

Alexander Búcsi

Pharmaceutical Physics



2025

Comenius University Bratislava

Faculty of Pharmacy, Comenius University Bratislava

Alexander Búcsi

Pharmaceutical Physics

2025

Comenius University Bratislava

© RNDr. Alexander Búcsi, PhD., 2025

Comenius University Bratislava, Faculty of Pharmacy

Reviewers

doc. Ing. Martin Pisárčik, CSc.

RNDr. Jozef Leja, PhD.



Publikácia je šírená pod licenciou Creative Commons CC BY-NC 4.0

(vyžaduje sa: povinnosť uvádzať pôvodného autora, len nekomerčné použitie). Viac informácií o licencií a použití diela: <https://creativecommons.org/licenses/by-nc/4.0/>



https://stella.uniba.sk/texty/FAF_koL_pharmaceutical_physics.pdf

Published by

Comenius University Bratislava

ISBN 978-80-223-6073-9 (online)

Contents

Forewords.....	6
1 Physical Quantities and Units.....	7
1.1 Base Units	7
1.2 Derived units.....	11
1.3 Non-SI Units.....	14
2 Vectors and Scalars	14
2.1 Operations with Vectors, Vector Addition	15
2.1.1 Vector Multiplication	16
3 Classical - Newtonian Mechanics.....	18
3.1 Kinematics.....	18
3.1.1 Uniform Linear Motion	18
3.1.2 Uniformly Accelerated Linear Motion.....	19
3.1.3 Freefall.....	21
3.1.4 Projectile Motion	21
4 Dynamics – Newton’s Laws of Motion.....	24
4.1 Conservation of momentum	25
4.2 Uniform Circular Motion	26
4.3 Centripetal Acceleration a_c , Centripetal Force F_c	28
5 Newton’s Law of Universal Gravity.....	29
6 Mechanical Energy (E), Mechanical Work (W), Power (P).....	29
7 Harmonic motion - Harmonic Oscillator	34
8 Component Motion - Solid Body Motion.....	37
8.1 Friction	38
9 Deformation of a Solid Body.....	39
10 Moments of Physical Quantities	40
10.1 Moment of Force – Torque (M).....	40
10.2 Centre of Mass	41
10.3 Moment of Inertia, Energy of the Rotating Object	43
10.4 Angular Analogues of Newton Laws of Motions	44
11 States of matter	45
11.1 Gibbs phase rules	46
11.1.1 Air Humidity.....	47
11.2 Fluids.....	49
11.2.1 Surface Tension	49

11.2.2 Capillary Elevation and Depression	50
11.2.3 Stalagmometry	51
11.2.4 Surface tension between two immiscible liquids.....	53
12 Fluid Mechanics, Hydrostatic Pressure.....	54
12.1 Atmospheric pressure.	54
12.2 Pascal's Principle	55
12.3 Archimedes Law.....	56
12.4 Density and Density Measurements	57
12.4.1 Pycnometry	59
12.5 Fluid Dynamics (Hydrodynamics).....	60
12.5.1 Ideal Liquid.....	60
12.5.2 Equation of Continuity, Bernoulli Equation.....	61
12.5.3 Toricelli's Theorem	63
12.6 Real Liquids	65
12.6.1 Viscosity of Newtonian Liquids.....	65
12.6.2 Hagen-Poiseuille's Law.....	66
12.6.3 Measuring of Viscosity	68
13 Thermal Physics	71
13.1 Thermal Expansion of Solid and Liquid Materials	71
13.2 Thermodynamics	72
13.2.1 Ideal Gas	73
13.3 Thermodynamics of Gases	73
13.3.1 Gay-Lussac I. Law, Charles' Law, Law of Volumes	74
13.3.2 Gay-Lussac II. Law, Pressure-Temperature Law.....	75
13.3.3 Boyle-Mariotte Law	75
13.3.4 Avogadro's Law	75
13.3.5 Ideal Gas Law – General Gas Equation.....	76
13.3.6 Dalton's Law	76
13.4 Real Gases, van der Waals Equation	76
13.5 Equipartition Theorem	77
13.6 Heat and Temperature, Zeroth Law of Thermodynamics	77
13.7 Thermodynamic Functions, Energy Conservation Law	78
13.8 Work of Ideal Gases.....	79
13.9 Internal Energy, First Law of Thermodynamics	79
13.10 Entropy	80
13.11 Enthalpy, Helmholtz Energy, Gibbs Free Energy	80

13.12 Thermodynamic Equilibrium	83
13.13 Isochoric Processes	85
13.14 Isobaric Processes	86
13.15 Isothermal (Isothermic) Processes	88
13.16 Irreversible processes	89
13.17 Adiabatic Processes	90
13.18 Efficacy of the Steam Engine and the Second Law of Thermodynamics	91
13.19 Third Law of Thermodynamics	94
13.20 Spontaneity of Processes	95
13.21 Heat Transfer and Diffusion	97
14 Electricity	99
14.1 Electric Field, Intensity of the Electric Field	101
14.1.1 Shell Theorem	103
14.1.2 Electric Dipole, Dipole Moment	103
14.2 Work and Potential Energy in an Electric Field	106
14.3 Electrical Resistance, Ohm's Law	108
14.4 Materials in Electric Field	109
14.4.1 Conductors, Semiconductors, Insulators	109
14.4.2 First and Second Type Electrical Conductors – Metals and Electrolytes	110
14.4.3 Semiconductors	111
14.5 Electromotive Force (EMF)	112
14.6 Galvanic Cells	113
14.7 Electrolysis	114
15 Magnetism	115
15.1 Magnetic Flux Density (Magnetic Induction, B), Magnetic Field Strength(H)	117
15.2 Lorentz Law	117
15.3 Magnetic Field Along a Straight Wire	118
15.3.1 Magnetic field of electric current in two parallel wires	119
15.4 Magnetic Flux – Electromagnetic Induction, Faraday Law, Lenz Law	119
16 Electromagnetic Radiation	120
16.1 Refraction and Reflection of Electromagnetic Radiation	123
16.1.1 Measurement of the Refractive Index	125
16.1.2 Polarisation of Electromagnetic Radiation	125
16.2 Absorption of Ultraviolet and Visible Light –UV-VIS Spectroscopy	127
16.3 Geometrical Optics	128

Forewords

Dear students,

one of the subjects that you must pass at the end of the first semester of your study is physics. During the years of teaching pharmacy students, I realised that for most of you, physics is not among your favourite subjects and many of you consider it futile for your professional life. Beyond the fact that it is a requirement of your study, there are several reasons to study basic physics, which offers this textbook. The physical properties of medicaments are important characteristics that must be taken into account in the process of their development, production, storage, and therapy. Moreover, physics (with mathematics) is considered to be the fundamental for other natural sciences, e.g. chemistry, that share the same basic principles and rules.

The topics in the textbook survey many fields of physics; however, not all, there are no quantum and relativistic physics. The available time for the lessons limits the material extent of the presented to a very brief introduction. You will find here descriptions of physical phenomena, their definitions, and mathematical prescriptions, i.e. formulas, with their meaning, often their origin or derivation, the conditions of possible application or restrictions of applications, and a few examples. The most important formulas are highlighted with red colour letters. In some cases are presented the mathematical rules, which were used in the preceding parts of the text, to help you understand the mathematical operations. However, basic maths skills of students are expected. Several experimental methods are described in more detail to offer a theoretical background for laboratory exercises from physics, which is also part of the subject physics and the exam from physics.

Finally, I express my gratitude to doc. RNDr. J. Gallová, CSc. for careful reading of the manuscript of this textbook and for her valuable suggestions and comments on it.

Author

1 Physical Quantities and Units

To formulate a quantity in everyday life, we use units of measurement. The amount of goods in stores is sometimes quantified in pieces (pair of shoes, clothes, etc.), but often we buy, for example, 1 kg of apples, oranges, potatoes, 2 litres of milk, 45 litres of petrol, etc. We plan our trips according to the length of the trip. We will not travel with a car 3500 km, prefer an aeroplane instead, but for a 50 km distance, the car is a better choice than the aeroplane.

These quantities, 1 kg, 2 litres, 45 litres, 3500 km, 50 km are clear and comprehensive values for people who share the same cultural background. In a different country or a different historical age, it can be different. For example, in the USA or in the UK, despite some efforts to use the metric system (SI – see later), in everyday life, the ‘old’ units are used, based on the Roman and Anglo-Saxon system. They use gallons instead of litres, feet, or yards instead of metres, miles instead of km, etc., and these units serve perfectly to express the required quantities as well. Problems arise in the case of communication of persons with different cultural backgrounds using different system of units. In the globalised world, there are efforts to unify units. The unified system helps people in everyday life, but probably more important is its impact on the economy and science, also, where the amount of goods must be expressed precisely as well as their properties, very often prescribed by local governments with laws, patents or other legal prescriptions. Despite the beneficial consequences of the unified units, the transfer from an old unit system to a new one used to be difficult, sometimes resulting resistance of a part of the population.

The system of units generally used, except for certain few countries (USA, Liberia, Myanmar), is the metric system, a decimal system of measure known as **The International System of Units (SI)** abbreviated from French *Système International (d’Unités)*. The SI system is defined with an international organisation, *The International Bureau of Weights and Measures (BIPM)*, which operates under exclusive supervision of the *International Committee for Weights and Measures*.

SI comprises a **coherent** system of units of measurement built on seven **base units** that are *second, metre (meter), kilogramme (kilogram), ampere, kelvin, mole, and candela*.

1.1 Base Units

SI involves seven base units, associated with seven base quantities listed in the following table.

Table 1. Base quantities of SI system with their symbols, units, and unit symbols.

Base quantity		Base unit	
Name	Typical symbol	Name	Symbol
time	t	second	s
length	$s, l, x, r, d, \text{etc.}$	metre	m
mass	m	kilogram	kg
electric current	I, i	ampere	A
thermodynamic temperature	T	kelvin	K
amount of substance	n	mole	mol
luminous intensity	I_v	candela	cd

The definitions of the units went through an evolution during history, where the first definitions were based on natural phenomena of general observations. Later, it was necessary to define them more

exactly. This requirement comes mainly from science and scientists and was completed in May 2019, when the base units were defined according to fundamental physical constants listed in Table 2:

Table 2. Fundamental constants of SI base units.

Definition constant	Symbol	Numerical value	Unit
hyperfine transition frequency of Cs	$\Delta\nu_{\text{Cs}}$	9 192 631 770	Hz, s^{-1}
speed of light in vacuum	c	299 792 458	m s^{-1}
Planck constant	h	$6.626\,070\,15 \times 10^{-34}$	J s
elementary charge	e	$1.602\,176\,634 \times 10^{-19}$	C
Boltzmann constant	k	$1.380\,649 \times 10^{-23}$	J K^{-1}
Avogadro constant	n_{A}	$6.022\,140\,76 \times 10^{23}$	mol^{-1}
luminous efficacy	K_{cd}	683	lm W^{-1}

Binding of the base units on fundamental constants means that: “... the realizations are separated conceptually from the definitions so that the units can, as a matter of principle, be realized independently at any place and at any time.”¹

Time (t) – second; s

Definition: The second, symbol s, is the SI unit of time. It is defined by taking the fixed numerical value of the caesium frequency $\Delta\nu_{\text{Cs}}$, the unperturbed ground-state hyperfine transition frequency of the caesium 133 atom, to be 9 192 631 770 when expressed in units of Hz, which is equal to s^{-1} .

$$1 \text{ s} = \frac{9\,192\,631\,770}{\Delta\nu_{\text{Cs}}} \quad \text{eq. 1}$$

Caesium 133 has an unpaired electron on the 6s orbital. The nucleus of this isotope has a non-zero nuclear spin (7/2). The unpaired electron has its spin orientated in the same direction as the nucleus (state with lower energy) or in the opposite direction (state with higher energy). The difference between the energies of these two states is constant, as well as the frequency ($\Delta\nu_{\text{Cs}} = 9\,192\,631\,770 \text{ s}^{-1}$) of electromagnetic radiation, which corresponds to this energy difference. The period ($T = 1/\Delta\nu_{\text{Cs}}$) equals $1/9\,192\,631\,770 \text{ s}$, so the electric and magnetic vector of electromagnetic radiation, which corresponds to this transition, makes exactly 9 192 631 770 oscillations during 1 s.

Historically, the definition of the second was related to the fraction 1/86400 of the mean solar day (one day is 86 400 s), but the length of the mean solar day is not constant, it is varying slightly in a period of a few years, but significantly during a long period.

Length (l, s, d) – metre (meter); m

Definition: Metre, symbol m, is the SI unit of length. It is defined by taking the fixed numerical value of the speed of light in vacuum c to be 299 792 458 when expressed in units m s^{-1} , where the second is defined in terms of the caesium frequency $\Delta\nu_{\text{Cs}}$.

$$1 \text{ m} = \left(\frac{c}{299\,792\,458} \right) \text{ s} \quad [c] = \text{m s}^{-1}$$

$$1 \text{ m} = \frac{c}{299\,792\,458} \frac{9\,192\,631\,770}{\Delta\nu_{\text{Cs}}} = 30.663319 \frac{c}{\Delta\nu_{\text{Cs}}} \quad \text{eq. 2}$$

In other words, the light travels during 1 second 299 792 458 m. This unit is based on a very basic physical constant, which is the speed of light in vacuum. Albert Einstein postulated in his Special Theory of Relativity that the speed of light in vacuum is constant for all observers, regardless of the motion of the light source.

¹ NIST Special Publication 330, 2019 EDITION, <https://doi.org/10.6028/NIST.SP.330-2019>

Historically, there were attempts to connect the metre with the length of equator and later with the length between two marks on a platinum-iridium alloy bar, specially prepared for this purpose. None of these methods was as precise as they claimed.

Mass (m) – kilogramme (kilogram); kg

Definition: Kilogramme, the symbol kg, is the SI unit of mass. It is defined taking the fixed numerical value of the Planck constant h to be $6.626\,070\,15 \cdot 10^{-34}$ when expressed in unit J s, which is equal to $\text{kg m}^2 \text{s}^{-1}$, where the metre and the second are defined in terms of c and $\Delta\nu_{\text{Cs}}$.

$$h = 6.626070 \cdot 10^{-34} \text{ kg m}^2 \text{s}^{-1}$$

Relation gives an exact expression for the kilogramme in terms of the three defining constants h , $\Delta\nu_{\text{Cs}}$ and c :

$$1 \text{ kg} = \left(\frac{h}{6.626\,070\,15 \times 10^{-34}} \right) \text{m}^{-2} \text{s}$$

$$1 \text{ kg} = \frac{(299\,792\,458)^2}{(6.626\,070\,15 \cdot 10^{-34})(9\,192\,631\,770)} \frac{h\Delta\nu_{\text{Cs}}}{c^2} \approx 1.475\,5214 \cdot 10^{40} \frac{h\Delta\nu_{\text{Cs}}}{c^2} \quad \text{eq. 3}$$

History. The unit of *kilogrammes* was based on a single artefact (platinum iridium cylinder; the international prototype of the kilogramme), which was reproduced several times and the copies served as national prototypes.

Electric current (I) – ampere; A

Definition: It is defined by taking the fixed numerical value of the elementary charge e to be $1.602\,176\,634 \times 10^{-19}$ when expressed in unit C (coulomb), which is equal to A s, where the second is defined in terms of $\Delta\nu_{\text{Cs}}$.

$$1 \text{ A} = \left(\frac{e}{1.602\,176\,634 \cdot 10^{-19}} \right) \text{s}^{-1}$$

$$1 \text{ A} = \left(\frac{e}{1.602\,176\,634 \cdot 10^{-19}} \right) \frac{\Delta\nu_{\text{Cs}}}{9\,192\,631\,770} \approx 6.789\,687 \cdot 10^8 \Delta\nu_{\text{Cs}} e \quad \text{eq. 4}$$

History. The ampere was that constant current, which, if maintained in two straight parallel conductors of infinite length of negligible circular cross section, and placed 1 metre apart in vacuum, would produce between these conductors a force equal to $2 \cdot 10^{-7} \text{ N/m}$.

Electric current is the amount of net positive charge moved from one point to another for one second. More often (but not always), the current is the flow of electrons. In that case, the direction of the electric current and the direction of the electron motion are opposite. The attractive force between the wires with current flowing in the same direction and the repulsive force of currents flowing in opposite directions comes from the generated magnetic fields of the moving charges along the wires.

Thermodynamic temperature (T) – kelvin; K

Definition: Kelvin, symbol K, is the SI unit of the thermodynamic temperature. It is defined by taking the fixed numerical value of the Boltzmann constant k to be $1.380\,649 \cdot 10^{-23}$ when expressed in the unit J K^{-1} , which is equal to $\text{kg m}^2 \text{s}^{-2} \text{K}^{-1}$, where the kilogram, metre, and second are defined in terms of h , c , and $\Delta\nu_{\text{Cs}}$.

$$1 \text{ K} = \left(\frac{1.380\,649 \cdot 10^{-23}}{k} \right) \text{kg m}^2 \text{s}^{-2}$$

or

$$1 \text{ K} = \frac{1.380\,649 \cdot 10^{-23}}{(6.626\,070\,15 \cdot 10^{-34})(9\,192\,631\,770)} \frac{\Delta\nu_{\text{Cs}} h}{k} \approx 2.266\,6653 \frac{\Delta\nu_{\text{Cs}} h}{k} \quad \text{eq. 5}$$

The effect of this definition is that one kelvin is equal to the change in thermodynamic temperature that results in a change of thermal energy kT by $1.308\,649 \cdot 10^{-23} \text{ J}$.

History. Kelvin was defined as the fraction $1/273.16$ of the thermodynamic temperature of the triple point of water. Figure 1 presents the phase diagram of chemically pure water, the pressure-temperature diagram, defining the conditions (coloured areas) of the existence of water in the solid (ice), liquid (water) and gaseous (vapour) phases and the transitions between them (the lines separating the coloured areas), the sublimation, melting, and evaporation. The point located at the intersection of these lines is the so-called triple point. There exists only one triple point for pure water, and its position is exactly defined (pressure: 610 Pa, temperature: 273.16 K). The $1/273.16$ fraction of the triple point temperature was the unit of the thermodynamic temperature. The starting point – zero point of the Kelvin scale was defined from the volume (or pressure) vs. temperature diagram of gases. The volume (or pressure) of gases decreases linearly with decreasing temperature when their behaviour is not significantly different from that of the ideal gases, when the interactions between the gas molecules/atoms are negligible, and when the volume occupied by the molecules/atoms is negligible in comparison with the overall volume of gas. This is usually valid at low pressures and for noble gases (He, Ne, Ar, etc.). By extrapolating the line connecting the points in volume-temperature or volume-pressure diagrams, the intercept of this line with the temperature axis is the zero point of this scale. At this point, the volume or pressure of the gas equals zero. This is only a hypothetical point; it is not possible to reach zero Kelvin.

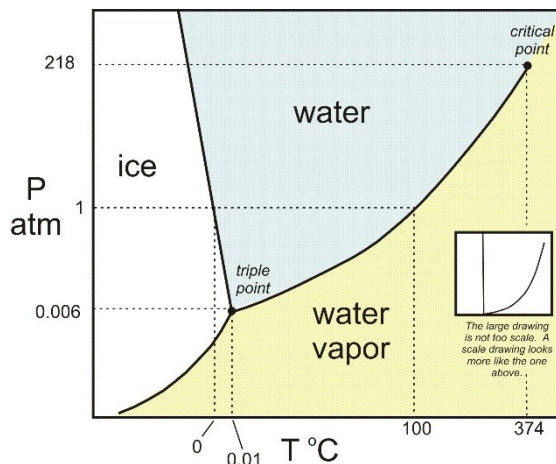


Figure 1 Phase diagram of water.²

Amount of the substance (n) – mole; mol

Definition: The mole, the symbol *mol*, is the SI unit of the amount of substance. One mole contains exactly $6.022\,140\,76 \cdot 10^{23}$ elementary entities. This number is the fixed numerical value of the Avogadro constant, N_A , when expressed in units mol^{-1} and is called the Avogadro number.

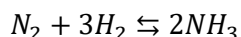
$$1\text{mol} = \left(\frac{6.022\,140\,76 \cdot 10^{23}}{N_A} \right) \quad \text{eq. 6}$$

Historically, mole was defined as the amount in substance of a system that contains as many elementary entities as there are atoms in 0.012 kilogrammes of carbon 12. ($N_A = 6.022 \cdot 10^{23}$ Avogadro's number)

Stoichiometric rule (Dalton's law, different from the rule on partial pressure of gas mixtures): When elements form compounds, the proportions of the components of these chemical compounds can be stated in small whole-number ratios. It is valid also for chemical reactions, where specific numbers of atoms and molecules interact; the ratio of these numbers of atoms, molecules, ions, etc. is a ratio of small whole numbers (integers) as well as the ratio of the amounts of the substance (number of moles). The ratio of their masses in reactions or compounds is usually not a whole number. For example, the

² Image Source: https://cdn.serc.carleton.edu/images/research_education/equilibria/h2o_phase_diagram_-_color.v2.jpg

reaction of nitrogen and hydrogen gases produces ammonia. The chemical equation for this reaction is as follows:



The ratio of the number of moles and the number of molecules of nitrogen and hydrogen in this reaction is 1: 3 and the moles and molecules of ratio of the nitrogen and ammonia and hydrogen and ammonia is 1:2 and 3:2. 1 mole of nitrogen reacts with 3 moles of hydrogen to produce 2 moles of ammonia. The ratios are expressed using small whole numbers (3,1 and 2). If we substitute the masses of the reactants and the product, 28.013 g of nitrogen reacts with 6.048 g of hydrogen to produce 34.061 g of ammonia. The ratio of the masses of the reactants is approximately 4.63 (28.013/6,048), the ratio of the masses of nitrogen and ammonia and hydrogen and ammonia are 0.822 (28.013/34.061) and 0.1776 (6.048/34.061), respectively. In the first case, the relations are ratios of small numbers, and these numbers define the ratios of the reacting particles. In the second case, the fractions contain the masses of the reactants, which directly carries no information about the reaction. In chemical reactions, atoms, and molecules (groups of atoms) react.

Luminous intensity (I) – candela; cd

Definition: The candela, symbol cd, is the SI unit of the luminous intensity in a given direction. It is defined by taking the fixed numerical value of the luminous efficacy of monochromatic radiation of frequency $540 \cdot 10^{12}$ Hz, K_{cd} , to be 683 when expressed in unit $lm W^{-1}$, which is equal to $cd sr W^{-1}$, or $cd sr kg^{-1} m^{-2} s^3$, where kilogramme, metre and second are defined in terms of h , c and $\Delta\nu_{Cs}$

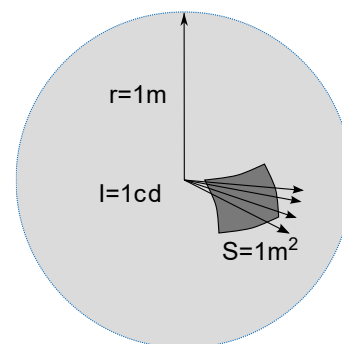


Figure 2. Scheme of luminous intensity definition.

$$1cd = \left(\frac{K_{cd}}{683} \right) kg m^2 s^{-3} sr^{-1}$$

or

$$1cd = \frac{1}{(6.626\,070\,15 \cdot 10^{-34})(9\,192\,631\,770)^2 683} (\nu_{Cs})^2 h K_{cd}$$

eq. 7

$$\approx 2.614\,830 \times 10^{10} (\nu_{Cs})^2 h K_{cd}$$

Historically, candela was defined as the luminous intensity in a given direction of a source that emits monochromatic radiation of frequency $540 \cdot 10^{12}$ hertz and has a radiant intensity in that direction of 1/683 watts per steradian.

In this definition, the power of light source (light green colour, wavelength ca. 555 nm) that passes or falls on the surface of a sphere with radius 1 m is $4\pi/683$ W (watt) or 1/683 W if the $1 m^2$ of this sphere of radius 1m is investigated. The power of 1/683W (watt) means that the energy of the radiation, which falls or passes through this area, is 1/683 J (joule) during one second.

1.2 Derived units

Other units of physical quantities, which are not listed among the base units, are defined by the combination (products of powers) of the base units. Up to now, no physical quantity is known, which is not possible to define with these base units or with their appropriate combination, i.e. the system of the base units is **complete**.

The **derived units** are derived from the definitions (formulas) of the physical quantities, where the units are substituted for the quantities.

Examples:

Speed (average) (v) is defined as the ratio of route (s) and time (t), $v=s/t$. $\rightarrow [v] = \text{m/s}$. The v in square brackets means that we are speaking about units of speed (v). m/s can also be written as $\text{m}\cdot\text{s}^{-1}$.

Average acceleration (a) is defined as the ratio of speed (v) change during time (t), $a = \Delta v/t \rightarrow [a] = \text{m}\cdot\text{s}^{-1}/\text{s}=\text{m}/\text{s}^2$.

Force (newton): mass times acceleration, $[F]=\text{kg m}/\text{s}^2 = \text{N}$

Pressure (pascal): force/area, $[P] = \text{N} / \text{m}^2 = \text{kg} / \text{ms}^2 = \text{Pa}$, etc.

Several derived units are named after famous physicists. For example, the unit of force is Newton (N). It is named after the famous English scientist Isaac Newton, who with his three basic laws of mechanics launched a revolution not only in physics but also in philosophy and society. His second law deals with force and the variation of momentum of motion in time.

$$F = \frac{d\,mv}{dt} \quad \text{eq. 8}$$

$$F = ma \quad \text{eq. 9}$$

Usually, the variation of momentum of motion is derivative of the speed with constant mass. The derivative of speed according to time is the acceleration (eq. 9). In relativistic physics, where the speed of the motion is comparable with the speed of light, this condition is not valid, the mass is not constant, and eq. 9 must be replaced by eq. 8. However, in the case, when the speed of the object is much less than the speed of light, eq. 9 works perfectly.

Table 3. Metric prefixes and their values.

Prefix	Label	Decimal value	Numerical factor
yocto	y	0.000 000 000 000 000 000 000 001	10^{-24}
zepto	z	0.000 000 000 000 000 000 000 001	10^{-21}
atto	a	0.000 000 000 000 000 000 001	10^{-18}
femto	f	0.000 000 000 000 001	10^{-15}
pico	p	0.000 000 000 001	10^{-12}
nano	n	0.000 000 001	10^{-9}
micro	μ	0.000 001	10^{-6}
milli	m	0.001	10^{-3}
centi	c	0.01	10^{-2}
deci	d	0.1	10^{-1}
----	---	1	10^0
deka	da	10	10^1
hecto	h	100	10^2
kilo	k	1 000	10^3
mega	M	1 000 000	10^6
giga	G	1 000 000 000	10^9
tera	T	1 000 000 000 000	10^{12}
peta	P	1 000 000 000 000 000	10^{15}
exa	E	1 000 000 000 000 000 000	10^{18}
zetta	Z	1 000 000 000 000 000 000 000	10^{21}
yotta	Y	1 000 000 000 000 000 000 000 000	10^{24}

The unit of force are derived from either equation eq. 8 or eq. 9 is $\text{kg}\cdot\text{m}\cdot\text{s}^{-2}$ substituting kg for the mass (m), m/s for the velocity (v) and s for time (t) to eq. 8, or kg and ms^{-2} to eq. 9, respectively. The unit $\text{kg}\cdot\text{m}\cdot\text{s}^{-2}$ is the definition of newton (N), the unit of force.

Using SI units in the calculation guarantees that the units of the result will also be in SI.

Metric prefixes. Multiple or fractional amounts of units are often specified with prefixes. The prefixes are decadic (based on ten), they denote the relation of the basic unit and the actual quantity with multiplication with a power of 10. The ratio between the subsequent prefixes is usually 1000 or 1 / 1000, except for the first two values next to the unit value. The names of the prefixes, their abbreviations and values are listed in the Table 3.

Planar angle in radians (rad). The angle in the plane can be defined using radians. The radian of an angle numerically equals the length of an arc of a circle of unit length, the endpoints of which are the intercepts of the circle and the lines of the angle (Figure 3).

The angle in radians of one complete revolution (360 degrees) is the length of the entire circumference divided by the radius, or $2\pi r/r$, or 2π . Thus, 2π radians is equal to 360 degrees, which means that one radian is equal to 57.295780 degrees.

The values of the angles in the radian are between 0 and 2π . The angle 2π rad is equal to the angle 0 rad. Trigonometric functions (sine, cosine, tangent, etc.) which use angles as arguments have regularly repeating function values. We call these functions periodic functions, because the value of the function $f(x)$ is the same when the independent variable (x) increases by the value of the period. For angles, the period is 2π (whole circle).

$$f(x + n \cdot 2\pi) = f(x) \quad \text{eq. 10}$$

n is an integer, 2π is the period, x is the independent variable, and $f(x)$ is the value of the function.

For example, the angle 20.3 rad is greater than 2π . We can rewrite it as the sum of a number smaller than 2π and the product of a whole number and 2π .

$$\begin{aligned} 20.3 \text{ rad} &= 1.45 \text{ rad} + 3 \cdot 2\pi \text{ rad} \\ \sin(20.3 \text{ rad}) &= \sin(1.45 \text{ rad}) = 0.9927 \end{aligned}$$

The arc length can be calculated as

$$l = \alpha r \quad [\alpha] = \text{rad} \quad \text{eq. 11}$$

An angle can also be defined using degrees (not degrees of Celsius!).

The scale in degrees has values between 0 and 360 and we denote it with the sign °. Degree is not involved in SI, but it is widely used.

We demonstrate a conversion of different units on the example of degree and radian. Let us take the angle from the previous example of 1.45 rad.

$$1.45 \text{ rad} \cdot \frac{360^\circ}{2\pi \text{ rad}} = 83.08^\circ \quad \text{eq. 12}$$

The idea of the calculation in (eq. 12) is that the quantity in 'old' units is multiplied by 1, expressed with a fraction, in which the numerator and denominator are equal. 360° is the same angle as 2π rad. Both define the angle of a whole circle. In the fraction, rad is in the denominator, making it possible to simplify the equation with it and obtain the result in the new unit, degree, placed in the numerator.

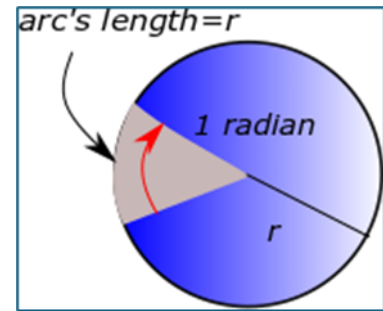


Figure 3. Definition of radian.

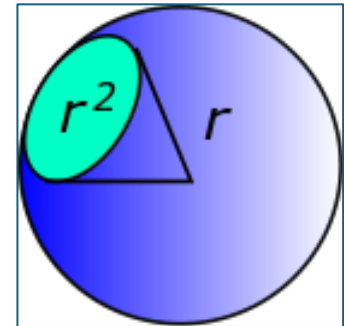


Figure 4. Definition of steradian.

The *steradian sr* is defined in a similar way to the radians applied to the spherical geometry (Figure 4). The angle (spherical angle - β) in the steradian is between values 0-4 π . 1 steradian is the angle that intercepts an area of 1m² from the surface a sphere of radius 1m.

