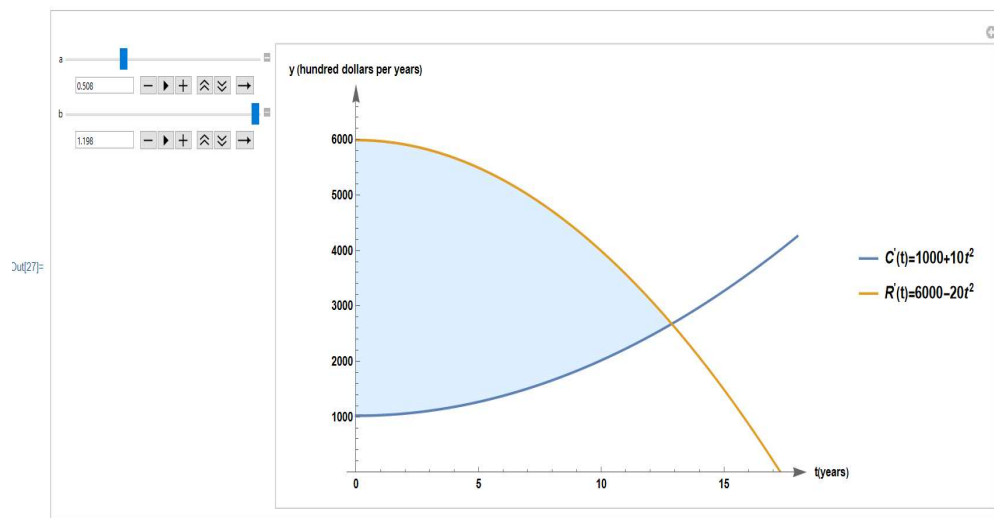
Activity 1: Use the provided Wolfram Mathematics application. Plot the Net Earnings $R'(t) = 5000 - 20t^2$ and Maintenance and Servicing $C'(t) = 2000 + 10t^2$ functions for t -years.

1. Consider the domain $x \geq 0$. Discuss the ranges of the functions.
2. Discuss the expectations for intersections.

3. Discuss the area between the curves, the y-axis and the line $x = x_0$, where x_0 is the intersection.
4. Conclude on resulting area being the subtraction of two areas.
5. Involve definite integral in calculations.
6. Compare the result with definite integral of the Profit function.

Activity 2: Use the provided Wolfram Mathematics application. Examine the change of the area between the curves according to the respective changes in the functions.

1. Try several parabolas with different y-axis intercepts.
2. Conclude of the resulting areas between the two curves.
3. Examine few more features of the powerful WM package.



Reference:

Calculus for Business Economics, and the Social and Life Sciences, Laurence Hoffmann, Gerald Bradley, McGraw-Hill Higher Education, Seventh edition 2000 and Tenth edition 2010.

STATISTICAL REASONING USING

JULIJE JAKŠETIĆ, MARJAN PRALJAK, ANA VUKELIĆ

1. INTRODUCTION

The main goal of statistical reasoning is to draw conclusions from available data and in this document we will describe ideas and methods behind this reasoning. The aim is not to explain concepts in depth, but to go rather quickly through the material and illustrate statistical reasoning using concrete and informative examples.

Important ingredient of this document is the use of a statistical software package. We will use R, an open source free software which is widely popular with extensive resources and manuals available online. Examples and text will be accompanied with key parts of the R code, while additional, more technical parts of the R code will be given in the Appendix. We will avoid describing unnecessary details of the code syntax, since we believe it is quite intuitive and easy to follow.

In the first part of this document we will deal with descriptive statistics whose goal is to give an initial sense of the information contained in the data. We will describe several graphical representations of the data which give us a first, visual impression of its distribution. We will also describe some important numerical measurements of data, like mean and standard deviation, which summarize important characteristics of the data like its expected values and variability.

Next we will proceed to build theoretical models for the data. In theoretical models the relative frequencies of outcomes observed in the data correspond to numbers called probabilities. While we can expect some discrepancies between observed frequencies and probabilities, for large samples, if the theoretical model is appropriate, these discrepancies should be small and tend to zero as the sample increases.

The last part of the document is devoted to statistical hypotheses testing. In essence, statistical hypothesis testing consists in checking whether the actual, observed data are in accordance with a theoretical model. Initial hypothesis that we want to test is restated in the form of a theoretical probability model. This theoretical model influences what kind of outcomes of measurements we would expect. If the observed data are unlikely to occur given this theoretical model, i. e. the more it diverges from the expected outcomes, we have more reasons to doubt the initial hypothesis and reject it as wrong.

2. DESCRIPTIVE STATISTICS

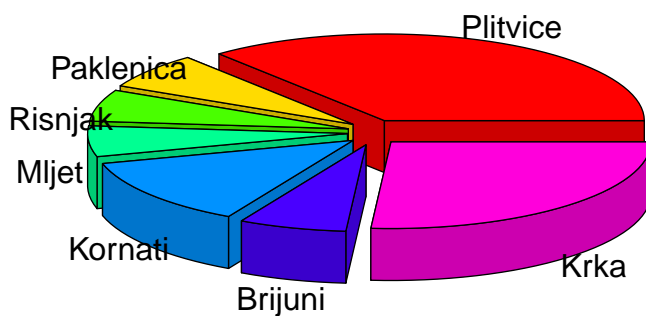
2.1. Charts.

Example 2.1. In the table below one can find the areas of the national parks in Croatia.

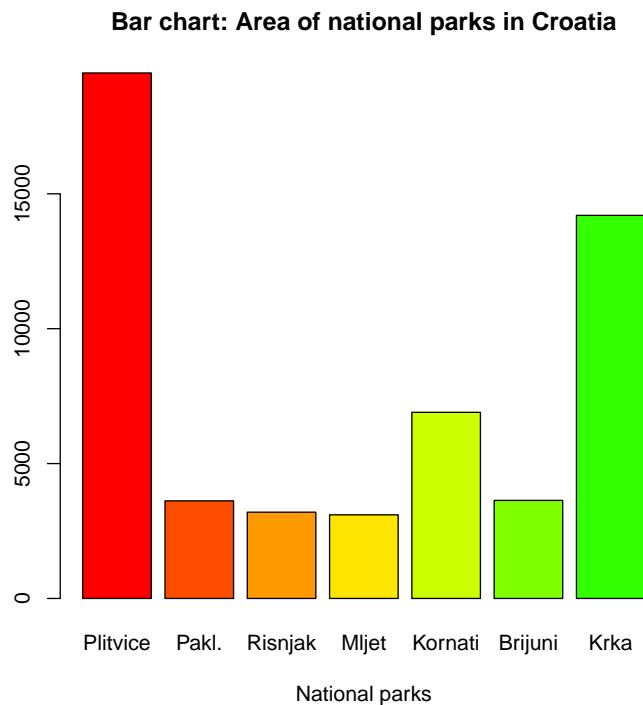
National park	Area (ha)
Plitvička jezera	19479
Paklenica	3617
Risnjak	3198
Mljet	3100
Kornati	6900
Brijuni	3635
Krka	14200

```
> library(plotrix)
> povrsine = c(19479,3617,3198,3100,6900,3635,14200)
> parkovi = c("Plitvice", "Paklenica", "Risnjak", "Mljet", "Kornati",
+ "Brijuni", "Krka")
> pie3D(povrsine,labels=parkovi,explode=0.1,
+       main="Pie chart: Area of national parks in Croatia ")
```

Pie chart: Area of national parks in Croatia



```
> parkovi = c("Plitvice", "Pakl.", "Risnjak", "Mljet", "Kornati",
+ "Brijuni", "Krka")
> barplot(povrsine, main = "Bar chart: Area of national parks in Croatia",
+ xlab = "National parks", names.arg = parkovi,col = rainbow(20))
```



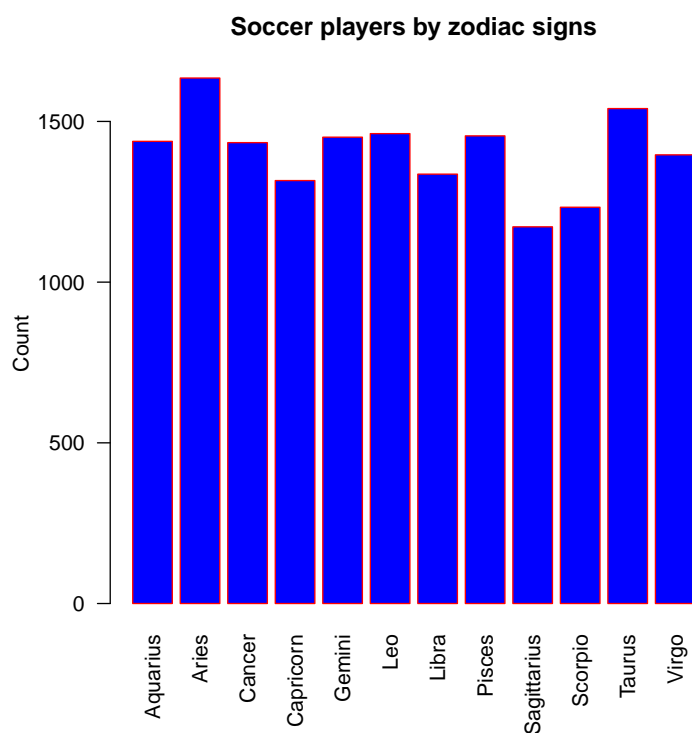
2.2. Bar chart. Example 2. Let us use file "person2020.csv". It contains list of 88937 famous people(for example: Julius Caesar, Muhammad, Marlon Brando, Donald Trump, Luka Modrić, Kiro Lazarov, James Dean) and for each of them there is 34 same facts such as birthrate, birthplace, etc.
Let us now extract just soccer players and associate them with their zodiac sign.

Air	Earth	Fire	Water
4225	4252	4269	4122

```
> library(xtable)
> library(anytime)
> library(DescTools)
> library(stringi)
> library(stringr)
> pantheon = read.csv("person2020.csv")
> nogometasi=subset( pantheon, pantheon$occupation == "SOCCER PLAYER")
> nogometasiA=nogometasi$birthdate
> nogometasiAB=anydate(nogometasiA)
> nogometasiAH=Zodiac(nogometasiAB)
> nogometasizod=stri_remove_empty(nogometasiAH, na_empty = TRUE)
> table1 = table(nogometasizod)
> mat = xtable(table1)
> colnames(mat) = c( "frequency")
> print(mat, sanitize.text.function = function(x){x})

> barplot(table1,
+         main="Soccer players by zodiac signs",
+         ylab="Count",
+         border="red", col="blue",las=2)
```

	frequency
Aquarius	1438
Aries	1635
Cancer	1434
Capricorn	1316
Gemini	1451
Leo	1462
Libra	1336
Pisces	1455
Sagittarius	1172
Scorpio	1233
Taurus	1540
Virgo	1396

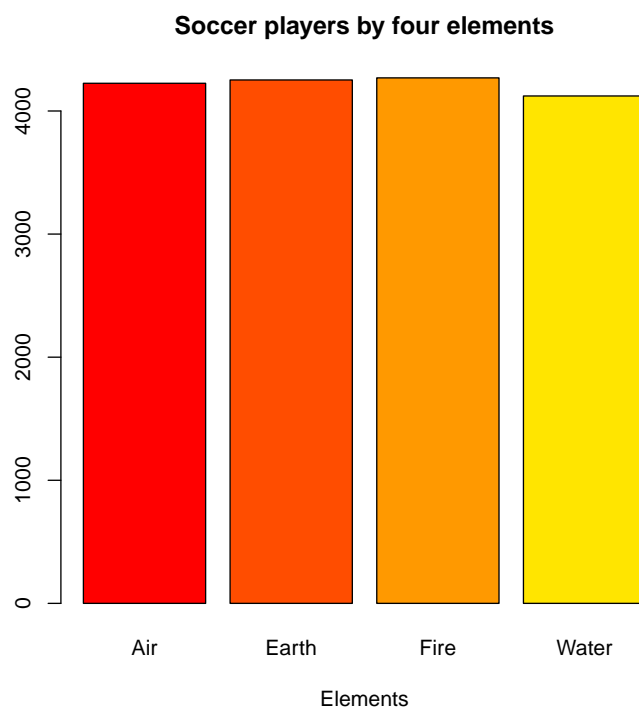


We can convert each zodiac sign to *four elements*(Fire, Water, Earth, Air) signs using conversion table:

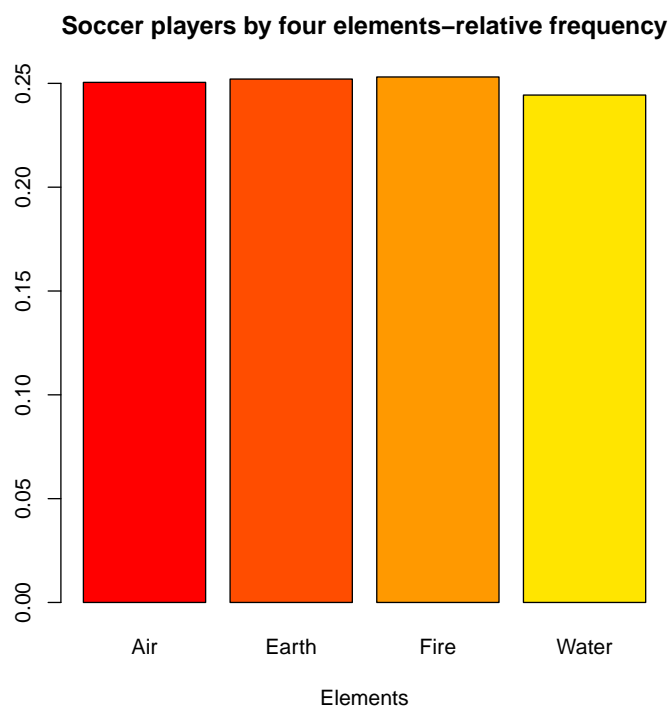
Fire (Aries, Leo, Sagittarius) Earth (Taurus, Virgo, Capricorn) Air (Gemini, Libra, Aquarius) Water (Cancer, Scorpio, Pisces).

```
> nogometasiAHA=str_replace_all(nogometasiAH, c("Leo" = "Fire", "Aries" = "Fire",
+ "Sagittarius" = "Fire", "Taurus" = "Earth", "Virgo" = "Earth",
+ "Capricorn" = "Earth", "Gemini" = "Air", "Libra" = "Air",
+ "Aquarius" = "Air", "Cancer" = "Water", "Scorpio" = "Water",
+ "Pisces" = "Water"))
> fourfreq = table(nogometasiAHA)
> signs = c("Air", "Earth", "Fire", "Water")
```

```
> barplot(fourfreq, main = "Soccer players by four elements",
+ xlab = "Elements", names.arg = signs,col = rainbow(20))
```



```
> frekr =fourfreq/sum(fourfreq)
> parkovi = c("Air", "Earth", "Fire", "Water")
> barplot(frekr, main =
+ "Soccer players by four elements-relative frequency", xlab = "Elements",
+ names.arg = parkovi,col = rainbow(20))
```



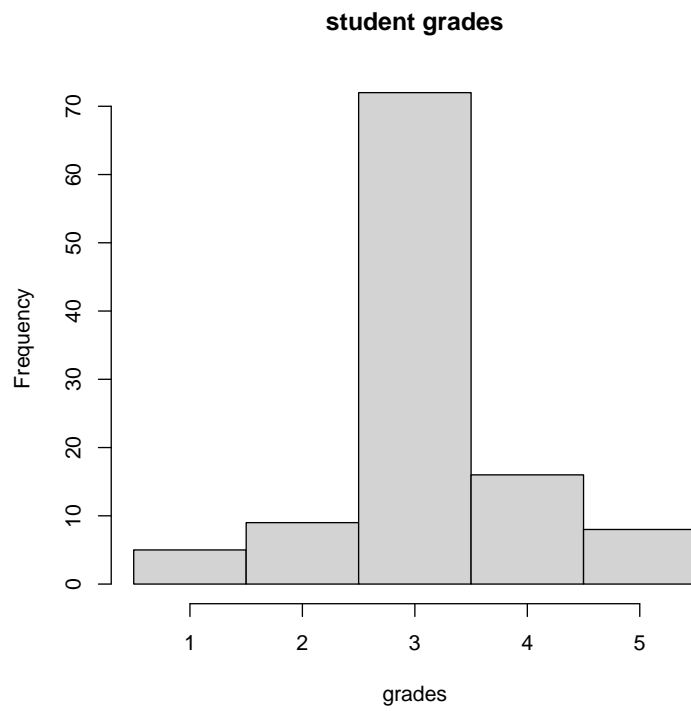
2.3. Relative frequencies-numerical valued data.

Example 2.2. *In the next table we have grades at an exam*

(1)

<i>grade</i>	1	2	3	4	5
<i>frequency</i>	5	9	72	16	8

```
> df= as.data.frame(cbind(grade= 1:5, frequency= c(5,9,72,16,8)))
> df.freq= as.vector(rep(df$grade, df$frequency))
> boje = c("red", "yellow", "green", "orange", "blue")
> hist(df.freq,breaks=c(0.5,1.5,2.5,3.5,4.5,5.5), main="student grades",
+       xlab="grades",freq=T)
```



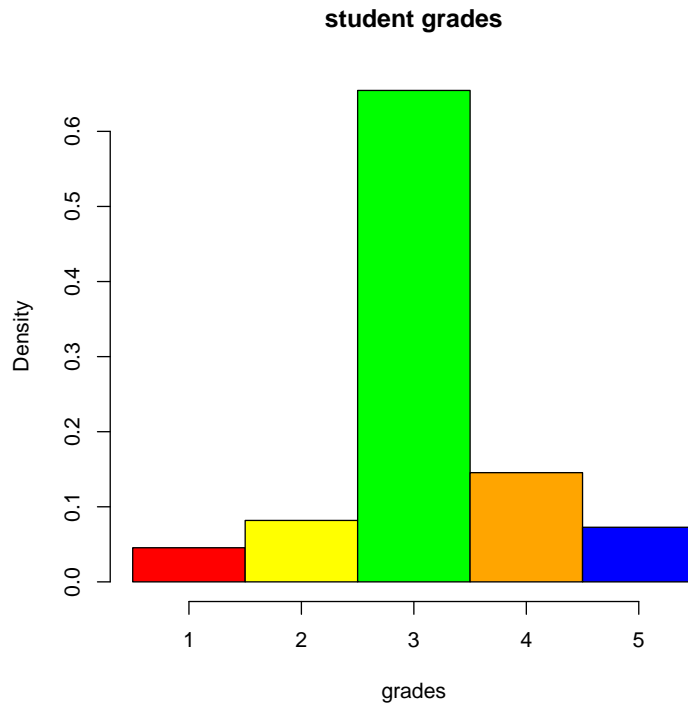
Observe, we have $n = 110$, grades (students). This information and the next table with *relative frequencies*

(2)

grade	1	2	3	4	5
frequency	$\frac{5}{110}$	$\frac{9}{110}$	$\frac{72}{110}$	$\frac{16}{110}$	$\frac{8}{110}$

is equivalent to the table (1),

```
> df= as.data.frame(cbind(grade= 1:5, frequency= c(5,9,72,16,8)))
> df.freq= as.vector(rep(df$grade, df$frequency))
> boje = c("red", "yellow", "green", "orange", "blue")
> hist(df.freq,breaks=c(0.5,1.5,2.5,3.5,4.5,5.5), main="student grades",
+       xlab="grades",freq=FALSE, col =boje)
```



2.4. **Mean and variance.** As we know

$$(3) \quad \bar{x} := \frac{1}{n} (x_1 + x_2 + \dots + x_n)$$

is the arithmetic mean of the numbers $\{x_1, x_2, \dots, x_n\}$.

If there is only $\{x_1, x_2, \dots, x_k\}$ different values with frequencies, respectively, $\{f_1, f_2, \dots, f_k\}$, $\sum_{i=1}^k f_i = n$, then (3) becomes

$$(4) \quad \bar{x} = \frac{f_1}{n} x_1 + \frac{f_2}{n} x_2 + \dots + \frac{f_k}{n} x_k$$

and now $\sum_{i=1}^k \frac{f_i}{n} = 1$.

Therefore, we can say, *expected* value

$$(5) \quad \mu = E[X] = \sum_{i=1}^k p_i x_i$$

of the random variable

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}$$

is the "mean" value of the random variable X .

2.5. Variance. We can differ between random variables using the notion of *variance*

$$(6) \quad \text{Var}(X) = E[(X - E[X])^2] = \sum_{i=1}^k p_i (x_i - \mu)^2,$$

and as the next animation illustrates.

3. PROBABILITY AND SIMULATIONS

3.1. Frequencies and probabilities.

Example 3.1. *Simulate 100 die throws of and calculate the relative frequency of the outcomes that were greater than or equal to 5.*

Let us first simulate, for example, 15 throws:

```
> Omega=1:6
> dice=sample(Omega, replace=TRUE, size=15)
```

The logical input `replace` tells whether the already sampled values can appear in the sample again. The sampled values were stored in the array `dice`, so the simulated die throws are

```
> dice
[1] 5 6 6 2 6 1 5 2 5 3 5 2 1 4 4
```

The following command checks which elements of `dice` are greater than or equal to 5.

```
> dice>=5
[1] TRUE TRUE TRUE FALSE TRUE FALSE TRUE FALSE TRUE FALSE TRUE FALSE
[13] FALSE FALSE FALSE
```

The answer to the posted question is obtained by the following commands

```
> Omega=1:6
> dice=sample(Omega, replace=TRUE, size=100)
> length(dice[dice %in% c(5,6)])/100
```

```
[1] 0.25
>
```

Example 3.2. *Two dice, one blue colored and the other red, are thrown simultaneously.*

- *Let A be the event that the number 4 was rolled on the blue die and the number 5 was rolled on the red die.*
- *Let B be the event that the sum of the two rolled numbers is equal to 9.*

Estimate the probabilities of A and B by using simulations with $n = 1000$ iterations.

```
> N=1000
> Omega=1:6
> diceRolls = replicate(2, sample(Omega, size=N, replace = TRUE))
```

The command `replicate`, as the name suggests, invokes twice the command `sample`. The results are stored in the two dimensional array `diceRolls`.

For example

```
> diceRolls[9,]
```

```
[1] 4 1
```

are the numbers rolled in the ninth throw. The first ten rolls of the blue die are

```
> diceRolls[1:10,1]
```

```
[1] 5 1 2 1 4 5 3 3 4 4
```

and the first ten rolls of the red die are

```
> diceRolls[1:10,2]
```

```
[1] 5 5 3 1 4 1 4 1 1 4
```

The estimated probability of A is

```
> sum(diceRolls[,1] == 4 & diceRolls[,2] == 5) / N
```

```
[1] 0.03
```

and the estimated probability of B is

```
> sum(diceRolls[,1]+ diceRolls[,2] == 9) / N
```

```
[1] 0.115
```

The table (2) can be put, in more probabilistic view, in terms of random variables

$$X \sim \left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ \frac{5}{110} & \frac{9}{110} & \frac{72}{110} & \frac{16}{110} & \frac{8}{110} \end{array} \right)$$

where

$X = \text{"M2 student grades in 2021"}$

Example 3.3. Give a simulation of, future, grades for the next generation of 100 students on the Mathematics 2. Represent simulated data in table and with histogram.