Surface area:

$$S = \beta r^2 \quad [\beta] = \text{sr} \quad \text{eq. 13}$$

$$\text{sphere surface area} = 4 \pi \cdot \text{radius}^2$$

1.3 Non-SI Units

In some cases, we use other units, such as SI, that are associated with SI. These have their equivalents in SI, but they usually make the definition of quantities simpler and more useful in everyday life. These are, for example, hour, litre, tonne, electronvolt, etc. The unit hour is not included in the SI system, the unit of time in SI is second, but it would be very strange to propose a meeting in 1800 seconds or 1.8 kiloseconds instead of a half an hour. Another example is the degree of Celsius for the unit of temperature.

Temperature in Celsius. We usually use the Celsius temperature scale, which is not part of the SI system. The Celsius temperature scale is defined according to the melting and boiling point of pure water. The zero point ($t = 0^\circ \text{C}$) on the Celsius temperature scale is the melting point of the water, the boiling point is 100°C . 1/100 of the interval between the boiling point and the melting point of the water is 1°C , the unit value of this scale.

The conversion of temperature from Celsius to Kelvin is done according to eq. 14. The differences between two temperatures in kelvins and degrees Celsius are equal because the unit values of the two scales are equal, but the beginning of the thermodynamic temperature is shifted with a constant value.

$$T(\text{K}) = t(^{\circ}\text{C}) + 273.15 \quad \text{eq. 14}$$

2 Vectors and Scalars

Certain physical quantities are expressed as numbers (magnitude) and units only. We can illustrate it by examples of temperature, mass, etc. These are scalars. On the other hand, there exist physical quantities that are completely defined with their magnitude, direction, and orientation. These are vectors. A force is completely defined if the magnitude (how big is the force), its direction (e.g., in the horizontal plane), and its orientation (e.g., from left to right) are known.

An example of a vector is depicted on the Figure 5. The usual

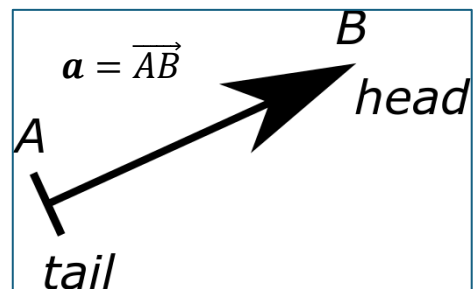


Figure 5. A general vector illustration.

notation for a vector is a bold letter, $\mathbf{a}$ or one or more letters with an arrow above the letter(s): $\vec{a}$, $\overrightarrow{AB}$. The starting point (origin) of the vector is called tail, the endpoint head, and the orientation of the vector is from tail to head. The length of the vector indicates the absolute value of the vector. Moving a vector in parallel with its original direction does not change the meaning of the vector.

In Figure 6 are three vectors. The lengths of the segments are equal, they are parallel, i.e., they directions are the same, and their orientation is the same, also. The three vectors in the picture are equal.

Usually, we place the vectors in the coordinate system and use the coordinate system to define the vector. In Figure 7 the OP segment defines a vector in the plane. It is placed in the Cartesian coordinate system with perpendicular axes. The tail of the vector O is located at the origin of the coordinate system. We can always do this by a parallel shift of any vector, as we explained above. The head of the vector (position) P is defined with its x- and y-coordinates. The length of the segment OP = r, the magnitude (absolute value) of the vector, is calculated from the Pythagorean theorem:

$$r = \sqrt{x^2 + y^2} \quad \text{eq. 15}$$

The angle between the x-axis and the vector is calculated according to:

$$\tan \theta = \frac{y}{x} \quad \text{eq. 16}$$

These equations are the transformation equations from the Cartesian coordinate system (after Descartes) to the polar coordinate system, where a point is defined with its distance from the origin of the coordinate system and with the angle between the x-axis and the line representing the vector.

Vector $\overrightarrow{OP}$ is represented with cartesian coordinates (x,y) or with polar coordinates (r,θ). The transformation from the polar coordinate system to the Cartesian coordinate system is as follows.

$$x = r \cos \theta \quad \text{and} \quad y = r \sin \theta \quad \text{eq. 17}$$

The absolute value of the vector is equal to the absolute value of the coordinate r in the polar coordinates:

$$|\overrightarrow{OP}| = \sqrt{r^2 \sin^2 \theta + r^2 \cos^2 \theta} \quad \text{eq. 18}$$

so

$$|\overrightarrow{OP}| = \sqrt{r^2} = |r|$$

Mathematical background: The sum of squares of the sine and cosine of an angle is equal to 1.

$$\sin^2 \theta + \cos^2 \theta = 1 \quad \text{eq. 19}$$

2.1 Operations with Vectors, Vector Addition

The procedure for adding two vectors is based on the rule that a vector moved in parallel is equal to the original vector. For example, we want to add vectors $\vec{A}$ and $\vec{B}$. The second vector's ($\vec{B}$) tail is shifted to the head of the first vector ($\vec{A}$), and the tail of the first vector and the head of the second vector is connected (Figure 8). The tail of the resultant vector $\vec{R}_1$ is the tail of the first vector and the head of the result vector is the head of the second vector.

$$\vec{R}_1 = \vec{A} + \vec{B} \quad \text{eq. 20}$$

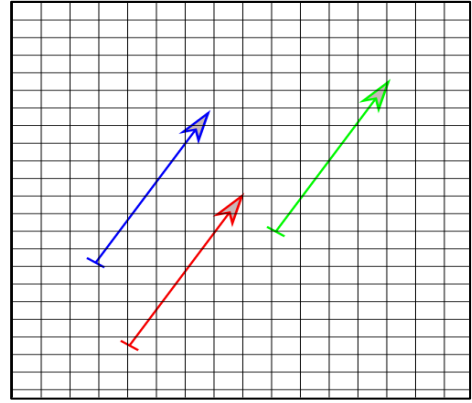


Figure 6. Three equal vectors.

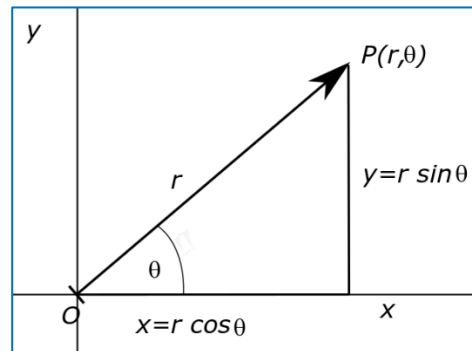


Figure 7. Relation of cartesian and polar coordinate systems.

The vector, which has the same magnitude, the same direction, but opposite orientation as the original vector $\vec{B}$ is the opposite vector $-\vec{B}$. To subtract a vector from another vector means adding to the first vector the opposite vector of the second vector (Figure 8).

$$\vec{R}_2 = \vec{A} - \vec{B} = \vec{A} + (-\vec{B}) \quad \text{eq. 21}$$

Let's have two vectors $\vec{A}, \vec{B}$ defined with their x, y components (A_x, A_y, B_x, B_y – the x,y coordinates of their head points, while the tails are in the origin) (Figure 9).

The result vector $\vec{R}$ is defined by its components (R_x, R_y) and its direction by the angle of the vector and the x axis.

$$\begin{aligned} R_x &= A_x + B_x & R_y &= A_y + B_y \\ \tan \alpha &= \frac{A_x + B_x}{A_y + B_y} \end{aligned} \quad \text{eq. 22}$$

The orientation of the result vector is defined with the segment, which begins at the origin of the coordinate system and ends at the point with coordinates R_x and R_y .

The magnitude of the result vector can be obtained from the magnitudes of the original vectors and from the angle between them using the cosine rule.

$$\begin{aligned} |\vec{R}|^2 &= |\vec{A}|^2 + |\vec{B}|^2 - 2|\vec{A}||\vec{B}|\cos(\pi - \theta) = \\ &= |\vec{A}|^2 + |\vec{B}|^2 + 2|\vec{A}||\vec{B}|\cos(\theta) \end{aligned} \quad \text{eq. 23}$$

In the cosine rule the $2|\vec{A}||\vec{B}|\cos\theta$ has negative sign, but the angle inside the triangle is $\pi - \theta$ and the angle between the vectors is θ (see mathematical background, Figure 10).

*Mathematical background: In trigonometry, the **law of cosines** (also known as the **cosine formula** or **cosine rule**) relates the lengths of the sides of a triangle to the cosine of one of its angles. Using the notation as in Figure 10, the law of cosine states and where $\pi - \theta$ denotes the angle contained between the sides of A and B.*

$$\cos \theta = -\cos(\pi - \theta) \quad \text{eq. 24}$$

$$\begin{aligned} (A + B)^2 &= A^2 + B^2 - 2AB \cos(\pi - \theta) = \\ &= A^2 + B^2 + 2AB \cos \theta \end{aligned} \quad \text{eq. 25}$$

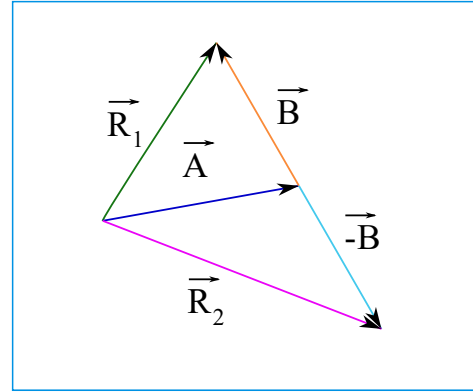


Figure 8. Vector addition

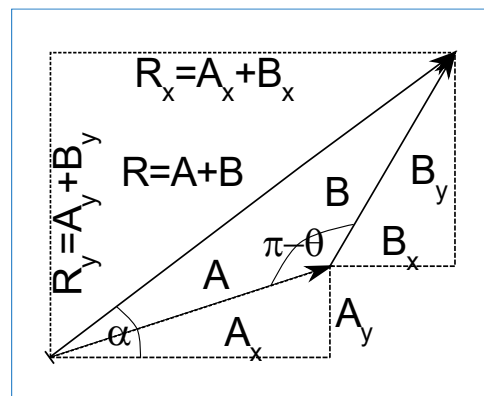


Figure 9. Calculating x and y coordinates of the resultant vector.

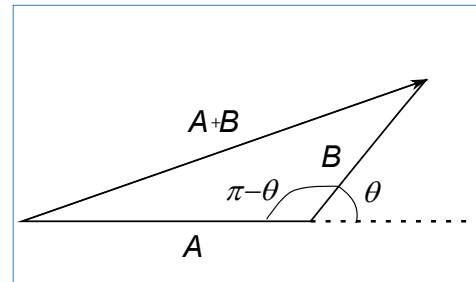


Figure 10. calculation of a triangle's side by cosine rule.

2.1.1 Vector Multiplication

a) Multiplication of a vector (**A**) with a constant (**k**). The result is a vector (**R**), with the same direction and orientation as the original vector. The vector x, y components, and magnitude are calculated by a simple multiplication of the original vector x, y components and magnitude with **k**, respectively.

$$|\vec{R}| = k|\vec{A}| \quad R_x = kA_x \quad R_y = kA_y \quad \text{eq. 26}$$

b) Scalar (dot) product of two vectors. The result (R) of a scalar product of two vectors is a *scalar*. The value of this scalar is given by:

$$R = |\vec{A}| \cdot |\vec{B}| \cos \theta \quad \text{eq. 27}$$

The operation notation is a dot, and the operation is also called the dot product. Another equation to calculate the scalar product is:

$$R = A_x B_x + A_y B_y \quad \text{eq. 28}$$

The condition for a scalar product of two perpendicular vectors (orthogonality) is:

$$A_x B_x + A_y B_y = 0 \quad \text{eq. 29}$$

c) Vector (cross) product of two vectors. The operation notation is a small „x“. The result is a vector.

$$\vec{R} = \vec{a} \times \vec{b} \quad \text{eq. 30}$$

The magnitude of the result vector (R) is as follows:

$$|\vec{R}| = |\vec{a}| |\vec{b}| \sin \theta \quad \text{eq. 31}$$

The direction of the result vector is perpendicular to the plane of the original vectors. The orientation is given by the right-hand rule (Figure 12); if the fingers of a right-hand are pointing in the direction from the first vector to the second vector, the thumb is pointing the orientation of the result vector. The cross product is anticommutative, changing the order of the vectors in operation results in a vector with opposite orientation.

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a} \quad \text{eq. 32}$$

When the original vectors are in the plane of the paper, the result vector will be perpendicular to the plane of the paper. To visualise any vectors, which are perpendicular to the plane of paper, a cross inside a circle is used; for the vectors pointing away from us, and with a dot in a circle for the vectors pointing toward us (Figure 13).

To calculate the Cartesian components of the result vector, we must solve a determinant (eq. 33), which has in the first row the unit vectors

$(\vec{i}, \vec{j}, \vec{k})$ in x, y and z directions, respectively. The second and third rows contain the corresponding components of the first and second vectors.

$$\vec{A} \times \vec{B} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \quad \text{eq. 33}$$

The solution of this determinant is as follows:

$$\vec{A} \times \vec{B} = \vec{i}(A_y B_z - A_z B_y) - \vec{j}(A_x B_z - A_z B_x) + \vec{k}(A_x B_y - A_y B_x) \quad \text{eq. 34}$$

The x, y, and z components of the result vector are $(A_y B_z - A_z B_y, -A_x B_z + A_z B_x, A_x B_y - A_y B_x)$.

The magnitude of the result vector is

$$|\vec{A} \times \vec{B}| = \sqrt{(A_y B_z - A_z B_y)^2 + (A_x B_z - A_z B_x)^2 + (A_x B_y - A_y B_x)^2} \quad \text{eq. 35}$$

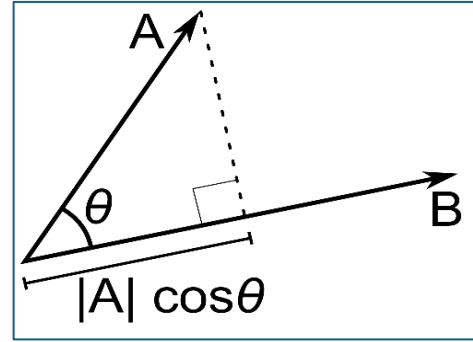


Figure 11. Angle between two vectors.

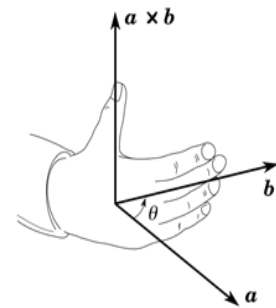


Figure 12. Right-hand rule.³



Figure 13. Drawing of perpendicular vectors to the plane.

³Image source <https://edurev.in/t/187555/Vector-product-of-two-vectors>

Mathematical background: the fastest way to solve a 3x3 determinant is to repeat the first two columns and multiply the numbers along the blue and red lines and make a sum of them. The products along the blue lines have a positive sign, and those along the red lines have a negative sign.

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - (a_{13}a_{22}a_{31} + a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33})$$

eq. 36

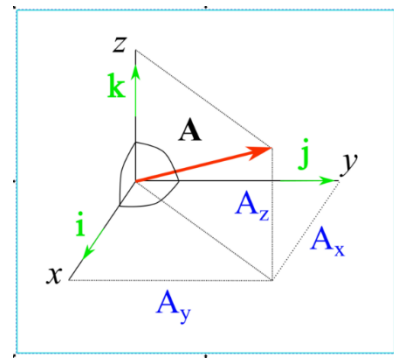


Figure 14. Three-dimensional vector.

3 Classical - Newtonian Mechanics

3.1 Kinematics

Isaac Newton (1642 – 1726/27)

Classical mechanics deals with the motion of macroscopic objects. Their speed is much lower than the speed of the light (no relativistic effects), so the mass of the object is constant. The time is running linearly, and neither of them depends on the speed of the motion. On the other hand, both the position of the object and the energy or momentum of the motion are possible to determine by unlimited precision independently on each other (no quantum effects) and the energy of the system may change its value continuously, without any restrictions.

At the beginning, we will deal with the motion of a point-like object. Using this approximation, the mass of the object is concentrated into one (dimensionless) point, and the trajectory, speed, acceleration, etc. of this point is studied.

3.1.1 Uniform Linear Motion

In the case of a uniform linear motion, the object is moving along a linear trajectory with constant speed. The route passed by the object is the same during equal time intervals. The basic requirement of the uniform linear motion is that acceleration (a) is zero or, equivalently, that the velocity is constant. *Velocity is the rate of the change of position with time, and acceleration is the rate of the change of velocity over time, in terms of both magnitude and direction.*

Here and in the following chapters, we consider the velocity as a vector and speed as a scalar; the magnitude of the velocity of a motion is the speed of the motion. If we consider a velocity constant, this condition involves a constant magnitude, direction, and orientation of the motion. On the other hand, a constant speed (magnitude of the vector), but varying direction means nonzero acceleration, not constant velocity, and in this case, it is not a linear uniform motion.

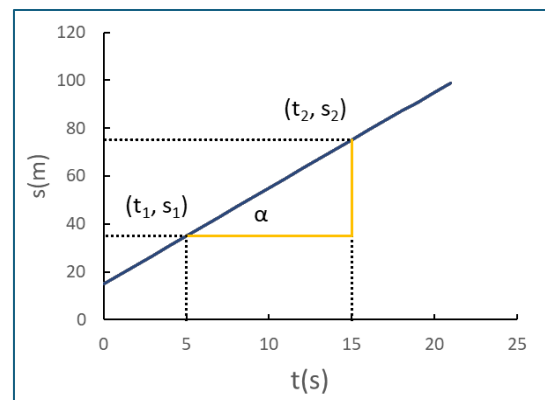


Figure 15. An example of the plot of the trajectory length as a function of time of uniform linear motion.

From the condition of zero acceleration, we can derive the equations for the speed (v) and the trajectory (s).

$$a = 0 \quad \text{eq. 37}$$

$$v(t) = \int a \, dt = at + \text{const} = 0 + \text{const} = v_0 \quad \text{eq. 38}$$

$$s(t) = \int v(t) \, dt = \int v_0 \, dt = s_0 + v_0 t \quad \text{eq. 39}$$

from the initial condition in $t = 0 \Rightarrow v(0) = v_0$, $s(0) = s_0$. The speed of the linear uniform motion is constant; the average speed is equal to the instantaneous speed at all times of the motion.

$$v_0 = \frac{ds}{dt} = \frac{\Delta s}{\Delta t} = \frac{s_2 - s_1}{t_2 - t_1} \quad \text{eq. 40}$$

The length of the route is a linear function of time (Figure 15); s_0 is the position of the object at the beginning (t_0). s_0 is the y intercept of the line $s=f(t)$, the speed v_0 is the slope, which determines the steepness of the line (Figure 15). (The key formulas are highlighted by red colour.)

Mathematical background:

The integration rule for power functions is as follows:

$$\begin{aligned} \int x^n \, dx &= \frac{x^{n+1}}{n+1} + \text{const} \\ \int k \, dt &= kt + \text{const} \end{aligned} \quad \text{eq. 41}$$

where k is a constant. If $k = 0$, then:

$$\int 0 \, dt = \text{const}$$

The value of the integration constant is found from an additional condition, e.g., from the initial speed, or position of the object, etc. If $t=0$ and $s=s_0$, then $\text{const} = s_0$ in eq. 39, $v=v_0$, $\text{const}=v_0$ in eq. 38.

The derivative and integral are inverse mathematical operations. Applying it to speed and acceleration; derivative of speed with respect to time is acceleration, and integral of acceleration with respect to time is speed:

$$a = \frac{dv}{dt} \quad \text{and} \quad v = \int a \, dt \quad \text{eq. 42}$$

The equation of a straight line in a plane is as follows.

$$y = ax + b \quad \text{eq. 43}$$

a and b are the parameters of the linear equation, slope, and y-intercept, respectively. The variables are x , y , independent, and dependent variables.

The angle between the line and the x axis α is defined (Figure 15, $x \equiv t$, $y \equiv s$):

$$\tan \alpha = a = \frac{y_2 - y_1}{x_2 - x_1} = \frac{dy}{dx} \quad \text{eq. 44}$$

Here, x_1, y_1, x_2, y_2 are the coordinates of two arbitrary points on the line.

3.1.2 Uniformly Accelerated Linear Motion

The trajectory of the object in the case of a uniformly accelerated linear motion is a straight line, but the length of the route does not increase linearly with time (Figure 16). The equations for speed (v) and route (s) are derived similarly from the acceleration as in the case of uniform linear motion, however, the acceleration is not zero.

$$v = \int a \, dt = at + \text{const} = v_0 + at$$

$$s = \int v dt = \int (at + v_0) dt \quad \text{eq. 45}$$

$$= s_0 + v_0 t + \frac{1}{2} at^2$$

The quantities with zero subscript are the initial speed and position of the object from the observer at the beginning of the motion, respectively. The speed of the uniformly accelerated motion is a linearly increasing ($a > 0$) or decreasing ($a < 0$) function. The slope of the line $v = f(t)$ (not shown) is the acceleration; the y-intercept is the initial speed (v_0). The path of the object is a parabola (blue curve Figure 16) with an y-intercept s_0 , which is the initial distance of the object from the reference point of the motion. The average speed is not equal to the instantaneous speed. The instantaneous speed equals the derivation of the route with time:

$$v = \frac{ds}{dt} = v_0 + at \quad \text{eq. 46}$$

It is the slope of the tangent (orange line, Figure 16) to the route (parabola, blue line) and its value differs for every moment.

The average speed of the uniformly accelerated motion is equal to the speed of a linear uniform motion, where the lengths of both routes are equal at the same times. The routes of the uniform linear motion (s_{lm}) and uniformly accelerated (s_{lam}) motion are as follows:

$$s_{lm} = s_0 + v_{lm}t$$

$$s_{lam} = s_0 + v_0 t + \frac{1}{2} at^2$$

$$s_0 + v_{lm}t = s_0 + v_0 t + \frac{1}{2} at^2 \quad \text{eq. 47}$$

$$v_{lm} = v_0 + \frac{1}{2} at$$

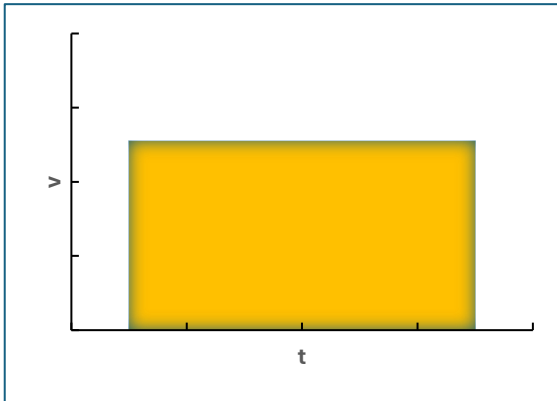


Figure 17. Speed of uniform linear motion as a function of time.

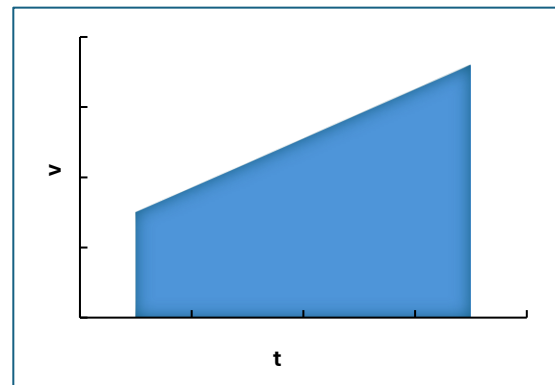


Figure 18. Speed of uniformly accelerated linear motion

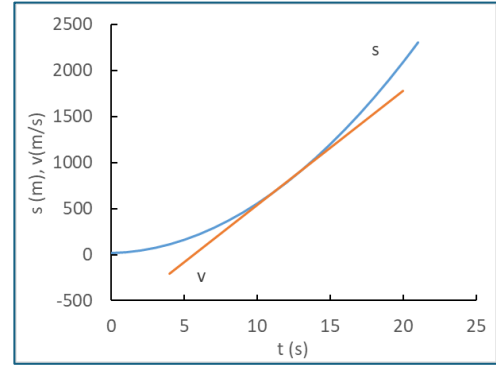


Figure 16. Plot of the trajectory as a function of time of uniformly accelerated linear motion (blue line) and speed as a tangent to an arbitrary point (orange line).

The average speed of the uniform accelerated motion depends linearly on time. A more general way is to use the definition of the route as a definite integral of the speed over time. The average speed is the route taken during a certain period divided by time. The route is equal to the area under the speed curve. On Figure 17 and Figure 18 the blue and yellow areas indicate the routes of the corresponding motions and are equal. The time intervals are also equal, so the speed of the uniform linear motion (the value at the top of the yellow rectangle (Figure 17) is equal to the average speed of the uniformly accelerated linear motion (Figure 18).

$$v_{average} = \frac{\int_{t_1}^{t_2} v(t) dt}{\int_{t_1}^{t_2} dt} = \frac{s_{total}}{t_{total}} \quad eq. 48$$

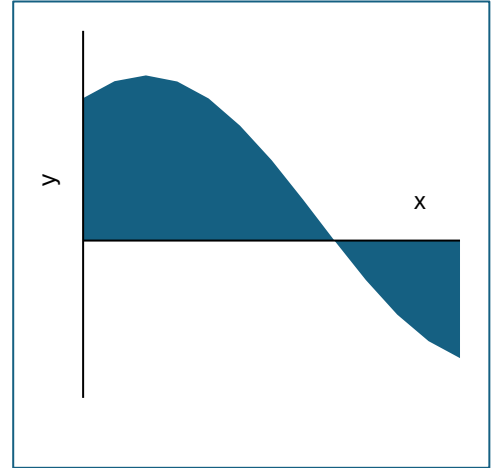


Figure 19. A general function with its integral as areas between the curve and the x-axis.

Mathematical background. The area (A) (eq. 49), which is between a curve $f(x)$ and the x-axis, its left and right edges are positioned to the lower (x_1) and upper (x_2) borders of the interval of a curve, is equal to the definite integral of the function. The part above the x-axis is positive, and the part below the x-axis is negative (Figure 19). $F(x_i)$ is the so-called primitive function (antiderivative) value at x_i .

$$A = \int_{x_1}^{x_2} f(x) dx = F(x_2) - F(x_1) \quad eq. 49$$

3.1.3 Freefall

Freefall is a special case of uniformly accelerated linear motion. The acceleration value (g , *gravitational acceleration*) is constant, slightly depending on the geographical position and altitude of the object above sea level. On Earth, its value is close to 9.81 ms^{-2} . On the Moon, the acceleration due to gravity is ca. $1/6$ of g on Earth, $g_{Moon} = 1.625 \text{ ms}^{-2}$. The equations are like equations of the linear uniformly accelerated motion:

$$\begin{aligned} v &= gt \\ h &= \frac{1}{2}gt^2 \end{aligned} \quad eq. 50$$

h is the altitude of the initial position above the surrounding surface.

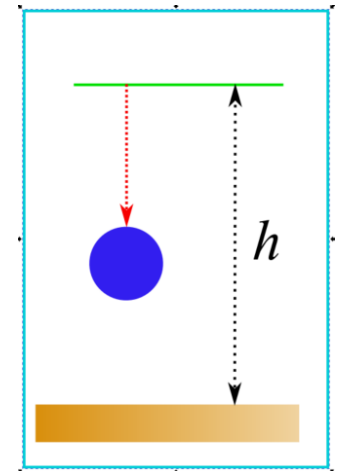


Figure 20. Freefall

3.1.4 Projectile Motion

Projectile motion is the case when a uniform linearly accelerated motion is combined with a freefall, i.e., the linear uniform motion is in the force field of the Earth. If the force field is a magnetic field and the moving object is a charged particle, the situation corresponds to mass spectroscopy (see the later). The object is launched (Figure 21) at an angle θ with a speed v_0 . Velocity ($\mathbf{v}$) and trajectory ($\mathbf{s}$) are vectors. The x- and y-components of the velocity are the following:

$$\begin{aligned} v_x &= v_0 \cos\theta \\ v_y &= v_0 \sin\theta - gt \end{aligned} \quad eq. 51$$

The horizontal (x) component of the velocity is a constant (we neglect air friction), and the x component of a trajectory is like the linear uniform motion.

$$s_x = \int v_0 \cos\theta \, dt = \quad \text{eq. 52}$$

$$= s_{x0} + v_0 t \cos\theta$$

The vertical (y) component of the velocity is a freefall with a non-zero upward initial velocity. The y component of the initial velocity has the opposite direction as the free-fall velocity. The y component of the velocity points at the beginning upward (dominance of the initial velocity over gravity), and at the top it is zero (equilibrium) later downward (dominance of gravity over initial speed).

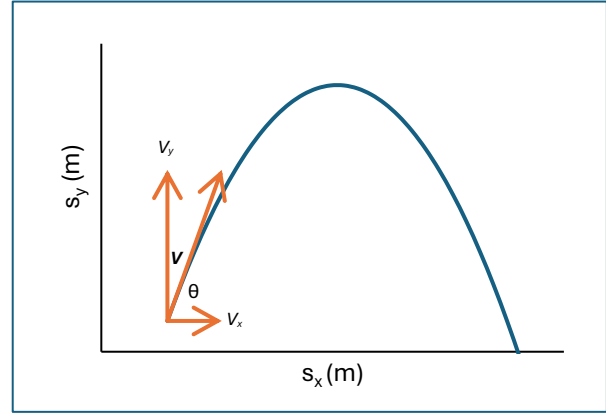


Figure 21. Projectile motion.

$$s_y = \int (v_0 \sin\theta - gt) \, dt = s_{y0} + v_0 t \sin\theta - \frac{1}{2}gt^2 \quad \text{eq. 53}$$

The time of flight is usually calculated from the boundary conditions; for example, in the moment of impact, s_y equals zero. Velocity (v) and trajectory (s) are calculated from the equations of its x and y components (eq. 51). The eq. 53 for s_y is a quadratic equation.

Boundary conditions.

a) The mathematical condition for the maximum (s_y) of a function is the following:

$$t = t_{peak} \Rightarrow \frac{ds_y}{dt} = v_y = 0 \quad \text{eq. 54}$$

$$\text{and} \quad \frac{d^2s_y}{dt^2} < 0$$

The highest point on the $s_y=f(t)$ curve is the moment (t_{peak}) when v_y , the component of the velocity, is zero and the acceleration a_y is negative (the second condition is valid for the whole time range).

From this condition we can find the position (time, t_{peak}) of the peak on the curve:

$$\frac{ds_y}{dt} = 0 = v_0 \sin\theta - gt_{peak}$$

$$t_{peak} = \frac{v_0 \sin\theta}{g}$$

$$\frac{d^2s_y}{dt^2} = -g < 0 \quad \text{eq. 55}$$

b) In the moment of impact of the object is on the ground (t_{max}), the motion is complete, and this point is the maximum distance of the motion. The s_y position is in that moment zero.

$$t = t_{max} \Rightarrow s_y = s_{y0} + v_0 t_{max} \sin\theta - \frac{1}{2}gt_{max}^2 = 0$$

rearranged

$$\frac{1}{2}gt_{max}^2 - v_0 t_{max} \sin\theta - s_{y0} = 0$$

The solution (of the quadratic equation) is as follows:

$$t_{max1,2} = \frac{v_0 \sin\theta \pm \sqrt{(-v_0 \sin\theta)^2 + 2gs_{y0}}}{g} \quad \text{eq. 56}$$

One of the roots of eq. 56 is positive, and one is negative or zero if $s_{y0}=0$. The negative root does not have physical meaning. The distance of the object in the horizontal direction in the moment of impact (distance in the x direction) is the following.

$$s_{xmax} = s_{x0} + v_0 \cos \theta \frac{v_0 \sin \theta + \sqrt{(-v_0 \sin \theta)^2 + 2gs_{y0}}}{g} \quad \text{eq. 57}$$

The velocity, its components, and the position of the object are available substituting times (t_{max} , t_{peak}) into the respective equations.

The constants s_{x0} , s_{y0} are the coordinates of the initial position of the object from the reference point. To simplify the equations, let us equalise them to zero. The equations for the positions are as follows:

$$\begin{aligned} s_x &= v_0 t \cos \theta \\ s_y &= v_0 t \sin \theta - \frac{1}{2} g t^2 \\ t_{max} &= \frac{2v_0 \sin \theta}{g} \\ s_{xmax} &= \frac{2v_0^2 \cos \theta \sin \theta}{g} = \frac{v_0^2}{g} \sin 2\theta \end{aligned} \quad \text{eq. 58}$$

The maximum distance of motion in the horizontal direction depends on the angle θ .

$$\begin{aligned} \frac{ds_{xmax}}{d\theta} &= 0 \\ \frac{d \frac{v_0^2}{g} \sin 2\theta}{d\theta} &= 0 \\ \frac{v_0^2}{g} 2 \cos 2\theta_{max} &= 0 \\ \cos 2\theta_{max} &= 0 \\ 2\theta_{max} &= \frac{\pi}{2} + n\pi \\ \theta_{max} &= \frac{\pi}{4} + n \frac{\pi}{2} \end{aligned} \quad \text{eq. 59}$$

The solution for the maximum angle is a periodic solution, with a period $\pi/2$ (90°), n is a whole number (integer). The solutions are $\pi/4$, $3\pi/4$, $5\pi/4$ (45° , 135° , 225° , etc.) From a physical point of view, the first solution is relevant, and the other means symmetric motions and curves reflected according to the x- and y-axes.

Mathematical background:

a) *Solution of a quadratic equation:*

$$ax^2 + bx + c = 0$$

is

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad \text{eq. 60}$$

A quadratic equation has roots x_1 and x_2 (real, imaginary, multiple), according to its discriminant (D):

$$D = \sqrt{b^2 - 4ac} \quad \text{eq. 61}$$

- $D > 0$; two real roots
- $D = 0$; a multiple (one) root
- $D < 0$; two imaginary roots, no real root.

b) *Extreme values of a function, maxima, and minima. In an interval where the function is defined and continuous, the function has a minimum or maximum, if its first derivative is*

zero, and in the case of a minimum, the second derivative is positive, and in the case of a maximum, negative.

$$\begin{array}{l} \frac{dy}{dx} = 0 \\ \text{Minimum} \quad \frac{d^2y}{dx^2} > 0 \\ \text{Maximum} \quad \frac{d^2y}{dx^2} < 0 \end{array} \quad \text{eq. 62}$$

c) Trigonometric relation used in eq. 58.

$$2\sin\alpha \cos\alpha = \sin 2\alpha \quad \text{eq. 63}$$

d) Derivatives of compound functions:

$$\frac{df(g(x))}{dx} = \frac{df}{dg} \frac{dg}{dx} \quad \text{eq. 64}$$

Example:

$$\frac{d\sin 2\alpha}{d\alpha} = 2\cos 2\alpha$$

4 Dynamics – Newton's Laws of Motion

The three laws of motion deal with the relation between the force acting on the object and the motion of the object. Their influence was significant not only on physics (science), but also on society and economy; certain economists tried to apply Newton's laws to economy and society. Without any assessment of their economic benefits, it shows the popularity and high respect to Newton's theories.

Newton's First Law: *Every object will remain at rest or in uniform motion in a straight line unless compelled to change its state by the action of an external force.*

Newton's Second Law: *The force is equal to the change in momentum (mv) per change in time.*

For a constant mass, the force equals mass times acceleration.

$$F = ma \quad \text{eq. 65}$$

Newton's Third Law: *For every action, there is an equal and opposite reaction.*

The first law (the law of inertia) states that to keep an object in motion or at rest, no force is required. Our everyday life experience is apparently in contradiction with it, because to travel from one location to another, we need the force of the engines of cars, trains, etc. The reason for this is the presence of forces that always act on us. These are the gravitational force, friction with the ground during a motion, air resistance, etc. The forces of the engines, the animals, are required to compensate for these forces around us. In Space, far from planets, stars, etc., objects moving in a straight line do not need a force to persist in motion. A body persists in its state of motion when no forces act on it or the forces acting on the body compensate each other; their sum is a zero vector. *Force is an influence that causes the change in the velocity of a moving object.*

A very common example of inertia is the situation that when travelling in a vehicle, a sudden brake might cause a fall of standing passengers. The reason is that the body of the traveller is in motion (relative to the street) and after the brake it tries to continue in this state. The traveller must compensate for the force caused by the braking with an opposite force to avoid falling. This force is proportional to its mass and to the deceleration; mass is the measure of inertia.

The unit of the force is newton.

$$[F] = 1 \text{ N} = 1 \text{ kg ms}^{-2} \quad \text{eq. 66}$$

Newton's second law defines the magnitude of this force:

$$F = ma = \frac{dmv}{dt} = m \frac{dv}{dt} \quad \text{eq. 67}$$

The product mv is the momentum and t is time, so the force is equal to the derivative (change) of the momentum in time. In the case that the speed is much less than the speed of light ($v \ll c$), the mass of the object is constant and the derivative (change) of the speed on time is the acceleration (eq. 68).

$$a = \frac{dv}{dt} \quad \text{eq. 68}$$

A constant force causes a constant acceleration (not a constant speed!) of the object. Force is a vector quantity (magnitude, direction, orientation), mass is a scalar, and acceleration (and velocity) is a vector. Force is the product of a vector (acceleration) with a scalar (mass). Its direction and orientation are defined with the direction and orientation of the acceleration (or velocity) and vice versa.

$$\vec{F} = \frac{dm\vec{v}}{dt} = m\vec{a} \quad \text{eq. 69}$$

Let us have an object (mass; m) that is accelerated with a constant force F to speed v_1 during time t_1 . According to Newton's second law:

$$\begin{aligned} F &= m \frac{dv}{dt} \\ \int_0^{t_1} F dt &= \int_0^{v_1} m dv \\ Ft_1 &= mv_1 \end{aligned} \quad \text{eq. 70}$$

The quantity on the right-hand side of the eq. 70 is the previously mentioned momentum of motion; on the left-hand side is the so-called impulse of force.

4.1 Conservation of momentum

In a closed system (no external influence on the system), the amount of momentum is constant.

$$\sum_i m_i v_i = 0 \quad \text{eq. 71}$$

The sign Σ stands for the sum of the momenta of all objects in the system. Let us have a closed system consisting of a boat (mass = m_b) and a person (mass = m_p), which are at rest (speed = 0) at the beginning (Figure 22).

The person jumps off the boat at speed v_p . The momentum of the system (boat+person) was zero at the beginning due to its zero speed. From the conservation law of the momentum, there also follows a zero momentum at the end of the process. However, the speed of the person was not zero (he jumped out), so his non-zero momentum is compensated with the momentum of the boat:

$$\begin{aligned} m_p v_p + m_b v_b &= 0 \\ m_b v_b &= -m_p v_p \end{aligned} \quad \text{eq. 72}$$

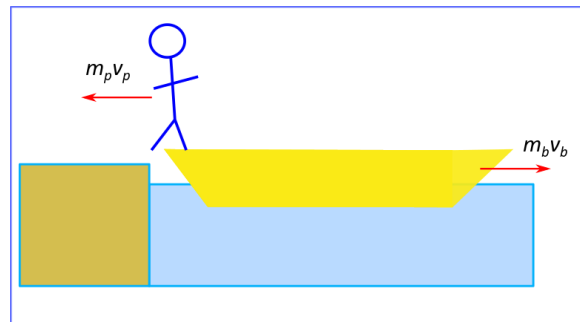


Figure 22. Conservation of momentum.⁴

⁴ Image source <https://www.wired.com/2013/10/a-closer-look-at-newtons-third-law/>

The momentum of the boat has the same magnitude (absolute value) as the momentum of the person, but opposite sign. The velocity (vector) of the boat is:

$$\vec{v}_b = -\frac{m_p \vec{v}_p}{m_b} \quad \text{eq. 73}$$

It means that the boat will move in opposite directions to the person.

The previous problem shows an example where the action invokes a reaction of equal magnitude but opposite direction.

Newton's third law of action and reaction:

$$F_1 = -F_2 \quad \text{eq. 74}$$

The person pushes the wall (Figure 23) with a force F_1 , and the wall is pushing the person with a force F_2 , which has the same magnitude and direction, but opposite orientation. These two forces are in equilibrium, and their resultant is zero.

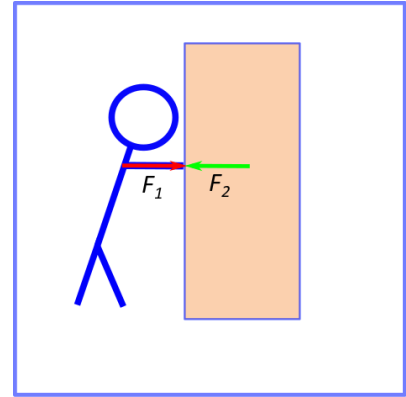


Figure 23. Action and reaction of forces.

4.2 Uniform Circular Motion

The trajectories of the motions, which were discussed in the previous parts, were connecting two points in the space with a straight line. There exist motions in which the trajectory is a closed curve; for example, a tramway can have a closed trajectory, parts of the engines circle, and the planets have an elliptic closed trajectory, etc. In the case of a uniform circular motion, the trajectory of the object is a circle. This motion is a periodic motion because the trajectory of the object repeats the same circle and regularly passes the same positions. The period of motion is $T[s]$, which is the time required to do one circle. Its reversed value, the frequency $f[s^{-1}]$, is the number of circles (rotations, turns) carried out during unit time. The angle $\theta[rad]$ is the angle caused by the radius of the circle during time t , the angular speed $\omega[rad s^{-1}]$ or $[s^{-1}]$ is the angle caused by the radius during unit time.

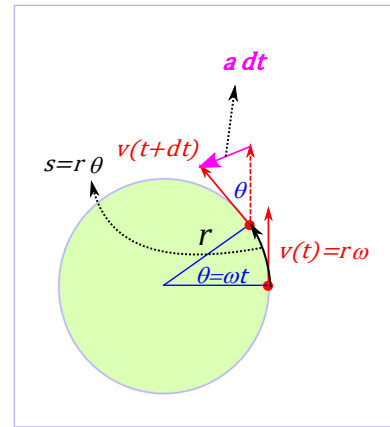


Figure 24. Circular motion

$$\omega = \frac{d\theta}{dt} \quad \text{eq. 75}$$

In the case of uniform circular motion, ω is constant and the angular acceleration (α) is zero. The formalism is like a uniform linear motion. The angle made by the radius during time t is:

$$\theta = \int \omega dt = \theta_0 + \omega t \quad \text{eq. 76}$$

Where θ_0 is the angle of the object from the reference point at the beginning of the motion. Choosing an appropriate reference point of the motion, it equals zero. Substituting the period T for time, the angle is equal to the angle of a whole circle ($2\pi \text{ rad}$, 360°):

$$2\pi = \omega T \Rightarrow \omega = \frac{2\pi}{T} \quad \text{eq. 77}$$

$$f = \frac{1}{T} \Rightarrow \omega = 2\pi f$$

The length of the trajectory, $s[m]$ performed by the object, is the length of the arc of the circle. Having the angle in radians, the length of the arc is the following:

$$s = \theta r \quad \text{eq. 78}$$

The motion of the object has a *tangential velocity* (v_T). The direction of its vector is perpendicular to the radius of the motion at every moment, so its direction changes permanently (Figure 24, red line). The magnitude of the tangential velocity is constant for uniform circular motion, but the change of direction causes the existence of non-zero acceleration (see later).

$$v_T = \frac{ds}{dt} = \frac{d\theta r}{dt} = \omega r \quad \text{eq. 79}$$

The relationship between the angular quantities (θ , ω) and their tangential equivalents (s , v_T) is obvious:

$$\text{tangential quantity} = \text{angular quantity} \cdot \text{radius}$$

In the previous definition, we omitted the vector aspect of the tangential velocity v_T . To define a vector, we must know its magnitude, direction, and orientation. The x and y components of the radius vector of the motion r are:

$$\vec{r}(r_x, r_y) = \vec{r}(r \cos(\omega t + \theta_0), r \sin(\omega t + \theta_0)) \quad \text{eq. 80}$$

similarly, as in Figure 7.

Tangential velocity, v_T , is defined as:

$$\begin{aligned} \vec{v}_T &= \frac{d\vec{r}}{dt} = \vec{v}_T \left(\frac{dr_x}{dt}, \frac{dr_y}{dt} \right) \\ v_x &= \frac{dr_x}{dt} = \frac{r d\cos(\omega t + \theta_0)}{dt} = -\omega r \sin(\omega t + \theta_0) \end{aligned} \quad \text{eq. 81}$$

and

$$v_y = \frac{dr_y}{dt} = \frac{r d\sin(\omega t + \theta_0)}{dt} = \omega r \cos(\omega t + \theta_0)$$

The magnitude of the tangential velocity vector is the following.

$$\begin{aligned} |\vec{v}_T| &= \sqrt{(-\omega r \sin(\omega t + \theta_0))^2 + (\omega r \cos(\omega t + \theta_0))^2} \\ |\vec{v}_T| &= \sqrt{\omega^2 r^2 [\sin^2(\omega t + \theta_0) + \cos^2(\omega t + \theta_0)]} = \pm \omega r \end{aligned}$$

The magnitude of the vector is a non-negative number, so:

$$|\vec{v}_T| = +\omega r = \omega |\vec{r}| \quad \text{eq. 82}$$

This result corresponds to eq. 79.

The relative angle between the radius and the tangential velocity is possible to find from the definition of the scalar product of two vectors.

$$\begin{aligned} \vec{r} \cdot \vec{v} &= |\vec{r}| \cdot |\vec{v}| \cos\theta \\ \vec{r} \cdot \vec{v} &= r_x v_x + r_y v_y \end{aligned}$$

So, the angle is as follows:

$$\cos\theta = \frac{r_x v_x + r_y v_y}{|\vec{r}| \cdot |\vec{v}|}$$

$$\cos\theta = \frac{r \cos(\omega t + \varphi)(-\omega r \sin(\omega t + \varphi)) + r \sin(\omega t + \varphi) \omega r \cos(\omega t + \varphi)}{r \omega r} = 0$$

$\cos\theta$ is zero, so θ is:

$$\theta = \frac{\pi}{2} + n\pi \quad \text{eq. 83}$$

where n is an integer, positive, negative whole number, or zero. The angle is a **right** angle (90°). The angle (difference) between the vector of the tangential velocity and the radius is $\pi/2$ (90°), the direction of the tangential velocity is perpendicular to the radius vector (eq. 83, Figure 25).

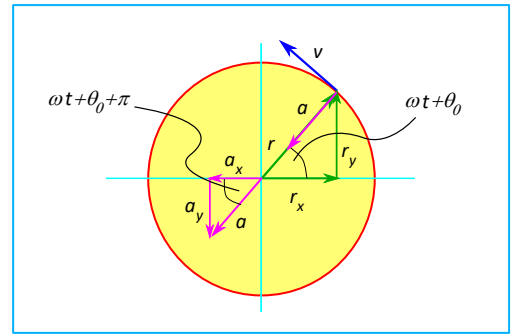


Figure 25. Circular motion with tangential velocity and centripetal

4.3 Centripetal Acceleration a_c , Centripetal Force F_c

To keep a body on a circular trajectory, it is necessary to exert a force (centripetal force $\vec{F}_c$) acting on the body toward the centre of the circle; otherwise, the body would leave its circular trajectory. The permanent change in the direction of the tangential velocity is related to a centripetal force $\vec{F}_c$ and centripetal acceleration $\vec{a}_c$. To define these vectors, the magnitude, direction, and orientation are required.

The x and y components of the centripetal acceleration $\vec{a}_c$ are:

$$a_{cx} = \frac{d^2 r_x}{dt^2} \quad a_{cy} = \frac{d^2 r_y}{dt^2}$$

Substituting for r_x and r_y from eq. 80 and performing the derivation:

$$a_{cx} = -r\omega^2 \cos(\omega t + \theta_0) \quad a_{cy} = -r\omega^2 \sin(\omega t + \theta_0) \quad \text{eq. 84}$$

The magnitude of the acceleration vector is:

$$|\vec{a}_c| = \sqrt{[-r\omega^2 \cos(\omega t + \theta_0)]^2 + [-r\omega^2 \sin(\omega t + \theta_0)]^2} \quad \text{eq. 85}$$

$$|\vec{a}_c| = r\omega^2 = \frac{v_T^2}{r}$$

The direction of the vector is the same as the direction of the radius because the respective components have the same trigonometric functions and arguments (eq. 80 and eq. 84). The orientation is opposite to the orientation of the radius, due to the negative signs of both components in the eq. 84, that is, the centripetal acceleration vector points at every moment from the object to the centre of the circle (Figure 25). The centripetal force is defined from Newton's second law (eq. 67):

$$\vec{F}_c = m\vec{a}_c$$

The direction and orientation are the same as for the centripetal acceleration and the magnitude is

$$F_c = mr\omega^2 = m \frac{v_T^2}{r} \quad \text{eq. 86}$$

To compensate for the centripetal force, there must exist another force (with a corresponding acceleration). It is called centrifugal force F_f (centrifugal acceleration); it has the same magnitude, the same direction, but opposite orientation to the centripetal force.

The centrifugal force is exploited in households (washing machine centrifuges) and in laboratories to separate components of suspensions (Figure 27). The effectivity of the centrifuges is expressed as the relative centrifugal force (RCF), which is the ratio between the centrifugal force and gravity.

$$RCF = \frac{F_f}{G} = \frac{m\omega^2 r}{mg} = \frac{4\pi^2 f^2 r}{g} \quad \text{eq. 87}$$

It is clear from the eq. 87, that the efficiency of the process can be improved by increasing the frequency (f) of the centrifuge or by increasing the radius (r) of the rotor.

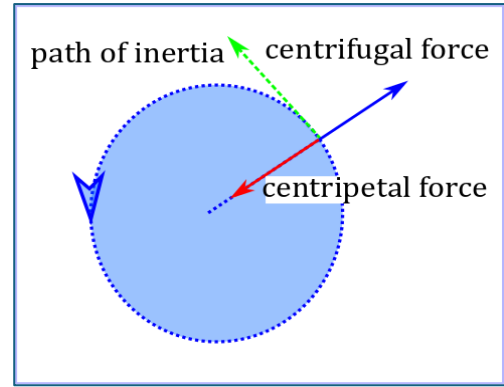


Figure 26. Centripetal and centrifugal forces.

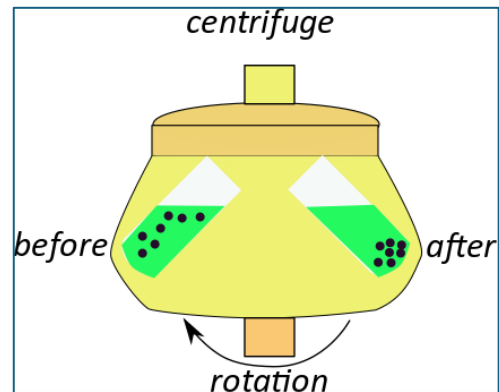


Figure 27. Scheme of a centrifuge.

5 Newton's Law of Universal Gravity

Newton's Law of Universal Gravity states that every object attracts every other object in the universe with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between their gravity centres.

$$F_G = \kappa \frac{m_1 m_2}{r^2} \quad \text{eq. 88}$$

The gravitational force is attractive and never repulsive. This is a difference compared to the electrostatic force between two charges, which is attractive (unlike charges) or repulsive (like charges). The gravitational force is decreasing quadratically with increasing distance of the objects. The constant κ is the universal gravitational constant; its value is:

$$\kappa = 6.673 \cdot 10^{-11} \text{Nm}^2 \text{kg}^{-2} \quad \text{eq. 89}$$

On Earth, the gravitational force is the following:

$$F_g = \kappa \frac{M_E m}{R_E^2} \quad \text{eq. 90}$$

where M_E is the mass of Earth ($M_E = 5.98 \cdot 10^{24} \text{kg}$) and its radius is R_E ($R_E = 6.38 \cdot 10^6 \text{m}$). The gravitational acceleration constant g is the following:

$$g = \kappa \frac{M_E}{R_E^2} = 9.81 \text{ms}^{-2} \quad \text{eq. 91}$$

The numerical value of g comes from the eq. 91, may also differ slightly, depending on geographic altitude and geographic location. The shape of the Earth is not an ideal sphere and R_E is varying. On other planets, the gravitational acceleration is different (according to the mass and the radius of the planet); for example, on the Moon it is approximately six times lighter than on the Earth. In this example, we can demonstrate the difference between the mass and the weight (gravitational force). The mass of the objects is the same on Earth or on the Moon; however, its weight on the Moon is six times less than on the Earth. 1 kg of sugar on Earth is 1 kg on the Moon; also, we can make sweet with it the same volume of tea; however, it would be six times easier to carry it home from a store on the Moon than on the Earth.

6 Mechanical Energy (E), Mechanical Work (W), Power (P)

The capacity of a system to perform work is the energy of the system. Different types of energy are known according to their origin or source (mechanical, thermal, chemical, nuclear, etc.). A similar classification of the work is valid as well. Energy is possible to change to work and work to energy, and one type of energy or work is possible to change to another type of energy or work. The energy of the Sun comes from the nuclear fusion of hydrogen, resulting in electromagnetic radiation, which heats the planet Earth. Radiation is absorbed by plants and converted to chemical compounds. The organic compounds of the plants were after a long period of time carbonised to coal, and the stored energy serves as a fuel or as a source to generate electricity in thermal power plants or as heat in households. The energy in mechanics is kinetic energy

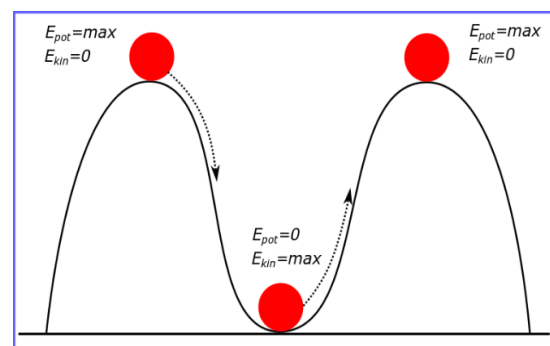


Figure 28. Transformation of potential energy to kinetic energy and vice versa.

or potential energy. The kinetic energy is the energy of the moving object; the potential energy is the positional energy in a force field (e.g., gravitational field). In an isolated system, the sum of the kinetic and potential energy is constant. In Figure 28 the ball, that rests at the top of the hill, has a maximum potential energy and zero kinetic energy. Rolling down from the hill, its potential energy, changes to kinetic energy and at the bottom, the kinetic energy is maximum, and the potential is zero. When we neglect energy loss (friction), the kinetic energy of the ball in the bottom position equals the potential energy on the top. The ball will go up the next hill to the same height as it was at the beginning of the process, and it will have the same potential energy as at the beginning or kinetic energy at the bottom of the hill. The total energy during the whole process is constant. The units of energy (E) and work (W) are joules.

$$[E] = [W] = 1J = 1Nm = 1kg\ m^2s^{-2}$$

Mechanical work is the scalar product of a force (F) that acts on an object and causes a displacement (s) of this object.

$$W = \vec{F} \cdot \vec{s} = Fs \cos\theta$$

eq. 92

Angle θ is the angle between the vector of the acting force and the displacement of the object (Figure 29). The work has the biggest value, if the force and the trajectory of the object have the same direction and orientation, but the work is equal to zero, when the force is perpendicular to the route. Angles can have values between 0° - 90° and 270° - 360° , when $\cos\theta$ is positive. If the angle is between 90° and 270° , the orientation of the trajectory and the acting force component (F) are opposite. In that case, the force will not move the object along that trajectory. Mechanical work is usually positive (when results in an increase of the kinetic energy), when the direction of the acting force and the motion directions are the same, but if they are opposed the work has a negative value (results in a decrease of the kinetic energy). For example, braking of a moving object by pulling it in opposite direction as its motion. The useful part of the force (component of force parallel to displacement, $F \cos\theta$) is less than the magnitude of the force if $\theta \neq 0$. Part of the force is a waste force (component of the force perpendicular to displacement, $F \sin\theta$) (Figure 29). The effectiveness of the force is the highest (the work has the highest value) if the trajectory of the movement and the force have the same direction ($\theta = 0$).

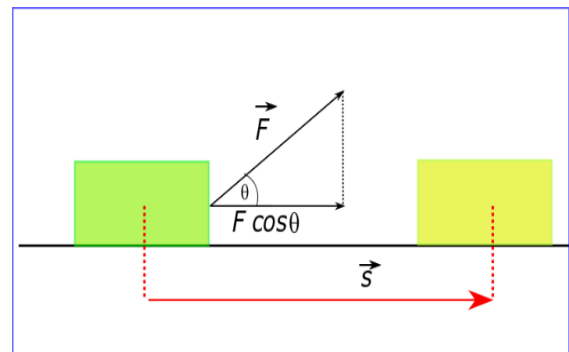


Figure 29. Mechanical work as scalar product of force and displacement of an object.

System in equilibrium and steady states. The examples in Figure 28 and Figure 30 allow us to define different steady states of the system. The ball, properly placed at the top of the hill, remains in its position until a force does not move it out of its steady state. The ball resting at the bottom of the hill will be in a steady state as well. However, there is a substantial difference between these two steady states. When the ball on the top is moved a little, it will not return to the top but rolls down while it finds a place on the bottom, where it will be in steady state as well.

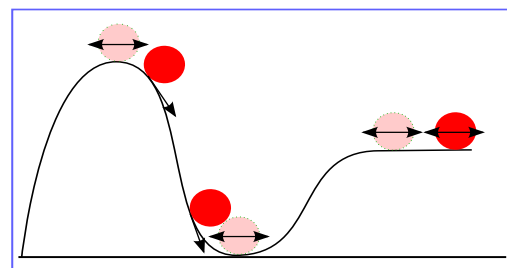


Figure 30. Different steady states. Pink balls represent the original steady state positions; red balls are out of steady state.

On the other hand, when the ball on the bottom is moved a little, it will return to its initial position. The ball on top is in an unstable steady state; a small disturbance causes the system to leave its initial state and find a new one. The ball on the bottom is in a stable steady state, and after a small disturbance, the system will relax back to the initial state. For completeness, there exists a third type of steady state; indifferent. Using the example of the ball, we can imagine that the ball lies on a flat horizontal plane (Figure 30, right position). A small disturbance moves the ball from its place to another position, and it will rest in this new position. It is apparently like the unstable steady state, but the difference between the unstable and indifferent steady states is that in the first case the ball new position will not depend on the force which moves it away but on the terrain around it (minimum position); in the case of the indifferent steady state, the terrain is uniform, the trajectory depends on the motion parameters (moving force, friction), and the new (indifferent) steady state may be at any point on the surface. The mathematical description of these states is based on the differential of the potential energy. In a steady state:

$$dE_{pot} = 0$$

In a one-dimensional (x) system:

eq. 93

$$\frac{dE_{pot}(x_{eq})}{dx} = 0$$

Geometrical interpretation: The slope of the tangent (black arrows on Figure 30) in the potential energy diagram in the steady state position (x_{eq}) is parallel to the x-axis (horizontal). This condition is valid for all three types of steady state. The steady state is unstable, when the potential energy has a maximum at the given initial position (x_{unst} , Figure 30 – left position):

$$\frac{d^2E_{pot}(x_{unst})}{dx^2} < 0 \quad \text{eq. 94}$$

and stable when the potential energy has a minimum at the given initial position (x_{st} , Figure 30 – middle position):

$$\frac{d^2E_{pot}(x_{st})}{dx^2} > 0 \quad \text{eq. 95}$$

This is an **equilibrium state**.

In the case of the indifferent steady state (Figure 30 - position right), all higher-order derivatives are zero.

$$\frac{d^n E_{pot}(x_{indif})}{dx^n} = 0 \quad \text{eq. 96}$$

where $n=1, 2, 3, \dots$ is the order of derivation.

Kinetic energy is the energy (work) required to increase the speed of the object from zero to v . The work performed by force F is equal to the kinetic energy E_{kin} . From Newton's second law:

$$Fdt = dp = m dv \quad \text{eq. 97}$$

and the infinitesimal work to increase the speed is (valid for $v \ll c$, c speed of the light):

$$\begin{aligned} dW &= Fds = Fvdt \\ Fvdt &= vdp = mvdv \\ dW &= dE_{kin} = mvdv \\ E_{kin} &= \int_0^v mvdv = \frac{1}{2}mv^2 \end{aligned} \quad \text{eq. 98}$$

The potential energy of the object is calculated as the work (performed by the acting force, F_{act}) required to move the object against an external force (F_{ext} , gravity).

$$dW = \overrightarrow{F_{act}} d\vec{x} = dE_{pot} = -\overrightarrow{F_{ext}} d\vec{x} \quad \text{eq. 99}$$

The external force in Figure 31 is gravity, the weight of the object. The potential energy of an object in height h above the ground is equal to the following:

$$E_{pot} = \int_0^h F_{act} dx = \int_0^h mg dx = mgh \quad \text{eq. 100}$$

The previous equations do not consider the 'speed' of the work, the time it takes to perform the action. The speed of performing a work is defined as the power (P). Power is the amount of work (energy) performed during the unit time.

$$P = \frac{W}{t} \quad \text{or} \quad P = \frac{dW}{dt} \quad \text{eq. 101}$$

In the first case in eq. 101 we have the average power and, in the second case, the instantaneous power (similarly to the average speed and instantaneous speed). The unit of power is watt (W):

$$[P] = 1 \text{ W} = 1 \text{ Js}^{-1}$$

Sometimes, only a part of the input energy (work) is useful; the remaining portion is lost due to friction of the spinning parts of the machines, air resistance, changed to waste heat of electrical spare parts, etc. The ratio of useful work (energy) and total input work (energy) is the efficiency (η) of the process.

$$\eta = \frac{W_{out}(E_{out})}{W_{in}(E_{in})} \quad \text{or} \quad \eta\% = \frac{W_{out}(E_{out})}{W_{in}(E_{in})} \times 100\% \quad \text{eq. 102}$$

Efficiency has no unit; it is either a number between zero and one or expressed as a percentage between 0% and 100%. For example, a tungsten filament bulb (Figure 32), with input energy of 100 J, releasing 10 J of light and 90 J of heat; has efficiency 0.1 (10%) if we use it for lighting. However, the same bulb will have a 0.9 (90%) efficiency if we use it for heating.

Problem: An airplane with a mass of 3000 kg reaches a flight altitude of 1 km and a speed of 180km/h after 1 minute of flight. Find the power of the plane engine! Neglect the air friction!
Solution.

A. The airplane has potential and kinetic energy. The sum of these energies equals the work of the engine.

$$E_{potential} = mgh = 3000 \times 9.81 \times 1000 = 29\,430\,000 \text{ J}$$

$$v = 180 \text{ km/h} = 50 \text{ m/s}$$

$$E_{kinetic} = \frac{1}{2}mv^2 = \frac{1}{2} \times 3000 \times 50^2 = 3\,750\,000 \text{ J}$$

$$P = \frac{E_{potential} + E_{kinetic}}{t} = \frac{29\,430\,000 + 3\,750\,000}{60} = 553\,000 \text{ W} = 553 \text{ kW}$$

The power of the engine was 553 kW.