We first simulate 100 grades, according to distribution given by the random variable X

```
> M2grades2022=sample(x = c(1: 5),
+                      prob = c(5/110, 9/110,72/110 ,16/110, 8/110),
+                      size = 100,
+                      replace = TRUE)
```

Now we can list simulated grades

```
> M2grades2022

[1] 3 3 3 3 1 3 3 4 3 3 1 3 3 4 5 3 1 3 3 4 3 3 3 3 1 3 3 3 4 3 3 3 1 3 3 4
[38] 3 3 2 3 2 3 3 2 3 3 5 4 3 3 3 3 3 1 3 3 4 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
[75] 3 3 3 2 3 3 3 4 2 3 3 2 3 2 3 2 3 3 4 2 2 3 4 2 3 3
```

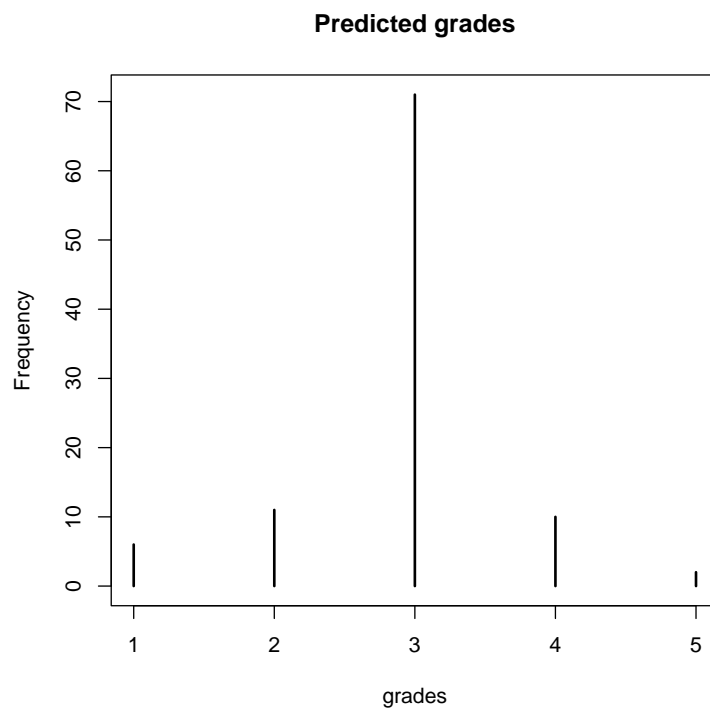
and put them into the frequency table

```
> table(M2grades2022)
```

```
M2grades2022
 1  2  3  4  5
 6 11 71 10  2
```

and then represent them visually

```
> plot(table(M2grades2022), xlab = 'grades', ylab = 'Frequency',
+       main = 'Predicted grades')
```



If we run once again the same lines of codes, we get a slightly different frequencies of numbers

```
> M2grades2022=sample(x = c(1: 5),
+                      prob = c(5/110, 9/110,72/110 ,16/110, 8/110),
+                      size = 100,
+                      replace = TRUE)
> table(M2grades2022)
M2grades2022
 1  2  3  4  5
4 10 62 15  9
```

3.2. **Cards.** We can easily simulate drawing card(s) from a deck.

The 52 deck is made of four suits, and each suit has 13 ranked cards: *ace, king, ..., deuce*.

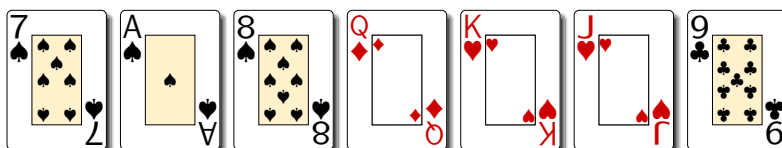
We stored these cards in the data frame `deck`

```
> deck = data.frame(rank=rep(c("A", "K", "Q", "J", "T", "9", "8", "7", "6", "5",
+                               "4", "3", "2"), 4),
+   suit=c(rep("Spade", 13), rep("Heart", 13), rep("Diamond", 13), rep("Club", 13)) )
```

If we want to simulate one hand of 7 cards, we can simulate 5 random numbers (without replication) between 1 and 52, and the simulated numbers are rows of `deck`

```
> x = sample(1:52, size=7)
> x
[1]  8  1  7 29 15 17 45
> hand = deck[x,]
> hand
```

	rank	suit
8	7	Spade
1	A	Spade
7	8	Spade
29	Q	Diamond
15	K	Heart
17	J	Heart
45	9	Club




Assume now that we are drawing a 5 card hand from a deck of 52 cards and let us denote the random variable

$X = \text{"number of spades in the hand"}$.

In the next example we estimate distribution of the X .

Example 3.4. *Simulate $N = 10000$ hands, count the number of spades in each hand.*

- Draw the histogram.
- Calculate the mean and standard deviation of the obtained data.

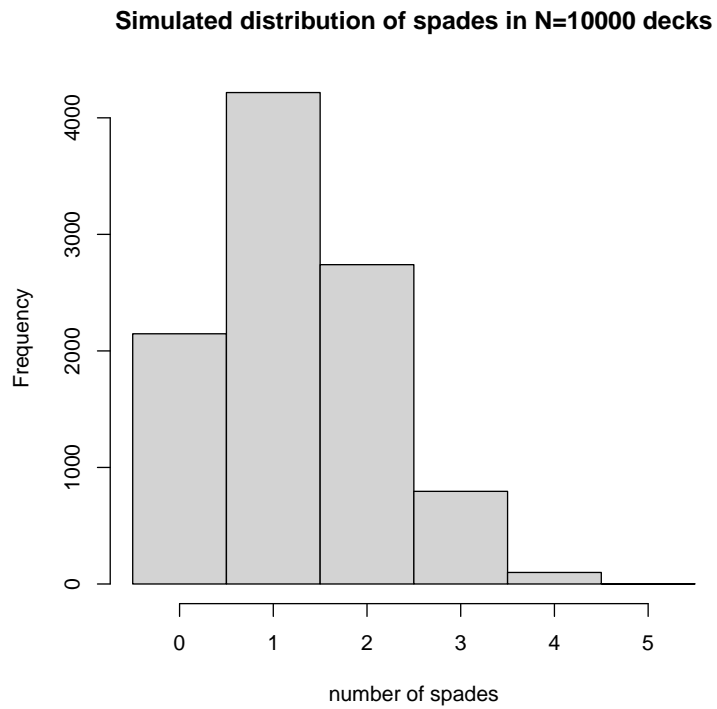
Let us first simulate in  $N = 10000$ hands, and make the table of obtained frequencies

```
> N=10000; dat = c();
> for (i in 1:N) {
+   x = sample(1:52, size=5); hand = deck[x,];
+   num_of_spades = sum(hand$suit=="Spade"); dat = append(dat, num_of_spades);
+ }
> table(dat)
```

dat	0	1	2	3	4	5
	2147	4217	2740	795	99	2

Now we form the histogram

```
> hist(dat, breaks=c(-0.5,0.5,1.5,2.5,3.5,4.5,5.5),  
+      main="Simulated distribution of spades in N=10000 decks",  
+      xlab="number of spades")
```



Estimated mean value of spades is

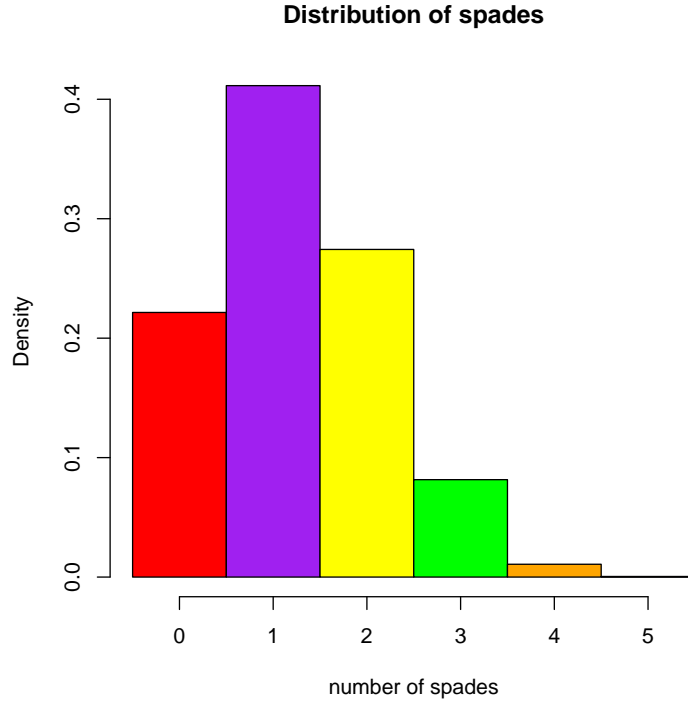
```
> mean(dat)
```

```
[1] 1.2488
```

and the standard deviation is

```
> sd(dat)
```

```
[1] 0.9149767
```

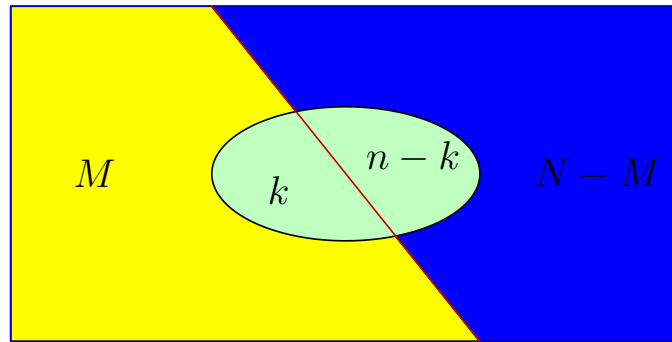


Remark 3.5. The distribution of the random variable X is well known *hypergeometric distribution* and

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}, \quad k = 0, 1, \dots, 5,$$

while

$$E[X] = \frac{n \cdot M}{N}, \quad \sigma = \sqrt{\text{Var}(X)} = \sqrt{\frac{n \cdot M}{N} \frac{N-M}{N} \frac{N-n}{N-1}}$$



4. TESTING OF HYPOTHESIS

The goal of statistical reasoning is to draw conclusions from the observed data. One way is to estimate certain parameters. Another approach, which is in fact an equivalent problem, but slightly reformulated, is to start with an initial hypothesis and then see if the observed data is in the accordance with this hypothesis or whether it diverges significantly.

The closest of the observed data to the initial, *null hypothesis* is done in the following way: we first calculate *test statistic* which is a value calculated from the observed data. The key thing about the test statistic is that we know its distribution under the null hypothesis, i.e. if the null hypothesis holds then we know what are the expected outcomes of the test statistic and with which probabilities. Then we are in the position to judge how likely or unlikely is the occurrence of the particular data we observed. The smaller the probability of the actual observed data (or worse), we have more reasons to doubt the null hypothesis and to reject it.

We will illustrate these concepts in more detail through the following example.

Example 4.1. *Traditional boat race between Oxford and Cambridge universities is held every year. So far Cambridge has won 85 races and Oxford 81 races. Can we conclude that both universities are equally successful, i.e. whether for each particular boat race both universities are equally likely to win?*


The test statistics that we will use is the number of wins one of these colleges, for example Cambridge.

$$X = \text{"number of wins of Cambridge"}$$

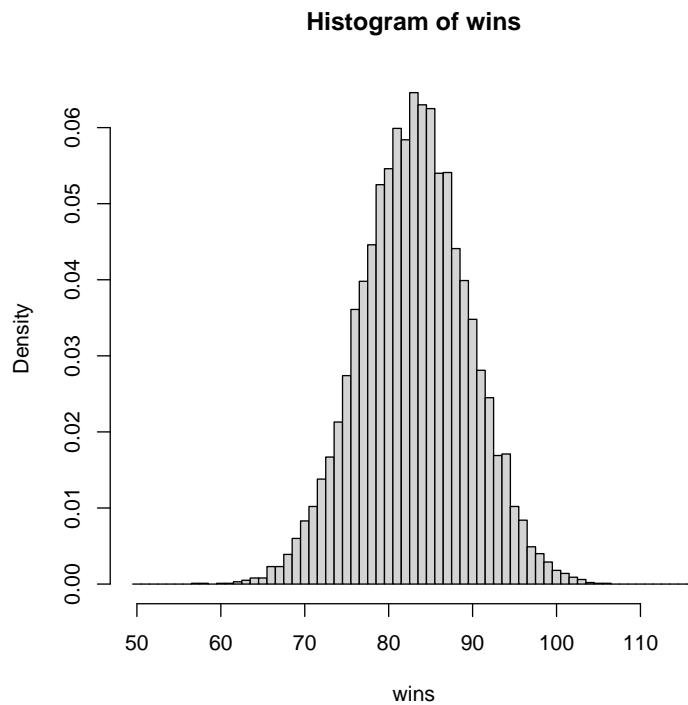
One approach to answer the question is to calculate the actual proportion of wins of Cambridge, $\hat{p} = 85/166 = 0.512$, and to judge whether this is close enough to one half, i.e. to hypothesized value of the parameter $p = 0.5$.

The other approach is to test whether the hypothesis that $p = 0.5$ is justified.

In each boat race we have two possible outcomes, the win of either Cambridge or Oxford, and the boat race was held 166 times. Note, this is equivalent to repeating an experiment with two possible outcomes, for example coin toss, for $n = 166$ times and in each of these experiments the probability of the observed event, for example the coin falls heads, is equal $p = 0.5$.

We can simulate in  the number of heads in 166 coin tosses (or wins of Cambridge in 166 boat races) with the following code.

```
> N=10000; wins = c();
> for (i in 1:N) {
+   wins_cambr=sum(sample(x = c(0: 1),
+     prob = c(0.5, 0.5),
+     size = 166,
+     replace = TRUE)); wins = append(wins,wins_cambr);
+ }
> hist(wins, breaks=seq(from=49.5, to=116.5, by=1),freq=FALSE)
```



Remark 4.2. More generally, if we repeat an experiment n times and in each of these experiments the probability of some observed random event is p , then the random variable X that counts how many times the observed event occurred in these n experiments has the so called **binomial** distribution:

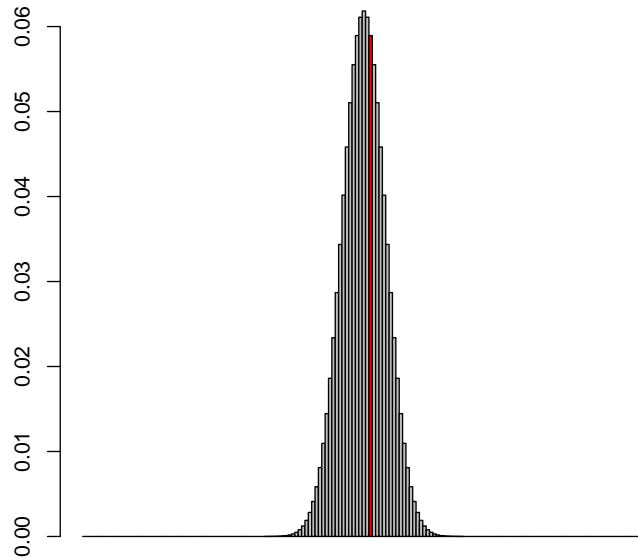
$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n;$$

$$E[X] = np, \quad \text{Var}(X) = np(1-p).$$

Let us denote the parameter p = "probability that Cambridge wins Oxford in one boat race".

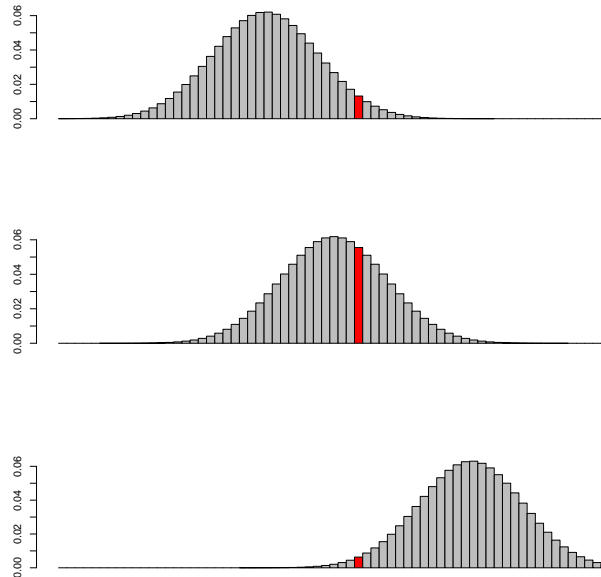
If the null hypothesis that $p = 0.5$ is true, then the distribution of the number of wins of Cambridge has binomial distribution with parameters $n = 166$, $p = 0.5$. The red highlighted bar is the probability of the actually observed value $X = 85$. Notice that the observed value 85 is in the range of quite likely outcomes.

```
> n = 166; p = 0.5;
> boje=rep("gray",n+1);
> boje[[86]] = "red"
> barplot(dbinom(0:n,n,p), col=boje, space=0)
> par(mfrow=c(3,1))
> barplot(dbinom(0:n,n,0.4), col=boje, space=0)
> barplot(dbinom(0:n,n,p), col=boje, space=0)
> barplot(dbinom(0:n,n,0.6), col=boje, space=0)
> par(mfrow=c(1,1))
>
```



For comparison, the next figure gives the distribution of wins of Cambridge under three different null hypotheses: that $p = 0.45$, $p = 0.5$, $p = 0.6$. The horizontal axes are aligned and red highlighted bar always represents the outcome 85 wins for Cambridge. Notice that, in the contrast to the case $p = 0.5$, the outcome of 85 wins is in the range of quite unlikely outcomes in the case of $p = 0.45$ (the first graph) and $p = 0.6$ (the third graph).

```
> n = 166; p = 0.5;
> n1 = 50; n2 = 116;
> boje = rep("gray", n+1);
> barplot(dbinom(0:n, n, p), col=boje, space=0)
> boje[[86]] = "red"
> barplot(dbinom(n1:n2, n, p), col=boje[n1:n2], space=0)
> par(mfrow=c(3,1))
> barplot(dbinom(n1:n2, n, 0.45), col=boje[n1:n2], space=0)
> barplot(dbinom(n1:n2, n, p), col=boje[n1:n2], space=0)
> barplot(dbinom(n1:n2, n, 0.6), col=boje[n1:n2], space=0)
> par(mfrow=c(1,1))
>
```



Let us now return to the originally tested hypothesis that $p = 0.5$. This hypothesis influences on what outcomes of the test statistic we expect- namely we expect Cambridge to win around half of the raises (i.e. X should be around 83 as one can see in Figure 4). What outcomes would cause us the doubt the null hypothesis? If Cambridge has too many wins compared to Oxford (i.e. we are on the right tail of the distribution of X) or if Cambridge has too few wins compared to Oxford (i.e. we are on the left tail of the distribution of X).

In general, how do we determine suspicious (critical) values of the test statistic for which we deem the null hypothesis as too doubtful and rejected? We first choose some (small) probability of error that we ready to tolerate. This probability is called the level of significance and is usually denoted by α . In our case let us take $\alpha = 0.05$ and as we have commented before, since too small or too high values of X are critical for null hypothesis, we will consider the $\alpha/2 = 0.025$ smallest values of X (i.e. on the left tail of its distribution) and the $\alpha/2 = 0.025$ largest values of X (i.e. on the right tail of its distribution) as the critical values of X for which we reject the null hypothesis.

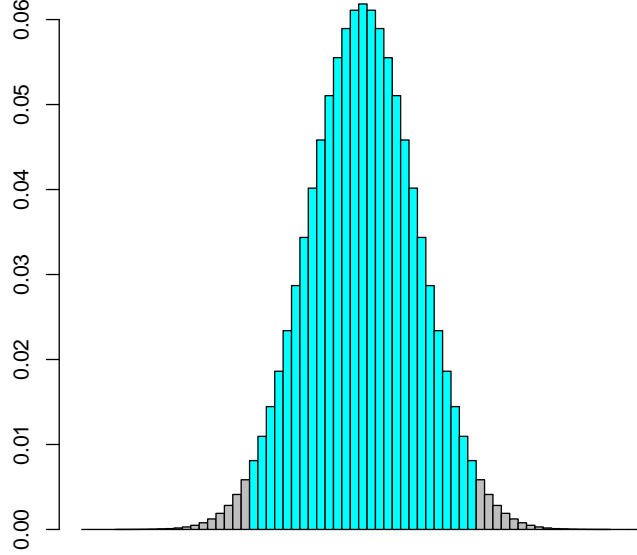
```
> alpha = 0.05;
> i1 = qbinom(alpha/2,n,p);
> i1

[1] 70

> i2 = qbinom(1-alpha/2,n,p);
> i2

[1] 96

> for (i in i1:i2) {boje[i] = "cyan"}
> barplot(dbinom(n1:n2,n,p), col=boje[n1:n2], space=0)
>
```



In our case the boundaries for the critical values were 70 and 96; i.e. if Cambridge had less than 70 or more than 96 wins, we would have rejected the null hypothesis.

Example 4.3. *Are there more left-handers among elite tennis players than in the general population? There are many studies about prevalence of the dominant hand and the estimates of the left-handed people closely vary around 10%. For the group that represents the elite tennis players we will take all tennis players who were ranked among top 100 in the ATP rankings at least once in the 2010s decade. There were altogether 310 such players of whom 46 were left-handed. The details on the source of the data and our code can be found in the Appendix.*

The test statistics that we use in this case

$X =$ "number of lefthanders in a group of 310 elite tennis players"

Similarly as in the previous example, for each tennis player in our sample we have two possible outcomes - he is either left-handed or not. This time the probability, i.e. the parameter p , is

$p =$ probability that an elite tennis player is left handed

We want to test the null hypothesis that $p = 0.1$. If the null hypothesis is true, then we would expect 10% of the players in our sample to be left-handed, i.e. $np = 0.1 \cdot 310 = 31$ player.

The way the question is formulated implies the null hypothesis is questionable only if there are too many lefthanders in our sample, i.e. if the value of X is too high. Therefore critical values of the test statistic X will be on its right tail.

If the null hypothesis holds, then the number of left handed tennis players X has binomial distribution with $n = 310$ and $p = 0.1$.

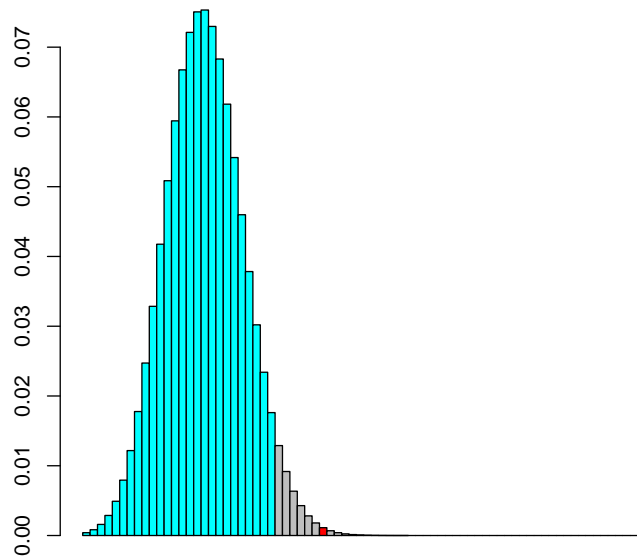
```

> n=310; p=0.1;
> alpha = 0.05;
> i1 = qbinom(1-alpha,size=n,prob=p);
> i1

[1] 40

> boje=rep("gray", n+1)
> for (i in 0:i1) {boje[i] = "cyan"}
> boje[[47]]="red";
> n1=15; n2=90;
> barplot(dbinom(n1:n2,n,p), col=boje[n1:n2], space=0)

```



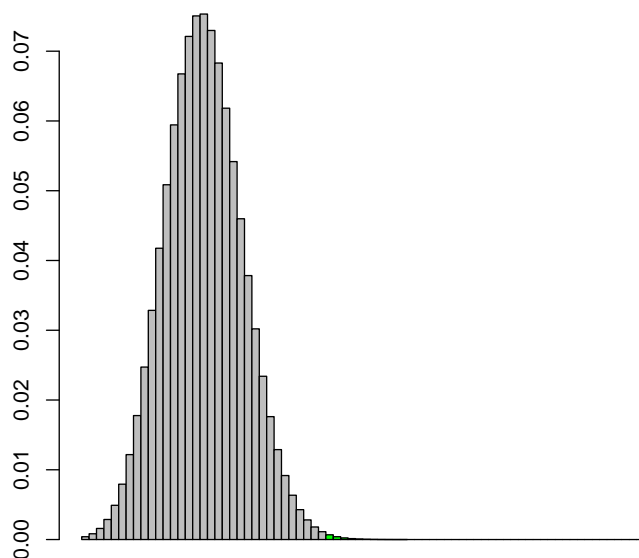
The blue colored bars represent the probabilities of the 95% smallest values of X (non-critical values), the gray colored bars represent the probabilities of the 5% highest values of X (critical values) and the (barely noticeable) red bar is the probability of the observed value $X = 46$. Since the observed value $X = 46$ is greater than boundary of the critical area 40. We therefore reject the null hypothesis and conclude that there is evidence of higher prevalence of lefthander among elite tennis players.