This is a quick and simple solution. Another way is to calculate the flight force of the engine, the path, and its angle with respect to the horizontal direction.

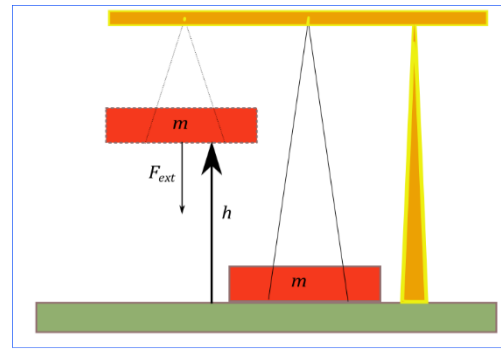


Figure 31. Potential energy of a load elevated by a crane.

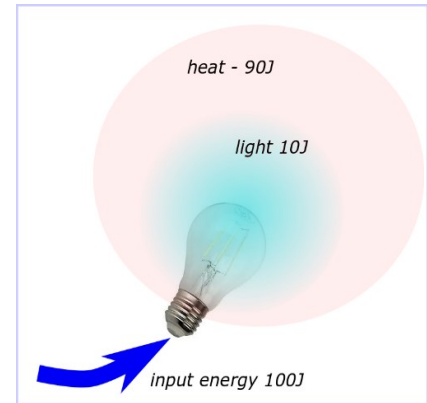


Figure 32. Distribution of energy of a filament bulb.

B. Acceleration is calculated from the final speed and the time of the motion.

$$a = \frac{v}{t} = \frac{50}{60} = 0.8\dot{3} \text{ ms}^{-2}$$

supposing a uniformly accelerated motion.

The route of this motion is the following.

$$s = \frac{1}{2}at^2 = \frac{1}{2} \times 0.8\dot{3} \times 60^2 = 1500 \text{ m}$$

The angle between the trajectory and the horizontal take-off area is calculated from the flight altitude ($s_y=1000 \text{ m}$), which is the component y of the trajectory, and from the trajectory ($s=1500\text{m}$).

$$\alpha = \arcsin \frac{s_y}{s} = \arcsin \frac{1000}{1500} = 41.810^\circ = 0.72973 \text{ rad}$$

The horizontal component of the trajectory s_x is:

$$s_x = s \cos \alpha = 1500 \cos 41.810^\circ = 1118.034 \text{ m}$$

The acceleration x and y components are the following:

$$a_x = a \cos \alpha = 0.62113 \text{ ms}^{-2}$$

$$a_y = a \sin \alpha = 0.5 \text{ ms}^{-2}$$

Gravitational acceleration acts on the airplane in the y direction, so the above calculated a_y component is the difference between the y component of the acceleration resulting from the engine (a_{engine}) and the gravitational acceleration (g). The a_x component is the same as the x component of the engine acceleration (we neglected air friction). The acceleration components that result from the engine are the following:

$$a_{x \text{ engine}} = 0.62113 \text{ ms}^{-2}$$

$$a_{y \text{ engine}} = a_y + g = 0.5 + 9.81 = 10.365 \text{ ms}^{-2}$$

$$a_{\text{engine}} = \sqrt{a_{x \text{ engine}}^2 + a_{y \text{ engine}}^2} = 10.38415 \text{ ms}^{-2}$$

The force of the engine is the following:

$$F = m a_{\text{engine}} = 31\,152.446 \text{ N}$$

The angle of the force action and the horizontal plane are:

$$\beta = \arcsin \frac{a_{y \text{ engine}}}{a_{\text{engine}}} = \arcsin \frac{10.365}{10.38415} = 86.571^\circ = 1.510945 \text{ rad}$$

The angle γ between the trajectory s and the force is the difference of β and α .

$$\gamma = \beta - \alpha = 86.571 - 41.810 = 44.760^\circ$$

The engine work is as follows.

$$W = F s \cos \gamma = 31\,152.446 \times 1500 \times \cos 44.760^\circ = 33\,180\,275.11 \text{ J}$$

$$P = \frac{W}{t} = \frac{33\,180\,275.11}{60} = 553\,004.59 \text{ W}$$

The results of A and B are almost the same; the difference is the rounding error of calculation B. Solution A is short and more precise than solution B. On the other hand, solution B provides a complete analysis and a detailed picture of the motion, providing a trajectory, force, and angles between them and their component.

7 Harmonic motion - Harmonic Oscillator

The vertical component of the trajectory of a uniform circular motion (Figure 33) or an oscillation of a load on a spring (if we neglect dumping, Figure 34) is a sine dependence on time. The parameters of the oscillation motion are period T , frequency f , amplitude A , phase shift φ . The equation of trajectory is as follows:

$$y(t) = A \sin(\omega t + \varphi) \quad \text{eq. 103}$$

where A is the amplitude; the maximum value of the dislocation from the reference point, ω – angular speed, φ – phase shift, the shift of the wave along the horizontal axis. The frequency f and period T are defined similarly to the case of a uniform circular motion.

$$\omega = 2\pi f \quad f = \frac{1}{T}$$

The period (T) is the distance of identical positions of the wave on the time axis (Figure 34). The speed of the motion is not constant, but it is a periodic function, too:

$$v(t) = \frac{dy(t)}{dt} = A\omega \cos(\omega t + \varphi) \quad \text{eq. 104}$$

as well as the acceleration:

$$a(t) = \frac{dv(t)}{dt} = -A\omega^2 \sin(\omega t + \varphi) = -\omega^2 y \quad \text{eq. 105}$$

From the properties of the sine and cosine functions, we can assess that the speed function is shifted by 90° according to the position ($y(t)$) function and the acceleration is shifted by 90° compared to the speed (Figure 35). The amplitude of the speed is $A\omega$. The acceleration is a sine wave as $y(t)$ with the same frequency and the same phase shift. The amplitude is $-A\omega^2$, the negative sign means that the sine curve of the acceleration is flipped upside down (along the time axis). The speed has a maximum value when the object is in the position of $y=0$ and equal to zero, when $y=A$ or $y=-A$. In the case of a harmonic oscillator (Figure 34, a load on a spring), the relationship between force and acceleration is according to Newton's second law (eq. 67) and from the calculated acceleration (eq 105):

$$F = ma = -m\omega^2 y \quad \text{eq. 106}$$

In the elastic deformation region of the string, the relationship between force and elongation obeys Hooke's law. Hooke's law states that the force (F) needed to expand or compress a spring by a distance y scales linearly with respect to that distance.

Hooke's law for a spring is often stated under the convention that F is the restoring force exerted by the spring if another force is pulling or pushing its free end.

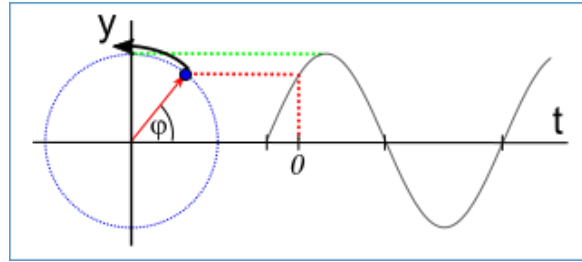


Figure 33. Harmonic motion as time function of circular motion's y component.

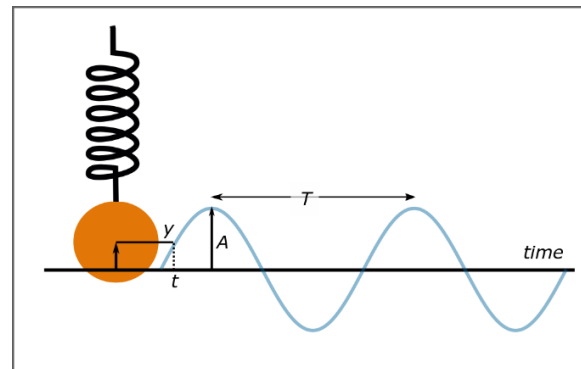


Figure 34. Harmonic motion of a body on a spring.

k is the spring constant, which is equal to the restoring force formed after unit length deformation of the spring. The restoring force and the displacement have opposite orientation (negative sign in eq. 107):

$$F = -ky \quad \text{eq. 107}$$

The combination of Hooke's law (eq. 107) and Newton's second law for a harmonic oscillator gives the relation between frequency and spring constant (eq. 106):

$$F = ma = m \frac{d^2y}{dt^2} = -ky$$

Or

$$\frac{d^2y}{dt^2} = -\frac{k}{m}y \quad \text{eq. 108}$$

The second derivation of the trajectory is proportional to the original function (y). There are several functions, which can be the solution of y (see Mathematical background below). One of them is $A\sin(\omega t + \varphi)$, where A , ω , φ are constants, the amplitude of the harmonic motion its angular speed and the phase shift, respectively:

$$\frac{d^2y}{dt^2} = -A\omega^2 \sin(\omega t + \varphi) = -\omega^2 y = -\frac{k}{m}y$$

$$\omega = \sqrt{\frac{k}{m}} \quad f = \frac{1}{2\pi} \sqrt{\frac{k}{m}} \quad \text{eq. 109}$$

A two-atomic molecule is possible to approximate with two spheres, connected with a spring (harmonic approximation, Figure 36). The spring represents the bond between the atoms, and the spheres are the individual atoms. Atoms vibrate around their equilibrium position. Atoms act mutually with equal forces in opposite directions (F_1 , F_2), their masses are m_1 and m_2 , and the accelerations a_1 and a_2 . From Newton's second (eq. 67) and third law (eq. 74):

$$F_i = m_i a_i \quad F_2 = -F_1 = F = \mu a_{rel}$$

$$m_2 a_2 = -m_1 a_1 \Rightarrow -a_1 = a_2 \frac{m_2}{m_1}$$

The relative acceleration (a_{rel}) is the difference of a_1 and a_2 :

$$a_{rel} = a_2 - a_1 = \frac{F}{\mu} \quad \text{eq. 110}$$

Where μ is the so-called reduced (effective) mass of the system.

$$\frac{F}{\mu} = a_2 + a_2 \frac{m_2}{m_1} = a_2 \left(1 + \frac{m_2}{m_1}\right) = a_2 m_2 \left(\frac{1}{m_2} + \frac{1}{m_1}\right) = F_2 \left(\frac{1}{m_2} + \frac{1}{m_1}\right)$$

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} \quad \text{eq. 111}$$

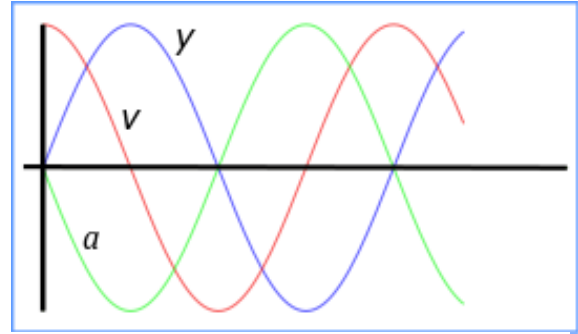


Figure 35. Relative positions of the trajectory (y , blue line), speed (v , red line) and acceleration (a , green line) of harmonic motion, omitting the possible different amplitude of y , v , and a .

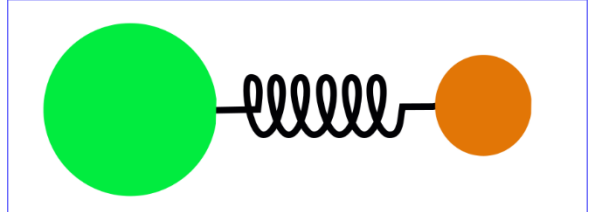


Figure 36. Harmonic representation of a two-atomic molecule.

This reduced mass is used to calculate the bond force constant k from the vibration frequency or vice versa. Masses m_1 and m_2 are the atomic masses calculated as the ratio of the molar masses and the Avogadro constant ($N_A = 6.022 \cdot 10^{23} \text{ mol}^{-1}$).

Problem. Hydrochloric acid is a two-atom molecule consisting of hydrogen ($M_w = 1.0078 \cdot 10^{-3} \text{ kg mol}^{-1}$) and chlorine ($M_w = 34.9688 \cdot 10^{-3} \text{ kg mol}^{-1}$). The harmonic vibration frequency of this molecule is $f = 8.66 \cdot 10^{13} \text{ Hz}$. The reduced mass is:

$$\mu N_A = \frac{1.0078 \cdot 10^{-3} \cdot 34.9688 \cdot 10^{-3}}{1.0078 \cdot 10^{-3} + 34.9688 \cdot 10^{-3}} = 9.7957 \cdot 10^{-4} \text{ kg mol}^{-1}$$

$$\mu = \frac{9.7957 \cdot 10^{-4} \text{ kg mol}^{-1}}{6.022 \cdot 10^{23} \text{ mol}^{-1}} = 1.6267 \cdot 10^{-27} \text{ kg}$$

The bond force constant k is (eq. 109):

$$k = (2\pi f)^2 \mu = (2\pi \cdot 8.66 \cdot 10^{13})^2 \cdot 1.6267 \cdot 10^{-27} = 481.6 \text{ N}$$

The energy of the harmonic oscillation is the sum of the potential and kinetic energies (see chapter Mechanical Energy (E), Mechanical Work (W), Power (P)). The values of these energies vary at every moment of the motion (see eq. 112 - eq. 114). The kinetic energy is:

$$E_{kin} = \frac{1}{2} m v^2 = \frac{1}{2} m (A \omega \cos(\omega t + \varphi))^2$$

$$k = m \omega^2 \Rightarrow E_{kin} = \frac{1}{2} k A^2 \cos^2(\omega t + \varphi) \quad \text{eq. 112}$$

The potential energy is defined as the integral of the displacement of the object (y) against the external force (F , eq. 107) acting on the object in the given position:

$$E_{pot} = - \int F dy = - \int (-ky) dy \quad \text{eq. 113}$$

$$= \frac{1}{2} k y^2 + \text{const}$$

The integration constant comes from the boundary condition, if the displacement of the object $y = 0$, then $E_{pot} = 0$, so the constant is zero. Substituting for y from eq. 103:

$$E_{pot} = \frac{1}{2} k A^2 \sin^2(\omega t + \varphi) \quad \text{eq. 114}$$

The total mechanical energy of the harmonic oscillator is the sum of the kinetic and potential energies. The kinetic energy (red line, Figure 33) and potential (blue line, Figure 33) change permanently with time, but their sum (green line, Figure 33) is constant. The total energy is one-half of the product of spring constant and the square of the oscillation amplitude.

$$E_{total} = E_{kin} + E_{pot} = \frac{1}{2} k A^2 [\cos^2(\omega t + \varphi) + \sin^2(\omega t + \varphi)] = \frac{1}{2} k A^2 \quad \text{eq. 115}$$

Mathematical background.

Solution of the differential equation $\frac{d^2 y}{dx^2} = b^2 y$ where b^2 is a constant.

The second derivative of y is proportional to the original function. Basically, it is valid for an exponential function:

$$y = a e^{bt}$$

$$\frac{d^2 y}{dt^2} = \frac{d^2 a e^{bt}}{dt^2} = b^2 a e^{bt} = b^2 y$$

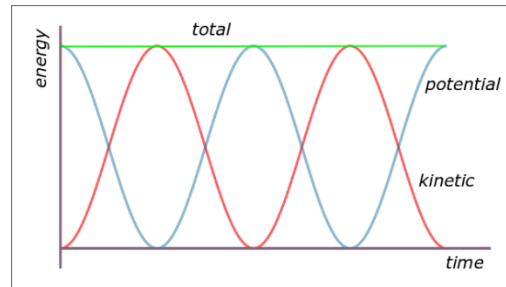


Figure 37. Time variation of kinetic, potential, and total mechanical energy of harmonic motion.

In our case, as comes from the eq. 107:

$$b^2 = -\frac{k}{m} \quad \Rightarrow \quad b = \pm i \sqrt{\frac{k}{m}} = \pm i\omega$$

b is a complex number. The solution is as follows.

$$y = A e^{i\omega t} + B e^{-i\omega t}$$

Using Euler's formulas for complex numbers:

$$e^{ix} = \cos x + i \sin x$$

$$e^{-ix} = \cos x - i \sin x$$

y is:

$$y = A (\cos \omega t + i \sin \omega t) + B (\cos \omega t - i \sin \omega t) = (A + B) \cos \omega t + i (A - B) \sin \omega t$$

Constants A and B are complex numbers. Trajectory y is a real quantity, which is matched if:

$$I. \quad A - B = 0 \quad \Rightarrow \quad A = B \quad \text{and} \quad A = \frac{C}{2}$$

$$\text{and } y = \left(\frac{C}{2} + \frac{C}{2}\right) \cos \omega t + i \left(\frac{C}{2} - \frac{C}{2}\right) \sin \omega t = C \cos \omega t \quad \text{or}$$

$$II. \quad A + B = 0 \quad \Rightarrow \quad A = -B \quad \text{and} \quad A = \frac{D}{2i}$$

$$\text{and } y = \left(\frac{D}{2i} - \frac{D}{2i}\right) \cos \omega t + i \left(\frac{D}{2i} + \frac{D}{2i}\right) \sin \omega t = D \sin \omega t$$

Constants C and D are the real amplitudes of the harmonic motion. Both solutions are valid, satisfy eq. 108, as well as their linear combination:

$$y = C \cos(\omega t + \varphi) + D \sin(\omega t + \varphi)$$

For completeness, to the solution was added the so-called phase shift (φ), a constant, which defines the initial position of the object ($t=0$). Its presence does not affect the solution of eq. 108, because its derivation is zero.

8 Component Motion - Solid Body Motion

In the previous chapters, we dealt with point-like particles. Real particles are not point-like, so their motions involve not only translational but also rotation, deformation, vibration, etc., and their combinations. For example, when we follow the motion of a car wheel, the centre of the wheel is moving along a linear trajectory (let us have a straight road). On the other hand, following a point on the tyre (Figure 38), the shape of the trajectory shows a quite complicated curve, a combination of a linear motion ($2\pi r f t = \omega r t$) and a circular motion. The x and y coordinates of the point followed are:

$$x = r \cos(-\omega t) + \omega r t$$

$$y = r + r \sin(-\omega t) \quad \text{eq. 116}$$

Where r is the radius of the wheel, t is time, ω – is the angular speed, and $v = \omega r$ is the speed of the linear motion. t – is the parameter of these equations; the other quantities are constants. The negative sign of the sine and cosine arguments results from the rotation direction of the wheel. The initial position of the followed point (X) is arbitrarily chosen with the coordinates $X[r, r]$. The route carried out by the car is the linear part of the movement,

$$s = 2\pi r f t = \omega r t = v t$$

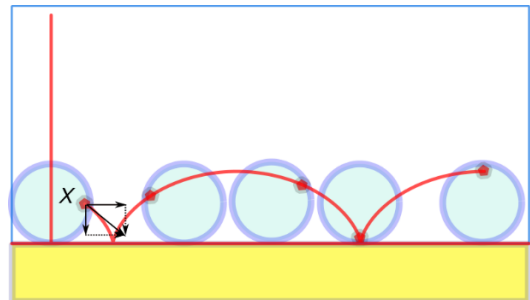


Figure 38. An example of solid body motion.

because the time average of the circular part of the motion in the x direction $r \cos(-\omega t)$ is varying between $-r$ and r , its contribution to the route, except at the beginning, is negligible. The value y changes from the interval $<0, 2r>$.

8.1 Friction

Newton's first law tells us that for a (uniform linear) motion, no force or energy is required. It is apparently(!) in contradiction with everyday experiences because it is necessary to supply energy and fuel to keep objects in uniform linear motion. This contradiction is only apparent; of course, Newton's law is valid, the energy supply during a uniform linear motion is consumed to compensate for the lost energy. One of the main sources of this loss is the friction of all moving objects. Friction is a force that resists the relative motion of solid surfaces and fluids that move around it. The result of friction is the waste energy, the transformation of the kinetic energy into heat (thermal energy) dissipated to the environment of the moving objects, deformation or abrasion of surfaces. A sliding object on a surface is in permanent contact with the surface of another object. Friction depends on the weight of the object and the quality of the contact areas of the surfaces. That dependence on weight, on gravity, causes our feeling of an energy requirement for permanent motion, because we have acquired our experiences in the environment of permanent gravitation field of Earth.

The friction force (F_{fr}) is proportional to the normal force N , which is perpendicular to the surface and equals in absolute values to the gravity force component in that direction. In eq. 117 are only the magnitudes of the friction and normal forces because the vectors do not have the same directions.

$$|\vec{F}_{fr}| = \mu |\vec{N}| \quad \text{eq. 117}$$

where μ is the friction coefficient (no units). If the ground is horizontal, the normal vector is the opposite vector of the gravity vector (Figure 39). If the object is in an inclined plane with angle θ (Figure 40), the normal vector is the opposite vector of the vector component $G \cos \theta$, it is always perpendicular to the surface. The friction is acting in the opposite direction of the motion. The vector component $G \sin \theta$ is parallel to the surface and is responsible for the sliding of the object on the ramp when the angle θ is greater than the limit angle.

Below this limit angle, the friction keeps the object at rest on the inclined plane. In the case of the limit angle, these two forces are balanced:

$$\begin{aligned} \mu N &= G \sin \theta \\ \mu G \cos \theta &= G \sin \theta \\ \mu &= \frac{\sin \theta}{\cos \theta} = \tan \theta \end{aligned} \quad \text{eq. 118}$$

The limit angle depends only on the friction coefficient and does not depend on the weight of the object. There is a difference between the friction coefficients when the object is moving already on a surface (kinetic friction coefficient) or is at rest (static friction coefficient). Static friction coefficients are higher than kinetic friction coefficients.

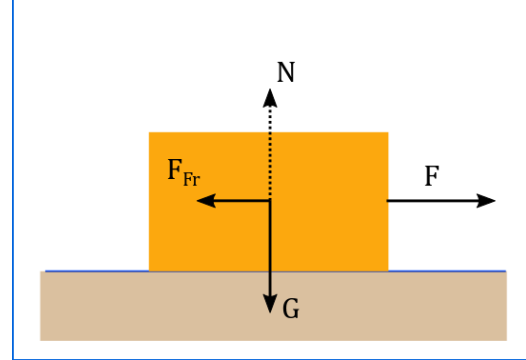


Figure 39. Friction on the horizontal surface.

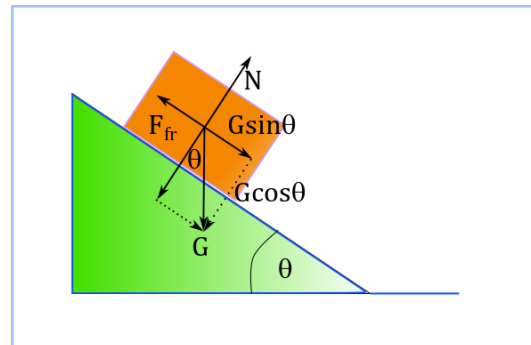


Figure 40. Friction on inclined plane.

The wheel, considered as one of greatest inventions of the mankind, was discovered in the late Neolithic or in Bronze age. Wheels also have a huge impact on human travelling and industry. The advantage of carrying heavy loads with wheels is that rolling friction is smaller than kinetic sliding friction. Rolling friction is as follows:

$$|\vec{F}_r| = \frac{\varepsilon}{r} |\vec{N}| \quad \text{eq. 119}$$

where ε is the rolling friction coefficient, r is the radius of the wheel and $\vec{N}$ is the normal vector. The rolling friction coefficient depends on the materials of the ground and the wheel. The friction depends on the radius r of the wheel, the larger the radius, the smaller the friction.

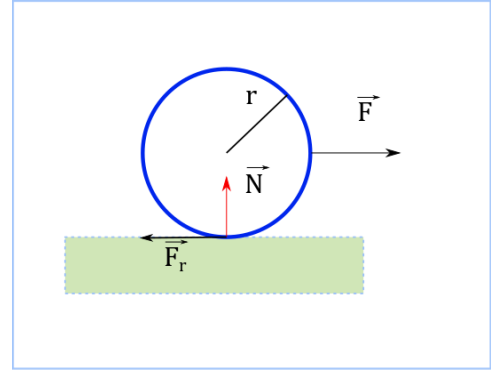


Figure 41. Rolling friction.

9 Deformation of a Solid Body

Up to this point, we considered the volume and the shape of objects invariable; an influence of a force caused only changes in motion, an acceleration. The collisions of the objects were elastic, which obeys the conservation law of momentum. However, a force acting on an object can deform the object, pulling an object might cause an elongation of the object (Figure 42), pushing contraction, etc. The strain ε is defined as the fraction between the elongation and the initial length:

$$\varepsilon = \frac{\Delta L}{L_0} \quad (\text{no units}) \quad \text{eq. 120}$$

Tensile stress (σ , units: pascals - Pa) is defined as the ratio of the deformation force (F) and the area (S), where the force acts:

$$\sigma = \frac{F}{S} \quad [\sigma] = 1 \frac{N}{m^2} = 1Pa \quad \text{eq. 121}$$

The strain is proportional to the tensile stress, and vice versa:

$$\sigma = \varepsilon E \quad \text{eq. 122}$$

E is the modulus of elasticity (Young's modulus, unit Pa), which is a material property. The length after the deformation is calculated as follows:

$$L = L_0(1 + \varepsilon) = L_0 \left(1 + \frac{\sigma}{E}\right) \quad \text{eq. 123}$$

or the deformation:

$$\Delta L = L - L_0 = L_0 \frac{\sigma}{E} \quad \text{eq. 124}$$

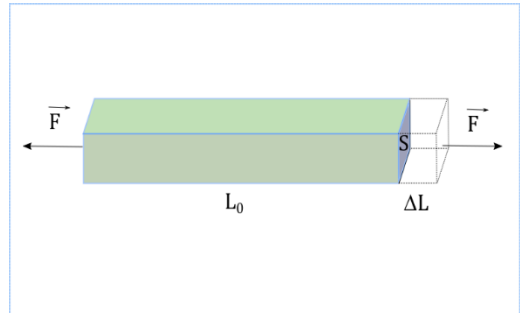


Figure 42. Deformation of a solid body.

The so-called strain-stress diagram demonstrates the behaviour of the object during mechanical tests (Figure 43). At the beginning of the curve, in the linear region, Hook's law is obeyed, the deformation is proportional to the deformation force. Beyond this region, the eq. 123 and eq. 124 do not describe the behaviour of the material. The deformation is reversible in the elastic region; that is, when the deformation force is removed, the object reverts to its original shape and size. On the other hand, deformation in the plastic region means irreversible deformation, lasting even after cancelling the deformation force. At the end of this region, we find the force that causes the material to rupture.

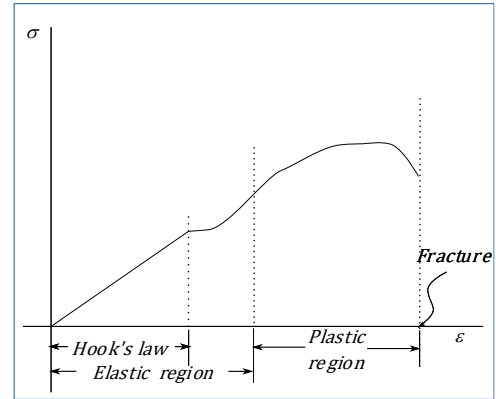


Figure 43. Graph of the object deformation on the force exerted on the object.

10 Moments of Physical Quantities

Generally, a moment of a physical quantity μ is the product of a distance r raised to a power n and the physical quantity q .

Distance means the distance of the points, where the quantity is acting and where the action is manifested. For example, in the case of torque (moment of force), the force acts through an arm to a point. The force acts beyond the object's geometrical centre; it is an *eccentric force*. According to the eq. 125, the moment increases with the length of the arm (see later) raised to the given power. The type of moment is

$$\mu = r^n q$$

characterised by the physical quantity. As examples we can list moment of force (torque), angular momentum, electric dipole moment, moment of inertia, etc.

When the quantity is not concentrated at one point but is dispersed along a line, the equation is as follows.

$$\mu = \int r^n \rho(r) dr \quad \text{eq. 126}$$

where ρ is the density of the quantity.

10.1 Moment of Force – Torque (M)

Torque, the rotational equivalent of linear force, is an example for the first-order moment, $n=1$. When we need to tighten or loose a nut (Figure 44), we instinctively know that holding the wrench at the end of its handle is more effective than holding it in the middle, applying the same force. The reason is that the torque $\vec{M}$, in the first case, is twice that in the second case.

$$\vec{M} = \vec{r} \times \vec{F} \quad \text{eq. 127}$$

M is the torque; r is the length of the arm (axis of rotation) of the force F . The unit of torque is:

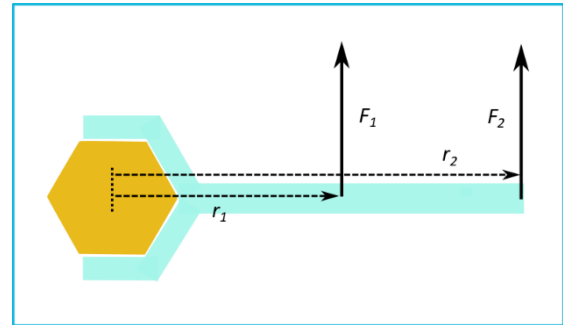


Figure 44. Nut and wrench working, based on the moment of force.

$$[M] = 1 \text{ Nm (not Joule!)}$$

eq. 128

The torque is a vector because it is defined as a vector product of two vectors: $\vec{r}$ and $\vec{F}$. The order of the vectors is important in the definition formula because the cross product of the vectors is anti-commutative (see the chapter Vector Multiplication). The magnitude of the resultant vector is as follows:

$$|\vec{M}| = |\vec{r}| |\vec{F}| \sin\theta \quad \text{eq. 129}$$

where θ is the angle between the vectors $\vec{r}$ and $\vec{F}$. The maximum torque is achieved when $\theta = 90^\circ$; when the force is perpendicular to the vector of the arm $\vec{r}$ and equals to zero when they are parallel.

Definition: The turning or rotational effect of an eccentric force, the torque, is equal to the product of a force and the perpendicular distance of the place of action and the place of its manifestation.

The direction of the torque vector is perpendicular to the plane of the original vectors, and the orientation is found by the right-hand rule (see Chapter Vector Multiplication).

The torque principle is exploited in several devices. For example, mechanical scales (Figure 45), children's toy seesaws, etc. These apparatuses consist of a supported arm, and the equilibrium is achieved compensating with weights on the ends of the arm. Equilibrium means that the torques (not the forces) are equal on both sides.

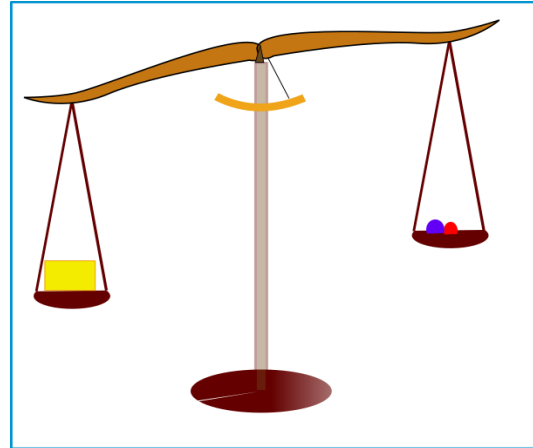


Figure 45. A scale based on the torque.

10.2 Centre of Mass

The equilibrium of mechanical balances is possible in the case of objects with compensating weights on both sides of the arms. The torque of the object on one side compensates for the torque of the object on the other side. The forces of these torques are the weights of the objects.

$$\begin{aligned} \sum M_i &= 0 \\ M_1 + M_2 &= 0 \\ m_1 r_1 g \sin\alpha_1 + m_2 r_2 g \sin\alpha_2 &= 0 \end{aligned} \quad \text{eq. 130}$$

Where m_1 and m_2 are the masses of objects, g is the gravitational acceleration, α_i is the angle between the arm of the torque and the force. The example in Figure 46 shows an equilibrium, where the two times larger mass (20 kg vs. 10 kg) is balanced with the two times larger length of the arm (1m vs. 2m, formula simplified with $g=9,81 \text{ m/s}^2$).

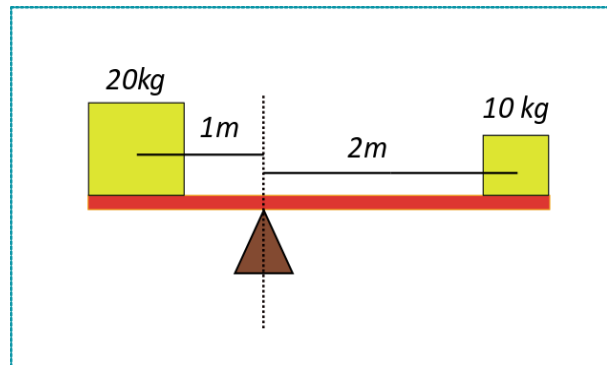


Figure 46. Equilibrium of torques.

$$20\text{kg} \cdot 1\text{m} \cdot \sin(-90^\circ) + 10\text{kg} \cdot 2\text{m} \cdot \sin(90^\circ) = 0$$

In Figure 46 the arm length (3m) was divided into a 1:2 ratio, which is the reversed ratio of the masses. The place on the bar that divides its length in that ratio is the centre of the mass of the system. Carrying a heavy load is easier if the load is held at the point where the load is in equilibrium, the sum of the torques is zero. This point is called the centre of the mass.

When the load is supported at the centre of the mass, it behaves in the same manner as a point-like object with the same mass located in the centre of the mass. The position of the centre of the mass (x_{cm}) in the case of a homogeneous object (constant density) is the following.

$$x_{cm} \int_m dm = \int_m x dm$$

$$x_{cm} = \frac{\int_m x dm}{\int_m dm} \quad \text{eq. 131}$$

$$x_{cm} = \frac{1}{m_{total}} \int_m x dm = \frac{1}{V\rho} \int_V x \rho dV = \frac{1}{V} \int_V x dV$$

where x is the position of a slice with mass dm and m_{total} is the total mass of this object. In the case of a system of point-like objects, where their dimensions are negligible compared with their distances, the integral is replaced by a sum:

$$x_{cm} = \frac{1}{m_{total}} \sum_i r_i m_i \quad \text{eq. 132}$$

Problem: The system of three objects A,B,C rotates around their centre of mass (CM) (Figure 47). Their masses are in the ratio $m_A:m_B:m_C=10:15:40$. Their positions from a reference point (RP[0,0]) are defined as Cartesian coordinates: A [2,3], B [3,20], and C[20,10]. Find the Cartesian coordinates of their centre of mass (the centre of their rotation), find the polar coordinates (r and φ) for each of them and the equations of their rotations around the centre of mass!

Solution: Using eq. 132

$$x_{CM} = \frac{m_A x_A + m_B x_B + m_C x_C}{m_A + m_B + m_C} = \frac{10 \cdot 2 + 15 \cdot 3 + 40 \cdot 20}{10 + 15 + 40} = 13.308$$

$$y_{CM} = \frac{m_A y_A + m_B y_B + m_C y_C}{m_A + m_B + m_C} = \frac{10 \cdot 3 + 15 \cdot 20 + 40 \cdot 10}{10 + 15 + 40} = 11.231$$

The coordinates of the centre of mass are CM[13.308,11.231] in arbitrary units. The distances of objects A,B,C from the CM are

$$r_A = \sqrt{(x_A - x_{CM})^2 + (y_A - y_{CM})^2} = \sqrt{(13.308 - 2)^2 + (11.231 - 3)^2} = 13.986$$

We calculate the distances of B and C in the same way; the result is as follows:

$$r_B = 13.533$$

$$r_C = 6.804$$

The angle between the line connecting the centre of mass with the object A and the x axis (horizontal line) is:

$$\sin \varphi_A = \frac{y_A - y_{CM}}{r_A} = \frac{-8.231}{13.986} = -0.5885$$

$$\cos \varphi_A = \frac{x_A - x_{CM}}{r_A} = \frac{-11.308}{13.986} = -0.8085$$

$$\varphi_A = \arcsin(-0.5885) = 216.05^\circ \quad \text{or} \quad 323.95^\circ$$

$$\varphi_A = \arccos(-0.8085) = 216.05^\circ \quad \text{or} \quad 143.95^\circ$$

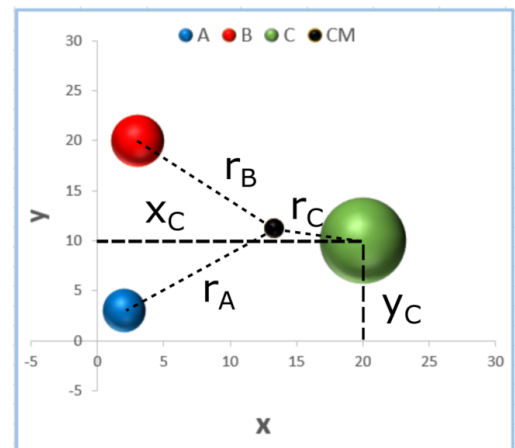


Figure 47. Rotation of objects (A, B, C) around their centre of mass (CM).

When the sine and cosine of an angle are negative, the angle is between 180° and 270° (π and $3/2\pi$, Figure 48), the solution is $216.05^\circ = 3.771 \text{ rad}$.

Angles ϕ_B and ϕ_C are calculated similarly.

$$\phi_B = 139.61^\circ = 2.437 \text{ rad}$$

$$\phi_C = 349.58^\circ = 6.101 \text{ rad}$$

The Cartesian coordinates of the objects are defined by the following equations:

$$x_i = r_i \cos(\omega t + \phi_i) + x_{cm}$$

$$y_i = r_i \sin(\omega t + \phi_i) + y_{cm}$$

$$A: \quad x_A = 13.986 \cos(\omega t + 3.771) + 13.308$$

$$y_A = 13.986 \sin(\omega t + 3.771) + 11.231$$

$$B: \quad x_B = 13.533 \cos(\omega t + 2.437) + 13.308$$

$$y_B = 13.533 \sin(\omega t + 2.437) + 11.231$$

$$C: \quad x_C = 6.804 \cos(\omega t + 6.101) + 13.308$$

$$y_C = 6.804 \sin(\omega t + 6.101) + 11.231$$

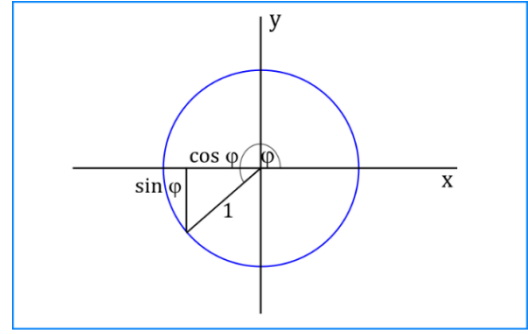


Figure 48. Sine and cosine of an angle in the third quadrant (180° - 270°).

10.3 Moment of Inertia, Energy of the Rotating Object

The kinetic energy of the system is the sum of the kinetic energies of its parts.

$$E_{kin} = \sum_i E_{kin i} = \sum_i \frac{1}{2} m_i v_i^2 = \sum_i \frac{1}{2} m_i (r_i \omega)^2 = \frac{1}{2} \sum_i J_i \omega^2 = \frac{1}{2} J \omega^2 \quad \text{eq. 133}$$

Here J_i is the moment of inertia (angular mass, rotational inertia) of the i-th rotating object. The definition of moment of inertia:

$$J = \sum_i J_i = \sum_i m_i r_i^2 \quad \text{eq. 134}$$

illustrates a second-order moment because the power of the arm vector is 2 (r^2). The moment of inertia is the rotational analogue of the mass; it is the measure of the inertia of the rotational motion. The units of the moment of inertia are as follows.

$$[J] = 1 \text{ kg m}^2$$

The overall moment of inertia of a system is the sum of the moments of inertia of individual components. In this case, the system consists of discrete objects. When we calculate the moment of inertia of a continuous system, we use an integral instead of the sum:

$$J = \int_m r^2 dm = \int_V r^2 \rho dV \quad \text{eq. 135}$$

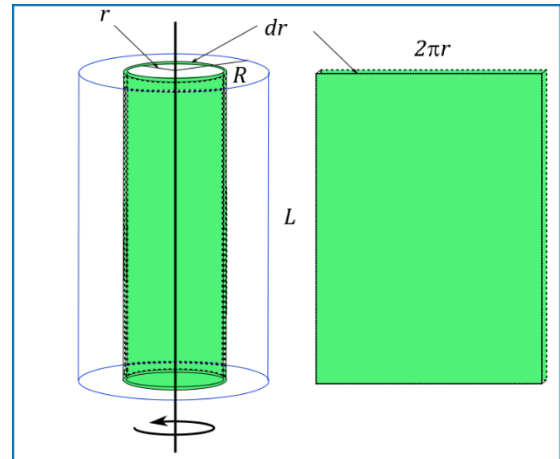


Figure 49. Division of a cylinder to slices.

The solution depends on the actual case. For example, the moment of inertia of a solid cylinder (Figure 49) rotating around its axial symmetry axis is calculated in such a way that the volume is divided into thin slices of the cylinder, which are parallelepipeds when opened, with sides L (length of the cylinder) and $2\pi r$ (the circumference of the slice) and dr (the thickness of the slice). The moment of inertia of a homogeneous solid cylinder with radius R , length L , and mass m is the following.

$$J = \int_0^V r^2 \rho dV = \rho \int_0^R r^2 2\pi r L dr = 2\pi \rho L \left[\frac{r^4}{4} \right]_0^R = \frac{\pi \rho L R^4}{2} = \frac{1}{2} (\pi R^2 L \rho) R^2 = \frac{1}{2} V \rho R^2 = \frac{1}{2} m R^2 \quad \text{eq. 136}$$

Angular momentum. The analogous value of the linear momentum of motion ($p=mv$) is the angular momentum L , the product of the moment of inertia and the angular speed.

$$L = J\omega \quad \text{eq. 137}$$

For a point-like object moving on a circular trajectory with radius r (Figure 50):

$$L = mr^2\omega = rmv = rp$$

$$\text{The unit is: } [L] = 1 \text{ kg m}^2 \text{ s}^{-1}$$

In vector form:

$$\vec{L} = \vec{r} \times \vec{p} \quad \text{eq. 138}$$

where $\vec{p}$ is the linear momentum of motion and $\vec{r}$ is the radius of the rotation. Similarly, as in the case of the linear momentum, the angular momentum obeys the conservation law. A well-known example is the case of a spinning figure skater around his/her axis with straightened arms and later with withdrawn arms close to his/her body. Conservation of the angular momentum causes an increase in the spin speed, compensating for the reduced radius of the spin. For example, if the radius $\vec{r}$ is halved the angular velocity ω is raised up four times:

$$L = mr^2\omega_1 = m\left(\frac{r}{2}\right)^2 \omega_2 \\ \omega_2 = 4\omega_1$$

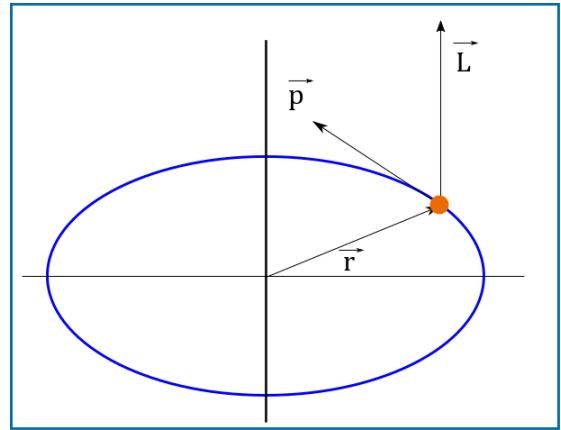


Figure 50. Vector of angular momentum (L) with corresponding momentum of motion and radius.

10.4 Angular Analogues of Newton Laws of Motions

Newton's laws of mechanics have analogues for rotational motions, where the quantities of linear motion are replaced with their rotational analogues.

1. An object at rest tends to remain at rest, and an object that is spinning tends to spin with a constant angular velocity, unless it is acted on by a nonzero net torque **or the object's mass distribution is changed.**

$$\sum_i M_i = 0 \quad \text{eq. 139}$$

Linear analogue:

$$\sum_i F_i = 0$$

2. The rate of change of angular momentum of a body is equal to the torque that causes it, and the change takes place in the direction in which the torque acts.

$$\vec{M} = \frac{d\vec{L}}{dt} \quad \text{eq. 140}$$

Linear analogue:

$$\vec{F} = \frac{d\vec{p}}{dt}$$

3. If object A exerts a torque on object B about a given axis, then B exerts an equal but opposite torque on A about the same axis.

$$\vec{M}_1 = -\vec{M}_2 \quad \text{eq. 141}$$

Linear analogue:

$$\vec{F}_1 = -\vec{F}_2$$

11 States of matter

Ancient Greek philosophers distinguished four elements, which built up the world; Earth, Water, Air (Wind) and Fire, and sometimes added a fifth element as Aether. They connected them with a pair of properties, the Earth with cold and dry, Water with cold and wet, Air hot and wet and Fire hot and dry. We know that materials are built up from atoms of elements, 90 of them are abundant in nature, natural, another 4-8 (according to criteria) occur on Earth as products or intermediate products of radioactive decay, and some of them are synthetic without natural occurrence. In this sense, the ancient Greek philosophers were not right; however, we can find similarities with the four states of matter and the ancient elements. The materials can exist in four states (Figure 51):

- Solid
- Liquid
- Gaseous
- Plasma

We can relate solid state with the element Earth, liquid to Water, gaseous to Air and plasma to Fire. However, we do not consider them as elements but rather as the state of the materials in which they exist according to the surrounding conditions. Practically, all kinds of material persist in all these states when the conditions are appropriate, usually defined by temperature and pressure. We can demonstrate this on a very common material of water (as a chemical entity, H_2O). Under normal pressure and low temperatures (below 0°C), the water is in the solid state as ice (snow). With increasing temperature, between 0°C and 100°C the water is in the liquid state, water. Above 100°C it evaporates and forms gaseous water, steam. That is the moment when the general experience ends. Elevating the temperature to the next step, where the molecules dissociate into hydrogen and oxygen, requires special treatments and apparatus, and occurs at very high temperatures. The resulting hydrogen and oxygen are gases. The plasma state for hydrogen begins around $10\,000\text{K}$, when electrons are released from atoms, which is typical of the stellar atmosphere. The formation of plasma is induced not only by temperature but also by electrical energy (arc, lightening). During these processes, a partially or fully ionised plasma is formed.

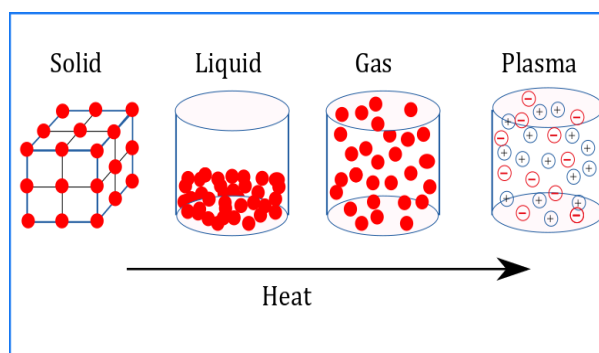


Figure 51. States of matter.

Solid-state materials possess well-defined shapes and volumes. Atoms or molecules are tightly packed and (usually) form a lattice. Atoms or molecules vibrate around the equilibrium position, but do not leave their positions. Translations or rotations are banned. With increasing temperature, the lattice will breakdown, and the solid material will melt.

Definition: The melting point is the temperature, at which the solid and liquid phases of a pure substance can exist in equilibrium. The melting temperature is a characteristic property of pure substances that can be used to identify the material. Amorphous materials melt at an interval of temperature. The melting point depends on the pressure of the environment, impurities in the material, and other defects in the crystal. The material above the melting point is in the liquid state, and the material below is in the solid state.

Liquid materials possess defined volumes, but their shape depends on the shape of the container in which they are. The molecules in liquids slide on each other, and translational and rotational motions of the particles are present as well. The density of a liquid is usually lower than that of solid material of the same chemical composition, but the particles are still close to each other, and the compressibility of the liquids is low. The interaction between the particles is significant. Liquids have viscosity (resistance against flowing) and form a surface (with surface tension) on contact places with other materials (gases, solids, and immiscible liquids). The molecules of the liquids move randomly in the volume of the liquid, causing numerous collisions with neighbouring particles. They provide translational, vibrational, and rotational motions.

Definition: The boiling point (temperature) of a liquid is the temperature at which the vapour pressure of the liquid is equal to the surrounding pressure.

Liquids exist between the melting point and the boiling point. Further heating of the liquid, at boiling point, does not cause an increase of temperature, but the supplied energy changes to the phase transition of the liquid to gas.

Gases fill in the available space uniformly, i.e., they do not have fixed volume and shape, but their volume and shape depend on the container. The density of a gas is (usually) lower than that of the liquids; the atoms or molecules are far from each other. All types of motion (translation, vibration, and rotation) do the particles, and frequent collisions occur between the particles themselves and between the particles and the walls of the container. In a fluid (gases and liquids) that is in thermodynamic equilibrium state and without convection, the translational motions of the particles are random, without any preferred direction, the average speed and the average density of the particles remain constant; however, the individual particles' speed can differ from it and density fluctuations are also present. This type of motion is called a random walk or *Brownian motion* (after Robert Brown, botanist). When the average energy of the particles is higher than the ionisation energy, electrons are released from the molecules (atoms), leaving positively charged fragments of molecules. Decomposition can continue through a mixture of bare nuclei and electrons to a mixture of nucleons, electrons, and other subatomic particles. This state of matter is called plasma.

11.1 Gibbs phase rules

The state behaviour of pure materials can be demonstrated on a phase diagram (Figure 52), where the state of the matter is displayed as a function of pressure and temperature. A phase diagram contains areas of different phases (solid, liquid, gaseous) divided with lines of the meaning of phase transitions (melting, sublimation, evaporation).

In Figure 52 horizontal lines show isobaric changes (constant pressure), which are the most common processes; the vertical lines show isothermal (constant temperature) processes. The line of the liquid-gas equilibrium (I-C) ends at the *critical point*, beyond that the liquid and gaseous phases have the same properties, it becomes a supercritical fluid.

The parameters of the system (in our case the temperature and pressure) determine the state of the system, whether it is in the solid, liquid, or gaseous phase. The system responds to the variation of the parameters differently. Let us have, for example, ice, at normal pressure and low temperature (below 0°C) represented with point A on the phase diagram. When the temperature raises a certain measure (A-B, B not included), it remains an ice, as well as in the case of a decrease in pressure (A-F, F not included). The state of the system still remains unchanged. We can change these *two parameters*; the system has degrees-of-freedom *two*. When the system is in the state of equilibrium between two phases (B, C, F, I), changing freely one parameter requires setting the other parameter to a specific value to keep the system in the given state; on the phase transition line. For example, the system is in state I, boiling (water-steam equilibrium) with temperature T_I . When the temperature is increased to T_D , corresponding to point D, the system will pass to the region of steam (water evaporates completely), and we have a different state of the system. To keep our system in the same state, we must increase the pressure from p_I to p_C (pressures of points I and C, respectively) to maintain the water – steam equilibrium. One of the parameters may vary freely, but the other does not. The degree of freedom of the system has been reduced to *one*. At the triple point (intercept of all three lines in Figure 52), where all three phases are in equilibrium, we may not change any of the parameters; otherwise, we come out of the equilibrium of the three phases. The triple point has zero degree of freedom (for a pure substance). The number of degrees of freedom is calculated according to the Gibbs phase rule.

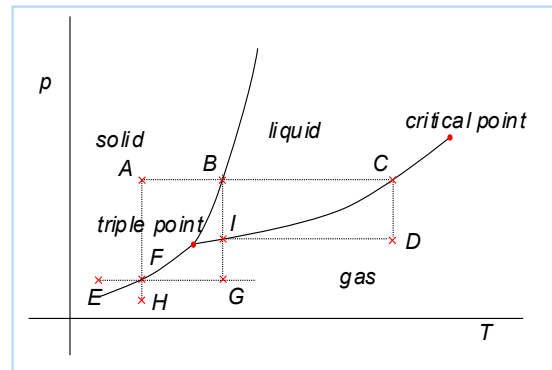


Figure 52. A general phase diagram with chosen states - points.

$$f = c - p + 2 \quad \text{eq. 142}$$

where f is the number of degrees of freedom, c is the number of components, p is the number of phases. For example, in the case of point B (Figure 52), we have one component (H_2O), two phases (solid, liquid), and the degree of freedom is:

$$f = 1 - 2 + 2 = 1$$

11.1.1 Air Humidity

The phase diagram in the previous chapter is a schematic diagram of a one-component system. The diagrams for two or more component systems are more complicated, and in the case of the two-component system the complete phase diagram is not a 2D (two-dimensional) diagram but a 3D diagram. The composition of the mixture is added to the temperature and pressure as a new parameter. Air is a mixture of gases, nitrogen 78.08 %, oxygen 20.95%, argon 0.93%, carbon dioxide 0.04 %, 0.4 % of water vapour, and several minor components. The complete air phase diagram, which considers the phase transitions of all components, is rather complicated, but most of the components are at normal pressure and temperature in the gaseous state, they are mixable without any limit and behave as one component. The only exception is the water vapour; water persists in all three phases under normal circumstances. Because of this, the air-water system can be simplified as a two-component system.

Parameters that can vary are pressure, temperature, and water content. The graph of the maximum water content (Φ_{max}) on temperature (Figure 53) and pressure (Figure 54) is a 2D projection of the air – water phase diagram, which expresses the border between the homogeneous phase (below the line) and the two-phase system (air - water droplets). The value of water saturation, that is, the amount of water vapour that can contain the unit volume of air without phase separation, increases exponentially with increasing temperature and pressure almost linearly. The temperature, when air reaches its saturation value, is called the dew point; if it is below the water freezing point, it is called the frost point.

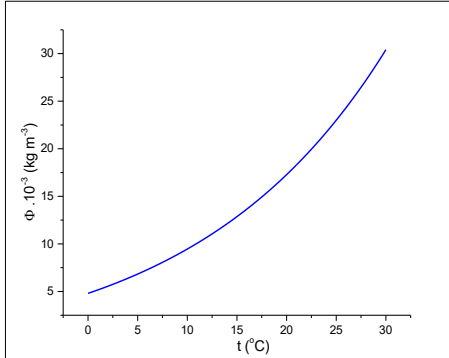


Figure 53. Absolute humidity graph on temperature.⁵

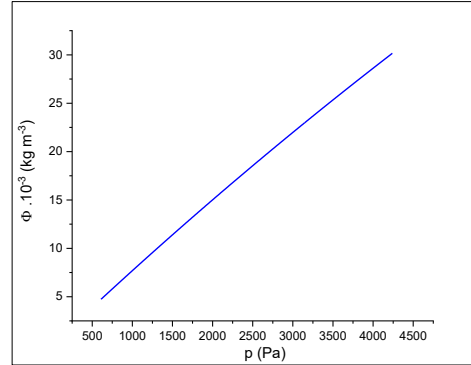


Figure 54. Absolute humidity graph on pressure.⁵

The water content is called absolute humidity of the air, the mass of the water vapours in the unit volume of the air.

$$\Phi = \frac{m}{V} \quad [kgm^{-3}] \quad eq. 143$$

The ratio of absolute and maximum humidity is called relative humidity; it is a dimensionless quantity, sometimes expressed as percentage.

$$\varphi (\%) = \frac{\Phi}{\Phi_{max}} (x100\%) \quad eq. 144$$

Air humidity can be measured with a psychrometer, which consists of two thermometers around which constant airflow is maintained with a fan. One of them measures the temperature of the environment, while the second is covered with a piece of wet cloth, and because of the evaporation of the water, its temperature is lower. Evaporation is proportional to the difference between the maximum air humidity at the given room temperature and the actual humidity in the room. Relative humidity can be calculated using the following empirical equations.

$$\varphi = \frac{p_2 - Ab(t_1 - t_2)}{p_1} \quad eq. 145$$

Where φ is the relative humidity, A , psychrometric constant (its value: 0.00066), b , atmospheric pressure, t_1 , t_2 , - temperature on the dry and wet thermometer, respectively, p_1 , p_2 - saturation pressure of water vapour in the air on the dry and wet thermometer, respectively.

⁵ Graphs drawn using data from the National Institute of Standards and Technology, NISTIR 5078, Table 1. Saturation (Temperature), <https://www.nist.gov/srd/nistir-5078>

11.2 Fluids

Liquids and gases are called fluids. Their molecules can easily slide on each other, and fluids do not have a defined shape, but take the form of the container. These properties are common. The difference between them is that a liquid has its own volume, which does not depend on the volume of the container and forms a surface.

11.2.1 Surface Tension

Molecules on the surface of a liquid are in an unbalanced state. The interactions between the liquid particles and the air are weaker than those between the particles of the liquid. The liquid molecules on the surface are in an energetically higher state, in comparison to the molecules inside the liquid, and they are attracted toward the bulk liquid.

The liquid attempts to keep its surface as small as possible to minimise its surface and its surface energy. In a place without a gravitational field, the liquids will form spherical droplets because a sphere has the least surface at given volume of all geometric forms. In a gravitational field, droplets are deformed by gravity and by interactions with other surfaces. A layer is formed on the surface of a liquid that is characterised by surface tension. This layer is observable in nature when it is used by certain small animals, mainly insects, that can stand, walk, or run on the surface of the water without sinking into it.

Surface tension (σ or γ) is defined as the energy necessary to increase the surface with unit area:

$$dE = \sigma dS \quad \text{eq. 146}$$

where dE – energy change, dS – surface change, σ – surface tension.

Let us have an experiment, where a bubble is stretched across a frame with three fixed sides and one mobile side (Figure 56). During the experiment, the mobile side of the frame is moved by length dx with a force F , and the bubble is stretched by the area dS . The contact area of the liquid and gaseous phases is on both sides of the frame, that is the reason of number 2 in eq. 147.

$$\begin{aligned} dS &= 2Ldx \\ dE &= \sigma 2Ldx \end{aligned} \quad \text{eq. 147}$$

The potential energy increases, as the frame is stretched against the force of the bubble, which attempts to decrease its surface.

$$dE = Fdx = 2\sigma Ldx \implies F = \frac{dE}{dx} = 2\sigma L \quad \text{eq. 148}$$

The units of surface tension are as follows.

$$[\sigma] = 1\text{Jm}^{-2} \quad \text{or} \quad [\sigma] = 1\text{Nm}^{-1}$$

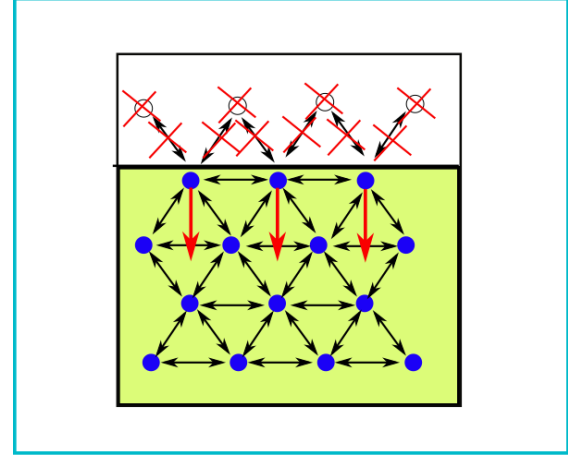


Figure 55. Interactions of liquid particles

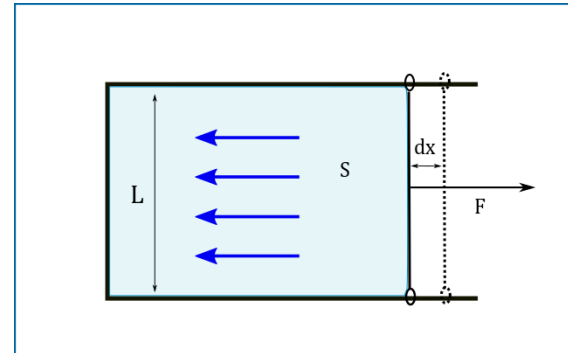


Figure 56. Surface tension experiment.

Polar materials, such as water, sulfuric acid, and formic acid have high surface tension because of the possibility of forming numerous hydrogen bonds in their structures, whereas n-hexane, with the lower surface tension in Table 4, is a nonpolar solvent, which is not able to form hydrogen bonds⁶.

Table 4. Surface tension values of common liquids⁷ at 20°C temperature and indication of the contact material.

liquid	$\sigma \cdot 10^3 \text{ N m}^{-1}$	in contact with
water	72.75	air
acetone	23.7	air or vapor
n-butanol	24.6	air or vapor
cyclohexane	25.5	air
ethanol	22.75	vapor
n-propanol	23.78	vapor
methanol	22.61	air
sulfuric acid (98.5%)	55.1	air or vapor
formic acid	37.6	air
n-hexane	18.43	air

11.2.2 Capillary Elevation and Depression

The particles of the liquid interact with each other and with the solid surface of the container. This interaction depends on the balance of the *cohesive* and *adhesive* forces. The cohesive forces are mutual attractive interactions among the liquid molecules, and the adhesive forces are attractive or repulsive interactions between the liquid molecules and the solid surface.

If the cohesion is stronger, it is energetically favourable to minimise the contact surface between the liquid and the solid surface; the liquid does not wet the solid surface. However, if the adhesion is stronger than the cohesion, the liquid wets the surface. In an open capillary, immersed in a liquid, the surface of the liquid is at a different level from the surface of the liquid outside the capillary (Figure 57). In the case of the wetting liquid, the surface in the capillary is higher than that outside, which is called *capillary elevation*. The surface of the liquid is concave (Figure 58, left). In the case of non-wetting liquids, the level of the liquid is lower than outside the capillary; it is called *capillary*

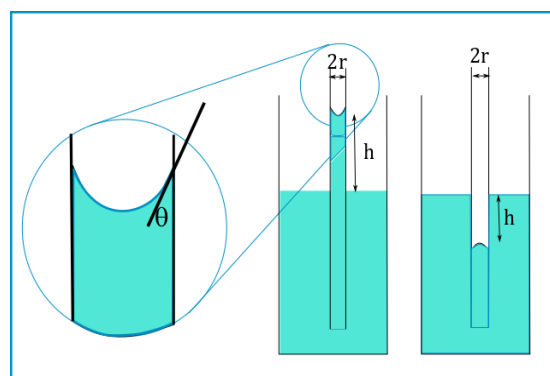


Figure 57. Capillary elevation and depression.