We will use this example to introduce another important concept in statistical hypothesis testing, the p -value. The p -value is the probability that, if the null hypothesis is true, we obtain the observed data or worse for the null hypothesis. The smaller the p -value indicates that it is more unlikely that the null hypothesis is true and that the observed data are obtained by chance, but rather that it is more likely that null hypothesis is false. Reaching a decision on a statistical hypothesis using the p -value is simple: if the p -value is small enough, i.e. less than some small significance level α we have chosen, then we deem the null hypothesis as too suspicious and we reject it. Otherwise, if the p -value is greater than α , we consider

that the null hypothesis holds.

In our example, the greater the number of left-handed tennis players in our sample means that the null hypothesis is less likely. Therefore, higher values of X are worse for the null hypothesis, so we will calculate the p -value on the right tail of the distribution of X .

```
> n=310; p=0.1;
> pvalue =sum(dbinom(47:n,size=n,prob=p));
> pvalue
[1] 0.002740603
> boje=rep("gray", n+1)
> for (i in 47:n+1) {boje[i] = "green"}
> n1=15; n2=90;
> barplot(dbinom(n1:n2,n,p), col=boje[n1:n2], space=0)
```



The green highlighted bars represent the probabilities of the outcomes of X of 46 or higher and their sum is the p -value. Since the p -value 0.0027406 is very small, we can safely reject the null hypothesis and conclude that there is significantly more lefthanders among elite tennis players.

5. PROBLEMS

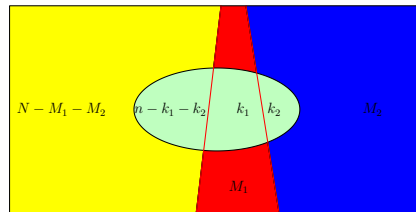
Problem 5.1. Load again the database `people2020.csv` into dataframe in R.

- a) Find all the people whose name contains string `Hopkins`
- b) Find the row number of the dataframe which contains information on Anthony Hopkins.
- c) Read the Twitter username and birth date Anthony Hopkins from dataframe.

Problem 5.2. Count the number of people from the database `people2020.csv` whose `occupation` is listed as: `"ACTOR"`, `"SOCCER PLAYER"`, `"POLITICIAN"` and `"RELIGIOUS FIGURE"`, and draw the pie-chart of these counts.

Problem 5.3. We draw a five card hand from the deck of 52 cards. Simulate $N = 10000$ such five card hand deals and estimate the probability of the following random events:

- a) the five card hand has exactly one ace;
- b) based on these 10000 simulations estimate the average number of aces in a five card hand
- c) the five card hand has exactly two hearts and exactly one spade
- d) calculate the exact probability of the random event from part c). Use the following figure as a hint.



Problem 5.4. Load the data from `oscars.csv` into an R dataframe `oscars`.

- a) Calculate the mean and standard deviation of the age of the recipients for the best actor and best actress award.
- b) Find all the recipients who won the award more than once.

Problem 5.5. a) Simulate 15 throws of the dice and calculate the mean and the standard deviation of these fifteen throws.

b) Repeat this procedure 1000 times, i.e. simulate 15 dice throws for 1000 times. For each repetition calculate the mean value and standard deviation and store them into arrays `mean_values` and `sd_values`

c) Calculate the mean values of the arrays `mean_values` and `sd_values`.

d) Draw the histograms of the arrays `mean_values` and `sd_values`.

e) Calculate the mean value and standard deviation for the theoretical model of dice throw, i.e. for the random variable

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

6. APPENDIX

*In the previous chapters we explained key features of **R** commands we invoked and in this appendix we will explain and comment more technical part of the code. As you have seen **R** is a command line software. For an easier use of **R**, it is more convenient to use with some **R** editors, and we recommend RStudio. RStudio allow us multiple windows to be simultaneously open and the user can, in addition to command line prompt, keep track of, for example, variables declared, data loaded, etc...*

*When **R** needs to load data or anything else, it first searches for it in the so called working directory. Working directory can be set with the command `setwd("Location")`, for example*

```
> setwd("C:/Users/Hogweed/Google disk/R/Vjezbe1")
```

The data can easily be loaded if it is stored in an appropriate form, for example in .csv or .tab or Excel file. For example, the command

```
> #pantheon = read.csv("./Datasets/person2020.csv")
>
```

loads data stored in the file `person2020.csv` which is located in the sub folder `Datasets` of the working directory. If the .csv file is already in the working directory one doesn't need to give the whole path to the file. The data is loaded in to the variable `pantheon`. This type of variable in the R syntax is called `dataframe`.

The data for Example 4.3 with left handed tennis players were taken from the database maintained by Jeff Sackmann on the web page <https://github.com/JeffSackmann/>

The extensive database contains information on tennis players, rankings, matches,... For our purposes we have downloaded the files `atp_players.csv` and `atp_rankings_10s.csv` which contain information on tennis players and their rankings in 2010s, respectively.

```
atp_players = read.csv('atp_players.csv')
atp_rankings_10s = read.csv('atp_rankings_10s.csv')
```

```
NA
```

```
atp_top_ranked = subset(atp_rankings_10s, rank <= 100)
```

Many players spend a lot of weeks ranked in top 100, so they would repeat many times in the dataframe `atp_top_ranked`. The R command `unique` returns unique values from an array and ignores repetitions, so the following command retrieves all the `player_ids` of players who were ranked in the top 100 ATP positions at least once and stores them in the array `atp_top_ranked_ids`

```
atp_top_ranked_ids = unique(atp_top_ranked$player)
```

The players in question can be retrieved with the command

```
subset(atp_players, player_id %in% atp_top_ranked_ids)
```

The information on the tennis player's dominant hand is given in the column `hand`, so the following `table` command returns frequencies of dominant hand for tennis players in `atp_top_ranked_ids`

```
table(subset(atp_players, player_id %in% atp_top_ranked_ids)$hand)
```

<i>A</i>	<i>L</i>	<i>R</i>	<i>U</i>
0	46	264	0

As we can see, there were all together 310 players who were ranked among top 100 at least once in the 2010s, of which 46 were left handed and 264 right handed.

Vectors and Vectors application
Summer course: STEM Ambassadors

Contents

1	Vector algebra	3
1.1	Vectors and scalars	4
1.2	Vectors in space	7
1.3	Operation with vectors	10
2	Products of two vectors	15
2.1	The scalar product of two vectors	15
2.2	The vector product of two vectors	18
3	Vector application	23
3.1	Lines in space	23
3.2	Planes	29
4	Exit ticket	35

Chapter 1

Vector algebra

Opening problem:



Figure 1.1:

A signpost gives information about distances and directions to towns or to other locations relative to the location of the signpost. Distance is a scalar quantity. Knowing the distance alone is not enough to get to the town; we must also know the direction from the signpost to the town. The direction, together with the distance, is a vector quantity commonly called the displacement vector. A signpost, therefore, gives information about displacement vectors from the signpost to towns.

Vectors are essential to science and engineering. Many fundamental physical quantities are vectors, including displacement, velocity, force, and electric and magnetic vector fields. In other words, vectors are a component part of STEM in much the same way as sentences are a component part of literature.

Vectors are Euclidean quantities that have geometric representations as arrows in one dimension (in a line), in two dimensions (in a plane), or in three dimensions (in space). They can be added, subtracted, or multiplied. In this chapter, we explore elements of vector algebra and its application in STEM.

1.1 Vectors and scalars

Quantities which can be specified completely by giving a single number and the appropriate unit are called scalars. For example, "the current temperature is 23 degrees", or the "This bag can hold a 5 kg of mass" are a scalar quantities. Scalar is a synonym of "number." Time, mass, distance, length, volume, temperature, and energy are examples of scalar quantities.

Scalar quantities that have the same units can be added or subtracted according to the usual rules of algebra for numbers. For example, if I have 70 kg weight, and you have 50 kg weight, together we have 120 kg weight. We can multiply one scalar with other scalar, and also, we can divide one scalar with another scalar (different than zero). The result is again scalar.

Many quantities, however, cannot be described completely by just a single number of units. For example, when a helicopter goes for a rescue mission, the rescue team must know not only the distance to the distress signal, but also the direction from which the signal is coming so they can get to its origin as quickly as possible. The quantities specified completely by giving a number of units (magnitude) and a direction are called vector quantities. Examples of vector quantities include displacement, velocity, position, force. In the language of mathematics, vector quantities are represented by mathematical objects called vectors. We can add or subtract two vectors, and we can multiply a vector by a scalar or by another vector, but we cannot divide by a vector. The operation of division by a vector is not defined.

Directed line segment presentation. We can represent a vector quantity using a directed line segment or arrow. The length of the arrow represents the vectors magnitude, and the arrowhead shows its directions.

Example 1.1.1. *Draw a scale diagram to represent a force of magnitude $6\sqrt{2}N$ in a north-west position. See Figure 1.2.*

Vector notation. Displacement is a general term used to describe a change in position, such as during a trip from the tent to the fishing hole.

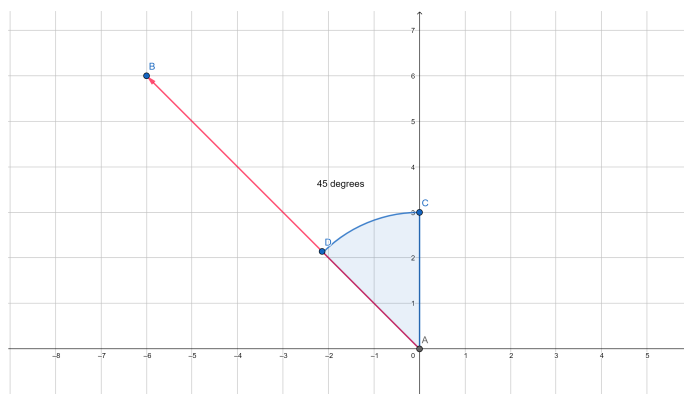


Figure 1.2:

Displacement is an example of a vector quantity. We will use the displacement to introduce the vector notation. If you walk from the tent (location A) to the hole (location B), as shown in Figure 1.3, the vector \vec{a} representing your displacement, is drawn as the arrow that originates at point A and ends at point B . The arrowhead marks the end of the vector. The direction of the displacement vector \vec{a} is the direction of the arrow. The length of the arrow represents the magnitude $|\vec{a}|$ of vector \vec{a} . Here, $|\vec{a}| = 6$ km. Since the magnitude of a vector is its length, which is a positive number, the magnitude is indicated by placing the absolute value notation around the symbol that denotes the vector. We can denote a vector with \vec{AB} and this is the

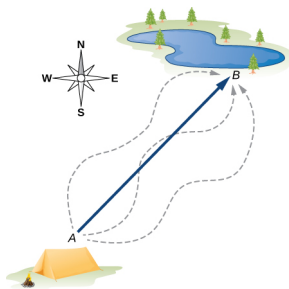


Figure 1.3:

vector that originates at A and terminates at B . The vector \vec{AB} is called displacement vector of B relative to A .

Vector equality. Two vectors are equal if they have the same magnitude

and direction. Equal vectors are parallel, they have the same direction and equal length. The arrows that represent them are translations of one another. In Figure 1.4 we have three vector presentations of the same vector \vec{a} .

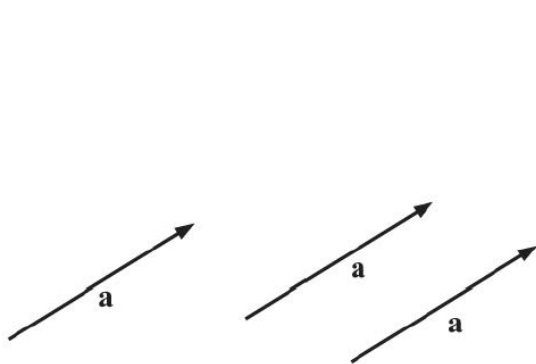


Figure 1.4:

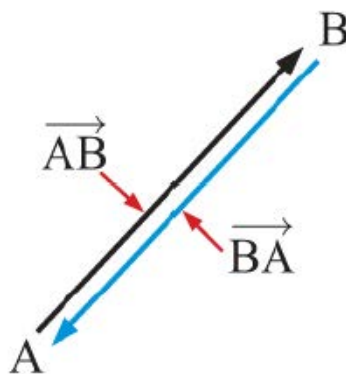


Figure 1.5:

We can introduce the negative vector of \vec{AB} and denoted with \vec{BA} , and that is the vector parallel to \vec{AB} , the same length, but the opposite direction, see Figure 1.5.

Vector addition and multiplication by scalar. Two vectors (or more) can be added and they give the resultant vector. Vector addition is commutative and associative. To construct a resultant vector of two vectors in a plane geometrically, we use the parallelogram rule, see Figure 1.6. When

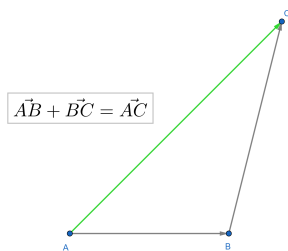


Figure 1.6:

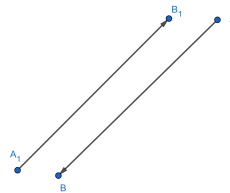


Figure 1.7:

a vector is multiplied by a scalar, the result is another vector of a different length than the length of the original vector. Multiplication by a positive scalar does not change the original direction; only the magnitude is affected. Multiplication by a negative scalar reverses the original direction, see Figure 1.7.

To make an exercise in addition of vectors and multiplication by scalar we use the interactive simulator for science and math, PheT developed by the University of Colorado, see on <https://phet.colorado.edu/en/>.

Example 1.1.2. Susan and Leo are pushing a heavy trolley containing groceries. Susan pushes the trolley with force 9.5N in the direction 18.4° , and Leo pushes the trolley with force 8.9N in the direction -26.6° . Use the scale diagram to estimate the resulting force from the two girls pushing.

We can see the resulting force on Figure 1.8.

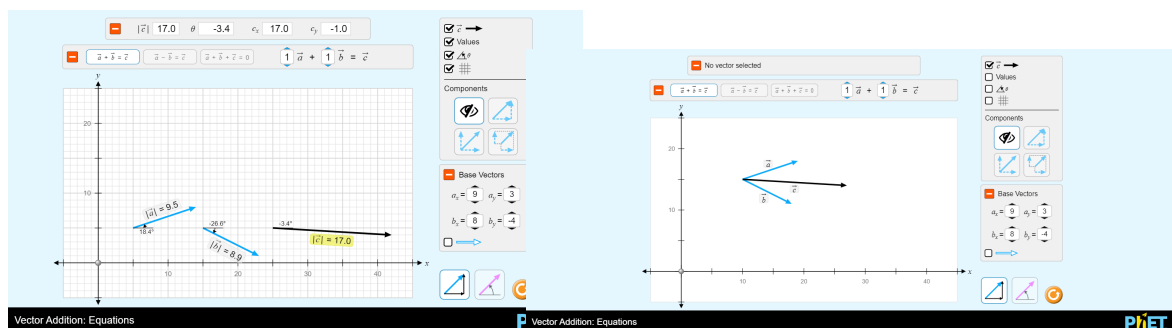


Figure 1.8:

Example 1.1.3. An adventurous cat strays from home, runs three blocks east, two blocks north, one block east, one block north, and two blocks west. Assuming that each block is about 100 m, how far from home and in what direction is the cat? Use a graphical method.

The resulting displacement of the dogs run is given on Figure 1.9. We can see that the dog is $\sqrt{13} \times 100$ meters far from home in north-east direction.

1.2 Vectors in space

In three-dimensional space, vector \vec{a} has three vector components: the x -component $\vec{a}_x = a_1\vec{i}$, which is the part of vector \vec{a} along the x -axis; the

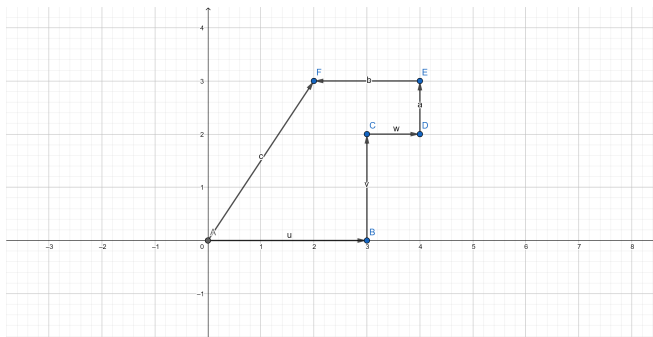


Figure 1.9:

y -component $\vec{a}_y = a_2\vec{j}$, which is the part of vector \vec{a} along the y -axis; and the z -component $\vec{a}_z = a_3\vec{k}$, which is the part of vector \vec{a} along the z -axis. A vector in three-dimensional space is the vector sum of its three vector components:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}.$$

Three unit vectors \vec{i} , \vec{j} and \vec{k} define a Cartesian system in three-dimensional space: $\vec{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ is the base unit vector in the x -direction, $\vec{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ is the base unit vector in the y -direction, and $\vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ is the base unit vector in the z -direction.

Example 1.2.1. On the Figure 1.10 we have presented the vector $\vec{v} = \begin{pmatrix} 5 \\ 2 \\ 5 \end{pmatrix} = 5\vec{i} + 2\vec{j} + 5\vec{k}$.

If we know the coordinates $A(x_1, y_1, z_1)$ of the origin point of a vector and the coordinates $B(x_2, y_2, z_2)$ of the end point of a vector, we can obtain the scalar components of a vector simply by subtracting the origin point

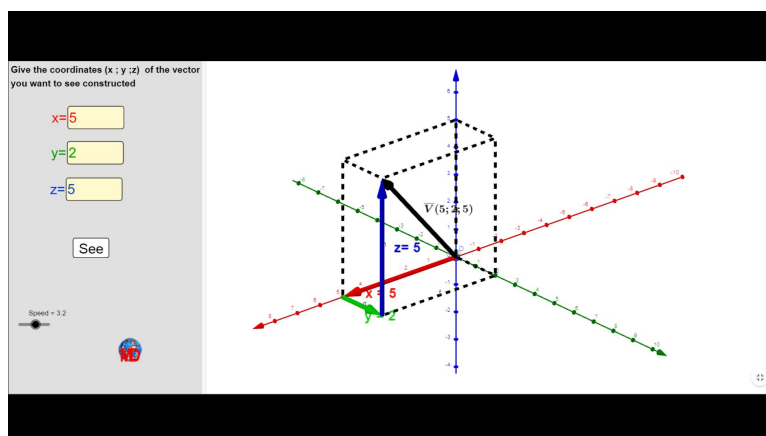


Figure 1.10:

coordinates from the end point coordinates, and the vector will be:

$$\vec{AB} = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{pmatrix}.$$

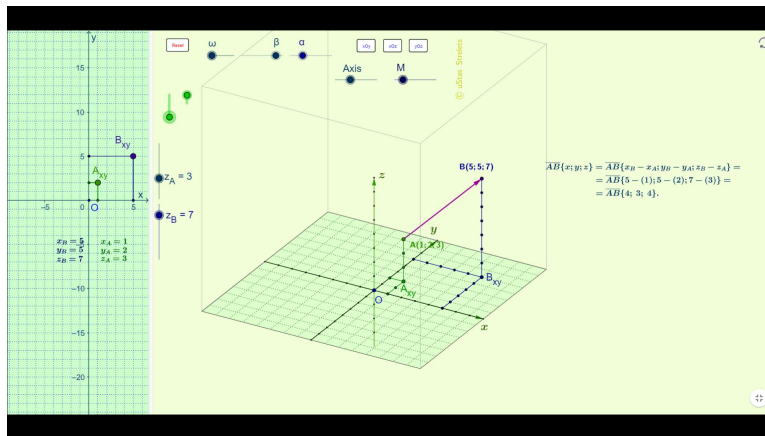
The magnitude (or sometimes we call it the length) of the vector $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ is:

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

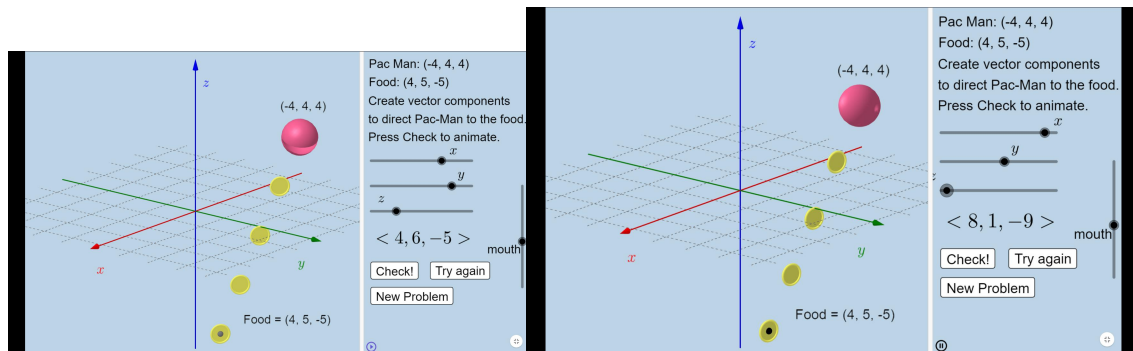
Example 1.2.2. A fly enters through an open window and zooms around the room. In a Cartesian coordinate system with three axes along three edges of the room, the fly changes its position from point $A(1, 2, 3)$ to point $B(5, 5, 7)$. Find the scalar components of the fly's displacement vector and express its displacement vector in vector component form. What is its magnitude?

The displacement vector will be:

$$\vec{AB} = \begin{pmatrix} 5 - 1 \\ 5 - 2 \\ 7 - 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 4 \end{pmatrix}.$$

$$|\vec{AB}| = \sqrt{(4)^2 + 3^2 + (4)^2} = \sqrt{16 + 9 + 16} = \sqrt{41}.$$


In Geogebra (see <https://www.geogebra.org/m/j28dfqtm>) there is a Pac Man demonstration on Vector Components (by Tim Brzezinski). It is an excellent exercise on creating the vector between two point, as you can see in the following videos.



Two vectors are equal when their corresponding scalar components are equal.

Resolving vectors into their scalar components (i.e., finding their scalar components) and expressing them analytically in vector component form allows us to use vector algebra to find sums or differences of many vectors analytically (i.e., without using graphical methods). For example, to find the resultant of two vectors \vec{a} and \vec{b} , we simply add them component by component. If $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ then:

$$\vec{a} + \vec{b} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix}.$$

If λ is a given scalar, then we have:

$$\lambda \vec{a} = \lambda \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix}.$$

Analytical methods for finding the resultant and, in general, for solving vector equations are very important in physics because many physical quantities are vectors. For example, we use this method in kinematics to find resultant displacement vectors and resultant velocity vectors, in mechanics to find resultant force vectors and the resultants of many derived vector quantities, and in electricity and magnetism to find resultant electric or magnetic vector fields.

Unit vector.

Given a non zero vector \vec{a} its magnitude $|\vec{a}|$ is a scalar quantity. If we multiply \vec{a} with the scalar $\frac{1}{|\vec{a}|}$ we obtain the parallel vector to $\frac{1}{|\vec{a}|}\vec{a}$. The magnitude (length) of this vector is 1:

$$\left| \frac{1}{|\vec{a}|}\vec{a} \right| = \frac{1}{|\vec{a}|}|\vec{a}| = 1,$$

so the vector $\frac{1}{|\vec{a}|}\vec{a}$ is a unit vector in the direction of \vec{a} , see Figure 1.12. We usually denote this vector with \vec{a}_0 .

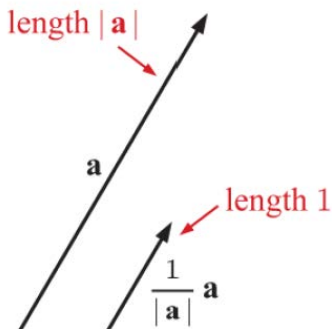


Figure 1.12:

Example 1.3.1. *At one point in space, the direction of the electric field vector is given in the Cartesian system by the unit vector $\vec{E}_0 = \frac{1}{\sqrt{5}}\vec{i} + \frac{2}{\sqrt{5}}\vec{j}$. If the magnitude of the electric field vector is $E = 400$ V/m, what are the scalar components E_x , E_y and E_z of the electric field vector \vec{E} at this point? What is the direction angle θ of the electric field vector at this point?*

Since the unit vector is given, we know the direction of \vec{E} , and we have $\vec{E} = |\vec{E}|\vec{E}_0 = 400\vec{E}_0 = 400 \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{pmatrix}$. For the direction angle (simple trigonometry) we have $\tan \theta = \frac{160}{80}$ and $\theta = \arctan 2$.

Example 1.3.2. *Two forces are given: $\vec{F}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and \vec{F}_2 . Second force has a magnitude 6 and direction same as the vector $\vec{a} = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$. Both of the forces act on one particle. Find the direction of the third force, so that the resulting force is 0?*

First we will determine the second force \vec{F}_2 . Using the fact that \vec{F}_2 has

the direction of the vector $\vec{a} = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$ we have:

$$\vec{F}_2 = \alpha \vec{a} = \begin{pmatrix} \alpha \\ 2\alpha \\ -2\alpha \end{pmatrix}.$$

Next, to determine α we use the known magnitude of \vec{F}_2 and we have:

$$\sqrt{\alpha^2 + 4\alpha^2 + 4\alpha^2} = 3|\alpha| = 6.$$

From this, $\alpha = 2$ and $\vec{F}_2 = \begin{pmatrix} 2 \\ 4 \\ -4 \end{pmatrix}$.

At the end, we want to determine the third force \vec{F}_3 . Then:

$$\vec{F}_1 + \vec{F}_2 + \vec{F}_3 = \vec{0},$$

and from this vector equation we have:

$$\vec{F}_3 = -\vec{F}_1 - \vec{F}_2 = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} - \begin{pmatrix} 2 \\ 4 \\ -4 \end{pmatrix} = \begin{pmatrix} -3 \\ -5 \\ 3 \end{pmatrix}.$$

Chapter 2

Products of two vectors

A vector can be multiplied by another vector but may not be divided by another vector. There are two kinds of products of vectors used broadly in physics and engineering. One kind of multiplication is a scalar multiplication of two vectors. Taking a scalar product of two vectors results in a number (a scalar), as its name indicates. Scalar products are used to define work and energy relations. For example, the work that a force (a vector) performs on an object while causing its displacement (a vector) is defined as a scalar product of the force vector with the displacement vector. A quite different kind of multiplication is a vector multiplication of vectors. Taking a vector product of two vectors returns as a result a vector, as its name suggests. Vector products are used to define other derived vector quantities. For example, in describing rotations, a vector quantity called torque is defined as a vector product of an applied force (a vector) and its distance from pivot to force (a vector). It is important to distinguish between these two kinds of vector multiplications because the scalar product is a scalar quantity and a vector product is a vector quantity.

2.1 The scalar product of two vectors

The scalar product of two vectors is also known as dot product or inner product. It is an operation between two vectors that results with a scalar. Be careful not to confuse the scalar product (a product of two vectors that give a scalar) with the scalar multiplication (a product of a scalar and a vector that gives a parallel vector).

If two vectors $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ are given, their scalar product is:

$$\vec{a} \cdot \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3.$$

Scalar multiplication of vectors is commutative, associative and obeys the distributive law. Other important property is that if we multiply \vec{a} with \vec{a} , then the scalar we obtain is the squared length of the vector \vec{a} , i.e.:

$$\vec{a} \cdot \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1^2 + a_2^2 + a_3^2 = |\vec{a}|^2.$$

We can use the scalar product of two vectors to determine the angle between them. Suppose θ is the angle between \vec{a} and \vec{b} , see the Figure 2.1. Using the

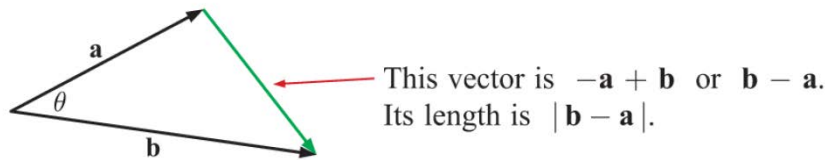


Figure 2.1:

cosine rule

$$|\vec{b} - \vec{a}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2|\vec{a}||\vec{b}|\cos\theta$$

we have that:

$$\begin{aligned} (b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2 &= a_1^2 + a_2^2 + a_3^2 + b_1^2 + b_2^2 + b_3^2 - 2|\vec{a}||\vec{b}|\cos\theta \\ 2a_1b_1 + 2a_2b_2 + 2a_3b_3 &= 2|\vec{a}||\vec{b}|\cos\theta \\ \vec{a} \cdot \vec{b} &= |\vec{a}||\vec{b}|\cos\theta \end{aligned}$$

From the last equation we see that the angle between two vectors can be found by following relation:

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|}.$$

Now, because of the properties of the cos-function we see that for the nonzero vectors \vec{a} and \vec{b} we have

- $\vec{a} \cdot \vec{b} = 0 \Leftrightarrow \vec{a}$ and \vec{b} are perpendicular.
- $\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \Leftrightarrow \vec{a}$ and \vec{b} are parallel vectors.
- If θ is the angle between \vec{a} and \vec{b} then for θ acute angle we have $\vec{a} \cdot \vec{b} > 0$ and for θ obtuse angle we have $\vec{a} \cdot \vec{b} < 0$.

Example 2.1.1. Find the angle between vectors $\vec{a} = 3\vec{i} - 4\vec{j}$ and $\vec{b} = 3\vec{i} + 4\vec{j}$.

We find the scalar product of the vectors, and their magnitudes.

$$\begin{aligned}\vec{a} \cdot \vec{b} &= 3 \cdot 3 + (-4) \cdot 4 = 9 - 16 = -7 \\ |\vec{a}| &= \sqrt{3^2 + (-4)^2} = 5 \\ |\vec{b}| &= \sqrt{3^2 + 4^2} = 5.\end{aligned}$$

The angle between the vectors is:

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|} = \frac{-7}{25}.$$

Angle between two forces

The angle between two forces is the angle between the vector of the forces.

Example 2.1.2. Two dogs are pulling on a stick in different directions. The first dog pulls with force $\vec{F}_1 = 10\vec{i} - 20.4\vec{j} + 2\vec{k}$ N, the second dog pulls with force $\vec{F}_2 = -15\vec{i} - 6.2\vec{k}$ N. What is the angle between \vec{F}_1 and \vec{F}_2 ?

We will use the scalar product of \vec{F}_1 and \vec{F}_2 to find the angle between the forces.

$$\vec{F}_1 \cdot \vec{F}_2 = 10 \cdot (-15) - 20.4 \cdot 0 - 2 \cdot 6.2 = -162.4.$$

On the other hand we have:

$$\begin{aligned}|\vec{F}_1| &= \sqrt{10^2 + (-20.4)^2 + 2^2} = \sqrt{520.16} = 22.81 \\ |\vec{F}_2| &= \sqrt{(-15)^2 + (-6.2)^2 + 0^2} = \sqrt{263.44} = 16.23.\end{aligned}$$

At the end,

$$\cos \theta = \frac{\vec{F}_1 \cdot \vec{F}_2}{|\vec{F}_1||\vec{F}_2|} = \frac{-162.4}{22.81 \cdot 16.23} = -0.437,$$

and the angle is $\theta = \arccos(-0.437) = 116^\circ$.

A work of a force. When force \vec{F} pulls on an object and when it causes its displacement \vec{d} , we say the force performs work. The amount of work the force does is the scalar product $\vec{F} \cdot \vec{d}$.

Example 2.1.3. *If the stick in the example before moves momentarily and gets displaced by vector $\vec{d} = -7.9\vec{j} - 4.2\vec{k}$ cm, how much work is done by the second dog?*

With the application of the dot product for the work done by the second dog we got:

$$\vec{F}_2 \cdot \vec{d} = (-15) \cdot 0 + 0 \cdot (-7.9) + (-6.2) \cdot (-4.2) = 26.04 \text{ Ncm}.$$

2.2 The vector product of two vectors

The vector product of two vectors is also known as cross product. It is an operation between two vectors that results with a third vector. The vector product is a vector that has its direction perpendicular to both vectors \vec{a} and \vec{b} . In other words, vector $\vec{a} \times \vec{b}$ is perpendicular to the plane that contains vectors \vec{a} and \vec{b} as shown in Figure 2.2. The magnitude of the vector product

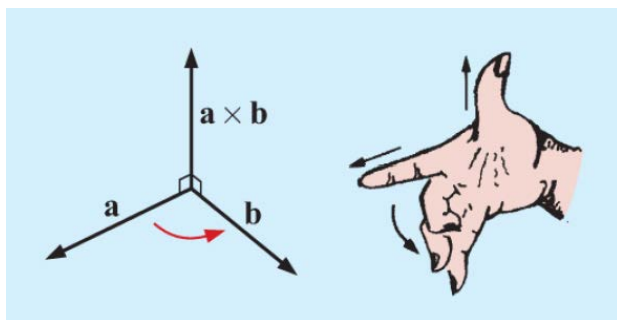


Figure 2.2:

is defined as $|\vec{a} \times \vec{b}| = |\vec{a}||\vec{b}| \sin \theta$, where θ is the angle between the vectors \vec{a}

and \vec{b} . On the line perpendicular to the plane that contains vectors \vec{a} and \vec{b} there are two alternative directions—either up or down, and the direction of the vector product may be either one of them. In the standard right-handed orientation, where the angle between vectors is measured counterclockwise from the first vector, vector $\vec{a} \times \vec{b}$ points upward. If we reverse the order of multiplication, so that now \vec{b} comes first in the product, then vector $\vec{b} \times \vec{a}$ must point downward. This means that vectors $\vec{a} \times \vec{b}$ and $\vec{b} \times \vec{a}$ are antiparallel to each other and that vector multiplication is not commutative.

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}.$$

Similar to the dot product, the cross product has the distributive property. For the base vectors we have following, see Figure 2.3.

$$\begin{aligned}\vec{i} \times \vec{j} &= \vec{k}, \vec{j} \times \vec{k} = \vec{i}, \vec{k} \times \vec{i} = \vec{j} \\ \vec{j} \times \vec{i} &= -\vec{k}, \vec{k} \times \vec{j} = -\vec{i}, \vec{i} \times \vec{k} = -\vec{j}.\end{aligned}$$

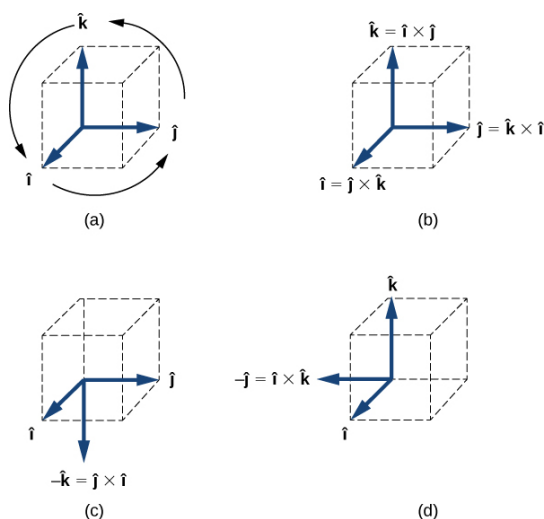


Figure 2.3:

If two vectors $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ are given, their vector

product (the component form) is:

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Using the definition above, we can easily conclude that two vectors \vec{a} and \vec{b} are parallel if and only if $\vec{a} \times \vec{b} = 0$. Also, this property can be easily proven from the relation $|\vec{a} \times \vec{b}| = |\vec{a}||\vec{b}|\sin\theta$, since \vec{a} and \vec{b} are parallel if and only if $\sin\theta = 0$, i.e., $\theta = 0$ or $\theta = \pi$. Special case is $\vec{a} \times \vec{a} = 0$.

The relation $|\vec{a} \times \vec{b}| = |\vec{a}||\vec{b}|\sin\theta$, is used very often to find the area of triangle or parallelogram. If the triangle has a defining vectors \vec{a} and \vec{b} , see Figure 2.4, then:

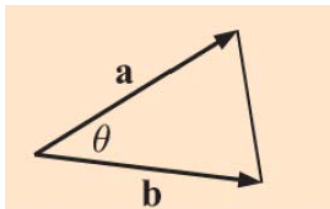


Figure 2.4:

$$P = \frac{1}{2}|\vec{a}||\vec{b}|\sin\theta = \frac{1}{2}|\vec{a} \times \vec{b}|.$$

From the last equation, if a parallelogram has a defining vectors \vec{a} and \vec{b} , its area will be $P = |\vec{a} \times \vec{b}|$.

Example 2.2.1. Given two vectors $\vec{a} = -\vec{i} + \vec{j}$ and $\vec{b} = 3\vec{i} - \vec{j}$ find $\vec{a} \times \vec{b}$, $|\vec{a} \times \vec{b}|$, the angle between \vec{a} and \vec{b} , and the angle between $\vec{a} \times \vec{b}$ and $\vec{c} = \vec{i} + \vec{k}$.

We have

$$\begin{aligned} \vec{a} \times \vec{b} &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -1 & 1 & 0 \\ 3 & -1 & 0 \end{vmatrix} = -2\vec{j}; \\ |\vec{a} \times \vec{b}| &= \sqrt{(-2)^2} = 2; \\ \cos\theta &= \frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|} = \frac{-3 - 1}{\sqrt{2}\sqrt{10}} = -\frac{4}{\sqrt{20}}; \\ \cos\alpha &= \frac{3}{2\sqrt{2}}. \end{aligned}$$

A Particle in a Magnetic Field

When moving in a magnetic field, some particles may experience a magnetic force. Without going into details let's acknowledge that the magnetic field \vec{b} is a vector, the magnetic force \vec{F} is a vector, and the velocity \vec{v} of the particle is a vector. The magnetic force vector is proportional to the vector product of the velocity vector with the magnetic field vector, which we express as $\vec{F} = \zeta \vec{v} \times \vec{b}$. In this equation, a constant ζ takes care of the consistency in physical units, so we can omit it here.

Example 2.2.2. *A particle moving in space with velocity vector $\vec{v} = -5\vec{i} - 2\vec{j} + 3.5\vec{k}$ enters a region with a magnetic field and experiences a magnetic force. Find the magnetic force \vec{F} on this particle at the entry point to the region where the magnetic field vector is $\vec{b} = 7.2\vec{i} - \vec{j} - 2.4\vec{k}$. find magnitude of the magnetic force and angle θ the force vector makes with the given magnetic field vector.*

We have:

$$\begin{aligned}\vec{F} = \vec{v} \times \vec{b} &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -5 & -2 & 3.5 \\ 7.2 & -1 & -2.4 \end{vmatrix} = 8.3\vec{i} + 13.2\vec{j} + 19.4\vec{k}; \\ |\vec{F}| &= \sqrt{8.3^2 + 13.2^2 + 19.4^2} = 24.9; \\ \cos \theta &= \frac{\vec{F} \cdot \vec{b}}{|\vec{F}||\vec{b}|} = \frac{8.3 \cdot 7.2 - 13.2 - 2.4 \cdot 19.4}{24.9 \cdot 7.6} = 0.\end{aligned}$$

Chapter 3

Vector application

There are many application of vectors in geometry. In three dimensional space, vector methods are very efficient, particularly when we consider the relationships between lines and planes.

3.1 Lines in space

We can determine the equation of a line in three dimensional space using its direction and any fixed point on the line. Suppose $R(x, y, z)$ is any point on the line, $A(a_1, a_2, a_3)$ is the known or fixed point on the line, and $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ is the direction vector of the line, see Figure 3.1.

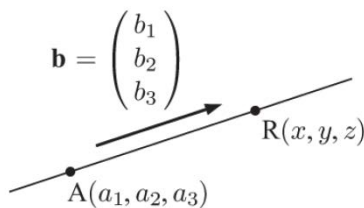


Figure 3.1:

The vector equation of the line is:

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \lambda \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

The parametric equation of the line is:

$$\begin{cases} x = a_1 + \lambda b_1 \\ y = a_2 + \lambda b_2 \\ z = a_3 + \lambda b_3 \end{cases}$$

where $\lambda \in \mathbb{R}$ is a parameter.

By equating λ values, we obtain the Cartesian equations of the line:

$$\frac{x - a_1}{b_1} = \frac{y - a_2}{b_2} = \frac{z - a_3}{b_3}.$$

Note that one line can be represented with different vectors (but collinear one) and different points, so the vector equation of a line is not unique.

Example 3.1.1. Find the parametric equations for the line through $A(2, -1, 4)$ and $B(-1, 0, 2)$.

We require the direction vector for the line, and this could be the vector \vec{AB} or \vec{BA} . We have that $\vec{AB} = \begin{pmatrix} -3 \\ 1 \\ -2 \end{pmatrix}$, and using point A the equations are:

$$\begin{cases} x = 2 - 3\lambda \\ y = -1 + \lambda \\ z = 4 - 2\lambda \end{cases}$$

We could use the point B to make the parametric equations of the line.

Angle between lines. The angle between two lines is the angle between its directional vectors. We can find the angle using the scalar product of the vectors. If \vec{b}_1 and \vec{b}_2 are directional vectors of the lines L_1 and L_2 , then the angle between them will be:

$$\cos \theta = \frac{\vec{b}_1 \cdot \vec{b}_2}{|\vec{b}_1||\vec{b}_2|}.$$

Example 3.1.2. Find the angle between the lines

$$\begin{aligned} L_1 : \quad \vec{r} &= \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ 1 \\ -2 \end{pmatrix} \\ L_2 : \quad \frac{1-x}{3} &= y = \frac{2-z}{2}. \end{aligned}$$

The line L_1 has a directional vector $\vec{b}_1 = \begin{pmatrix} -3 \\ 1 \\ -2 \end{pmatrix}$ and the line L_2 has a directional vector $\vec{b}_1 = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$. Since the vectors are colinear, the lines are parallel and the angle between them is 0.

Constant velocity problems. An object moving with a constant velocity will travel in a straight line. To model the position using vectors:

- the velocity vector of the motion gives the direction vector of the line,
- time is the parameter,
- the initial position of the object gives a fixed point in the line.

The speed of the object is the magnitude of the velocity vector.

Example 3.1.3. A mermaid is initially at the point $A(15, 8, -1)$ and she swims with the velocity vector $\vec{v} = 3\vec{i} - 5\vec{j} - 4.5\vec{k}$ m/s. Find: the position of the mermaid after t seconds, the speed of the mermaid, the time when the mermaid reaches 28 meters depth.

Since we know one point of the line (initial coordinates) and the direction vector (velocity vector), after t seconds the mermaid will be at the position:

$$\vec{r}(t) = \begin{pmatrix} 15 + 3t \\ 8 - 5t \\ -1 - 4.5t \end{pmatrix}.$$

The speed of the mermaid is $|\vec{v}| = \sqrt{3^2 + (-5)^2 + (-4.5)^2} = 7.37 \text{ m/s}$. The mermaid reaches 28 meters of depth when $z(t) = -28$. We have that $-1 - 4.5t = -28$ and from here $t = 6$ seconds.

A shortest distance from a point to the line. We will consider this item through a following example.

Example 3.1.4. Consider point $P(-1, 2, 3)$ and the line with parametric equations $L_1 : x = 1 + 2t, y = -4 + 3t, z = 3 + t$. Find the shortest distance from P to L_1 .

Let us have a point A at the given line. Then its coordinates will be $A(1 + 2t, -4 + 3t, 3 + t)$. Then the vector \vec{AP} will be:

$$\vec{AP} = \begin{pmatrix} -1 - 1 - 2t \\ 2 + 4 - 3t \\ 3 - 3 - t \end{pmatrix} = \begin{pmatrix} -2 - 2t \\ 6 - 3t \\ -t \end{pmatrix}.$$

If A is the closest point of the line to P , then $\vec{AP} \perp \vec{b}$ and their scalar product equals 0. Since $\vec{b} = 2\vec{i} + 3\vec{j} + \vec{k}$, from the condition $\vec{AP} \cdot \vec{b} = 0$, we have:

$$\vec{AP} \cdot \vec{b} = \begin{pmatrix} -2 - 2t \\ 6 - 3t \\ -t \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = -4 - 4t + 18 - 9t - t = 0.$$

At the end we got that $t = 1$ and $A(3, -1, 4)$. The wanted distance is the magnitude of the vector \vec{AP} ,

$$|\vec{AP}| = \left| \begin{pmatrix} -4 \\ 3 \\ -1 \end{pmatrix} \right| = \sqrt{16 + 9 + 1} = \sqrt{26}.$$

Relationships between lines. Lines in the space are coplanar if they are in the same plane. In this case they can be intercepting, parallel or coincident, see Figure 3.2. Let two lines L_1 and L_2 be given with their direction vectors \vec{b}_1 and \vec{b}_2 respectively. We can check if the lines are coplanar or not by evaluating the dot product of the vector \vec{AB} (A is a point on L_1 and B is a point on L_2), and the vector $\vec{b}_1 \times \vec{b}_2$. If this product is 0, then the lines are coplanar, if not the lines skew. We calculate the angle between two lines in space using the scalar product of the direction vectors of the lines. If the lines are parallel, the angle between them is 0, otherwise, the angle between them is θ .

If the lines are not coplanar they are skew, see Figure 3.3. If the lines are skew, then the angle between them is again the angle between the directional vectors.

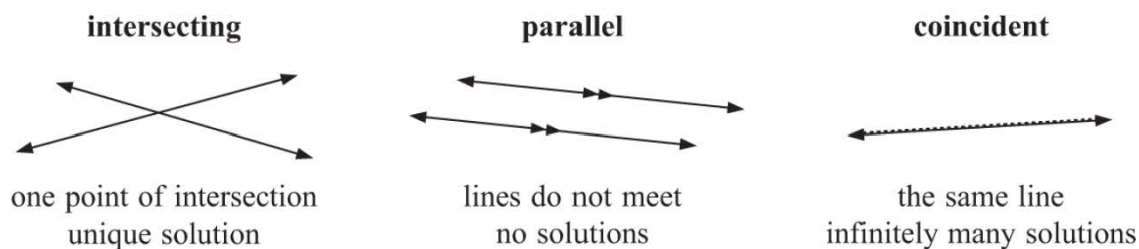


Figure 3.2:

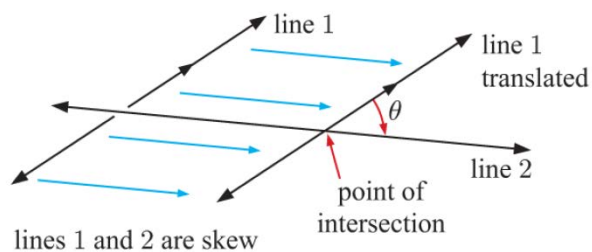


Figure 3.3:

Example 3.1.5. Consider the lines:

$$L_1 : \quad \vec{r} = \begin{pmatrix} 5 \\ 4 \\ 4 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ -3 \\ 4 \end{pmatrix}$$

$$L_2 : \quad \vec{r} = \begin{pmatrix} 1 \\ -4 \\ 14 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}.$$

Show that these lines intersect, and give the coordinates of the intersection point.

The direction vectors of the lines are: $\vec{b}_1 = \begin{pmatrix} -2 \\ -3 \\ 4 \end{pmatrix}$ and $\vec{b}_2 = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}$.

We have:

$$\vec{b}_1 \times \vec{b}_2 = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -2 & -3 & 4 \\ 0 & -2 & 2 \end{vmatrix} = 2\vec{i} + 4\vec{j} + 4\vec{k}.$$

The vector $\vec{AB} = -4\vec{i} - 8\vec{j} + 10\vec{k}$, so for the dot product we have:

$$\vec{AB} \cdot \vec{b}_1 \times \vec{b}_2 = 2(-4) + 4(-8) + 4 \cdot 10 = 0.$$

The lines are coplanar, but not parallel (since the direction vectors are not colinear), and they intercept. The point of interception is on the both lines, but it is got for a diferent parameter. That is why we have the sysstem:

$$\begin{aligned} 5 - 2\lambda &= 1 \\ 4 - 3\lambda &= -4 - 2\mu \\ 4 + 4\lambda &= 14 + 2\mu \end{aligned}$$

We got that $\lambda = 2$ and $\mu = -1$, so the interception point is $(1, -2, 12)$.