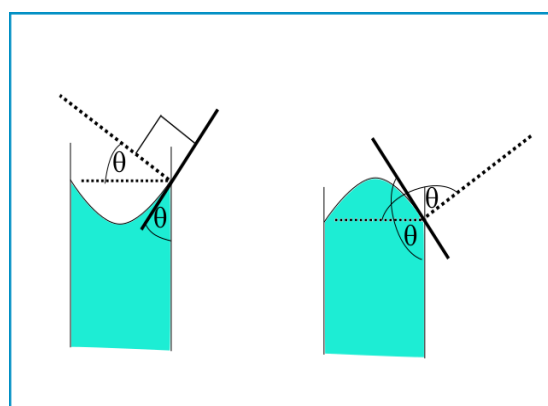


Figure 58. Contact angles between the liquid and container in the case of capillary elevation and depression.

⁶ Yeagle, P. L.: *The Membranes of Cells*, p. 19. 2016, Academic Press, ISBN 9780128004869, doi:10.1016/b978-0-12-800047-2.00001-2

⁷ CRC Handbook of Chemistry and Physics. A Ready-Reference Book of Chemical and Physical Data. 60th Edition 1979-1980, ed. Robert Weast, ISBN: 0849304601

depression. The liquid surface is convex (Figure 58, right). Capillary elevation is an important phenomenon in nature, primarily for the transport of water in soil, plants, porous materials, etc. The angle between the liquid and the solid surface is the contact angle θ (Figure 58). The weight of the liquid column above (below) the outer liquid surface is balanced with the force of the surface tension. In the case of a capillary with circular cross section of radius r , with liquid density ρ and contact angle (θ) between the meniscus of the liquid surface and the wall of the capillary, the capillary force equals:

$$F = 2\pi r \sigma \cos \theta \quad \text{eq. 149}$$

The weight of the liquid column above the level outside the capillary is balanced with the capillary force:

$$\begin{aligned} G = mg = V\rho g = \pi r^2 h \rho g \\ 2\pi r \sigma \cos \theta = \pi r^2 h \rho g \end{aligned} \quad \text{eq. 150}$$

The height of the liquid column above (below) the outer liquid surface is h :

$$h = \frac{2\sigma \cos \theta}{r \rho g} \quad \text{eq. 151}$$

The height of the capillary elevation (depression) depends on the surface tension and is inversely proportional to the radius of the capillary. The thinner the capillary, the larger the h . The contact angle is measured between the tangent of the meniscus at the contact point of the meniscus with the wall and the vertical wall line, or the horizontal line at the contact point of the capillary wall and the meniscus, and the perpendicular line to the tangent of the meniscus at the contact point (Figure 58). If the contact angle θ is from the first quadrant (Figure 59, between 0° and 90°), the $\cos \theta$ and h are positive, the liquid in the capillary is higher than outside. In the case of non-wetting liquids, the angle is from the second quadrant (between 90° and 180° , Figure 59) where the cosine is negative, so h is also negative.

Mathematical background: In planar geometry, the Cartesian axes divide the plane into four regions, four quadrants (I-IV, Figure 59,). The angle is always measured from the positively orientated part of the x-axis, as shown in the case of β . The signs of the sine and cosine functions are found in positive or negative directions on the x-axis or the y-axis.

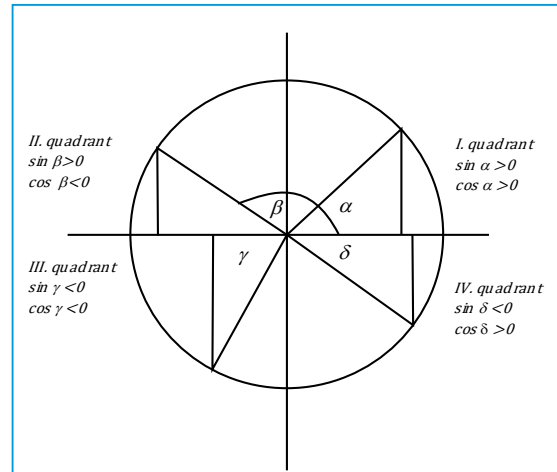


Figure 59. Angles and their sine and cosine in four quadrants.

11.2.3 Stalagmometry

Stalagmometry is one of the methods to measure surface tension. When a drop forms at the end of a capillary (of radius r , Figure 60), the weight ($m_{drop}g$) of the drop depends on the surface tension (σ). The surface tension of the liquid holds the drop at the end of the capillary. As the drop grows, the weight of the drop grows as well, and at the moments of breaking down, these forces are in equilibrium. The liquid and the capillary contact area is the circumference of the capillary, and the force is:

$$\begin{aligned} F = 2\pi r \sigma = m_{drop}g = G \\ m_{drop} = \frac{2\pi \sigma r}{g} f \end{aligned} \quad \text{eq. 152}$$

Part of the drop remains at the bottom of the capillary, so the mass of a real drop is less than the calculated one. Usually, it falls around 2/3 of the calculated mass. In eq. 152 is the parameter $f(\approx 2/3)$, which takes into account the remaining part of the drop. The value of this parameter is only approximate. The surface tension measurement method, which is based on the formation of drops on the end of a capillary, is called stalagmometry. The method uses a reference liquid (usually distilled water) and compares the weights of the same number of drops or the number of drops of the same volume of the sample and a reference liquid to avoid the uncertainty of the factor f . In the method of same number of the drops (N_{const}):

$$\frac{N_{const} m_{drop}}{\sigma} = \frac{2\pi r N_{const}}{g} f \quad \text{eq. 153}$$

The right-hand side of the equation is constant under the given experimental conditions, independent of the liquid. The mass of the N drops of the sample is m_s , that of the reference liquid (distilled water) is m_{ref} .

$$\frac{m_s}{\sigma_s} = \frac{m_{ref}}{\sigma_{ref}} \quad \text{eq. 154}$$

Another method is to count the drops (N) of the same volume (V_{const}) of the sample and the reference liquid.

$$V_{const} = \frac{N m_{drop}}{\rho} = \frac{N}{\rho} \frac{2\pi r \sigma}{g} f = \frac{\sigma N}{\rho} k \quad \text{eq. 155}$$

where k is a constant ($2\pi r f/g$). eq. 156

$$\frac{\sigma_s N_s}{\rho_s} = \frac{\sigma_{ref} N_{ref}}{\rho_{ref}}$$

Bubble pressure method. A capillary immersed in a liquid will be partially filled with the liquid due to capillary elevation (Figure 61). The pressure required to remove the liquid from the capillary equals the capillary pressure (first term on the right-hand side of eq. 157) and the hydrostatic pressure of the liquid outside the capillary (second term on the right-hand side of eq. 157). The applied pressure is provided with a water column (left-hand side of eq. 157) and is measured with a U-tube manometer filled with water. In equilibrium, the pressure on the U-tube manometer is equal to the pressure required to remove the sample from the capillary.

$$h_v \rho_v g = \frac{2\pi r \sigma \cos \theta}{\pi r^2} + h \rho g \quad \text{eq. 157}$$

$$\sigma = \frac{r g}{2 \cos \theta} (h_v \rho_v - h \rho) \quad \text{eq. 158}$$

The reference liquid, distilled water, is used to find the capillary radius and contact angle.

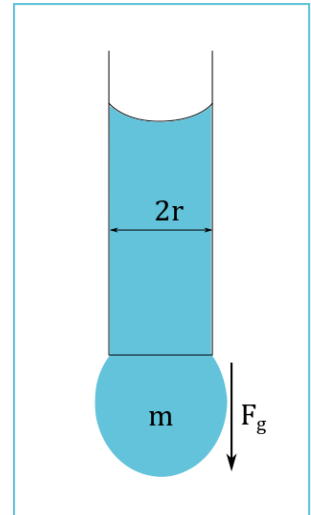


Figure 60. Forming of a drop on the bottom of a capillary.

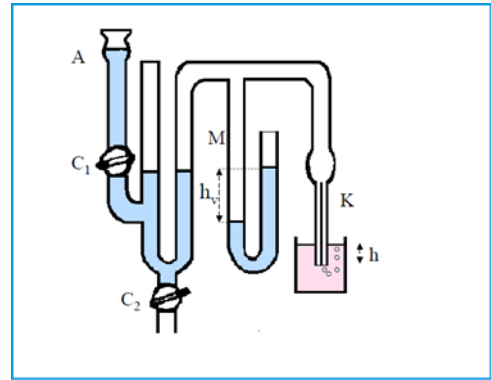


Figure 61. Surface tension measurement by bubble-pressure method.

11.2.4 Surface tension between two immiscible liquids

Immiscible liquids are liquids that do not form a homogeneous phase after mixing. The mutual solubility of these liquids is very low because of their different chemical nature. Solubility, in general, is followed by an increased disorder of the system (increased entropy; see later in the chapter Thermodynamics), because in solutions the particles of the component are uniformly dispersed and mixed. The increase in entropy is a spontaneous process and acts as a driving force in the formation of the solution. Immiscible liquids are formed in the case when the entropy gain is superseded by another phenomenon, the energetically unbeneficial interaction between the components of the solution. As an example, we can present the case of mixing water and oil. Water has polar molecules with numerous hydrogen bonds. However, oil consists of molecules with nonpolar hydrocarbon chains. It is energetically more beneficial for water molecules to interact with other water molecules and for oil molecules with other oil molecules than it is to form water-oil pairs or clusters. The entropy factor is superseded with the interaction energy factor (enthalpy, see later in Thermodynamics).

Adding oil to water, they will separate from each other and on the oil-water contact area a surface is formed. The liquid with higher density (usually water) gathers on the bottom and the liquid with lower density (usually oil) above it. When the amount of oil is small, the oil can form a film that covers the water surface or form small droplets (Figure 62). The case which occurs depends on the surface tensions of all contact areas. There will be a surface tension of air-water, air-oil, and oil-water. The force originated from the surface tension of the air-water, trying to stretch the oil drop to minimise the air-water surface. The forces of the air-oil and oil-water surface tensions try to minimise the oil-air and oil-water surfaces. The latter two forces are not in line with the air-water vector and with each other. In equilibrium, the sum of the air-oil and water-oil surface tension compensates for the air-water surface tensions. In the case where for these three vectors the so-called triangle inequality is valid (Neumann triangle, Figure 63, left), that is, the sum of arbitrary two of them is greater than the third vector, the oil-water and oil-air surface tensions are able to compensate for the air–water surface tension, and an oil droplet is formed on the water surface. On the contrary, when the air-water surface tension is greater than the sum of the air-oil and oil-water surface tensions, it is not possible to find a triangle for these three vectors, the water-air surface tension overwhelms the other two surface tensions, and oil spreads on the water surface and forms a film (Neumann triangle, Figure 63, right).

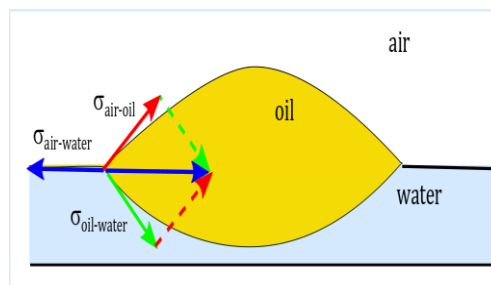


Figure 62. An oil drop on water and the surface tensions.

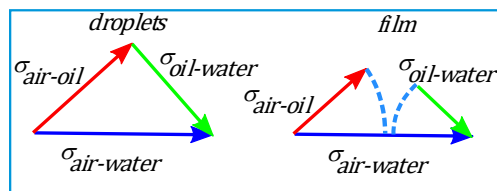


Figure 63. Triangle inequality of surface tensions.

12 Fluid Mechanics, Hydrostatic Pressure

Pressure (p) is the force applied perpendicularly to the unit surface of an object.

$$p = \frac{F}{S}$$

eq. 159

Unit of pressure is pascal (Pa):

$$1\text{Pa} = 1 \frac{\text{N}}{\text{m}^2} = 1\text{kgms}^{-2}\text{m}^{-2} = 1\text{kgm}^{-1}\text{s}^{-2}$$

Liquids are practically incompressible, and their molecules slide on each other, so an external force spreads through a liquid in all directions. In the gravitational field of Earth, all the liquid layers act on the layers below them. This force is the origin of the hydrostatic pressure. In deep water (ocean, see) it is a serious problem; the pressure increases by 9810 Pa with every metre of increasing depth. Exploring the deepest regions of the oceans is so dangerous and technically challenging that we still have very limited knowledge about these regions. Pressure in a liquid is the sum of the atmospheric pressure and hydrodynamic pressure.

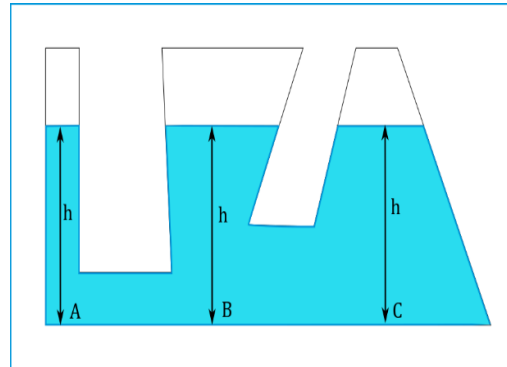


Figure 64. Hydrostatic pressure of a liquid in vessels of various shapes.

$$p = p_{atm} + p_{hydr}$$

eq. 160

The atmospheric pressure is the result of the weight of the air; its value is around 100kPa and changes with the altitude and with weather, temperature, wind, etc. However, during a usual experiment and in calculations (usually), it is constant, and we do not consider it. The net hydrostatic pressure comes from the weight of the water per surface:

$$p_{hydr} = \frac{F}{S} = \frac{mg}{S} = \frac{V\rho g}{S} = \frac{hS\rho g}{S} = h\rho g$$

eq. 161

Substituting 1000 kgm^{-3} for the density of water and 9.81 ms^{-2} for the gravitational acceleration gives the formerly mentioned 9810 Pa for every metre of depth.

The hydrostatic pressure depends on the density of the liquid and the height of the liquid column, and considering the gravity acceleration as constant, the hydrostatic pressure does not depend on the shape of the vessel. The hydrostatic pressure is the same in point A, in point B, and in point C in Figure 64.

12.1 Atmospheric pressure.

Atmospheric pressure is the pressure of the air. It depends, similarly to the pressure of the liquid, on the height of the gas column above the location of measurement. However, while the density of water is around 1000 kgm^{-3} , the density of the air is approximately 1.225 kgm^{-3} , only. Another difference between the air and water, gases and liquids, is in their compressibility. Liquids are relatively incompressible; gases are easily compressible; the density varies significantly with altitude.

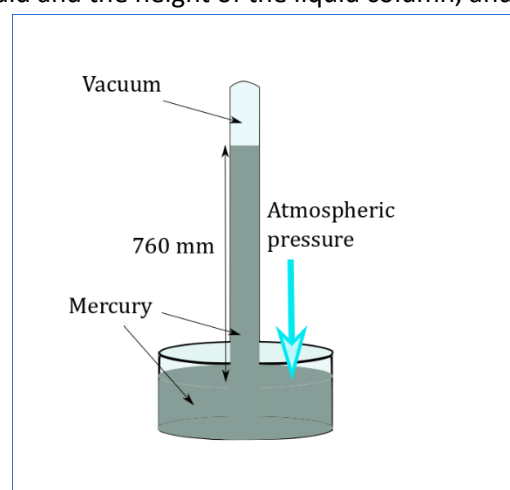


Figure 65. Measuring of air pressure.

The first modern measurement of the air pressure was done with Evangelista Torricelli in 1644, when he demonstrated; that the atmospheric pressure equals the pressure of 760 mm height mercury column. In his experiment, he filled with mercury a closed tube from one side and put it with its open side down into a vessel filled with mercury (Figure 65). The mercury leaked out and vacuum formed on the closed side of the tube, which stopped the leakage, when the pressure of the mercury column balanced the atmospheric pressure. In honour of Torricelli, the old units for the pressure were torrs. The conversion of torrs to pascals ($760 \text{ Torr} = 101\,325 \text{ Pa}$) is:

$$1 \text{ Torr} = 133.32 \text{ Pa}$$

12.2 Pascal's Principle

Definition: Pascal's principle (after Blaise Pascal, French scientist, 1623-1662)

When an additional pressure is exerted on a confined liquid, the pressure is transmitted equally to all parts of the liquid. (Figure 66)

The reason is the (almost) incompressible character of liquids; the pressure exerted on a liquid spreads in all directions and will have the same magnitude at every point of the liquid. Pascal's rule is exploited in hydraulic press (hydraulic jack), which consists of two connected vessels filled with liquid (technical oil, Figure 67). The diameters of the vessels are different (S_1, S_2), and to maintain constant pressure, the ratio of the forces acting on both vessels is the same as the ratios of the cross-sectional areas of the vessels (pistons).

$$p = \frac{F_1}{S_1} = \frac{F_2}{S_2}$$

$$\frac{F_2}{F_1} = \frac{S_2}{S_1} = \frac{h_1}{h_2}$$

eq. 162

h_1, h_2 are the vertical displacements of the pistons.

The ratio of the piston cross sections is called the hydraulic ratio and is also the ratio of the forces on both sides of the hydraulic jack. The benefit of a hydraulic press is that the force on one side is magnified on the other side. In this way, it is possible to balance heavy loads with smaller forces. However, in the case that we are lifting a load, the height of the lifting is in the reversed ratio of the cross-sectional areas. The proof is very simple; the volume of the hydraulic liquid pressed down on one side is the same as the volume pressed up on the other side:

$$V = S_1 h_1 = S_2 h_2 \quad \text{eq. 163}$$

It comes also from the work (energy) conservation law; the work received on one side is equal to the work input on the other side:

$$W = F_1 h_1 = F_2 h_2 \quad \text{eq. 164}$$

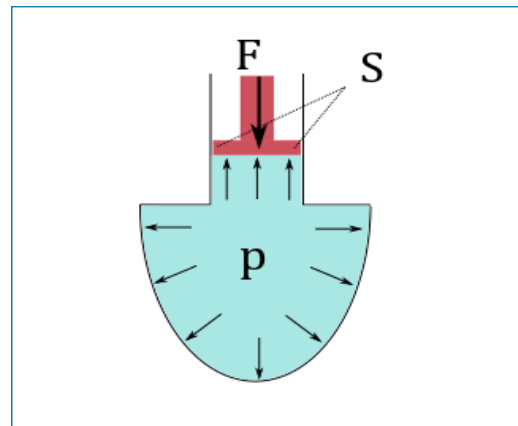


Figure 66. Spreading of pressure in a liquid.

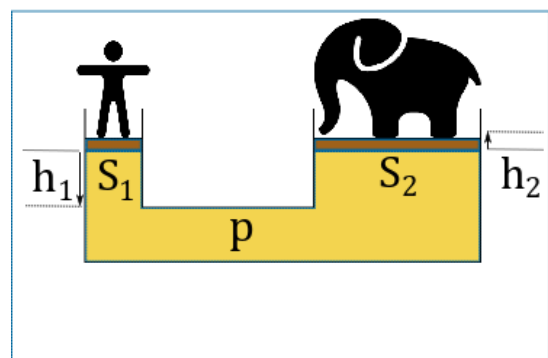


Figure 67. Hydraulic press scheme.

Problem. A person of 90 kg weight stands on the small piston of a hydraulic press and balances an elephant of weight 5000 kg on the large piston (Figure 67). The radius of the big piston (circular cross sections) is 1m. a) Find the radius of the small piston! b) How much should be pushed down the small piston to lift the elephant by 1 cm?

Solution: Using the equation:

$$\frac{F_2}{F_1} = \frac{S_2}{S_1}$$

substituting weights for the forces and areas of circles:

$$\frac{m_2 g}{m_1 g} = \frac{\pi r_2^2}{\pi r_1^2}$$

$$r_1 = \sqrt{\frac{m_1}{m_2}} r_2 = \sqrt{\frac{90}{5000}} 1 = 0.134m$$

$$\frac{F_2}{F_1} = \frac{h_1}{h_2}$$

$$h_1 = h_2 \frac{m_2 g}{m_1 g} = 1 \frac{5000}{90} = 55.55cm$$

Answer: The radius of the small piston is 0.134 m. To lift the elephant by 1cm, the small piston should be pushed down 55.56 cm.

12.3 Archimedes Law

Archimedes from Syracuse (c.287 - c.212 BC)

The pressure in a liquid, at a given point, acts uniformly in all directions (Pascal rule). An object submerged in a liquid is exposed to the pressure of the liquid at every point on its surface (Figure 68). This pressure depends on the height of the liquid column above this point. The pressure around the sides of the object is balanced with the same pressure on the opposite side of the object, which is at the same depth. The pressure on the top of the object is different from the pressure on the bottom, and these pressures are not balanced. The pressure at the bottom of the object is higher than on the top of the object and acts in the opposite direction to the gravity force. The net force acting on the object is calculated as the difference between the pressures and surface areas on the bottom and on the top of the object. The higher the difference of the hydrostatic force (pressure) at the bottom (F_B) and the force (pressure) at the top (F_T) of the object, the higher is the net lifting force, which is called the buoyancy force (F_L).

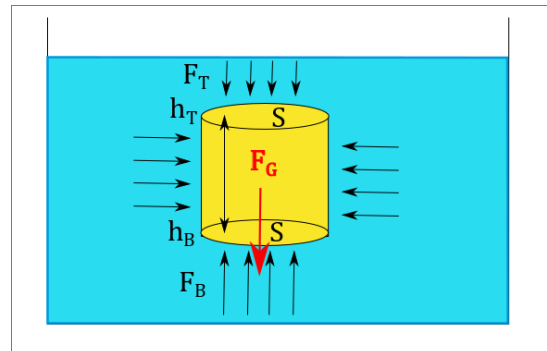


Figure 68. Hydrostatic forces in a liquid.

$$\begin{aligned} F_L = F_B - F_T &= \rho_{\text{liquid}} g h_B S - \rho_{\text{liquid}} g h_T S = \\ &= \rho_{\text{liquid}} g (h_B - h_T) S = \rho_{\text{liquid}} g V = m_{\text{liquid}} g \end{aligned} \quad \text{eq. 165}$$

Eq. 165 is the mathematical definitions of the *Archimedes law*:

Any object, wholly or partially immersed in a fluid, is buoyed by a force equal to the weight of the fluid displaced by the object.

The difference between the heights of the liquid columns at the bottom and at the top of the object is the height of the object ($h = h_B - h_T$, Figure 68), multiplied by the base area (S), is the volume (V) of the

object. The volume of the object multiplied by the density of the liquid and the gravity acceleration are the mass and weight of the liquid replaced by the object, respectively.

The net lifting force is the difference between the buoyancy (F_L) and the gravity force (F_g) of the object. The result of these two vectors (F):

$$F = F_g - F_L = V\rho_s g - V\rho_L g = Vg(\rho_s - \rho_L) \quad \text{eq. 166}$$

Problem: A homogeneous wooden cylinder of density $\rho_{\text{wood}}=740\text{kgm}^{-3}$ is immersed in water ($\rho_{\text{water}} = 1000 \text{ kgm}^{-3}$, Figure 69). Calculate what portion of the cylinder is above the water surface!

Solution: The height of the cylinder is h and the portion above the surface is x . x is between 0 and 1. The weight of the cylinder is balanced with the buoyancy, that is, the weight of the replaced water.

$$\begin{aligned} Sh\rho_{\text{wood}}g &= S(1-x)h\rho_{\text{water}}g \\ (1-x) &= \frac{\rho_{\text{wood}}}{\rho_{\text{water}}} = \frac{740}{1000} = 0.74 \\ x &= 0.26 \end{aligned}$$

$x=0.26$ (26%) of the wooden cylinder is above the surface

of the water. It is obvious from the solution that the portion above (below) the surface depends only on the ratio of the densities of the solid object and the liquid and is independent of the shape of the object.

The Archimedean law deals with fluids, not only liquids but also gases. However, the buoyancy force of the air is much less than the buoyancy force resulted from liquids, e.g., water. The ratio of forces is equal to the ratio of the densities of the liquid and gas. Let us have an object of volume (V) fully immersed in the fluid. The density of water (ρ_{water}) is 1000 kgm^{-3} and the density of air (ρ_{air}) is 1.23 kg m^{-3} . The buoyancy in water:

$$F_{\text{water}} = V\rho_{\text{water}}g = V \cdot 1000 \cdot 9.81 = V \cdot 9810 \text{ N}$$

In the air

$$\begin{aligned} F_{\text{air}} &= V\rho_{\text{air}}g = V \cdot 1.23 \cdot 9.81 = V \cdot 1.206 \text{ N} \\ \frac{F_{\text{water}}}{F_{\text{air}}} &= \frac{\rho_{\text{water}}}{\rho_{\text{air}}} \cong 813 \end{aligned}$$

The ratio of the two forces is more than eight hundred.

When we measure the mass (more precisely, the weight) of an object, our measurement is not precise because of the buoyancy of the air. The systematic error due to buoyancy in air is calculated from the ratio of correct weight (F_{cor}) and the measured weight in air (F_{air}):

$$\begin{aligned} \text{rel. error (in \%)} &= \frac{F_{\text{cor}} - F_{\text{air}}}{F_{\text{cor}}} \cdot 100\% = \\ &= \frac{V\rho_{\text{object}}g - V(\rho_{\text{object}} - \rho_{\text{air}})g}{V\rho_{\text{object}}g} \cdot 100\% = \frac{\rho_{\text{air}}}{\rho_{\text{object}}} \cdot 100\% \end{aligned}$$

The error depends on the density of the air and the material. Typical densities for solids are two to three orders of magnitude higher than the density of air; wood ca. 700 kgm^{-3} , plastics ca. 1100 kgm^{-3} , iron 7870 kgm^{-3} etc., gives the relative errors of 0.176%, 0.11%, and 0.016%, respectively.

12.4 Density and Density Measurements

Density is the mass of the object of unit volume.

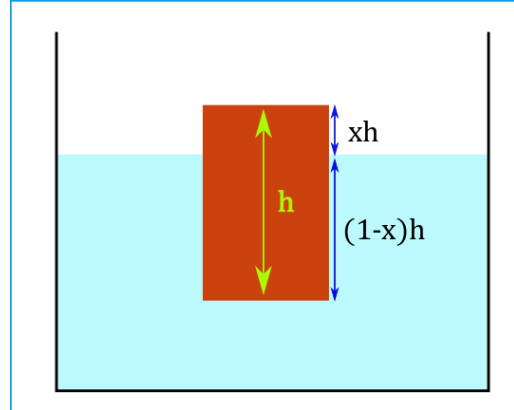


Figure 69. Partially immersed object in a liquid.

$$\rho = \frac{m}{V} \quad \text{eq. 167}$$

The basic unit of the density is $\text{kg}\cdot\text{m}^{-3}$. Other units are $\text{kg}\cdot\text{dm}^{-3}$ and $\text{g}\cdot\text{cm}^{-3}$. The density is an indicator of the purity of simple materials. Its value depends on the temperature, because the volume of an object depends on the temperature, too. The volume dependence on temperature is as follows:

$$V = V_0(1 + \beta\Delta T)$$

$$\rho = \frac{m}{V} = \frac{m}{V_0(1 + \beta\Delta T)} = \frac{\rho_0}{1 + \beta\Delta T} \quad \text{eq. 168}$$

Or

$$\rho = \frac{\rho_0}{(1 + \beta\Delta T)} \frac{(1 - \beta\Delta T)}{(1 - \beta\Delta T)} = \frac{\rho_0(1 - \beta\Delta T)}{1 - \beta^2\Delta T^2} \approx \rho_0(1 - \beta\Delta T)$$

where ρ_0 is the density at the reference temperature, β is the thermal expansion coefficient of the volume. Due to its small value ($\approx 10^{-5}$ - 10^{-6} K^{-1}), the $\beta^2\Delta T^2$ term is negligible in the denominator of eq.168. The density decreases with increasing temperature.

An important exception is the anomaly of water. Water is the most common liquid and solvent in nature, and it has a few very unusual properties. For example, it has a relatively high boiling temperature, a relatively high surface tension, and a density anomaly. The unusual properties relate to the structure of the water. Water forms hydrogen bonds between the hydrogen and oxygen atoms of neighbouring molecules, creating short-range structures.

A general rule which is valid for the materials is that the density decreases with increasing temperature. This is true for water above ca. 4°C . The water has a maximum density at temperature 3.98°C . The consequence of this anomaly is that water freezes from its surface. When the temperature decreases and reaches ca. 3.98°C , the cooler water falls to the bottom because it is denser than the warmer (and cooler as well) water. It continues until the temperatures is greater than 3.98°C . When the water temperature reaches 3.98°C , the circulation stops, and the upper layer will continue to cool and freeze. Ice has a lower density than water. The upper water layer and the later formed ice (partially) insulate the water layer on the bottom, and the water freezes from its top, not from bottom. Fortunately, because freezing of lakes and rivers from the bottom would have fatal consequences for living organisms in water. On the other hand, during hot periods, the warmer water (with smaller density than at 3.98°C) remains at the surface and keeps the layer below cooler. Increasing the temperature of water decreases the solubility of the air; therefore, too warm water is dangerous to organisms living in the water as well. The warmer top layer of water insulates the cooler bottom layer.

Densitometer. The density of liquids can be measured with different methods. Measurement with a densitometer is the fastest and easiest way, but the least precise. Densitometers consist of a glass vessel, closed on both sides. On the bottom of the densitometer are lead bullets, fixed with a layer of wax (Figure 70). The upper part of the densitometer is calibrated. The method is based in the principle presented on the previous problem (Figure 69), that is, the depth of immersion of an object in

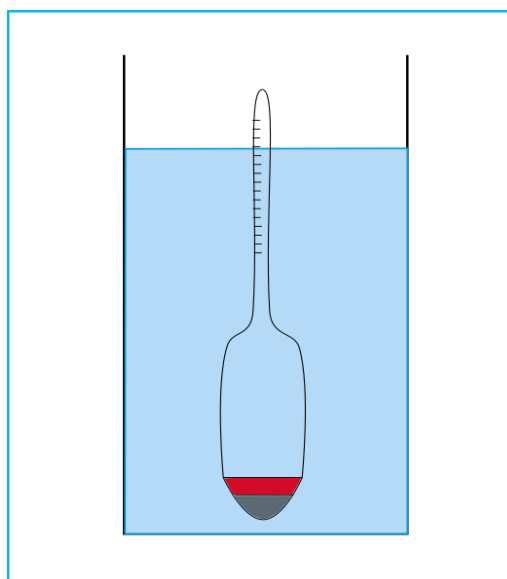


Figure 70. Densitometer.

a liquid depends only on the ratio of the (average) density of the solid object and the measured liquid. The density is read directly from the scale on the densitometer, at the level of the liquid surface. Another apparatus for density measurement is the *Mohr-Westphal scales*, which consists of scales with a glass cylinder hung on the end of the arm of the scales. The scales are equalised at the beginning of the experiment when the glass cylinder is not submerged in water or any liquid. After that, the whole volume of the glass cylinder is sunk into the liquid. The scales show a decrease in the weight of the cylinder due to buoyancy (F_B) in the liquid.

$$F_B = V_{cyl} \rho_{liq} g \quad eq. 169$$

The buoyance, proportional to the density of the liquid (ρ_{liq}), is compensated for by weights. The equipment is calibrated so that the weights used for compensation and their position on the scales arm directly show the density of the measured liquid.