The shortest distance between lines. If lines are parallel, we choose any point at one of the lines and find the shortest distance to the other line, see Figure 3.4. If the lines intercept, the distance between them is zero. If

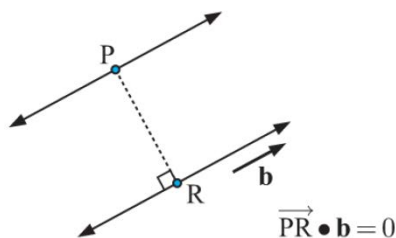


Figure 3.4:

the lines are skew, we need to find two points A and B , one on each line, so the vector \vec{AB} is perpendicular to both skew lines, see figure 3.5. It is clear that \vec{AB} is parallel to $\vec{b}_1 \times \vec{b}_2$, and we have $\vec{AB} = k \cdot \vec{b}_1 \times \vec{b}_2$ where \vec{b}_1 and \vec{b}_2 are the direction vector of the lines. The distance is the length of \vec{AB} .

Example 3.1.6. Find the shortest distance between the skew lines $x = t, y = 1 - t, z = 2 + t$ and $x = 3 - t, y = -1 + 2t, z = 4 - t$.

The direction vectors of the lines are: $\vec{b}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ and $\vec{b}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$.

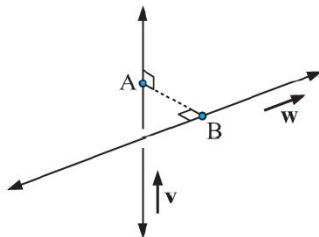


Figure 3.5:

For the vector product we have:

$$\vec{b}_1 \times \vec{b}_2 = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & -1 & 1 \\ -1 & 2 & -1 \end{vmatrix} = -\vec{i} + \vec{k}.$$

Let A and B be two points, one on each line, so the distance $|\vec{AB}|$ is the shortest between the lines. Now $\vec{AB} \parallel \vec{b}_1 \times \vec{b}_2$, i.e.:

$$\begin{pmatrix} 3 - \mu - \lambda \\ -1 + 2\mu - 1 + \lambda \\ 4 - \mu - 2 - \lambda \end{pmatrix} = k \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

From this system we have that $\lambda = 3$, so $A = (3, -2, 5)$ and $\mu = \frac{-1}{2}$, so $B(\frac{7}{2}, -2, -\frac{9}{2})$. The shortest distance between the skew lines is:

$$|\vec{AB}| = \left| \begin{pmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{pmatrix} \right| = \sqrt{\frac{1}{4} + \frac{1}{4}} = \frac{1}{\sqrt{2}}.$$

3.2 Planes

Let $A(a_1, a_2, a_3)$ is a point from a given plane, and $P(x, y, z)$ is any point from that plane. Then the vector $\vec{AP} = \begin{pmatrix} x - a_1 \\ y - a_2 \\ z - a_3 \end{pmatrix}$ lies entirely inside the

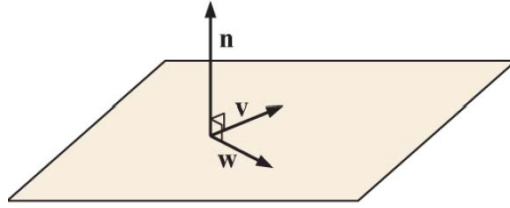


Figure 3.6:

plane. Let \vec{n} be a vector with direction perpendicular to that of the plane, see Figure 3.6. Since \vec{n} and \vec{AP} are perpendicular, their scalar product is zero. We have:

$$\vec{n} \cdot \vec{AP} = \begin{pmatrix} A \\ B \\ C \end{pmatrix} \cdot \begin{pmatrix} x - a_1 \\ y - a_2 \\ z - a_3 \end{pmatrix} = A(x - a_1) + B(y - a_2) + C(z - a_3) = 0,$$

is the plane equation.

Example 3.2.1. Find the equation of the plane with a normal vector $\vec{n} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ which passes through the point $A(-1, 2, 4)$.

Since $\vec{AP} = \begin{pmatrix} x + 1 \\ y - 2 \\ z - 4 \end{pmatrix}$, for the equation of the plane we have:

$$\vec{n} \cdot \vec{AP} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} x + 1 \\ y - 2 \\ z - 4 \end{pmatrix} = (x + 1) + 2(y - 2) + 3(z - 4) = 0,$$

i.e. $x + 2y + 3z - 15 = 0$. When a line is given, it can meet a plane in some point, it can be parallel to the plane, or it can entirely lie in the plane.

Example 3.2.2. Find the parametric equations of the line through $A(-1, 2, 3)$ and $B(2, 0, -3)$. Hence find where this line meets the plane with the equation $x - 2y + 3z = 26$.

$\vec{AB} = \begin{pmatrix} 3 \\ -2 \\ -6 \end{pmatrix}$ and the parametric equations of the line through A and B is $x = -1 + 3t, y = 2 - 2t, z = 3 - 6t$. The plane meets the line when:

$$\begin{aligned} x - 2y + 3z &= 26 \\ -1 + 3t - 2(2 - 2t) + 3(3 - 6t) &= 26 \\ 4 - 11t &= 26 \\ t &= -2. \end{aligned}$$

Substituting $t = -2$ in the line equation, we got the intersection point of the line and the plane, $(-7, 6, 15)$.

Distance from a point to a plane. Let d be the distance from a point $P_1(x_1, y_1, z_1)$ to a given plane $Ax + By + Cz = D$, see Figure 3.7. Let Q be

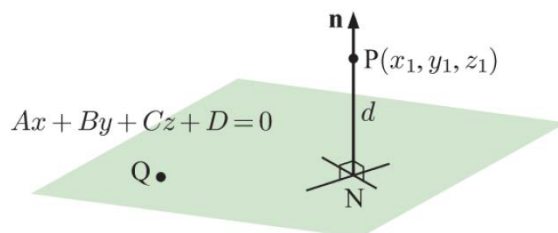


Figure 3.7:

any point from that plane. Then $d = \frac{|\vec{QP} \cdot \vec{n}|}{|\vec{n}|}$. From this equation we have:

$$d = \frac{|A(x - x_1) + B(y - y_1) + C(z - z_1)|}{\sqrt{A^2 + B^2 + C^2}} = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}},$$

and this is how we evaluate the distance between a point and a plane.

Example 3.2.3. Show that the line $x = 2 + t, y = -1 + 2t, z = -3t$ is parallel to the plane $11x - 4y + z = 0$, and find its distance from the plane.

When a line is parallel to a plane or it lies entirely in the plane its directional vector \vec{b} is perpendicular to the normal vector \vec{n} of the plane. In this case $\vec{b} \cdot \vec{n} = 1 \cdot 11 + 2 \cdot (-4) + (-3) \cdot 1 = 0$, so this condition is fulfilled.

To check whether the line lies in the plane or it is parallel to it, we take any point from the line and check if it belongs to the plane. So,

$$\begin{aligned} 11x - 4y + z &= 0 \\ 11(2+t) - 4(-1+2t) - 3t &= 0 \\ 22 &= 0. \end{aligned}$$

The last equation is not possible, so there is no point from the line that meets the plane. The line is parallel to the plane.

In order to find the distance, we pick any point from the line, let say for $t = 0$ we have $A(2, -1, 0)$ and using the distance formula we have:

$$d = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}} = \frac{|11 \cdot 2 - 4 \cdot (-1) + 0|}{\sqrt{11^2 + (-4)^2 + 1^2}} = \frac{26}{\sqrt{138}}.$$

Angles in space. The angle between line and a plane is the angle between the direction vector of the line and the normal vector of the plane. The angle between two planes in space is the angle between the normal vectors of the planes. We can use the scalar product to find these angles.

Example 3.2.4. Find the angle between the planes with equations $x + y - z = 8$ and $2x - y + 3z = -1$.

Since $\vec{n}_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ and $\vec{n}_2 = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ we have:

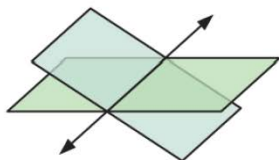
$$\cos \theta = \frac{\vec{n}_1 \cdot \vec{n}_2}{|\vec{n}_1||\vec{n}_2|} = \frac{|2 - 1 - 3|}{\sqrt{3}\sqrt{14}} = \frac{2}{\sqrt{3}\sqrt{14}}.$$

Intercepting planes. Two planes in space can intercept, be parallel, or coincident, see Figure 3.8. Three planes in space could have one of the following eight arrangements, see Figure 3.9. To determine the relationship between the planes, we have to solve linear system of three equation. We will illustrate this on the following example, using row operations.

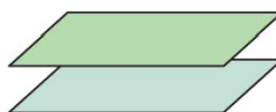
Example 3.2.5. Solve the following system of linear equations. Give a geometric interpretation of the system.

$$\begin{cases} 2x - y + z &= 5 \\ x + y - z &= 2 \\ 3x - 3y + 3z &= 8 \end{cases}.$$

(1) intersecting



(2) parallel



(3) coincident

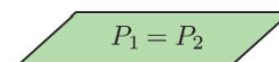
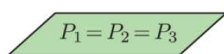
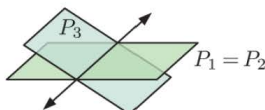


Figure 3.8:

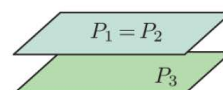
(1) all coincident



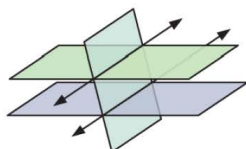
(2) two coincident and one intersecting



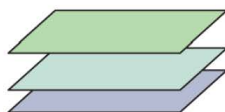
(3) two coincident and one parallel



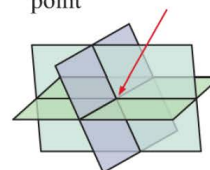
(4) two parallel and one intersecting



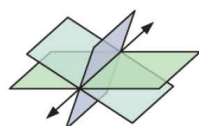
(5) all three parallel



(6) all meet at the one point



(7) all meet in a common line



(8) the line of intersection of any two is parallel to the third plane.

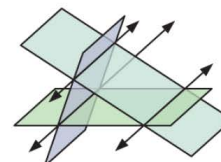


Figure 3.9:

In augmented matrix form the system is:

$$\begin{pmatrix} 1 & 1 & -1 & | & 2 \\ 2 & 1 & -1 & | & 5 \\ 3 & -3 & 3 & | & 8 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & -1 & | & 2 \\ 0 & -3 & 3 & | & 1 \\ 0 & -6 & 6 & | & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & -1 & | & 2 \\ 0 & 1 & -1 & | & -\frac{1}{3} \\ 0 & 0 & 0 & | & 0 \end{pmatrix}.$$

The row of zeroes indicates infinitely many solutions. If we let $z = t$ in the second equation, we have $y = t - \frac{1}{3}$. From the first equation we got $x = 2 - y + z = 2 - t + \frac{1}{3} + t = \frac{7}{3}$. The solution has form $x = \frac{7}{3}, y = t - \frac{1}{3}, z = t$, which is obviously a line in space. So, the system represents three planes that intersect in a line.

Chapter 4

Exit ticket

1. Which of the following is a vector: a person's height, the altitude on Mt. Everest, the velocity of a fly, the age of Earth, the boiling point of water, the cost of a book, Earth's population, or the acceleration of gravity?
2. Give a specific example of a vector, stating its magnitude, units, and direction.
3. Suppose you add two vectors \vec{a} and \vec{b} . What relative direction between them produces the resultant with the greatest magnitude? What is the maximum magnitude? What relative direction between them produces the resultant with the smallest magnitude? What is the minimum magnitude?
4. Is it possible for two vectors of different magnitudes to add to zero? Is it possible for three vectors of different magnitudes to add to zero? Explain.
5. Can a magnitude of a vector be negative?
6. If two vectors are equal, what can you say about their components? What can you say about their magnitudes? What can you say about their directions?
7. If three vectors sum up to zero, what geometric condition do they satisfy?

8. A delivery man starts at the post office, drives 40 km north, then 20 km west, then 60 km northeast, and finally 50 km north to stop for lunch. Use a graphical method to find his net displacement vector.
9. A small plane flies 40 km in a direction 60° north of east and then flies 30 km in a direction 15° north of east. Use a graphical method to find the total distance the plane covers from the starting point and the direction of the path to the final position.
10. Explain why a vector cannot have a component greater than its own magnitude.
11. If two vectors have the same magnitude, do their components have to be the same?
12. Suppose you walk 18.0 m straight west and then 25.0 m straight north. How far are you from your starting point? What is your displacement vector? What is the direction of your displacement? Assume the x -axis is horizontal to the right.
13. Given two vectors $\vec{a} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, and $\vec{b} = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}$, find the vectors $\vec{a} + \vec{b}$ and $\vec{a} - 2\vec{b}$. Determine the magnitudes of these vectors.
14. In the control tower at a regional airport, an air traffic controller monitors two aircraft as their positions change with respect to the control tower. One plane is a cargo carrier Boeing 747 and the other plane is a Airbus a380. The Boeing is at an altitude of 2500 m, climbing at 10° above the horizontal, and moving 30° north of west. The Airbus is at an altitude of 3000 m, climbing at 5° above the horizontal, and cruising directly west. (a) Find the position vectors of the planes relative to the control tower. (b) What is the distance between the planes at the moment the air traffic controller makes a note about their positions?
15. Find the angle between vectors $\vec{a} = 3\vec{i} - 4\vec{j} + \vec{k}$ and $\vec{b} = 3\vec{i} + 4\vec{j} - \vec{k}$.
16. Find the angle that the vector $\vec{d} = \vec{i} - 2\vec{j} + \vec{k}$ makes with the x , y and z axes.

17. Vectors \vec{a} and \vec{b} have the magnitudes of 5 units. Find the angle between them if $\vec{a} + \vec{b} = 5\sqrt{2}\vec{j}$.
18. If the cross product of two vectors vanishes, what can you say about their directions?
19. If the dot product of two vectors vanishes, what can you say about their directions?
20. What is the dot product of a vector with the cross product that this vector has with another vector?
21. If the velocity vector of a polar bear is $\vec{v} = -18\vec{i} - 13\vec{j}$ km/h, how fast and in what geographic direction is it heading? Here, \vec{i} and \vec{j} are directions to geographic east and north, respectively.
22. Given two vectors $\vec{a} = -\vec{i} + 2\vec{j} - \vec{k}$ and $\vec{b} = 3\vec{i} - \vec{k}$ find $\vec{a} \times \vec{b}$, $|\vec{a} \times \vec{b}|$, the angle between \vec{a} and \vec{b} , and the angle between $\vec{a} \times \vec{b}$ and $\vec{c} = \vec{i} + 2\vec{j}$.
23. Starting at the island of Moi in an unknown archipelago, a fishing boat makes a round trip with two stops at the islands of Noi and Poi. It sails from Moi for 4.76 nautical miles (nmi) in a direction 37° north of east to Noi. From Noi, it sails 69° west of north to Poi. On its return leg from Poi, it sails 28° east of south. What distance does the boat sail between Noi and Poi? What distance does it sail between Moi and Poi? Express your answer both in nautical miles and in kilometers. Note: 1 nmi = 1852 m.
24. A particle moving in space with velocity vector $\vec{v} = -5\vec{i} - 2\vec{j} + 3.5\vec{k}$ enters a region with a magnetic field and experiences a magnetic force. Find the magnetic force \vec{F} on this particle at the entry point to the region where the magnetic field vector is $\vec{b} = 4\vec{k}$. find magnitude of the magnetic force and angle θ the force vector makes with the given magnetic field vector.
25. Show that $\vec{a} \cdot (\vec{b} \times \vec{c})$ is the volume of the parallelepiped, with edges formed by the three vectors.
26. Find the volume of the parallelepiped, with edges formed by the following vectors: $\vec{a} = -\vec{i} + \vec{j} + 3\vec{k}$, $\vec{b} = 4\vec{j}$ i $\vec{c} = -3\vec{i} + 2\vec{j} + 1\vec{k}$.

27. Find the equations for the line through $A(-2, -1, 4)$ and $B(1, 0, 3)$.
28. A helicopter at $A(6, 9, 3)$ moves with constant velocity. Ten minutes later it is at $B(3, 10, 2.5)$. Distances are in kilometers. Find: \vec{AB} , the speed of the helicopter, and determine the line equation that represents the helicopter's path. The helicopter is traveling directly to its landing position with a z coordinate 0. Find the total time taken for the helicopter to land. At what angle to the horizontal is the helicopter flying?
29. A diver swims from $(12, 25, -20)$ back to his boat at $(7, -15, 0)$, at speed of 0.9 m/s . Find the velocity vector of the diver. An octopus watches the diver from his home at $(12, -8, -5)$. At what time the diver is closest to the octopus? Find the shortest distance from the octopus to the diver.
30. Suppose we have two particles in space. Particle A's position after t seconds is given by $x_A(t) = 5 - 2t, y_A(t) = 4 - 3t, z_A(t) = 4 + 4t$. Particle B's position after t seconds is given by $x_B(t) = 1, y_B(t) = -4 - 2t, z_B(t) = 14 + 2t$. All distance units are meters. Find the initial position of each particle. Find the velocity vector of each particle. Will the particles collide? Explain your answer.
31. Find the shortest distance between the skew lines $x = 1 + 2t, y = -t, z = 2 + 3t$ and $x = y = z$.
32. Find the equation of the plane through $A(-1, 2, 0), B(3, 1, 1)$ and $C((1, 0, 3))$.
33. Find the parametric equations of the line through $P(1, -2, 4)$ and $Q(2, 0, -1)$. Hence find where this line meets the yz plane, and plane with the equation $y + z = 2$.
34. Find the equations of two planes parallel to $2x - y_2z = 5$ and two units from it.
35. The planes $x + 5y - 3z = 8$ and $2x + 2y + kz = 9$ are perpendicular. Find the value of k .

36. Solve the following system of linear equations. Give a geometric interpretation of the system.

$$\begin{cases} x + 3y - z = 15 \\ 2x + y + z = 7 \\ x - y - 2z = 0 \end{cases}.$$

Poglavlje 1

Dynamic programming

One of the most important algorithm design techniques is dynamic programming (DP for short). The technique is among the most powerful for designing algorithms for optimization problems that have certain well-defined clean structural properties. DP is used in problems which solution can be interpreted as a result of a sequence of decisions (from vertex u go to u_1 or u_2 , put some item in the knapsack or don't, ...).

- "Resemblance" with divide-and-conquer method - breaks problems down into smaller subproblems.
- "Big difference" - in divide-and-conquer algorithms subproblems are disjoint, while in DP the subproblems overlap (one subproblem is "subset" of some previous subproblem).

Solving problems using dynamic programming rely on two important structural qualities, optimal substructure and overlapping subproblems.

- OPTIMAL SUBSTRUCTURE ... this property (sometimes called the principle of optimality) states that for the global problem to be solved optimally, each subproblem should be solved optimally.
- OVERLAPPING SUBPROBLEMS ... while it may be possible to subdivide a problem into subproblems in exponentially many different ways, these subproblems overlap each other in such a way that the number of distinct subproblems is reasonably small, ideally polynomial in the input size. An important issue is how to generate the solutions to these subproblems. There are two complementary (but essentially equivalent) ways of viewing how a solution is constructed.
 - TOP-DOWN ... solution to a DP problem is expressed recursively. This approach applies recursion directly to solve the problem. However, due to the overlapping nature of the subproblems, the same recursive call is often made many times. An approach, called MEMOIZATION, records the results of recursive calls, so that subsequent calls to a previously solved subproblem are handled by table look-up.

- BOTTOM-UP ... Although the problem is formulated recursively, the solution is built iteratively by combining the solutions to small subproblems to obtain the solution to larger subproblems. The results are stored in a table.

MEMOIZATION ... we create table of results of recursive calls, where results obtained for smaller subproblems are used for computing bigger subproblems. We can use those results multiple times.

Example 1. *Fibonacci numbers*

$$F(n) = F(n-1) + F(n-2), \quad n \geq 2,$$

$$F(0) = 1, F(1) = 1.$$

Problem: for given number $n \in \mathbb{N}$ compute $F(n)$.

PROCEDURE FIBONACCI (n)

BEGIN

IF ($n = 0$) *THEN RETURN*(0)

ELSEIF ($n = 1$) *THEN RETURN*(1)

ELSE RETURN (*FIBONACCI*($n-1$) + *FIBONACCI*($n-2$))

END;

Algorithm complexity is:

$$T(n) = T(n-1) + T(n-1) + O(1)$$

$$T(n) \geq T(n-1) + T(n-2), \quad \text{za } n \geq 2$$

→ T grows at least as fast as the Fibonacci numbers, that is exponentially quickly.

What is the problem? We call recursively many times the same thing, which is then computed every time.

For instance, for *FIBONACCI* (6) five times is called *FIBONACCI* (2), three times *FIBONACCI*(3), ...

MEMOIZATION idea:

PROCEDURE FIBONACCIDP1 (n)

BEGIN

$F = [0, 1, \infty, \infty, \dots, \infty]$

FOR $i = 2$ *TO* n *DO*

$F[i] = F[i-1] + F[i-2]$

END;

RETURN($F[n]$)

END;

Fibonacci numbers aren't an optimization problem, but they are a good example, where we see the idea of memoization. We don't have many same recursion calls, but rather we remember value for the first call, put it in table, and use it from the table.

1.1 Examples

Basic graph concepts have been presented in earlier courses, and so we will present a very quick review of the basic material for today's lecture. A graph $G = (V, E)$ is a structure that represents a discrete set of objects V , called vertices or nodes, and a set of pairwise relations E between these objects, called edges. Edges may be directed from one vertex to another or may be undirected.

Observe that multiple edges between the same two vertices are not allowed, but in a directed graph, it is possible to have two oppositely directed edges between the same pair of vertices.

Many classical problems can be represented using graphs and problems on them, for instance, travelling salesman problem ... "Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city?" It is an NP-hard problem in combinatorial optimization.

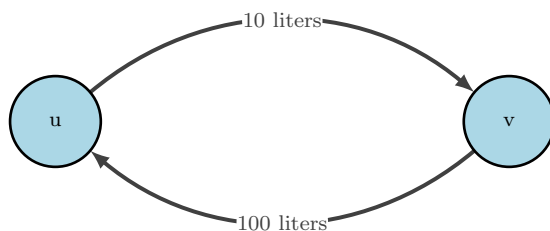
1.1.1 Shortest path

In graph theory, the shortest path problem is the problem of finding a path between two vertices (or nodes) in a graph such that the sum of the weights of its constituent edges is minimized.

The problem of finding the shortest path between two intersections on a road map may be modeled as a special case of the shortest path problem in graphs, where the vertices correspond to intersections and the edges correspond to road segments, each weighted by the length of the segment.

There are n cities given (marked by numbers from 1 to n). Between some cities there are direct connection, but some cities are not directly connected. We can represent this by graph G , where vertices or nodes are cities, and edges are direct connections between cities (set E).

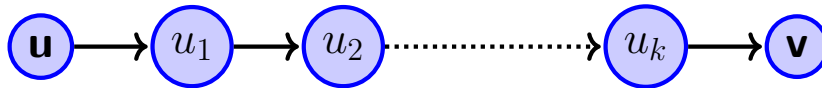
- One-way edges are allowed, that is ordered-pair (u, v) means "road goes from city u to city v ".
- Every direct connection has a length (price) $l(u, v)$. For instance, gas spent from city u in the mountain (with altitude 2000 meters) to the city v with altitude 100 meters (obviously, $l(u, v) \ll l(v, u)$).
→ directed graph



We have two problems:

- For cities u i v , is there a path from u to v ? That is a problem "is graph connected", which can be solved using greedy algorithms. We assume that graph is connected.
- Shortest path from u to v . That means, find sequence of cities u_1, u_2, \dots, u_k that are directly connected and constitute the path from u to v such that length of this path is minimum possible (obviously, between two cities can exist many different paths). Finding solution of this problem can be interpreted as sequence of decisions:

- in which city u_1 go from u ?
- in which city u_2 go from u_1 ?
- \vdots



Decisions are obviously not independent.

1.1.2 0 – 1 Knapsack

Imagine that a burglar breaks into a museum and finds n items. Let v_i denote the value of the i -th item, and let w_i denote the weight of the i -th item. The burglar carries a knapsack capable of holding total weight M . The burglar wishes to carry away the most valuable subset items subject to the weight constraint. For example, a burglar would rather steal diamonds before gold because the value per pound is better. But he would rather steal gold before lead for the same reason. We assume that the burglar cannot take a fraction of an object, so must make a decision to take the object entirely or leave it behind.

We have a knapsack which can carry weight M (that is knapsack of capacity M). We have n items with its weights and values.

We have to fill this backpack with items.

We need to choose for every item shall we bring it or not, so knapsack is not overfull, and the profit is maximal.

More formally, given $\{v_1, v_2, \dots, v_n\}$ and $\{w_1, w_2, \dots, w_n\}$ and $M > 0$ we wish to we wish to determine the subset $T \subseteq \{1, 2, \dots, n\}$ (of objects to "take") that maximizes

$$\sum_{i \in T} v_i$$

subject to constraint

$$\sum_{i \in T} w_i \leq M.$$

This is 0 – 1 knapsack problem.

$x_i \in \{0, 1\}, i = 1, \dots, n$, means that is we either take one item or we don't (0 denotes we didn't take it).

- Hard problem, at the moment doesn't exist polynomial algorithm for solving this problem.
- Decisions for solving: take i -th item or not. Number of possibilities.: $(1 + 1)(1 + 1) \cdots (1 + 1) = 2^n$.
- Idea: by smart choosing reduce number of possibilities that need to be checked.

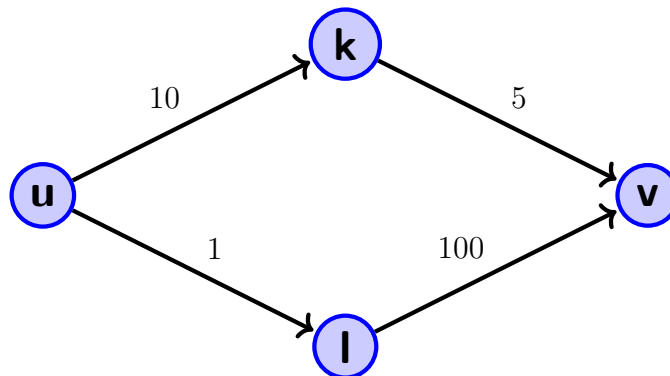
1.2 DP - ideas

Greedy algorithms - in each step we make the "best" decision and thus find the optimal sequence of decisions.

For many problems, it will not be possible to make decisions in individual steps based only on local information so that the entire sequence of decisions is optimal.

Shortest path We are looking for the shortest path from vertex u to vertex v . Let us denote by $A(u)$ the set of all cities, i.e. vertices we can reach directly from the city u .

Which of the cities from $A(u)$ should be next on the shortest path to v ? There is no way to decide this locally without looking at the rest of the graph, and to guarantee that future decisions lead to the optimal sequence of cities.



If we expand the problem and search for the shortest paths from u to all other cities we can reach, we can make a correct decision (for example, Dijkstra's greedy algorithm). However, if we look greedily only from u to v , we will not be able to make a correct decision.

0 – 1 ruksak Look at the following example for 0 – 1 knapsack.

i	w_i	p_i
1	4	3
3	2	2

Shall we take the first item or not? However we decide, we can make mistake. We don't know until we see what happened to the other items.

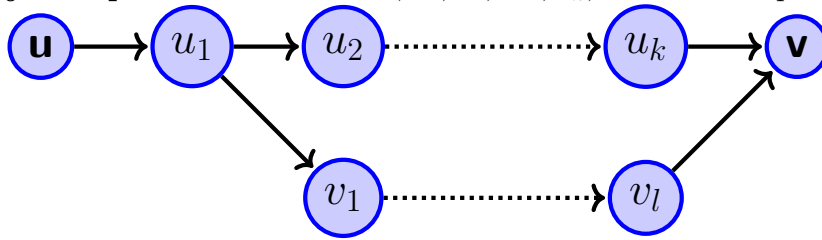
- a) $M = 5$ and we take the first item, i.e. $x_1 = 1$. You can't take anything else ($M - w_1 = 1$), so the total profit is 3. It is better to take the second and third item, so $x_1 = 0, x_2 = x_3 = 1$ and the total profit is $p_2 + p_3 = 4$.

- b) $M = 4$ and we take the first item. Obviously, we did not make a mistake in this case.

The optimal sequence of decisions (or more of them, if they exist) is reached by constant direct use of the so-called principle of optimality, if we find such a principle for our problem (which needs to be proven, but is usually obvious).

Theorem 1 (Principle of optimality, Bellman, 1957 g). *An optimal sequence of decisions has the property that for any initial state and initial decision in that state, the remaining decisions must form an optimal sequence as seen from the state after the first decision.*

Najkraći put Let's assume $u, u_1, u_2, \dots, u_k, v$ is shortest path from u to v .



If $u_1, v_1, v_2, \dots, v_l, v$ would be optimal path from u_1 to v (more "optimal" than u_1, u_2, \dots, u_k, v), then $u, u_1, v_1, v_2, \dots, v_l, v$ would be more "optimal" path from u to v than $u, u_1, u_2, \dots, u_k, v$, and that is contradiction.

0 – 1 ruksak It is obvious that principle holds.

Remark 1.

If the first decision was wrong, further optimality will not lead to an optimal solution. The issue of optimizing the first decision still remains, but only according to the optimal sequences after that decision.

Global sequences that have non-optimal subsequences are certainly not optimal if principle of optimality is found.

1.3 Shortest path

We have shown that the principle of optimality is valid for this problem and is formulated with the length of the shortest path. DP, i.e. the principle of optimality rejects all non-optimal paths from u_1 to v , where

$$u \rightarrow u_1 \rightarrow \dots \rightarrow u_k \rightarrow v.$$

Our algorithm for finding the shortest path from u to v therefore looks like this:

- cities u_1 may be more and all those initial decisions should be checked, because we do not know which one is optimal.
- from any u_1 we need to find the shortest path to v .

- take the shortest total path through all u_1

Let's denote by $ld(u, v)$ the length of the optimal path from u to v , for any two cities, i.e. vertices u, v (ld , stands for least distance).

$$\rightarrow ld(u, v) = \min_{u_1 \text{ such that } (u, u_1) \in E} \{l(u, u_1) + ld(u_1, v)\}.$$

Remark 2. • *Finding the shortest path from u_1 to v does not depend on the decision in the first step, it is a problem that we solve independently without knowing that we have chosen u_1 from u — PRINCIPLE OF INVARIANCE.*

- *The problem of finding the shortest path from u_1 to v is a problem of the same type as the starting problem, only with other cities (" $ld(u_1, v) \subseteq ld(u, v)$ ") — OVERLAPPING PRINCIPLE. We can apply the same algorithm recursively for the subproblem.*

If we denote $u_0 = u$, in every step following is valid

$$ld(u_k, v) = \min_{u_{k+1} \text{ such that } (u_k, u_{k+1}) \in E} \{l(u_k, u_{k+1}) + ld(u_{k+1}, v)\}.$$

Our algorithm will be solving the problem backward, from v (more natural), i.e. using BOTTOM-UP method.

At the end of the paths are those vertices that are directly connected to v , i.e. for which

$$ld(u_k, v) = l(u_k, v)$$

is valid, and that is the given data.

There are at most $n - 1$ steps (because there are n cities), and we stop when we backward come across $u_k = u$.

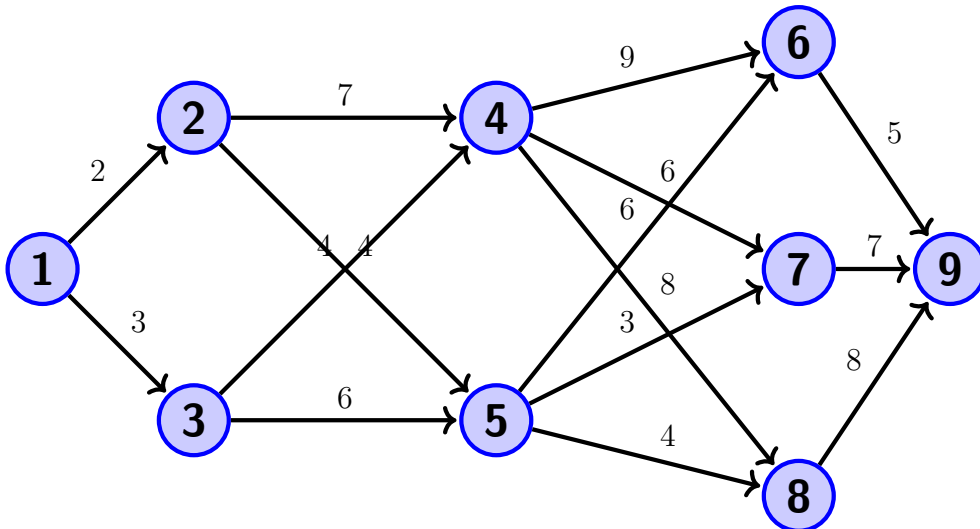
Obviously, there must be no repetition of cities on the path, i.e. cycles in the graph (due to the minimality of the path).

During realization, we assume that the given directed graph is layered, i.e. the vertices are divided into $m + 1$ disjoint sets, $m \geq 1$.

We denote those sets by V_0, V_1, \dots, V_m .

Direct connections exist only between vertices in adjacent layers V_k, V_{k+1} (and not every pair has to be connected).

We will look at the following example



Sources ... nodes from V_0 .

Sinks ... nodes from V_m .

The problem in the layered graph states: for a given source $u \in V_0$ and sink $v \in V_m$ find shortest path $ld(u, v)$.

The advantage of layers: each path from the source to the sink has exactly m sections, i.e. one node in every layer.

The recursive equation becomes

$$ld(u_k, v) = \min_{u_{k+1} \in V_{k+1} \text{ such that } (u_k, u_{k+1}) \in E} \{l(u_k, u_{k+1}) + ld(u_{k+1}, v)\}$$

for $k = 0, 1, \dots, m-2$, with distances for last two layers

$$ld(u_{m-1}, v) = l(u_{m-1}, v), \text{ for all } u_{m-1} \text{ such that } (u_{m-1}, v) \in E.$$

Example 2. In example find $ld(1, 9)$.

Rješenje. $1 \rightarrow 2 \rightarrow 5 \rightarrow 7 \rightarrow 9, ld(1, 9) = 16$. ■

1.3.1 Algorithm

We use one vertex from the first and last layer, so we can assume $|V_0| = |V_m| = 1$.

- Vertices in layers are numbered in order by layer from 1 to n , where the source number is 1, the sink number is n , and the vertices from layer V_{k+1} have bigger number then vertices from layer V_k .
- Beside length of the optimal path, we want additionally from the algorithm vertices of which that path consists.
- We are looking for $ld(1, n)$.

Input is layered graph with $m+1$ layers, n vertices indexed in ascending order by layers. Obviously, it is sufficient to specify E set of edges along with $d: E \rightarrow \mathbb{R}_0^+$, length of vertices.

The output is a field P with m elements with indices of vertices (cities) on the shortest path from 1 to n , and l the length of the shortest path.

ld is the field of values shortest paths from node j to node n .

Algorithm - pseudocode

FUNCTION MIN-PATH ($E, m, n, d: E \rightarrow \mathbb{R}_0^+$)

BEGIN

$ld(n) = 0;$

FOR $j = n-1$ TO 1 DO

BEGIN

find node r such that $(j, r) \in E$ and $d(j, r) + ld(r)$ is minimal over all such r

$ld(j) = d(j, r) + ld(r)$... shortest path from j to n

$ind(j) = r$... number of following node from node j

in the shortest path from j to n

END;

$l = ld(1);$

```

    P(1) = ind(1);
    FOR j = 2 TO m DO
    BEGIN
        P(j) = ind(P(j - 1));
    END;
    RETURN(l);
    RETURN(P);
END;

```

Example 3. *In example:*

$n = 9, m = 4.$

$ld(9) = 0$

$j = 8 \rightarrow$	$ld(8) = 8$	$ind(8) = 9$
$j = 7 \rightarrow$	$ld(7) = 7$	$ind(7) = 9$
$j = 6 \rightarrow$	$ld(6) = 5$	$ind(6) = 9$
$j = 5 \rightarrow$	$ld(5) = 4 + ld(8) = 12$	
	$3 + ld(7) = 10$	$ind(5) = 7$
	$6 + ld(6) = 11$	
$j = 4 \rightarrow$	$ld(4) = 8 + ld(8) = 16$	
	$6 + ld(7) = 13$	$ind(4) = 7$
	$9 + ld(6) = 14$	
$j = 3 \rightarrow$	$ld(3) = 6 + ld(5) = 16$	$ind(3) = 5$
	$4 + ld(4) = 17$	
$j = 2 \rightarrow$	$ld(2) = 4 + ld(5) = 14$	$ind(2) = 5$
	$7 + ld(4) = 17$	
$j = 1 \rightarrow$	$ld(1) = 3 + ld(3) = 19$	
	$2 + ld(2) = 16$	$ind(1) = 2$

$l = ld(1) = 16$

$P(1) = ind(1) = 2$

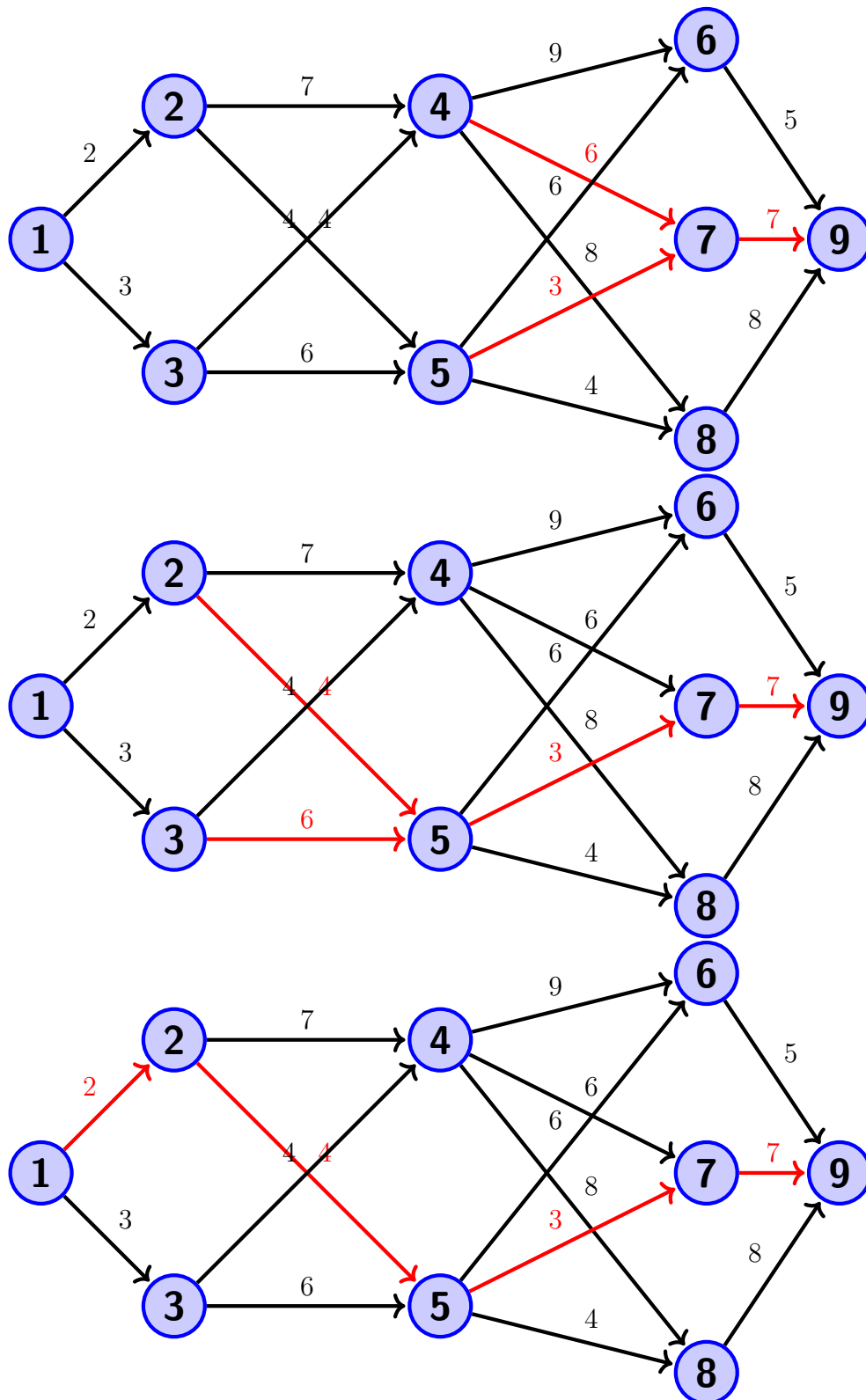
$P(2) = ind(P(1)) = ind(2) = 5$

$P(3) = ind(P(2)) = ind(5) = 7$

$P(4) = ind(P(3)) = ind(7) = 9$

$\rightarrow P = [2, 5, 7, 9], l = 16$

In the following, we can see optimal path from vertices in layers (red edges, from sink v to source u)



1.3.2 Complexity

Finding node r so that $l(j, r) + ld(r)$ is minimal:

- If we represent graph, i.e. it's edges, with adjacency list, we search for all

vertices connected to j . The duration is proportional to the degree of that vertex.

- As we go through all vertices, the total duration is proportional to the number of edges $|E| = e$, i.e. $O(e)$.
- This search goes into a loop. We add $O(n)$ time for the $\rightarrow O(n + e)$.
- At the end (the last FOR loop) we spend $O(m)$ time, but since $m < n$ it is already covered by $O(n)$.
- The total time is

$$O(n + e).$$

Remark 3. *The principle of optimality can be formulated from front*

$$A(v) = \{w \in V : (w, v) \in E\}.$$

$$ld(u, v) = \min_{w \in A(v) \text{ such that } (w, v) \in E} \{ld(u, w) + l(w, v)\}.$$

1.4 Unbounded integer and 0 – 1 knapsack

We have n items with weights w_i and values (profit) p_i .

We have a knapsack with a capacity of M .

We consider two cases (in both we are not allowed to cut items, i.e. $x_i \in \mathbb{N}_0$): we have infinitely many of each object and we have only one of each object (we can take it or not).

Dynamic programming is usually described in 4 steps.

1.4.1 Unbounded integer knapsack

When we say integer knapsack, we assume we are talking about unbounded integer knapsack.

1. step - optimal choice of subproblems

Obviously, the optimal choice of the subproblem is the problem of an unbounded integer knapsack, but with a smaller capacity (smaller knapsack \rightarrow slightly larger $\rightarrow \dots$).

If (x_1, x_2, \dots, x_n) is the optimal solution for capacity M , then $(x_1, \dots, x_{i-1}, x_i - 1, x_{i+1}, \dots, x_n)$ is optimal solution for $M - w_i$ (principle of optimality). Note that $(x_1, \dots, x_{i-1}, x_i - 1, x_{i+1}, \dots, x_n)$ means that we have taken one object x_i (i.e. put it in a knapsack), so the capacity has dropped to $M - w_i$.

2. step - recursive formula

We need to find a recursive formula for the value of the optimal solution.

We denote:

$KNAP(M)$... the optimal value for the capacity of M .

$$KNAP(M) = \max_{i \in \{1, \dots, n\} \text{ such that } w_i \leq M} \{KNAP(M - w_i) + p_i\}.$$

Note that if we have taken the subject i , we must add p_i to the value $KNAP(M - w_i)$. Obviously, for the initial value we take $M = 0$, i.e. $KNAP(0) = 0$ (there is no subject i i.e. $w_i \leq 0$).

It makes no sense to put an item worth 0 in your knapsack, we don't look at such items.

3. step - algorithm using dynamic programming

Algorithm - pseudocode

FUNCTION NEO-KNAPSACK (M, n, W, P)

{ the input is the capacity M , number of elements n , weight of the elements in the field W , and value of elements in the field P }

BEGIN

{ first initialization of field K }

$K(0) = 0$;

FOR $i = 1$ TO M DO ... capacity from the smallest 1 to the largest M

BEGIN

$K(i) = 0$; ... for each capacity we initially set the optimal value 0

FOR $j = 1$ TO n DO ... as we have infinitely many of each item in every step of the for loop we check for each item

BEGIN

IF $w_j \leq i$ DO

{ if item j is heavier then the capacity, we don't take it in the knapsack }

$K(i) = \max\{K(i), K(i - w_j) + p_j\}$

{ if we get a bigger value by taking item j , we have to update $K(i)$ }

{ note, if $K(i)$ is not updated, that means we didn't take item j }

END;

RETURN($K(M)$);

END;

Complexity: obviously $O(n * M)$.

4. step - additional information

The algorithm can also return some additional information. In this case, it is natural to return which items are taken in the knapsack for an optimal solution.

The output is additionally field *ITEMS* of items taken in the optimal solution.

FUNCTION NEO1-KNAPSACK (M, n, W, P)

{ the input is the capacity M , number of elements n , weight of the elements in the field W , and value of elements in the field P }

BEGIN

{ first initialization of fields K and *ITEMS* }

$K(0) = 0$;

ITEMS(0) = \emptyset ; ... initially for capacity 0 we didn't take any item in the optimal

solution

```

FOR  $i = 1$  TO  $M$  DO ... capacity from the smallest 1 to the largest  $M$ 
BEGIN
   $K(i) = 0$ ; ... for each capacity we initially set the optimal value 0
  FOR  $j = 1$  TO  $n$  DO ... as we have infinitely many of each item in every step
of the for loop we check for each item
  BEGIN
    IF  $w_j \leq i$  DO
      { if item  $j$  is heavier then the capacity, we don't take it in the knapsack }
       $K(i) = \max\{K(i), K(i - w_j) + p_j\}$ 
      { if  $K(i)$  is updated, i.e. we took item  $j$ , then  $ITEMS(i) = ITEMS(i -$ 
 $w_j) \cup \{j\}$  }
    END;
  RETURN( $K(M)$ );
  RETURN( $ITEMS(M)$ );
END;
```

Example 4. Let the capacity of the knapsack be $M = 5$ and we have 4 items available: $w_1 = 1, p_1 = 1, w_2 = 2, p_2 = 4, w_3 = 3, p_3 = 6, w_4 = 4, p_4 = 7$, t_j .

i	w_i	p_i
1	1	1
2	2	4
3	3	6
4	4	7

$$\begin{aligned}
i = 1 \quad j = 1, w_1 = 1 \leq 1 \quad & K(1) = \max\{K(1), K(1-1) + p_1\} = 1 \\
& ITEMS(1) = ITEMS(0) \cup \{1\} = \{1\} \\
& j = 2, w_2 = 2 > 1 \\
i = 2 \quad j = 1, w_1 = 1 \leq 2 \quad & K(2) = \max\{K(2), K(2-1) + p_1\} = 2 \\
& ITEMS(2) = ITEMS(1) \cup \{1\} = \{1, 1\} \\
& j = 2, w_2 = 2 \leq 2 \quad K(2) = \max\{K(2), K(2-2) + p_2\} = 4 \\
& ITEMS(2) = ITEMS(0) \cup \{2\} = \emptyset \cup \{2\} = \{2\} \\
& j = 3, w_3 = 3 > 2 \\
i = 3 \quad j = 1, w_i = 1 \leq 3 \quad & K(3) = \max\{K(3), K(3-1) + p_1\} = 5 \\
& ITEMS(3) = ITEMS(2) \cup \{1\} = \{2, 1\} \\
& j = 2, w_2 = 2 \leq 3 \quad K(3) = \max\{K(3), K(3-2) + p_2\} = 5 \\
& j = 3, w_3 = 3 \leq 3 \quad K(3) = \max\{K(3), K(3-3) + p_3\} = 6 \\
& ITEMS(3) = ITEMS(0) \cup \{3\} = \emptyset \cup \{3\} = \{3\} \\
& j = 4, w_4 = 4 > 3 \\
i = 4 \quad j = 1, w_1 = 1 \leq 4 \quad & K(4) = \max\{K(4), K(4-1) + p_1\} = 1 \\
& ITEMS(4) = ITEMS(3) \cup \{1\} = \{3, 1\} \\
& j = 2, w_2 = 2 \leq 4 \quad K(4) = \max\{K(4), K(4-2) + p_2\} = 8 \\
& ITEMS(4) = ITEMS(2) \cup \{2\} = \{2, 2\} \\
& j = 3, w_3 = 3 \leq 4 \quad K(4) = \max\{K(4), K(4-3) + p_3\} = 8 \\
& j = 4, w_4 = 4 \leq 4 \quad K(4) = \max\{K(4), K(0) + p_4\} = 8 \\
i = 5 \quad j = 1, w_1 = 1 \leq 5 \quad & K(5) = \max\{K(5), K(5-1) + p_1\} = 9 \\
& ITEMS(5) = ITEMS(4) \cup \{1\} = \{2, 2, 1\} \\
& j = 2, w_2 = 2 \leq 5 \quad K(5) = \max\{K(5), K(5-2) + p_2\} = 10 \\
& ITEMS(5) = ITEMS(3) \cup \{2\} = \{3, 2\} \\
& j = 3, w_3 = 3 \leq 5 \quad K(5) = \max\{K(5), K(5-3) + p_3\} = 10 \\
& j = 4, w_4 = 4 \leq 5 \quad K(5) = \max\{K(5), K(5-4) + p_4\} = 10 \\
& \rightarrow K(5) = 10, ITEMS(5) = \{2, 3\}.
\end{aligned}$$

In addition, we only need a one-dimensional field (or, for example, a row matrix) from which we read the previous optimal solutions (for subproblems).

i	0	1	2	3	4	5
$ITEMS(i)$	\emptyset	$\{1\}$	$\{1, 1\}$ $\{2\}$	$\{2, 1\}$ $\{3\}$	$\{3, 1\}$ $\{2, 2\}$	$\{2, 2, 1\}$ $\{3, 2\}$

LESSON 1: INTRODUCTION TO ORTHOGONAL AXONOMETRY METHOD

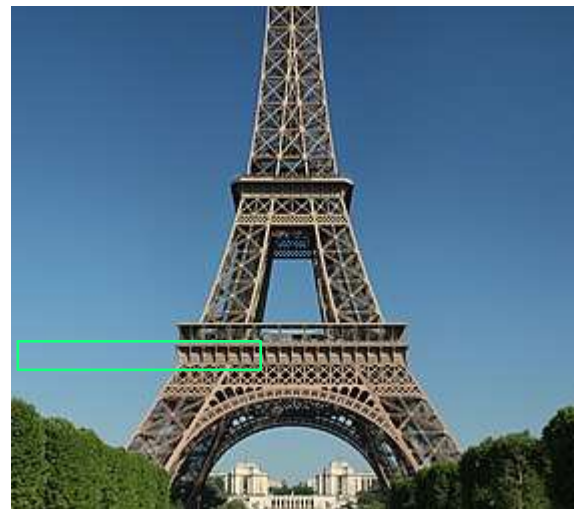
INTRODUCTION TO DESCRIPTIVE GEOMETRY

The Monge method

In descriptive geometry, objects are projected onto 2D surface where one can extract their significant properties from 3D space. On the other hand, using 2D projections, one can visualize its 3D shape. In modern technologies, these procedures have additional advances since by a fast changing of parameters of projections one can simulate moving an object and see all its sides from different views, all on 2D surface of a monitor.