12.4.1 Pycnometry

In density measurements, the mass determination is relatively simple and precise; however, the determination of the volume is rather difficult, especially of objects with irregular shape. Pycnometry is based on the precise determination of the volume of the sample. Pycnometer is a vessel with a tap and with accurately defined volume. Inside the tap is a capillary, which enables to leak out the extra liquid volume from the pycnometer. In the case of liquid samples, the pycnometer is filled with the liquid, and the mass of the liquid inside the pycnometer is found from the difference of the full and empty pycnometer. The volume of the liquid is either taken directly from the pycnometer specification, or a reference liquid, usually distilled water, is used. In the latter case, the exact volume of the pycnometer is calculated from the mass of water in the pycnometer and from the tabulated density of the water at the given temperature. The density of the liquid sample is found as follows:

$$\rho_{sample} = \frac{m_{sample}}{m_{water}} \rho_{water} \quad eq. 170$$

The density of a solid sample can also be found by pycnometry. During the experiment, the weights of the solid sample, the weights of the pycnometer filled with water, and the pycnometer with the solid sample and water are determined. The volume of the solid sample is determined from the volume (mass and density ratio) of the replaced water. The weight of the replaced water is determined from the weight of the water in the pycnometer added to the weight of the solid sample and the mass of the solid sample with the water in the pycnometer (Figure 71).

$$m_{replaced\ water} = \quad eq. 171$$

$$= m_{water} + m_{sample} - m_{water+sample}$$

The volume of the sample is the following:

$$V_{sample} = \frac{m_{replaced\ water}}{\rho_{water}} \quad eq. 172$$

Density is calculated in the usual way as the ratio of the mass of the sample to its volume.

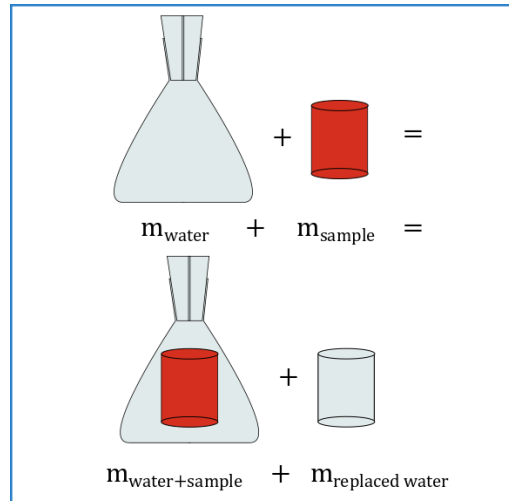


Figure 71. Determination of the volume of the sample in pycnometry.

$$\rho_{\text{sample}} = \frac{m_{\text{sample}}}{V_{\text{sample}}} \quad \text{eq. 173}$$

Important! What is the difference between weight and mass? These two quantities are closely related but not the same. The weight is the gravitational force of an object, defined as the product of the mass and the gravitational acceleration.

$$F_G = mg$$

In a certain geographical site, the gravitational acceleration is constant, so the weight and the mass are directly proportional. With scales, we measure the gravitational force of the object, not its mass, although we use the term mass in our results (unit kg), as well as the scales are calibrated to show mass and not the weight.

Problem. Find the density of solid beads with pycnometry. The mass of the empty pycnometer is 24 g. The mass of the water-filled pycnometer is 33.9725 g. The mass of the pycnometer with beads is 49.0542 g and with beads and water is 56.0801 g (the density of water is 997 kg m⁻³).

Solution: The mass of the replaced water is:

$$m_{\text{replaced water}} = (33.9725 - 24) + (49.0542 - 24) - (56.0801 - 24) = 2.9466 \text{ g}$$

The volume of the beads (=replaced water) is as follows:

$$V_{\text{sample}} = \frac{2.9466}{0.997} = 2.9555 \text{ cm}^3$$

and the density:

$$\rho_{\text{sample}} = \frac{49.0542 - 24}{2.9555} = 8.4772 \text{ gcm}^{-3} = 8477.2 \text{ kgm}^{-3}$$

The density of solid beads is 8477.2 kgm⁻³.

12.5 Fluid Dynamics (Hydrodynamics)

Fluid dynamics, formerly named hydrodynamics, is part of fluid mechanics which deals with the motion of fluids (liquids and gases) and the forces, pressures, etc. acting on these systems, as well as inside these systems. As solid objects differ from gases and liquids, aerodynamics (fluid dynamics of air and gases) and hydrodynamics (fluid dynamics of liquids, not only water) differ from solid object dynamics. In the case of solid objects, the shape and volume of the object are (usually) unchanged, the object is moving as a whole, and the relative distances of its parts are unchanged during the motion. The connections among the parts and components of a solid object are strong, whereas the connections between the components of gases and liquids are weak. The molecules of fluids slide on each other permanently changing their relative positions. Because of that, the interactions between the molecules change, they form and disappear. Even when they are in equilibrium, local fluctuations are present; the equilibrium is always a dynamic equilibrium, and the properties of the system, which characterise it, are average properties, which can slightly differ from the local properties.

12.5.1 Ideal Liquid

The ideal liquid is a model system that does not really exist. It helps to understand the basic rules that control the processes in real systems. It uses certain simplification, but at the cost of less precise values of the properties and processes. When an ideal system is used (ideal gas, ideal liquid, ideal crystal, etc.), it is necessary to consider the definition of the system and assess the impact of the simplifications on the result, to estimate how far the real properties of the system are from the definition of the ideal

one. If these conditions are close to the investigated system, the ideal system is suitable for our purposes, but if not, we can expect erroneous results.

The definition of an ideal liquid is as follows.

- its viscosity is zero, it has zero resistance against flow, during the flow no turbulences occur, and the flow is laminar.
- It is incompressible; its density is the same at every point of the liquid; it has steady flow rate, the velocity of flow rate being constant in time (but it can vary in different parts of the liquid).

12.5.2 Equation of Continuity, Bernoulli Equation

An ideal liquid flowing in a pipeline obeys the equation of continuity. It comes directly from the incompressibility of the ideal liquid; the amount of liquid that flows into an arbitrary cross-section equals the amount that flows out. There is no accumulation or diminution in the liquid at any point of the flowing liquid since its density is constant in space and time. The volume and mass of the liquid, which flows through a cross-section (S) with speed (v) during time (t), are:

$$V = Svt \quad m = \rho Svt$$

and the volume and mass flow rates are:

$$Q_V = \frac{V}{t} = Sv \quad Q_m = \frac{m}{t} = \rho Sv \quad \text{eq. 174}$$

The steady flow means that Q_v or Q_m are constant. The equation of continuity is as follows:

$$Sv = \text{const} \quad \text{or} \quad S_1v_1 = S_2v_2 \quad \text{eq. 175}$$

When an ideal liquid is flowing through a pipeline with a broader and narrower part, the decrease in the cross section of the pipeline causes an increase of the speed to keep its product constant (Figure 72).

Bernoulli equation (Daniel Bernoulli, 1700-1782). Bernoulli equation is valid for an ideal liquid (non-viscous and incompressible), where laminar steady flow is present and in its simplest form is valid for incompressible flows. It is based on the energy conservation law for an isolated system, and so is valid for an ideal liquid flowing through a pipeline (Figure 72). The flowing liquid is in equilibrium, steady-state flow is present, with constant flow rates (v_1, v_2) without applying external pressure, and in thermal equilibrium condition, also. The total energy is constant; its variation in time and space is zero. It consists of pressure energy, kinetic energy, and potential energy. The constant energy does not mean that its components are constants, also, but that the change of one component is compensated with a change of another component in the opposite sense. The mathematical formulation is as follows.

$$E_{\text{total}} = E_{\text{pressure}} + E_{\text{kinetic}} + E_{\text{potential}} = \text{constant}$$

or

$$dE_{\text{total}} = dE_{\text{pressure}} + dE_{\text{kinetic}} + dE_{\text{potential}} = 0$$

The individual energies are defined as follows:

$$E_{\text{pressure}} = pV$$

$$E_{\text{potential}} = mgh = \rho Vgh$$

$$E_{\text{kinetic}} = \frac{1}{2}mv^2 = \frac{1}{2}\rho Vv^2$$

$$dE_{\text{pressure}} = Vdp + pdV$$

$$dE_{\text{potential}} = \rho gh dV + \rho Vg dh$$

$$dE_{\text{kinetic}} = \frac{1}{2}\rho v^2 dV + \rho Vv dv$$

eq. 177

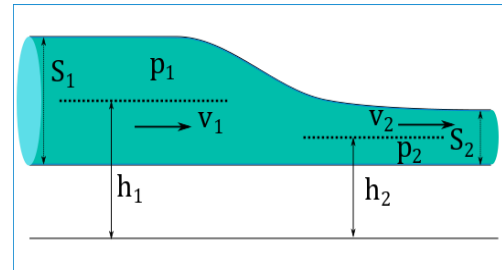


Figure 72. Ideal liquid flowing through a pipeline with reduced cross section.

The meanings of kinetic and potential energies are clear from the previous analogies in solid object mechanics. The pressure energy **is the energy contained in each unit of the fluid** because of the thermal kinetic motions of the atoms lessened by the attractive forces of the fluid molecules on each other. Formally, we can derive the relation between the energy and pressure from the definition of mechanical work (energy is the capability of systems to perform work).

$$W = Fs = pSs = pV \quad \text{eq. 178}$$

F is the force acting on the object, s is the trajectory of the motion and p is the pressure of the force acting on the surface S . Non-compressibility means that dV is zero, substituting eq. 177 to eq. 176:

$$\begin{aligned} Vdp + \rho Vv dv + \rho Vg dh &= 0 \\ dp + \rho v dv + \rho g dh &= 0 \end{aligned} \quad \text{eq. 179}$$

After dividing the equation by volume, all terms of the eq. 179 have the meaning of pressure (energy per unit volume); for example, the member from the potential energy (ρgh) is the hydrostatic pressure. The solution of eq. 179 for the system of two points (1,2) of the pipeline with pressures p_1, p_2 , flow speeds v_1, v_2 and average height (vertical position) h_1, h_2 gives the Bernoulli equation:

$$p_1 + \frac{1}{2}\rho v_1^2 + \rho gh_1 = p_2 + \frac{1}{2}\rho v_2^2 + \rho gh_2$$

Or

$$p + \frac{1}{2}\rho v^2 + \rho gh = \text{constant} \quad \text{eq. 180}$$

Problem: Find the changes in the flow speed and pressure of an ideal liquid ($\rho=1000\text{kgm}^{-3}$), which flows with a speed 0.3 ms^{-1} through a pipeline of 8 cm radius, pressure 100 kPa and in the same average vertical position, when the cross-section radius is reduced to 4 cm (Figure 73)!

Solution: We use the equation of continuity (eq. 175) to find the flowing speed in the narrower part:

$$v_2 = \frac{S_1 v_1}{S_2} = \frac{\pi r_1^2 v_1}{\pi r_2^2} = \frac{8^2 \times 0.30}{4^2} = 1.2\text{ms}^{-1}$$

Notice that we did not change the radius to metres; despite that, the result is correct, because the centimetres occur in the numerator and denominator, also.

Using a conversion to metres, the fraction can be simplified with the conversion factor (0.01^2).

$$\frac{0.08^2 \times 0.3}{0.04^2} = \frac{8^2 \times 0.01^2 \times 0.3}{4^2 \times 0.01^2} = \frac{8^2 \times 0.30}{4^2}$$

The speed in the narrower part of the pipeline is 1.2ms^{-1} .

The pressure is calculated from the Bernoulli equation (eq. 180):

$$\begin{aligned} p_2 &= p_1 + \frac{1}{2}\rho(v_1^2 - v_2^2) \\ &= 100\,000 + \frac{1}{2} \times 1000 \times (0.3^2 - 1.2^2) \\ p_2 &= 99325\text{ Pa} \end{aligned}$$

The pressure in the narrower part is 99.3 kPa. It means that the narrowing of the pipeline caused an increase of the speed, and the increase of the speed caused a lowering of the pressure.

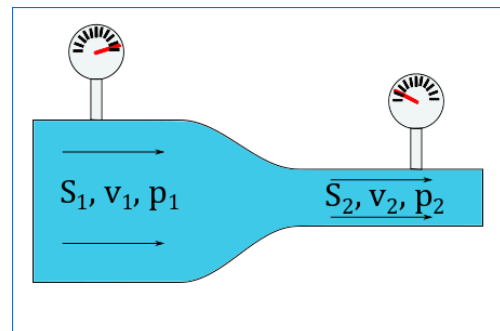


Figure 73. Change of pressure due to change of flow-speed of a liquid.

The previous problem showed that narrowing the pipeline causes an increase in the flowing speed because the same amount of liquid must flow through the narrower part as through the wider part; otherwise, an accumulation of liquid occurs on the narrow side and depletion on the wide side. It is a consequence of the incompressibility and is the result of the equation of continuity. The increase in flow speed resulted in a decrease in pressure from 100 kPa to 99.3 kPa. The pressure drop occurs not only in liquids but also in gases. It is used in several devices, in pressure drop sprayers, when fast-flowing air strikes the droplets of a liquid from its container because of the pressure drops and spreads it. Another important example is the airplane wing profile, the so-called aerofoil, which means a flat surface on the bottom side and a rounded surface on the top side. This type of wing produces a lift force due to the higher flowing speed on the top side than on the bottom side. The wing splits the air layers during its motion and the layers join each other approximately behind the wing in the same position. This means, that the air flowing speed on the top side is higher than on the bottom side, generating pressure drop and a net lifting force acting on the aircraft (Figure 74).

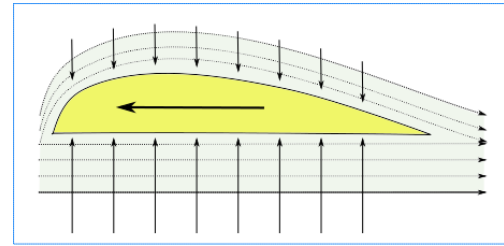


Figure 74. Airflow around an airplane wing.

In the previous problem, the vertical position of the pipeline was constant. Let us have a similar system, but instead of the cross-section change, the pipeline will continue at a different height (Figure 75).

Problem: Find the change of the pressure of an ideal liquid ($\rho=1000\text{kg m}^{-3}$), which is flowing with a speed 0.3 ms^{-1} through a pipeline of radius 8 cm, pressure 100 kPa and in the average vertical position 1m, when the altitude of the pipeline position is changed to 2m (Figure 75)!

Solution: The cross section is unchanged, so from the continuity equation it follows that the flowing speed is unchanged, also. Change in altitude changes pressure in the pipeline.

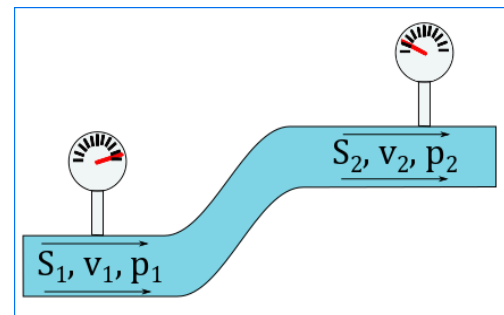


Figure 75. Pressure change due to pipeline position.

$$p_2 = p_1 + \rho g(h_1 - h_2) = 100\,000 + 1000 \times 9.81 \times (1 - 2) = 90\,190\text{ Pa}$$

The pressure is reduced to 90.190 kPa. The change in the average height h_1 to h_2 decreases the pressure to 9810 Pa for 1 m of lifting the position of the pipeline.

12.5.3 Toricelli's Theorem

Evangelista Torricelli (1608 –1647)

The Torricelli theorem is a special case of the Bernoulli equation. It concerns the flow speed of a liquid from a container with holes at different height (Figure 76). In previous cases, the liquid in a pipeline was in a closed vessel. In this case, the vessel is open and the pressures p_1 , p_2 are the same. On the left-hand side, we have the potential energy of the liquid (hydrostatic pressure), which is transformed to the kinetic energy on the right-hand side:

$$\rho gh = \frac{1}{2} \rho v^2 \quad \text{eq. 181}$$

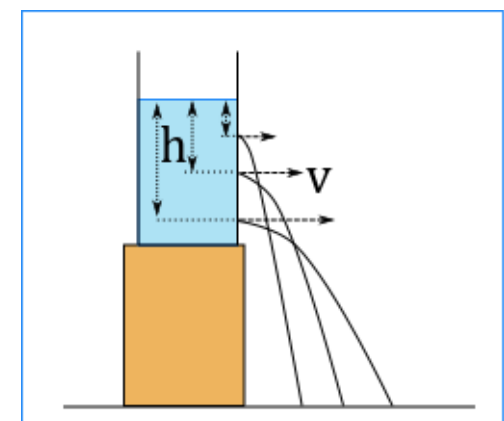


Figure 76. Outflow speed profile due to hydrostatic pressure.

The speed of liquid leaking from the vessel is:

$$v = \sqrt{2gh}$$

Mathematical background: A quadratic equation, like Bernoulli's equation, and therefore Torricelli's theorem, has two roots:

$$v = \pm\sqrt{2gh}$$

The negative root: that is, a negative speed is not correct from the physical point of view, so we do not consider it.

Problem: From a container with circular base of radius (r_v) 1m, water flows through a hole of diameter (r_h) of 4 cm (Figure 76). The height of the water column above the hole is constant (h_0), 80 cm. Find the speed of the water stream next to the hole!

Solution: We use the Torricelli theorem:

$$v = \sqrt{2gh} = \sqrt{2 \times 9.81 \times 0.8} = 3.96 \text{ ms}^{-1}$$

The speed of the water stream is 3.96 ms^{-1} .

The calculated speed is a function of the height of the water column. When this column height is kept constant, with the same amount of water supplied to the vessel, the speed will also be constant. If not, then the height of the column will decrease in time, as well as the speed of the water flow. Let us change the previous problem in this sense and calculate the time dependence of the water stream speed and the water column height!

Solution: The water volume, that leaks out through the hole, is the product of the speed and the hole area:

$$dV_h = \pi r_h^2 v dt = \pi r_h^2 \sqrt{2gh} dt$$

The volume decrease in the vessel is the product of the vessel base and the water column height decrease:

$$dV_v = -\pi r_v^2 dh$$

These two volumes are equal:

$$\pi r_h^2 \sqrt{2gh} dt = -\pi r_v^2 dh$$

$$\frac{dh}{\sqrt{h}} = -\frac{r_h^2}{r_v^2} \sqrt{2g} dt$$

Let us define a constant c to simplify the equation:

$$c = \frac{r_h^2}{r_v^2} \sqrt{2g}$$

$$\frac{dh}{\sqrt{h}} = -c dt$$

The solution after integration is as follows:

$$2\sqrt{h} = ct + \text{const.}$$

The integration constant comes from the initial condition:

$$t = 0 \Rightarrow h = h_0$$

$$\text{const} = 2\sqrt{h_0}$$

$$2\sqrt{h} = -ct + 2\sqrt{h_0}$$

$$h = \left(-\frac{1}{2}ct + \sqrt{h_0} \right)^2 = \left(-\frac{1}{2} \frac{r_h^2}{r_v^2} \sqrt{2g} t + \sqrt{h_0} \right)^2$$

$$h = (-0.00354t + 0.894427)^2$$

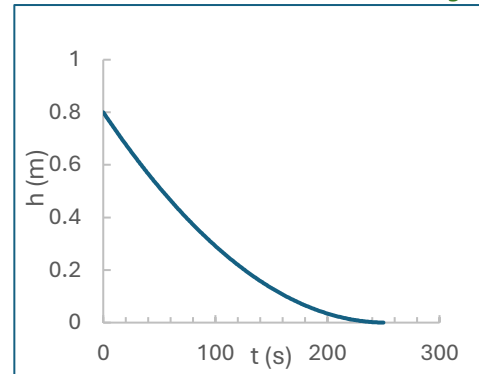


Figure 77. Decrease of liquid column height with time.

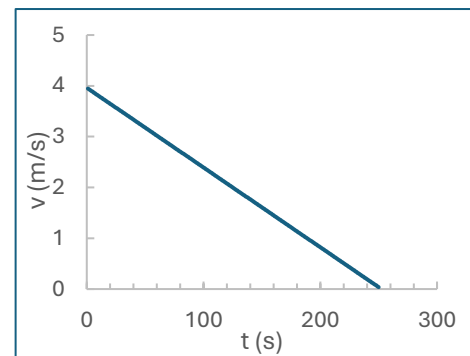


Figure 78. Decrease of liquid outflow speed with time

$$v = \sqrt{2gh} = \pm \left(-\frac{r_h^2}{r_v^2}gt + \sqrt{2gh_0} \right)$$

$$v = (-0.015696t + 3.9618)$$

The speed has two roots (see $\pm$ before the parentheses); the $+$ root means a linearly decreasing speed with time, while the $-$, a linearly increasing speed, which is not correct from the physical point of view. The volume of the water outflow is possible to find from the water depletion in the vessel or from the amount of outflow:

$$dV_{outflow} = \pi r_v^2 dh = \pi r_h^2 v dt$$

Substituting for v :

$$dV_{outflow} = \left(-\frac{\pi r_h^4 g}{r_v^2} t + \pi r_h^2 \sqrt{2gh_0} \right) dt$$

After integration:

$$V_{outflow} = -\frac{\pi r_h^4 g}{2r_v^2} t^2 + \pi r_h^2 \sqrt{2gh_0} t$$

The integration constant is zero because $V_{outflow}=0$ for $t=0$. For this special case:

$$V_{outflow} = -3.94483 \times 10^{-5} t^2 + 1.99144 \times 10^{-2} t$$

The whole process takes about 250 s, the height of the water column decreases in a parabolic manner, the speed of the water stream decreases linearly, and the total volume of the water outflow is an increasing parabolic function (Figure 77-Figure 79). The maximum value of the outflow volume in the following interval is 2.513 m^3 , what is the total volume of the 80 cm height water column in the vessel:

$$V_{total} = \pi r_v^2 h_0 = 2.51327 \text{ m}^3$$

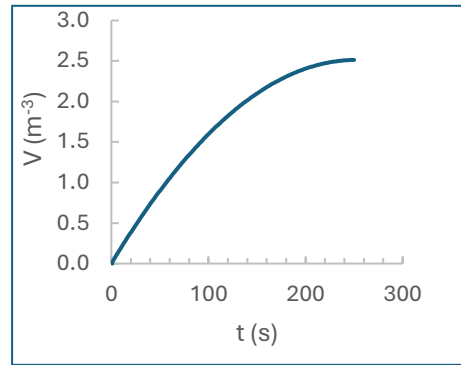


Figure 79. Increase of liquid volume leaked out from the container.

12.6 Real Liquids

12.6.1 Viscosity of Newtonian Liquids

Dynamic viscosity is the measure of the resistance of liquids to flow. The layers of liquid molecules can slide on each other, unlike molecules of solid materials which are fixed to their positions. Friction between the sliding layers of liquids causes viscosity. The viscosity of liquids varies significantly for different liquids (Table 5); well-known examples are water, relatively low viscosity, oil, and honey, high-viscosity liquids.

Viscosity is the result of intermolecular interactions between the molecules in the liquid. It is dependent on temperature, and with increasing temperature the viscosity is decreasing.

Let us have an experiment in which a liquid is between two plates of area S (Figure 80). One of the plates is fixed; the other is mobile. To keep the plate in motion at constant speed, we must exert a constant force F on the mobile plate. The liquid between the plates can be divided into several layers that will move at different speeds depending on the distance from the mobile plate; the closer they are to it, the higher is their speed. The force necessary to keep the mobile plate in motion with constant speed is proportional to the contact area of the liquid and the plate and to the speed gradient $\left(\frac{dv}{dy}\right)$

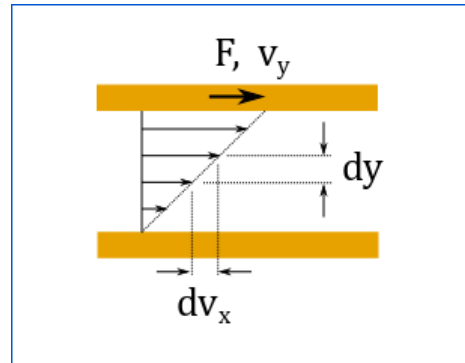


Figure 80. Gradient of the speed in a liquid between a mobile and static plate.

across the liquid layers. *The dynamic viscosity equals the force per unit surface, which keeps unit speed gradient across a flowing liquid.*

The tension between the layers, due to their different speeds, is the shear stress (τ in Pa):

$$F = \tau S = \eta S \frac{dv}{dy}$$

The equation

eq. 182

$$\tau = \eta \frac{dv}{dy}$$

is the mathematical definition of a Newtonian liquid, a liquid with constant viscosity, independently of applied shear at a given temperature.

The units of dynamic viscosity are Pa s:

$$\eta = \frac{F dy}{S dv} \Rightarrow [\eta] = 1 \frac{N}{m^2} \frac{m}{ms^{-1}} = 1 Pa s$$

The reversed value of the dynamic viscosity is the fluidity – φ :

$$\varphi = \frac{1}{\eta} \quad [\varphi] = 1 Pa^{-1} s^{-1} \quad \text{eq. 183}$$

Its meaning is the ease of flow of liquids. The ratio of dynamic viscosity and liquid's density (ρ) is kinematic viscosity (ν):

$$\nu = \frac{\eta}{\rho} \quad [\nu] = 1 \frac{Pa s}{kg m^{-3}} = 1 \frac{kg m^{-1} s^{-1}}{kg m^{-3}} = 1 m^2 s^{-1} \quad \text{eq. 184}$$

Viscosity is an important property of liquids; its value significantly affects their behaviour. In the definition of ideal liquids, viscosity was one of the properties that was neglected. This is a useful approximation for understanding some basic features of the hydrodynamics; however, real liquids have nonzero viscosity and, with its increasing value, the deviations of the behaviour of real liquids from the predicted behaviour become more and more significant.

Table 5. Examples of dynamic viscosities of liquids⁹ at 20°C.

Liquid	$\eta \cdot 10^3$ (Pa s)	Liquid	$\eta \cdot 10^3$ (Pa s)
Acetone ⁸	0.332	n-Butanol	2.948
Water	1.002	Oil, linseed	33.1
Methanol	0.597	Oil, olive	84.8
Ethanol	1.2	Oil, soja bean	69.3
n-Propanol	2.256	Glycerine	1490

12.6.2 Hagen-Poiseuille's Law

In the case of a laminar flow of a liquid in a tube, the liquid can be divided into thin layers that flow with different speeds, and the shear stress is proportional to the speed between the layers. The layer at the wall has zero speed, and the layer at the middle has maximum speed. The friction force acting between the layers *opposite* to the force, which keeps the liquid moving (Figure 81):

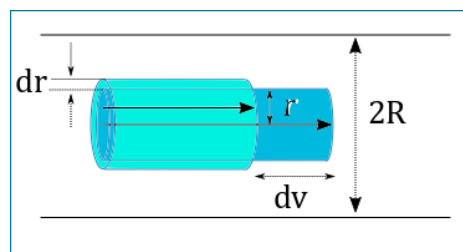


Figure 81. Flow of a real liquids.

$$F_{fr} = S_{contact} \tau = L 2\pi r \eta \frac{dv}{dr} \quad \text{eq. 185}$$

⁸ Lange's Handbook of Chemistry, 10th Edition, 1967, McGraw-Hill Book Co., New York

⁹ CRC Handbook of Chemistry and Physics (n7)

where r is the distance of the layer from the tubes center (radius of the cylinder), L is the length of the tube and $S_{contact}$ is the contact area of the two neighbouring layers. The moving force of the liquid is the following:

$$F_{moving} = S_{cross} \Delta p = \pi r^2 \Delta p \quad eq. 186$$

where S_{cross} is the area of the cross section and Δp is the pressure difference between the end points of the tube (length L). In the case of a steady flow, the two forces are balanced:

$$\begin{aligned} -L 2\pi r \eta \frac{dv}{dr} &= \pi r^2 \Delta p \\ dv &= -\frac{\Delta p}{2L\eta} r dr \end{aligned}$$

The negative sign is due to the opposite direction of forces. After integration:

$$v(r) = -\frac{\Delta p}{4L\eta} r^2 + const$$

The integration constant is found from the boundary condition; the flow speed on the wall ($r=R$, $r=-R$) is zero:

$$\begin{aligned} 0 &= -\frac{\Delta p}{4L\eta} R^2 + const \\ v(r) &= \frac{\Delta p}{4L\eta} (R^2 - r^2) \end{aligned} \quad eq. 187$$

The speed profile is a quadratic function of the tube radius; it is a parabola with a maximum value in the middle of the tube at:

$$r = 0 \Rightarrow v(0) = \frac{\Delta p}{4L\eta} R^2 \quad eq. 188$$

and minimum value:

$$r = R \text{ or } r = -R \Rightarrow v(R) = 0$$

The flow volume rate is defined as

$$Q_V = \int_0^S v(r) dS = \int_0^R \frac{\Delta p}{4L\eta} (R^2 - r^2) 2\pi r dr$$

Notice that:

$$\begin{aligned} dS &= d(\pi r^2) = 2\pi r dr \\ Q_V &= \frac{\pi \Delta p}{2L\eta} \int_0^R (R^2 - r^2) r dr = \frac{\pi \Delta p}{2L\eta} \left[R^2 \frac{r^2}{2} - \frac{r^4}{4} \right]_0^R = \frac{\pi \Delta p}{2L\eta} \left[\frac{R^4}{2} - \frac{R^4}{4} \right] \\ Q_V &= \frac{\pi R^4 \Delta p}{8L\eta} \end{aligned} \quad eq. 189$$

The volume flow rate depends on the fourth power of the tube radius. That is the reason for dramatic changes in blood pressure and/or blood flow rate in cases of arteriosclerosis or vasodilatation, when the effective radius of the arteries becomes smaller or larger.

Problem: a) The blood pressure is 120mmHg and the flow rate is 100 cm min⁻¹. The radius of an artery decreased by 12% due to arteriosclerosis. How much must change the blood pressure to keep the flow rate at the same level?

Solution: the initial and final radii are r_1 and r_2 , respectively. The 12% decrease means:

$$r_2 = 0.88r_1$$

We must keep the flow rate constant; so

$$\frac{Q_V 8L\eta}{\pi} = \Delta p r^4 = constant$$

$$\Delta p_1 r_1^4 = \Delta p_2 r_2^4 = \Delta p_2 (0.88 r_1)^4$$

$$\Delta p_2 = \frac{\Delta p_1}{0.88^4} = \frac{120}{0.6} = 200 \text{ mmHg}$$

Blood pressure should be 200 mmHg to maintain the same blood flow, which is dangerous for the patient.

b) What will be the flow rate in the previous problem if we keep the pressure constant?

Solution:

$$\frac{Q_V}{r^4} = \frac{\pi \Delta p}{8 \eta L} = \text{constant}$$

$$\frac{Q_{V1}}{r_1^4} = \frac{Q_{V2}}{r_2^4}$$

$$Q_{V2} = Q_{V1} \frac{r_2^4}{r_1^4} = Q_{V1} \left(\frac{0.88 r_1}{r_1} \right)^4 = 0.6 Q_{V1} = 60 \text{ cm}^3 \text{ min}^{-1}$$

The blood flow rate will decrease from $100 \text{ cm}^3 \text{ min}^{-1}$ to $60 \text{ cm}^3 \text{ min}^{-1}$.

12.6.3 Measuring of Viscosity

Viscosity can be measured with different methods and different equipments; rotational, capillary (Ubelohde), falling ball (falling sphere, Höppler) viscometers, etc.