Original method of descriptive geometry is named after French mathematician Gaspard Monge (1746-1818) who developed it and is nowadays known as the Monge method.

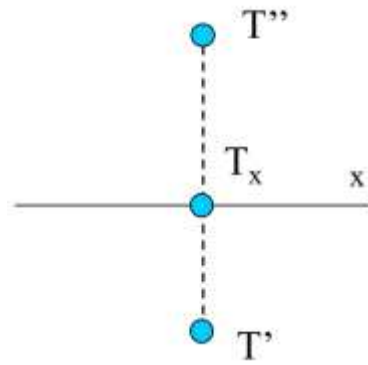
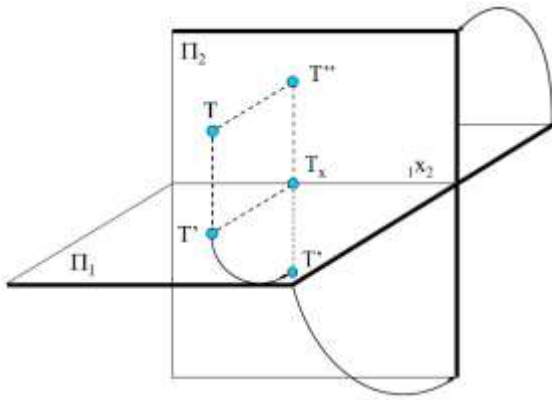
In recognition of his contribution, his name is inscribed among 72 names of French scientists, engineers and mathematicians on the base of the Eiffel Tower.



The

Monge method is implemented in almost all 3D-CAD programmes since it is very applicable in computer modelling of various objects and is a compulsory part of the education of the future engineers.

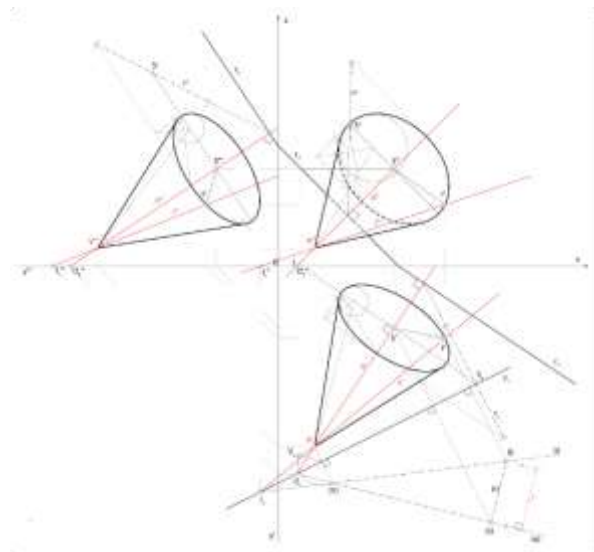
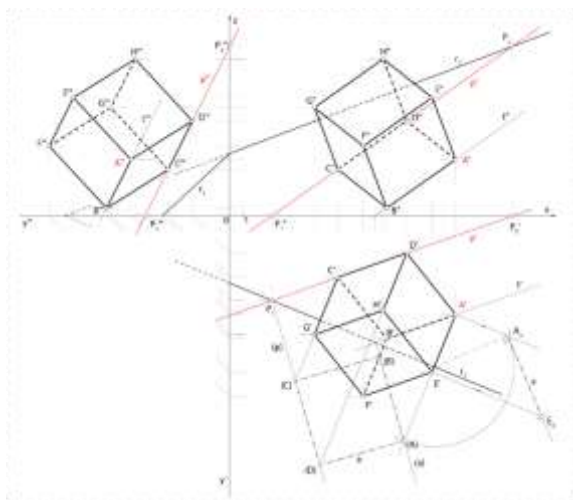
The method is based on the orthogonal projection of an object in 3D space onto two or more planes of projection.



Here we see the horizontal (top view) and the vertical (front view) planes of projections. The vertical one assumes the role of the 2D surface for the space interpretation: all the contents of both planes is finally presented in the vertical one, by a 90° rotation of the horizontal one. Notice that all of the rays are mutually *parallel* and are *orthogonal* on the planes of projections.

We can also add the side view orthogonal projection in the third plane of projections, which is perpendicular to the horizontal and to the vertical one.

To illustrate the scope of the method we give examples of simple solids with their three orthogonal projections including its top, front and (right) side view.

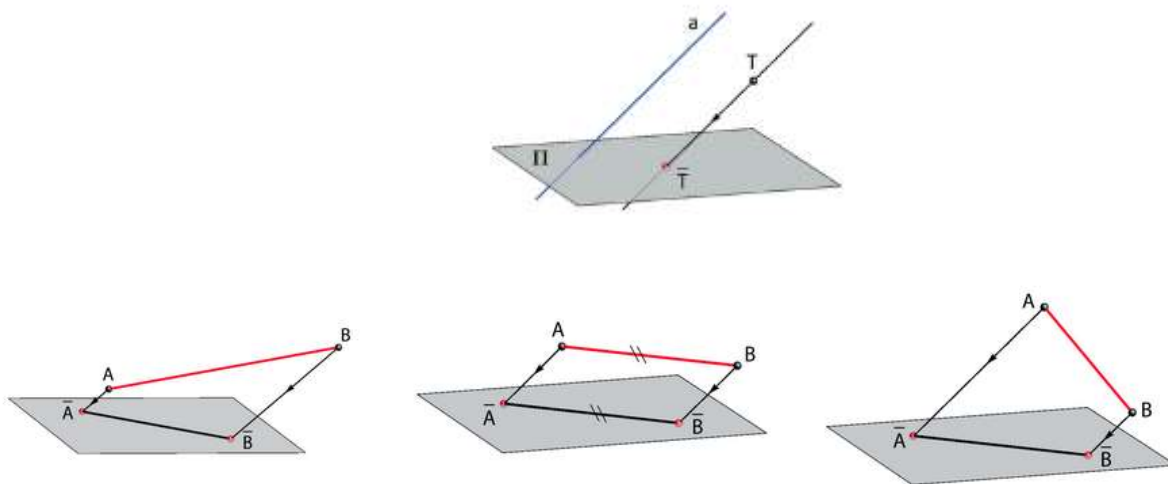


As we can see, the method develops the whole variety of procedures which lead to a final goal of the construction and a variety of spatial properties are constructed according to the strict laws of the method itself.

Axonometry methods

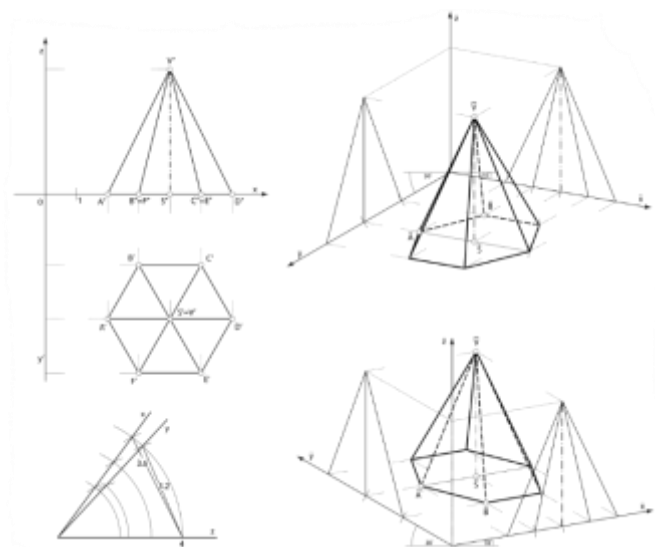
With the Monge method as the starting point, many other methods in descriptive geometry are developed. Axonometric methods *e.g.* present an object together with its coordinate system and by parallel rays project both onto the 2D surface. Thus we get a complete 3D image of an object on a unique plane of projection. Axonometric methods have their specific advances in engineer practice, but also originate from the Monge method and then branch in different sub-methods.

In general, rays of projection can be laid in different angles toward the plane of projection but are always mutually parallel:



Thus we

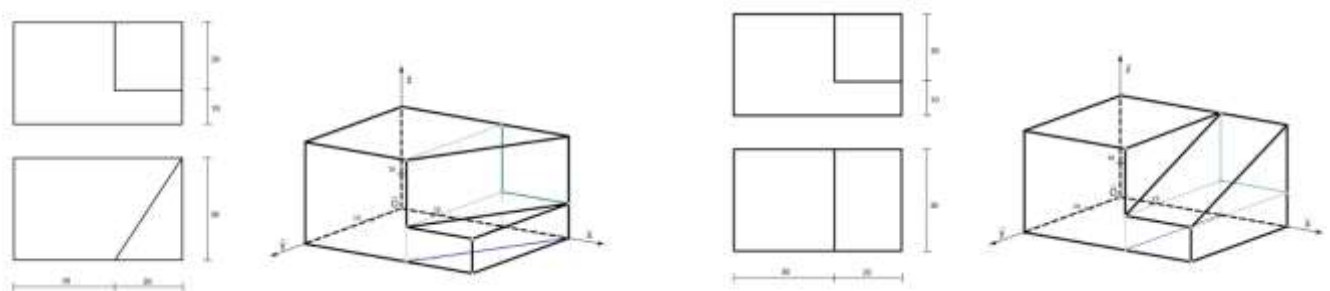
get an *oblique parallel* projection of an object together with its coordinate system:



The orthogonal coordinate system with the orthogonal projections of the pyramid serves as a starting point in constructing its 3D image in the oblique axonometry. The coordinate system follows the pyramid in the axonometric image and loses its orthogonality, but parallel lines stay parallel.

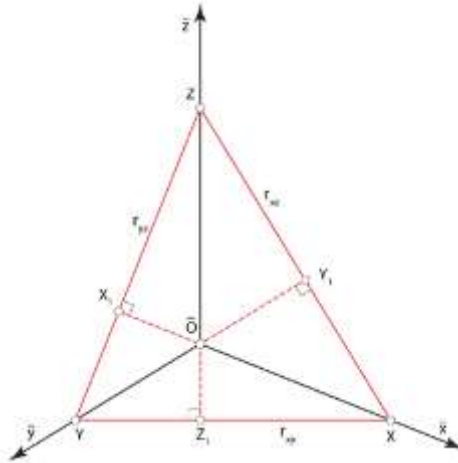
As we can see, we observe the pyramid with a right side view and vary the axes of view from above and from below.

Here are the oblique axonometric images of two different objects observed from above.

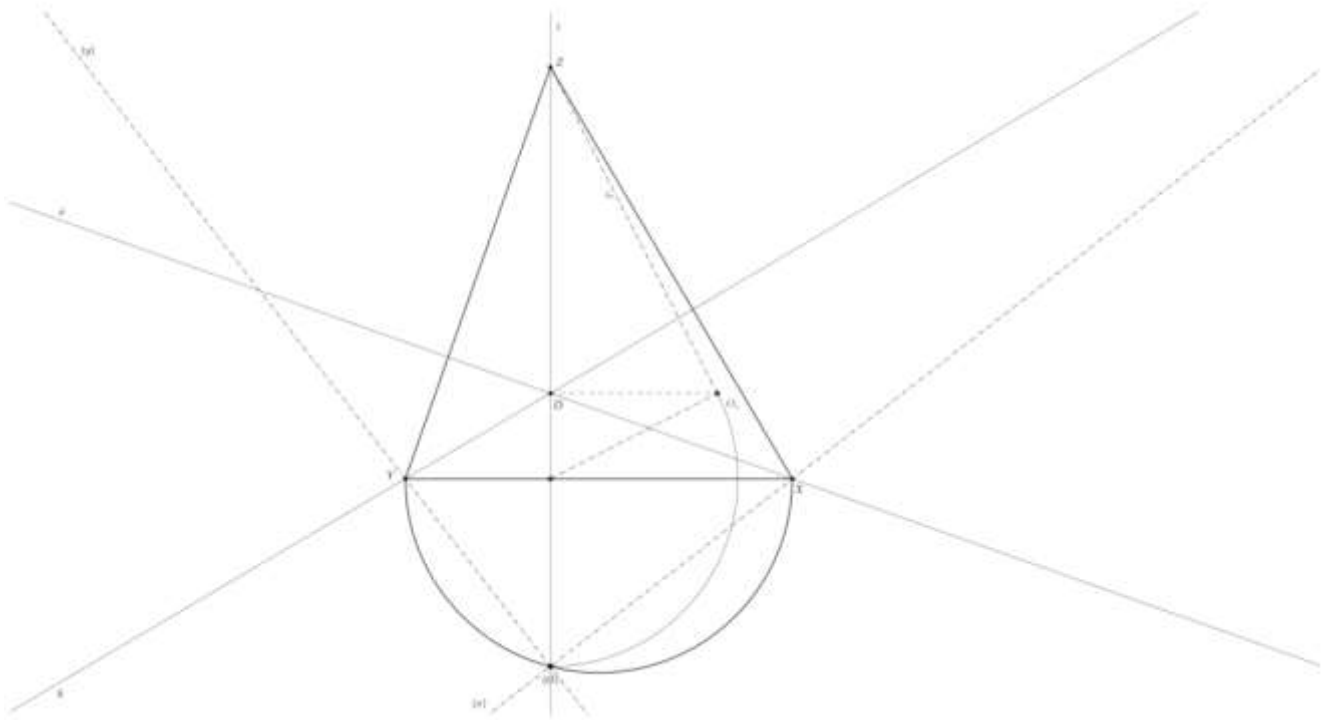


ON ORTHOGONAL AXONOMETRY

If the rays of the parallel projection are *orthogonal* to the plane of projection, axonometry method is called *orthogonal axonometry*. The main task here is to project the coordinate system orthogonally to the plane of projection. Former planes of projection that are determined by the x , y and z axes, for the top, the front and the side view in the Monge method are now together orthogonally projected on the single plane of projection in axonometry.

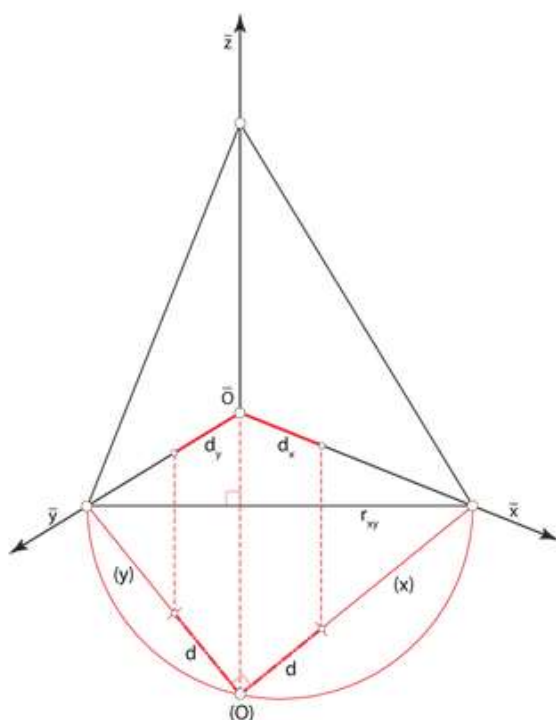


Since these planes of projection are brought into the new coordinate system, lines rx , ry and rz are their traces in the plane of projection in orthogonal axonometry. The specific triangle $\triangle XYZ$ is called a *trace triangle*. The orthogonal projections of the axes are orthogonal on the traces rx , ry and rz . Hence we have the axes as the heights and the origin as the orthocenter of the constructed trace triangle.

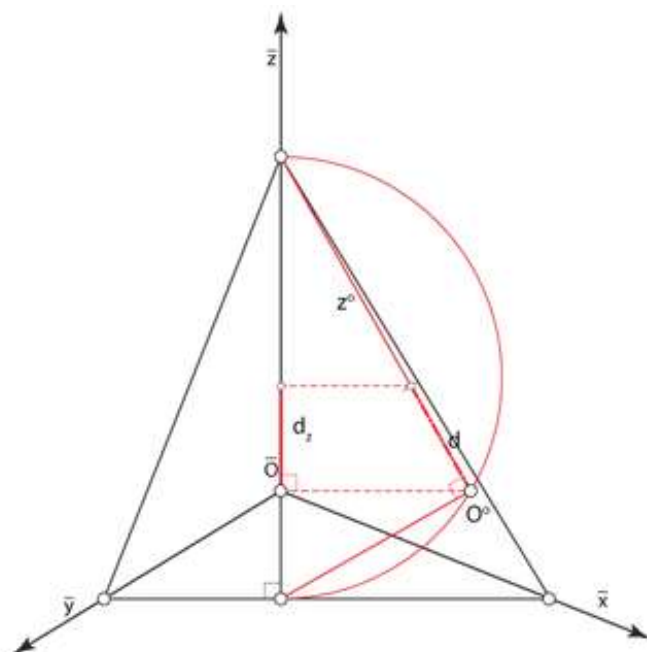


If we want to see true dimensions of an object in order to have its orthogonal axonometry 3D image, standard procedure implies revolution of the plane in which one of the basis of the object lays. The revolution is done into the plane of projection, around the trace of the plane in which the basis is layed.

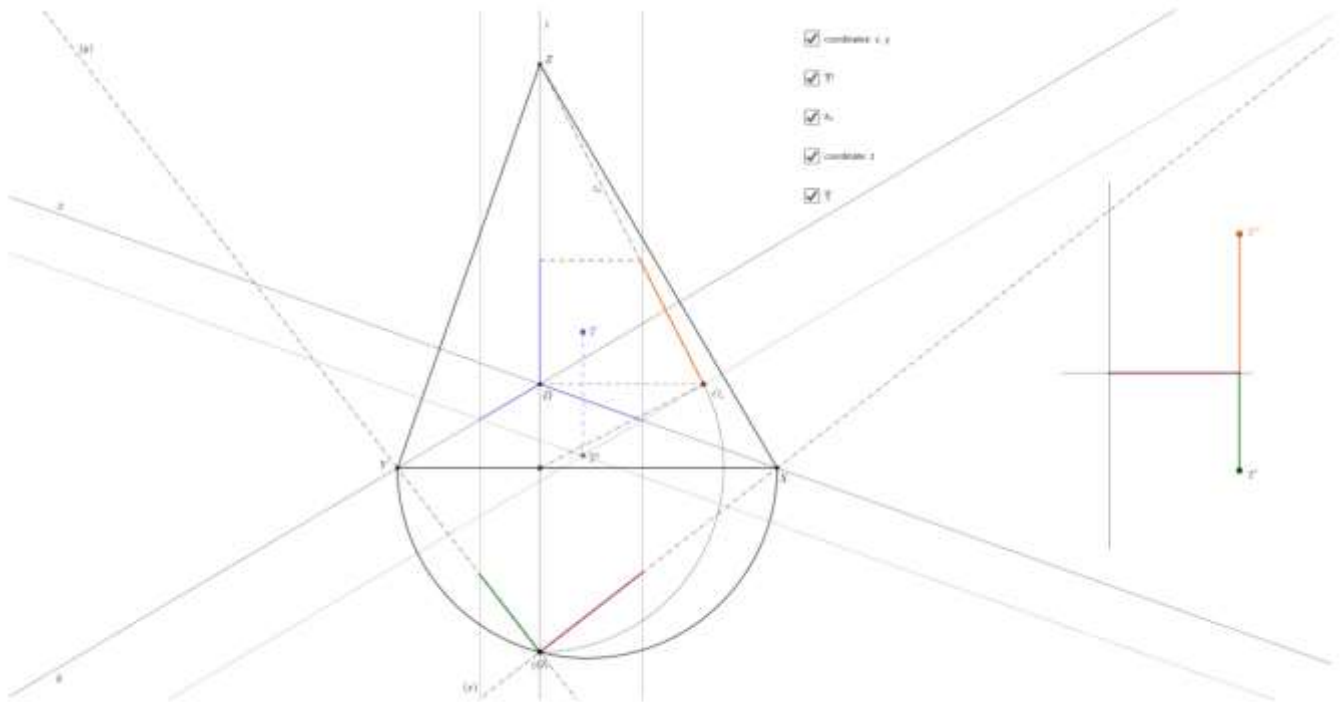
Since we locate the bases of objects in one of the three coordinate planes, we have the following variants of the revolution.



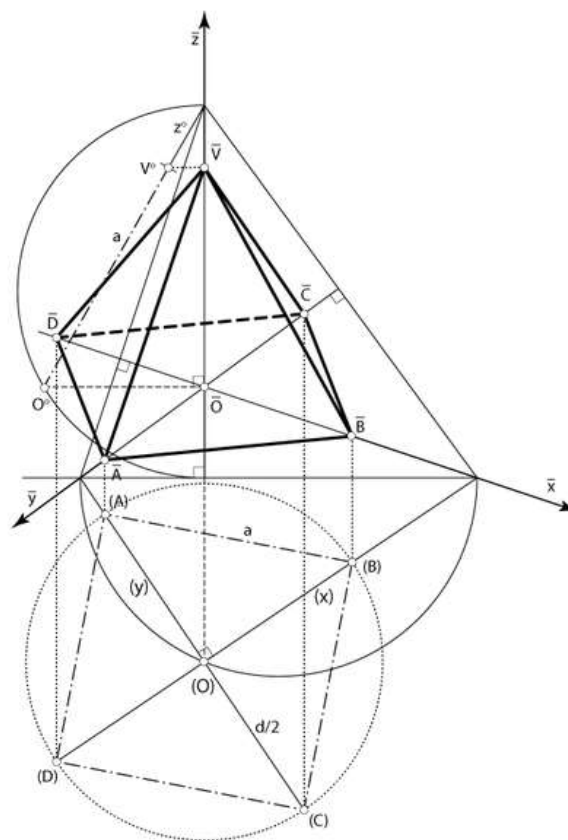
Sometimes it is enough to do 90° rotation into the plane of projection, for example, if we want to construct the true length of the height of an object.



We can carry out the step-by-step construction of an orthogonal axonometric image of a point in GeoGebra.



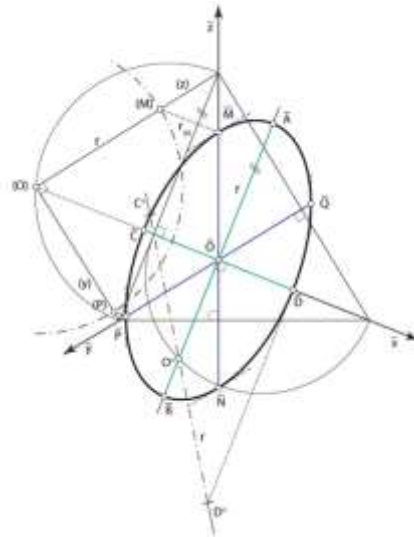
Similar procedures are used for the construction of a quadrilateral pyramid given in the following picture. Here the height of the pyramid is obtained by the rotated position of the z-axis.



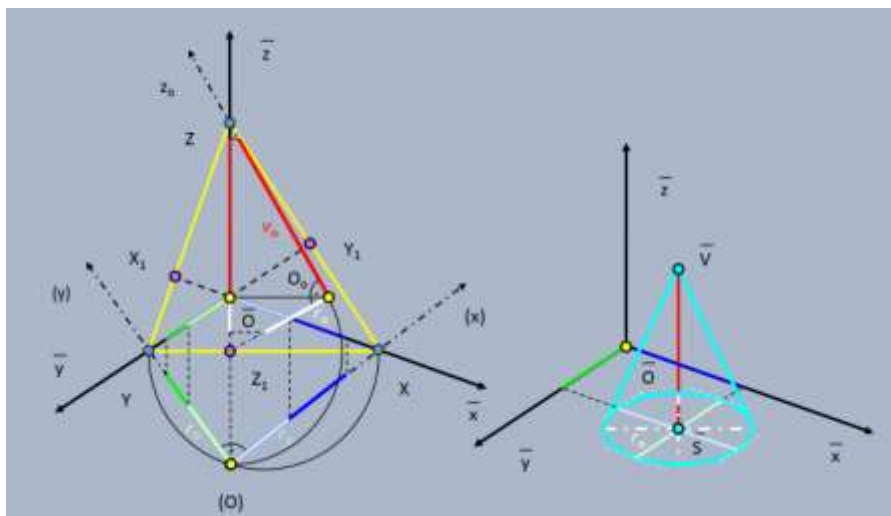
Orthogonal axonometric image of a circle demands a specific way of construction. In parallel projections of different types, circles degenerate into ellipses, in general. In some special cases, when circles are laid parallel to the plane of projection, they are projected as circles.

Ellipse, as an orthogonal axonometric image of a circle, demands the construction of its major and its minor axis. The major axis originates from the diameter of the circle which is parallel to the plane of projection and the minor axis originates from the diameter which is perpendicular to the first one.

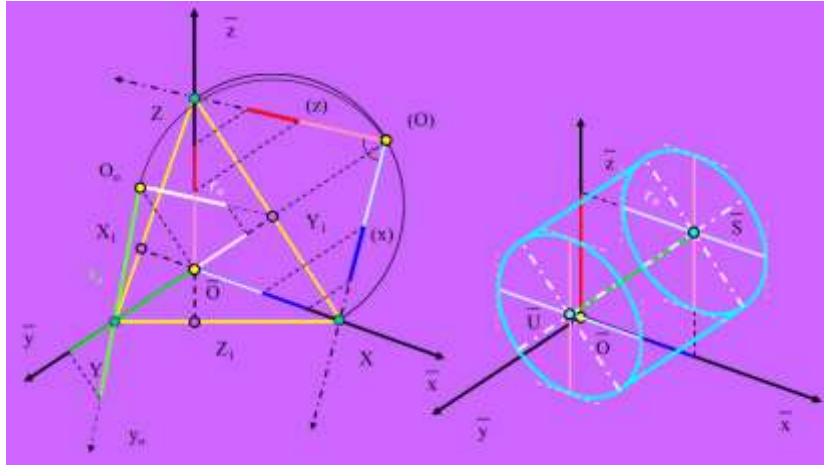
Besides the axes of an ellipse we always use the pair of its conjugate diameters which are parallel with the coordinate axes in the corresponding plane of the circle. Regarding the position of the circle in the following picture, we use diameters parallel to y and z axes.



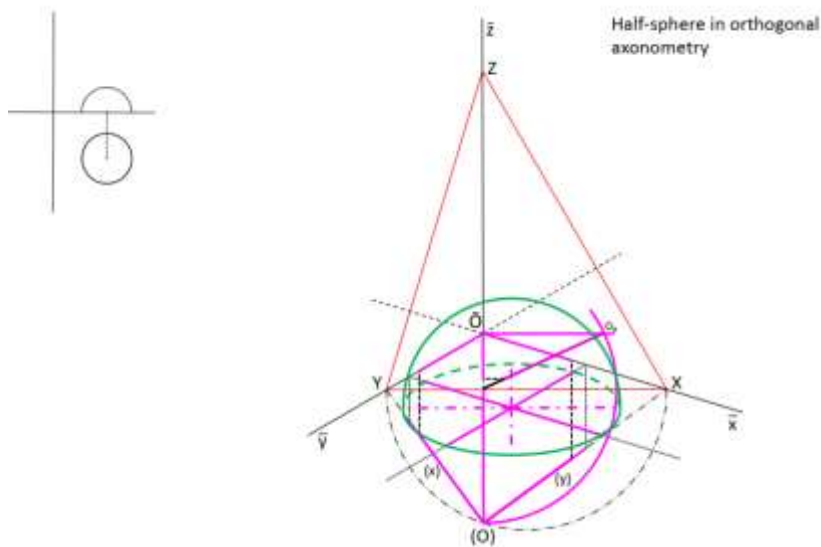
Similarly, as it was done with the pyramid, one can construct a solid with the circle in the basis, *e.g.* a conus, as in the following example. The circle is located in the xy - plane.



On the other hand, one of the bases of a cylinder is located in xz - plane.



Finally, one can easily construct a half-sphere using the similar construction of the main circle section of the sphere.



The contour circle of the half-sphere has the same radius as the main circle itself which makes the construction in orthogonal axonometry simplified.

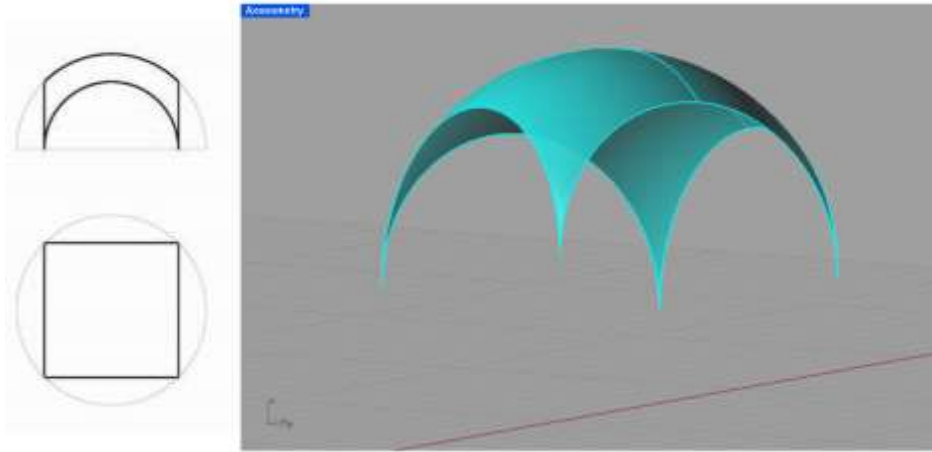
That is the reason why this method is often applied to constructions of the domes in architecture.

Three main types of domes in architecture

Taking into account that the contour circle of the sphere stays with the same radius in orthogonal axonometry as in its true length, it is much easier to construct all of the spherical elements and also the transitional ones when leaning the various types of dome vaults down onto the quadratic (rectangular) basis. Three main types of domes in architecture are *spherical*, *bohemian* and *pendentive* domes. These are simultaneously presented by their pairs of orthogonal projections where the geometric elements of their architectural concept can be properly recognized.

The *spherical dome* is constructed over the square basis of an object.

Given a trace triangle and two orthogonal projections of a spherically vaulted object with a square basis one can construct the orthogonal axonometric image of the corresponding spherical dome.

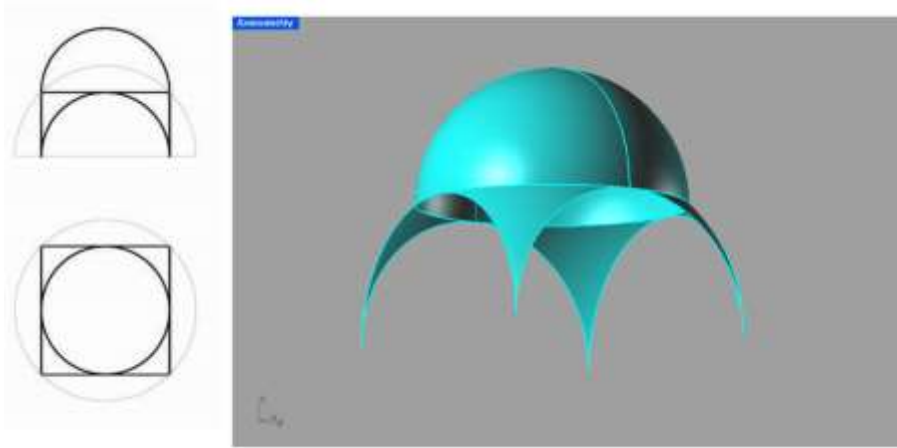


The procedure is following. Firstly we rotate around the first trace of the trace triangle in the plane of projections and determine the true dimension of the basis square. According to the rules of orthogonal axonometric projection, we carry out the projection of the basis square, as a parallelogram. Then we draw the contour circle of the dome with the true length of the given sphere radius. What follows is the construction of ellipses as projections of circles in the vertical planes (parallel in pairs): major axis, minor axis, the highest point and two points on the square for each (half)-ellipse in the vertical plane. The final step is visualization and drawing the arcs of the spherical dome, according to its spatial visibility.

National library of France in Paris serves as an example of an implementation of spherical domes in architecture.



A *pendentive dome* is a spherical dome upgraded by a half-sphere and supported by pendentives or spherical triangles. It can be obtained by upgrading a spherical one by an additional half-sphere.



The center of the horizontal circle section of the half-sphere is located on the z – axis. The true length of the radius of the half-sphere is determined. The half-sphere is constructed as was described in the text. The accent is put on the common points of the half-sphere and the spherical dome, i.e. on the constructed *spherical triangles-pendentives*. The visibility of the constructed pendentive dome is finally determined.

Numerous pendentive domes were built in Istanbul and the most famous example is Hagia Sophia.



Finally, a *bohemian dome* is constructed over the rectangular basis of an object. Given a trace triangle and two orthogonal projections of a spherically vaulted object with a rectangular basis, one can construct the orthogonal axonometric image of the corresponding bohemian dome.



A nice example of a bohemian dome is the Church of Saint Matthew in Split, Croatia.



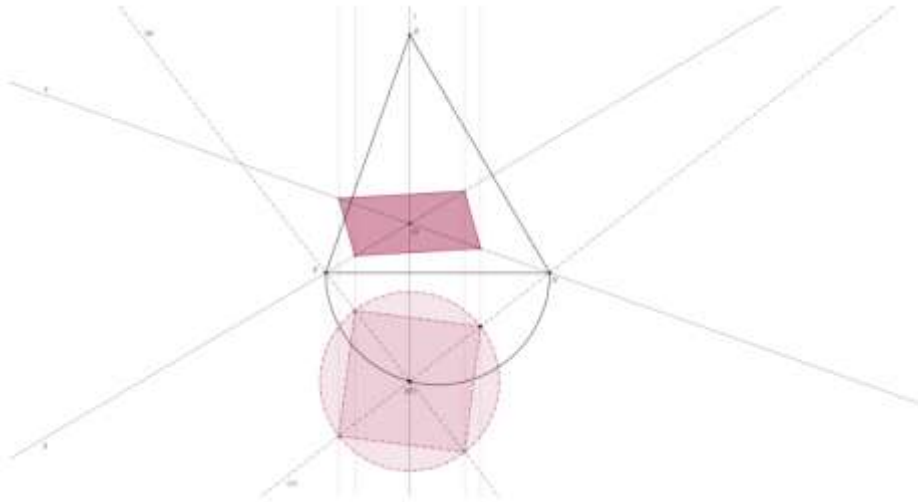
LESSON 2: ORTHOGONAL AXONOMETRY OF VARIOUS OBJECTS

The second lesson on orthogonal axonometry is a continuation of the previous one. It is primarily based on the construction process and the discussion of the number of possible solutions, which is one of the main purposes of descriptive geometry: how to see and understand 3D relations and their appearance on 2D surface.

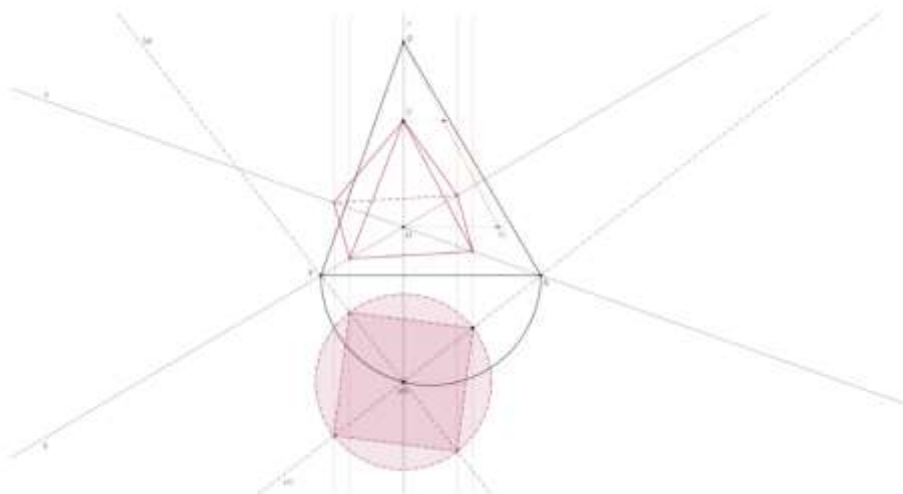
Making use of formerly described elementary techniques related to this method, step-by-step constructions are carried out in GeoGebra.

Construction of a square. Upgrading a pyramid and a prism over the square basis

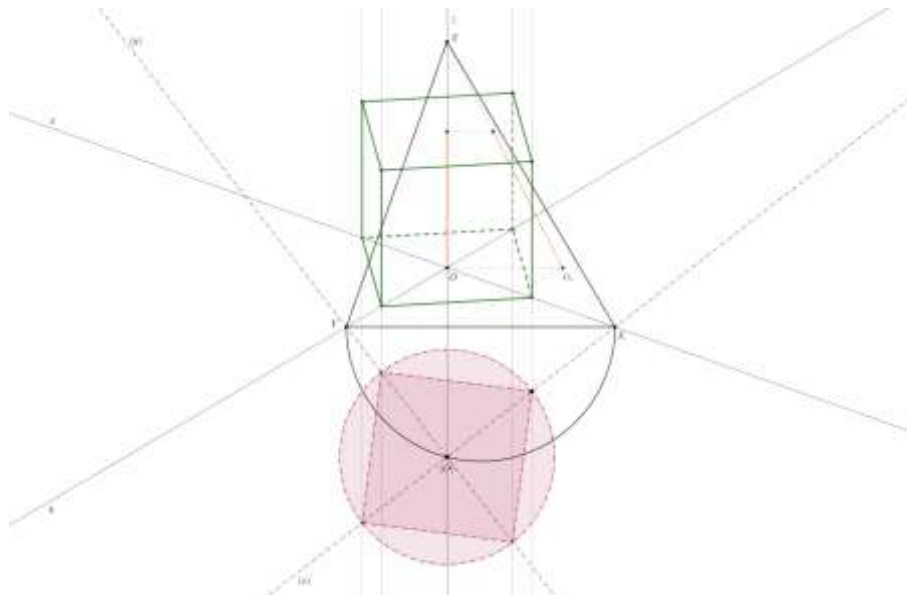
A square serves as the first example of an object which demands its true dimensions in one of the rotated planes (here xy - coordinate plane) and is afterwards projected as a parallelogram in orthogonal axonometry.



Upgrading the square by the height of a solid (another previously described construction) leads to a regular quadrilateral pyramid. Here the 90° rotation of the z -axis is used.



Another possibility for upgrading a solid over the same square base is a regular quadrilateral prism. Both solids demand the construction of the height, but differ in the process of determining visibility of their sides with the similarly given view axis.



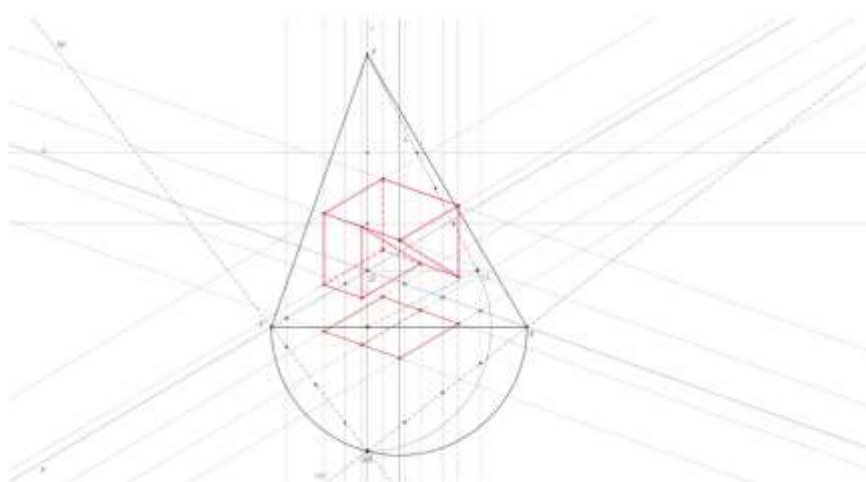
Construction of an angular object in GeoGebra

A more complex task is the construction of the orthogonal axonometric image of a non-regular angular object, which is given in its front and its top view of orthogonal projections.

The technique of the construction demands similar steps as before, but here the accent is on how accurately one can determine or visualize the true shape of the object.

Once the object is defined considering the visibility of its edges, *i.e.* its sides in 3D image, the construction of its orthogonal axonometry may begin.

The step-by-step construction in GeoGebra is done in detail.



The final step is analysing the number of solutions. Namely, often with only two orthogonal projections of an object one can derive more than one possible 3D shapes that correspond to them. Moreover, some tasks can provide numerous solutions. In this manner, students can significantly develop their space visibility skills.

QUIZ ("Exit tickets")

1. Using the template in GeoGebra, students are asked to draw the orthogonal axonometric 3D image of an object given by two orthogonal projections.
2. Using the template in GeoGebra students are asked to draw the right side view of an object given by two orthogonal projections and to find as many as possible solutions to the given task.
3. Students are given photos of a few examples of domes in the world architecture and are asked to determine the type of domes (spherical, pendentive or bohemian) for each example.

EXIT TICKET

ORTHOGONAL AXONOMETRY

1. For each given object choose the type of dome which it contains:



S



Daut Pasha Hamam, Skopje, N. Macedonia

- a) pendentive dome
- b) spherical dome
- c) bohemian dome



Saint Clement of Ohrid, Skopje, N. Macedonia

- a) pendentive dome
- b) spherical dome
- c) bohemian dome



- a) pendentive dome
- b) spherical dome
- c) bohemian dome

Macedonian Philharmonic, Skopje, N. Macedonia

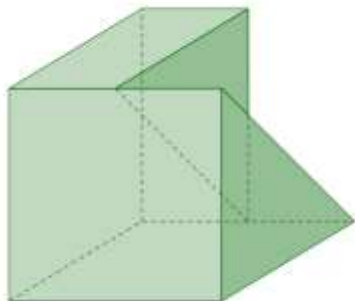


- a) pendentive dome
- b) spherical dome
- c) bohemian dome

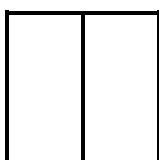
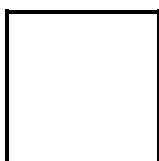
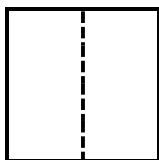
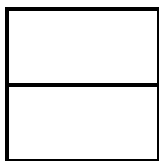


Mustafa Pasha Mosque, Skopje, N. Macedonia

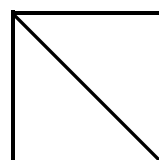
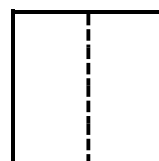
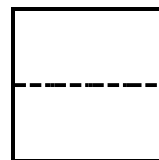
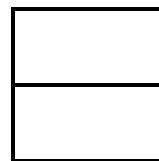
2. For a given object, choose the correct top and front view obtained by orthogonal projections:



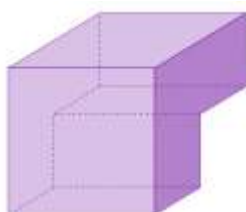
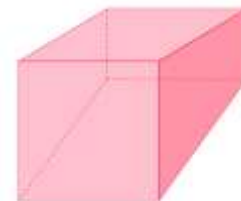
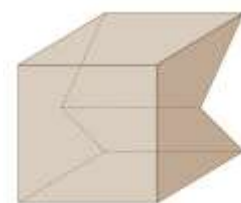
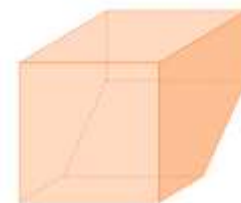
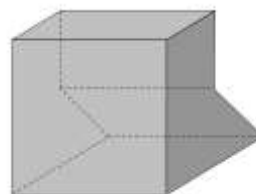
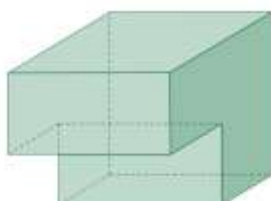
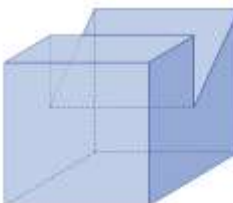
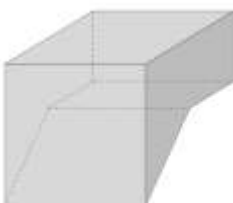
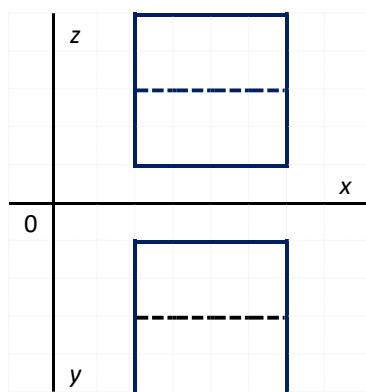
Top view:



Front view:



3. Select the objects to which the given top and front orthogonal view correspond:



REFERENCES:

- S. Gorjanc, E. Jurkin, I. Kodrnja, H. Koncul: Descriptive geometry, web textbook, Faculty of Civil Engineering, Zagreb, (2019)
www.grad.hr/geometrija/udzbenik
- V. Szivovicsa, E. Jurkin: Descriptive geometry, CD textbook, Croatian Society for Geometry and Graphics & Faculty of Civil Engineering , Zagreb (2005)
- KoG, Scientific and Professional Journal of Croatian Society for Geometry and Graphics, No. 7, Zagreb, (2003)
- Photographs: online resources and private albums