Höppler viscometer. The viscometer consists of a tube with circular cross section, closed with taps on both sides. The tube is filled with the measured liquid and a ball falls in the tube (Figure 82). The diameter of the ball is almost the same as the diameter of the tube. The ball, after the initial stage, falls at constant speed. Measurement is carried out during this constant-speed period. The constant speed means that the result of all forces that act on the ball is zero (consider Newton's first law), the forces are in equilibrium. There are three forces present: gravity (F_G), buoyancy (F_B), and friction force (Stokes force, F_S). The orientation of gravity is opposite, as is the orientation of the buoyance and friction force.

$$F_G - F_B - F_S = 0 \quad \text{eq. 190}$$

F_G is the weight of the falling ball, its mass - m , volume - V , density - ρ_S and radius - r .

$$F_G = mg = V \rho_S g = \frac{4}{3} \pi r^3 \rho_S g \quad \text{eq. 191}$$

The buoyance is equal to the weight of the replaced liquid of density - ρ_L .

$$F_B = \frac{4}{3} \pi r^3 \rho_L g \quad \text{eq. 192}$$

Friction (Stokes law) is proportional to the dynamic viscosity of the liquid (η), radius of the ball (r) and its speed (v).

$$F_S = 6 \pi \eta r v \quad \text{eq. 193}$$

$$\frac{4}{3} \pi r^3 \rho_S g - \frac{4}{3} \pi r^3 \rho_L g - 6 \pi \eta r v = 0 \quad \text{eq. 194}$$

$$v = \frac{2r^2(\rho_S - \rho_L)g}{9\eta}$$

During the experiment, the time (t) that the ball falls between two marks on the tube is measured. The motion is a simple uniform linear motion; the trajectory (s) of the ball is as follows:

$$s = vt = \frac{2r^2(\rho_S - \rho_L)g}{9\eta} t \quad \text{eq. 195}$$

The falling speed of the ball is proportional to the second power of its radius, the density difference of the ball, and the liquid, and inversely proportional to the dynamic viscosity. Viscosity is expressed as

$$\eta = \frac{2r^2(\rho_s - \rho_L)g}{9s} t = K(\rho_s - \rho_L)t \quad \text{eq. 196}$$

where

$$K = \frac{2r^2g}{9s}$$

K is the constant of the apparatus, which is found from a measurement of a liquid of known viscosity, usually distilled water. Eq. 194 is also valid for the settlement of dispersions in solutions. The calculated settling velocity (v) of a particle depends on its radius (r), on the difference between the densities of the particle and the liquid, and the viscosity of the liquid. When optimising the parameters (particle radius, liquid viscosity, solid and liquid density), it is possible to suppress or enhance sedimentation according to the actual requirements. In the case of drugs prepared in suspension form, sedimentation of dispersed solid particles is not desired. With a lower radius of the particle, higher viscosity of the liquid, and comparable densities of the solid and liquid phases, it is possible to avoid sedimentation. However, in some cases, it is necessary to separate the liquid and solid phases (e.g., blood test). In these cases, we can use a centrifuge to improve sedimentation. In the centrifuge, the sample is spinning at high speed, so the gravity acceleration (g) in the equation is replaced by the centrifugal acceleration (a_c):

$$a_c = \omega^2 R = (2\pi f)^2 R$$

$$v = \frac{2r^2(\rho_s - \rho_L)\omega^2 R}{9\eta} = \frac{2r^2(\rho_s - \rho_L)(2\pi f)^2 R}{9\eta} \quad \text{eq. 197}$$

Where ω is the angular speed, R is the distance of the spinning sample from the rotation axis, f is the rotation frequency. The settling velocity increases linearly with the rotor radius (R) but quadratically with the rotation frequency.

Capillary viscometer (Ubbelohde viscometer, Figure 83). The main part of this viscometer is a glass tube with a capillary (6), through which the measured liquid flows. Above the capillary, there is a broadened part (5) of the tube, a reservoir, for the measured liquid. The tube has two marks (A, B), and the time of the liquid depletion between these two marks is registered to ensure the measurement of the same volume of the liquid outflow. The outflow volume is proportional to time:

$$V = Q_V t = \frac{\pi \Delta p r^4}{8\eta L} t = \frac{\pi h \rho g r^4}{8\eta L} t \quad \text{eq. 198}$$

The pressure is the hydrostatic pressure of the measured liquid:

$$\Delta p = h\rho g \quad \text{eq. 199}$$

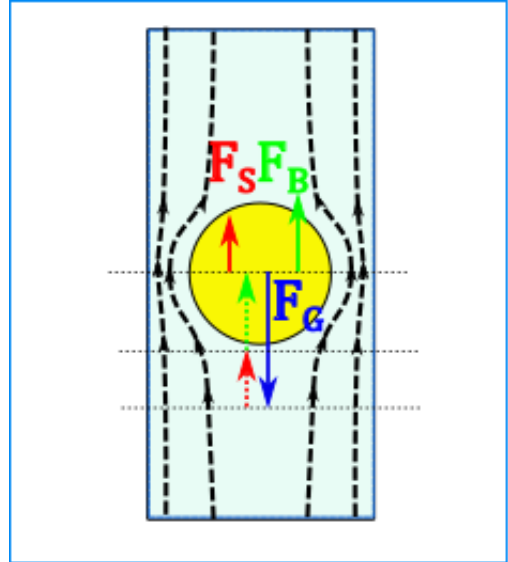


Figure 82. Forces in Höppler's viscometer.

The height of the liquid column (h) is changing with time, so it is not constant in this meaning; however, it changes in the same way for every measurement.

$$\frac{\eta}{\rho} = \frac{\pi r^4 h g}{8LV} t \quad \text{eq. 200}$$

$$\frac{\eta}{\rho} = \nu = At - \frac{B}{t}$$

where ν is the kinematic viscosity, A , B are constants. The second term (with constant B) on the left-hand side is a correction to obtain more precise values. In this method, a reference liquid, usually distilled water, is used, and the kinematic viscosity is found from the ratios of the outflow times. It is interesting that while the Höppler viscometer measures the dynamic viscosity, the Ubbelohde viscometer measures the kinematic viscosity.

The flow of real liquids differs from that of ideal liquids. The latter have zero viscosity, so they flow without any friction. The streamlines do not cross each other (they are parallel), and the flow is laminar. The flow rate is constant throughout the area of the tube. Real liquids have nonzero viscosity, the layer at the wall of the tube is attached to the wall, and the layers closer to the middle are flowing faster. The speed profile is parabolic for the laminar flow (see eq. 187, Figure 84). The flow is characterised by a dimensionless number; the so-called Reynolds number (Re):

$$Re = \frac{\text{inertial force}}{\text{viscous force}} = \frac{\nu \rho d}{\eta} = \frac{\nu d}{\nu}$$

Where ν and d are the characteristic speed and diameter of the flow, ρ is the density of the liquid, η and ν are the dynamic and kinematic viscosities, respectively. The flow is laminar if the Reynolds number is less than approx. 2000 and turbulent above ca. 4000. Between 2000 and 4000 is a transition regime. Re is proportional to the flow speed, the density of the liquid and the diameter of the pipe, while inversely proportional to the viscosity.

The change in flowing character from laminar to turbulent or vice versa is accompanied by a change in the intensity of the flow noise. During blood pressure measurements, the cuff squeezes the artery. If the external pressure is higher than the blood pressure in the artery, there will be no blood flow and no sounds. After that, the external pressure is continuously reduced. When external pressure decreases below systolic blood pressure, the artery opens for this time and turbulent flow occurs with the so-called Korotkoff sounds. External pressure continues to decrease, and in the moment when it is lower than diastolic blood pressure, the cuff does not squeeze the artery, the blood flow is laminar and silent. The appearance and disappearance of these sounds indicate the values of the systolic and diastolic blood pressures.

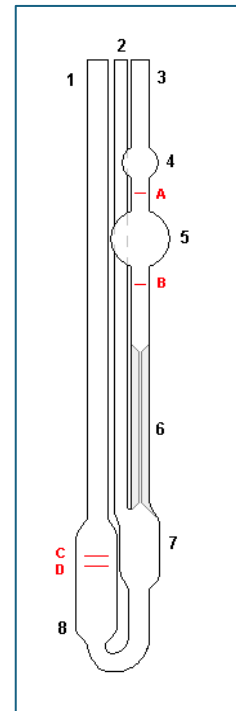


Figure 83. Ubbelohde's viscometer.¹⁰

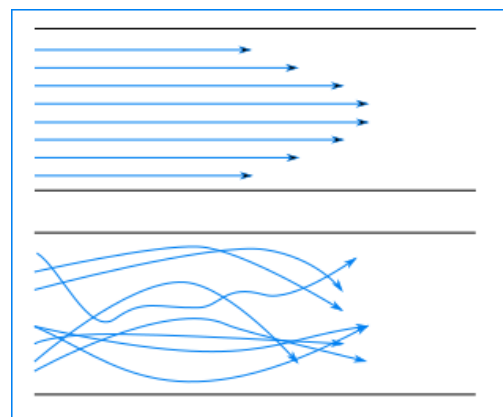


Figure 84. Laminar (top) and turbulent (bottom) flow

¹⁰Image source https://commons.wikimedia.org/wiki/File:Ubbelohde_lepkosciomierz.PNG

13 Thermal Physics

Temperature is a property of objects that we notice permanently; the weather (surrounding atmosphere) is warm or cold, the meal, the drink is hot or cold, etc. It is an intensive property; it means that it is independent of the amount of material, on the size of the system. Temperature is an important physical quantity, and it very often affects the extent or direction of physical and chemical processes. Measuring the temperature is a common task in everyday life; however, we do not directly measure the temperature, but we do measure another property affected by it. The most common thermometers are filled with a coloured liquid (alcohol), formerly with mercury, and the temperature is determined according to the volume expansion of the liquid in a capillary. Different thermometers work on the temperature dependence of other physical properties, e.g. electrical resistance, pressure of the gases, infrared radiation, etc. The measured value is expressed in unit kelvin, or in degrees Celsius (centigrade).

Kelvin, symbol K, is the SI unit of temperature and is one of the SI base units. In physics, in formulas and equations, kelvins are used, in everyday life rather Celsius. The conversion is as follows:

$$T \text{ (K)} = t \text{ (}^{\circ}\text{C)} + 273.15 \quad \text{eq. 201}$$

The unit values of the two temperature scales are equal; therefore, the differences between two temperatures have the same value in kelvins or degrees of Celsius. In the case where in the formula there is a difference of temperature, substitution in degrees of Celsius in a calculation is correct, but in other cases usually leads to erroneous results.

Later, we will see that the temperature cannot fall to and below 0 K (-273.15 ° C); negative kelvin temperatures are not defined.

13.1 Thermal Expansion of Solid and Liquid Materials

The volumes of solid materials depend on their temperature. In general, the higher their temperature, the larger their volume. In the case that one of their dimensions is significantly bigger than the remaining two (length vs. width and height), the change in dimension is simplified to a linear thermal expansion. A common example is the elongation of rail tracks. To avoid dangerous situations, the trails regularly contain gaps, which allow the material to expand without deformation of the path. The linear expansion is calculated as follows:

$$l = l_0(1 + \Delta l) = l_0(1 + \alpha \Delta T) \quad \text{eq. 202}$$

Where l is the length at the final temperature, l_0 is the initial length, ΔT is the difference between the final and the initial temperature, and α (unit K^{-1}) is the linear expansion coefficient. The expansion occurs in the other dimensions, also, but as the height and width of the trails (a few centimetres) are negligible in comparison with the length (hundreds and thousands of kilometres), the expansions of the width and height are negligible, also. In the previous example, we saw thermal expansion as a potential source of danger which should be considered in projects of an object of huge dimensions; however, it is exploited, as in the case of liquid (ethanol, mercury) thermometers, where the temperature is measured due to the thermal expansion of the liquid. Another example is the bimetallic switch, which is used to convert the temperature change to a mechanical action, usually to turn off a process to avoid overheating or other undesirable processes. It consists of metal strips with different expansion coefficients. Metal strips are fixed to each other along their length. The increase in temperature causes a bending of the strips to the side of the material having a lower expansion coefficient to compensate for the difference in their length at elevated temperature. When these metal

strips are used as a switch in an electrical circuit, bending of the bimetallic strip causes the switch to be turned off.

The thermal coefficient values for solid materials are of the order of 10^{-5} - 10^{-6} K⁻¹. Liquid materials have (in general) higher values than solid materials. It is a consequence of the crystalline lattice of the solid materials and the restricted mobility of the particles in the lattice. These motions are mainly vibrations around the equilibrium position. The particles in liquids provide also translational and rotational motions, and the interparticle distances can also vary in a larger interval. The volume expansion of an object (all dimensions) is as follows.

$$\begin{aligned} V &= (l + \Delta l)^3 = (l_0 + 3l_0^2\Delta l + 3l_0\Delta l^2 + \Delta l^3) \cong \\ &\cong l_0^3 \left(1 + \frac{3\Delta l}{l_0}\right) = V_0(1 + 3\alpha\Delta T) = V_0(1 + \beta\Delta T) \end{aligned} \quad \text{eq. 203}$$

where from eq. 202:
$$\frac{\Delta l}{l_0} = \frac{l-l_0}{l_0} = \alpha\Delta T$$

Eq. 203 is valid for small Δl , so Δl^2 and Δl^3 are negligible. β is the thermal expansion coefficient, its value is three times the value of the linear thermal expansion coefficient.

13.2 Thermodynamics

Thermodynamics is a branch of physics that deals with thermodynamic functions such as heat, energy, work, etc. of a system and with their changes during the processes.

A system in thermodynamics is the content of a macroscopic volume that is separated from its environment by walls. The state of this system is defined with the thermodynamic variables (temperature, pressure, volume, etc.). The processes involve changes within the system and between the system and its environment, exchange of energy (work), and matter. The systems in thermodynamics are categorised according to the allowance to exchange energy and matter with the environment. A system which can exchange matter and energy with its environment is an *open system*. In the case of a *closed system*, the exchange of energy is allowed but the exchange of matter is not. The third type is an *isolated system* when no energy nor matter exchange is allowed between the system and its environment.

Thermodynamic variables that describe the state of the system are temperature, pressure, volume, number of moles, etc. Some of them depend on the size of the system; these are *extensive* variables (volume, number of moles, number of particles, internal energy, enthalpy), while the others; *intensive* variables, which are independent of the size of the system (pressure, temperature). Extensive variables can be transformed to intensive variables. For example, the volume (extensive) of a matter divided by its number of moles results in the molar volume of this matter, which is an intensive variable.

Thermodynamic functions are divided into two groups. The so-called *state functions* are defined as the functions whose values and changes are given only with their state and do not depend on the path along which they reached this state. Among these functions are the internal energy, the enthalpy, the entropy, etc. The second type of functions is the *path functions*; their values and changes depend on the path through the system passes to reach the actual state. Path functions are, for example, work and heat. The mathematical condition of the state function is the exact differential of the function (see later in chapter Enthalpy, Helmholtz energy, Gibbs Free Energy).

Processes that result in a transition of the system from one state to another are followed by changes in their variables. These processes are *reversible* or *irreversible* processes. In the case of reversible processes, the system is in equilibrium with its environment during the process, and it is possible to

revert the direction of the process with a small action. In the case of irreversible processes, these requirements are not met. Reversible processes are idealised processes; real processes are usually irreversible. However, the changes of the state functions for irreversible processes are equal to the changes during the reversible processes if they have the same initial and final states and that simplifies their determination.

13.2.1 Ideal Gas

Ideal gas is defined as a gaseous system, where:

- a. the size of the gaseous particle is neglected.
- b. There are no interactions between the particles except for the moments of their collisions.
- c. The collisions of the particles are always elastic; that is, the sum of the kinetic energy of the colliding particles is conserved during the collision and it is not transformed into other types of energy (potential energy, bonding energy, etc.).
- d. Ideal gas molecules consist of monoatomic molecules (such as noble gases); the particles perform merely translational motions, without vibrations or rotations.

Real systems usually do not follow these conditions; noble gases resemble the ideal gas state at low pressures. Real gases 'violate' the requirements of the ideal gas state, and their particles often consist of more than one atom (CO_2 , O_2 , N_2 , etc.); due to this, they perform not only translational motions, but also vibrations and rotations as well. The particles have nonzero volume, so the volume available for other particles is reduced by the occupied volume, and between the particles there are certain interactions (e.g., electrostatic). The ideal gas is a useful approximation that simplifies the processes to understand them or to receive a solution, although not entirely precise, when the problem is too complicated to solve. In the case of real gases, the extent of the error depends on the dissimilarity of the properties of the definition of the real gas and the ideal gas.

The particles of a material are in permanent motion. It is valid for all states; solid, liquid, gaseous, and plasma, also. However, the measure and the restrictions of these motions decide about the state of the material. In the case of fluids (gas and liquid), the particles move randomly (*Brownian motion*) from one place to another and collide with other particles along their path. The length of the path between two collisions depends on the density of the fluid; usually, it is larger for gases than for liquids. This motion is responsible for local density fluctuations even in a homogeneous fluid in equilibrium. This is basically a translational motion, but particles consisting of more than one atom also perform vibrations and rotations. In the case of solid materials, the particles cannot move away from their lattice positions, so their motions are usually restricted to vibrational motions.

13.3 Thermodynamics of Gases

The following chapters deal with the thermodynamics of gases, mainly ideal gases. The investigated system contains a gas, which is defined with its properties (pressure, temperature, volume, amount of substance, etc.), and it is surrounded by an environment. We consider an equilibrium state of the system, usually, and according to the conditions, different processes occur between the system and the environment, which result in changes in the state of the system and its properties. These processes can be very complex, and to simplify it, we usually fix a parameter or property of the system. The processes are classified according to this condition, for example, if the temperature is constant, the processes will run under isothermal conditions, in the case of constant pressure under isobaric conditions, etc.

13.3.1 Gay-Lussac I. Law, Charles' Law, Law of Volumes

All names in the title mean the same law, which is valid in *isobaric* conditions, that is, the **pressure** in the system is **constant**. The system consists of an ideal gas surrounded by an environment. In the system, no chemical reactions occur, the system is closed but not isolated; no exchange of matter is allowed between the system and the environment, but energy (heat) exchange is possible. The part of the wall between the environment and the system is mobile (a piston), which can move without resistance (friction), to maintain constant pressure inside the system (Figure 85). Most real processes are run under isobaric conditions, where the pressure is the atmospheric pressure.

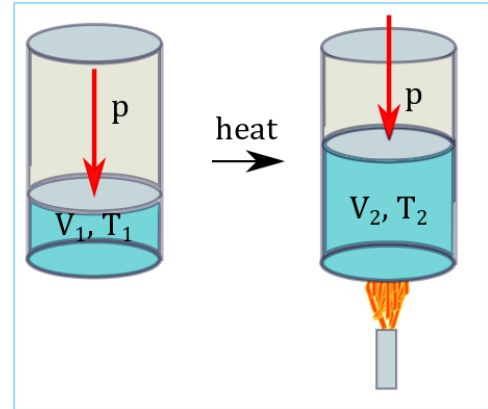


Figure 85. Isobaric gas expansion.

$$p = \text{const}$$

The amount of gas inside the system is constant because no chemical *reactions* take place in the system, as well as no matter exchange is allowed with the environment (closed system). The volume and temperature relation of the system during the processes is as follows:

$$\frac{V}{T} = \text{const}$$

$$\frac{V_1}{T_1} = \frac{V_2}{T_2}$$

eq. 204

or

$$V = k T$$

where k is a constant.

The volume of an ideal gas in equilibrium conditions and under constant pressure is proportional to its temperature. At very low temperatures (close to 0 K), it is not possible to apply this law; there is a significant deviation of the behaviour of the system from it.

When the graph of volumes of isobaric systems is plotted against the temperature (Figure 86), increasing lines are obtained with increasing temperature. Extrapolating these lines to the low temperature region, the lines intercept the temperature axis at one point, at the value -273.15°C ($=0\text{ K}$), which is the so-called absolute zero temperature (zero kelvin). At this temperature, the volume should be zero, which is impossible.

Let us have an ideal gas at $273,15\text{ K} = 0^\circ\text{C}$ of volume V_0 . The increase in its volume when the temperature increases by one degree is:

$$\frac{V_1}{273,15 + 1} = \frac{V_0}{273,15}$$

$$V_1 = V_0 \left(\frac{273,15 + 1}{273,15} \right) = V_0 \left(1 + \frac{1}{273,15} \right)$$

$$\Delta V = V_1 - V_0 = \frac{V_0}{273,15}$$

The volume of an ideal gas at $273\text{ K} - 0^\circ\text{C}$ increases (decreases) with its $1/273$ fraction when the temperature is increased (decreased) by one degree. It is also approximately valid for real gases.

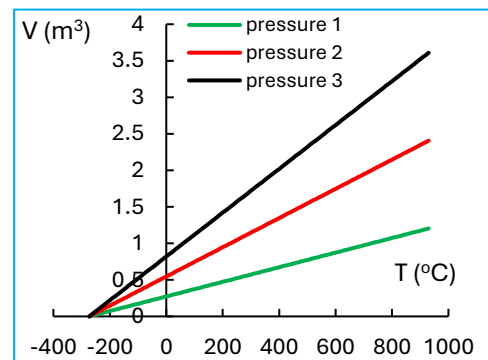


Figure 86. Isobaric volume-temperature graphs.

13.3.2 Gay-Lussac II. Law, Pressure-Temperature Law

The basic condition for the *isochoric* processes is the **constant volume**. The system consists of an ideal gas in a container. This container has a constant volume, no chemical reactions occur in the system, and no material exchange is allowed between the environment and the system. The system is not isolated, so energy exchange is allowed.

$$V = \text{const} \quad \frac{p_1}{T_1} = \frac{p_2}{T_2} = \text{const} \quad p = k T \quad \text{eq. 205}$$

This means that the pressure of an ideal gas in equilibrium is proportional to its temperature. The graph of pressure on the thermodynamic temperature is also linear (similar to Figure 86). Extrapolating the pressure to zero pascal, the lines intercept the temperature axis at zero kelvin, similarly as in the case of the isobaric processes.

13.3.3 Boyle-Mariotte Law

When a process is run at **constant temperature**, that is, *isothermal process*, then the Boyle-Mariotte law is valid. At a constant temperature, the product of the pressure and volume of a given mass of an ideal gas in a closed system is always constant. The system consists of an ideal gas; no material exchange is allowed between the system and the environment, but the walls of the system allow a quick heat exchange, and the system is kept under constant temperature, which is the temperature of the environment. One of the walls of the container is mobile and can move without friction.

$$T = \text{const} \quad p_1 V_1 = p_2 V_2 \quad p = \frac{k}{V} \quad pV = \text{const} \quad \text{eq. 206}$$

When the pressure is doubled, the volume decreases to one-half of its value, and when the pressure is halved, the volume doubles. The pressure of an ideal gas in equilibrium is inversely proportional to its volume and vice versa. The graph $p=f(V)$ is a hyperbola. Neither pressure nor volume can have zero value because it would require an infinity value of the other parameter.

$$0 = \frac{k}{\infty}$$

13.3.4 Avogadro's Law

Under the same temperature and pressure, equal volumes of different gases contain an equal number of molecules, regardless of their chemical nature and physical properties. The volume of a gas depends on the number of molecules or the number of moles.

$$V = k n = k' N \quad \text{eq. 207}$$

where k and k' are constants, and n and N are the number of moles and number of molecules, respectively. The molar volume of gases is defined as

$$V_m = \frac{V}{n} \quad V = n V_m \quad \text{eq. 208}$$

Under standard conditions (pressure 101 325 Pa, temperature 273.15 K), the molar volume of the gases is 22.41 dm³ (L) and the number of molecules in this volume is the *Avogadro number*

$$N_A = 6.022 \cdot 10^{23} \text{ mol}^{-1}$$

13.3.5 Ideal Gas Law – General Gas Equation

The combination of previous laws results in the ideal gas law, which involves all of them. It is valid for ideal gases, but the behaviour of real gases is like that of a certain measure.

$$pV = nRT \quad \text{eq. 209}$$

where R is the universal gas constant ($= 8.314 \text{ J K}^{-1} \text{ mol}^{-1}$). If the number of moles n is replaced by the number of particles N , the equation is as follows:

$$pV = NkT \quad \text{eq. 210}$$

where k is the Boltzmann constant ($= 1.3806 \cdot 10^{-23} \text{ J K}^{-1}$) and the relationships are:

$$R = N_A k \quad N = N_A n \quad \text{eq. 211}$$

The universal gas constant is a macroscopic constant on the scale of moles, and the Boltzmann constant is a microscopic constant on the scale of the particles.

13.3.6 Dalton's Law

In a mixture of nonreacting gases, the total exerted pressure is equal to the sum of the partial pressures of the individual gases. The total amount of gases is the sum of the individual amounts:

$$n = \sum_i n_i \quad \text{eq. 212}$$

where n is the total number of moles of gases and n_i is the individual number of moles of the gases. The total pressure is the sum of the individual pressures:

$$p = \frac{nRT}{V} = \sum_i \frac{n_i RT}{V} = \sum_i p_i \quad \text{eq. 213}$$

Each gas component in a mixture of gases exerts the same pressure as if it were in the container alone.

13.4 Real Gases, van der Waals Equation

Dalton's law considers the gases components as entities that do not affect each other. This is true for a certain measure. When we consider, for example, the molar volumes of a liquid and its vapours (gas) at normal pressure, the volume of the gaseous phase is larger by several orders, which means that the interparticle distances in gases are significantly larger than those of liquids. During the Brownian motion of gaseous particles, they move randomly and are not bound with interparticle forces to a certain position, as in a solid lattice, and they do not form a surface, as in the case of the liquids, which hold the liquid molecules inside this volume. However, there are certain differences between ideal gases and real gases. In real gases, there are weak interparticle interactions (attractions, repulsions) that affect the pressure of the system, and the particles occupy a volume which is not available for other particles. To consider these differences, the ideal gas law was modified to the *van der Waals equation*.

$$\left(p + \frac{n^2 a}{V^2} \right) (V - nb) = nRT \quad \text{eq. 214}$$

where a and b are characteristic constants for the given gas. In the first parentheses is the correction of the pressure ($\frac{n^2 a}{V^2}$), due to the interparticle interactions, in the second, it is the correction of the volume due to the occupied volume (nb) with other particles.

13.5 Equipartition Theorem

Real gases do not necessarily fulfil one more requirement of the definition of ideal gases; they are not monoatomic particles (except for a few cases of noble gases). Real gases are generally diatomic, sometimes with three or even more atomic molecules. In the kinetic theory of gases, the internal energy of the gases is correlated with the number of degrees of freedom of motion. A monoatomic particle makes translational motions in the three directions of the space; no vibrations or rotations are present. According to the equipartition theorem, the energy is evenly distributed between the degrees of freedom of motion, and in the case of the ideal gas the value is:

$$\epsilon = n_{df} \frac{1}{2} k_B T \quad \text{eq. 215}$$

where ϵ is the energy of one particle, n_{df} number of degrees of freedom, k_B Boltzmann constant, T thermodynamic temperature. For the ideal gas $n_{df}=3$, because the particles motion vector is an unrestrained combination of vector components in three different directions (in x, y, and z), so:

$$\epsilon = \frac{3}{2} k_B T \quad \text{eq. 216}$$

for one particle. For the N particle:

$$\epsilon_N = \frac{3}{2} N k_B T = \frac{3}{2} n N_A k_B T = \frac{3}{2} n R T \quad \text{eq. 217}$$

This is the kinetic energy of the ideal gas, and because the potential energy of the ideal gas is zero, this energy is the internal energy (U , see later) of the ideal gas.

In the case of diatomic molecules, we must consider the rotation around the axes, which are perpendicular to the bond. These two additional motions alter the energy of diatomic molecules to the following:

$$\epsilon_N = \frac{5}{2} n R T \quad \text{eq. 218}$$

Values $3/2R$ and $5/2R$ are the isochoric molar heat capacities of monoatomic and diatomic gases, the energy necessary to increase the temperature of 1 mole of gas by one kelvin. Isochoric means constant volume, so the heat capacity is measured under the conditions of constant volume when the gas does not perform work and the added heat is transformed into a change of the internal energy only. The heat capacity value for an ideal gas is ca. $12.5 \text{ J mol}^{-1} \text{ K}^{-1}$, which is close to the values of noble gases. The molar isochoric heat capacities of the two atomic gases (CO , N_2 , H_2 , etc.) is approximately $20.6 \text{ J mol}^{-1} \text{ K}^{-1}$, which is $\frac{5}{2}R$. At high temperatures (10^3 - 10^4K), vibrational excitations should also be considered, and the heat capacity increases with the corresponding degrees of freedom.

13.6 Heat and Temperature, Zeroth Law of Thermodynamics

Heat is the manifestation of thermal energy. The unit of heat is *joule* (J). Heat and temperature are related to each other:

$$dQ = c dT \quad \text{eq. 219}$$

Here, dQ is the infinitesimal change in the heat (received or released) by an object, and dT is the corresponding temperature change. c is the heat capacity of the object (unit: $J \text{ K}^{-1}$), which is an extensive property, depends on the character and size (mass, number of moles, etc.) of the object. Because of that, we usually use the intensive equivalents of it, the molar (c_m) or specific (c_s , mass) heat capacities, which are independent of the amount of the material.

$$dQ = n c_m dT \quad dQ = m c_s dT \quad \text{eq. 220}$$

where n and m are the number of moles and the mass of the substance, respectively. The unit of molar heat capacity is $J\ mol^{-1}K^{-1}$, and the specific heat capacity is $J\ kg^{-1}K^{-1}$.

The molar (specific) heat capacity is the amount of heat necessary to increase the temperature of one mole (one kilogramme) of the material by 1 K. The heat tends to be distributed evenly, which means that if two objects of different temperatures are in contact, heat flows from the hotter object to the cooler until their temperatures are the same. Heat **never flows spontaneously** from a cooler object to a hotter one! When two objects have the same temperature, they are in thermal equilibrium, and no heat flow occurs between them. The thermal equilibrium is transitive:

When two objects are in thermal equilibrium with a third one, then they are in thermal equilibrium with each other. This is the so-called **Zeroth Law of Thermodynamics**. It also allows us to measure the temperature with contact thermometers when from the thermal equilibrium of the measured object and the thermometer we also deduce the thermal equilibrium with each object of the same temperature.

13.7 Thermodynamic Functions, Energy Conservation Law

As was shown in eq. 217, the overall kinetic energy of the particles of the ideal gas is the internal energy of the ideal gas.

$$U = \frac{3}{2}nRT = \frac{3}{2}NkT \quad \text{eq. 221}$$

The units of internal energy are *joules (J)* and depends only on the amount of ideal gas and the temperature. When the temperature increases, the internal energy of the ideal gas also increases linearly. At zero kelvin its value would be (theoretically) zero. When the system receives heat ($Q>0$) the temperature of the gas increases and usually (but not always) performs a work. The measure of the temperature growth and the performed work correlates with the received heat that; the higher the amount of heat is absorbed by the gas, the higher its temperature and the more work will be performed. Heat (thermal energy) is transformed into an equivalent amount of internal energy and work. During the process, no energy was created, no energy was destroyed, but it was changed to another form (internal energy, work). This equivalence is expressed in the conservation law of energy, which was formulated by several scientists in a different way.

- *Energy can be neither created nor destroyed (J.R. Mayer).*
- *Inside a closed (isolated) system, i.e., a system into which neither energy is transported nor taken out of it, in which arbitrary processes (mechanical, thermal, electrical, and chemical) occur, the total energy remains conserved (H. Helmholtz).*
- *Different forms of energy (electrical, mechanical, and thermal) could be transformed into each other (J.P.Joule).*
- *In a closed system, in which arbitrary mechanical and thermal processes occur, the total energy is constant (R.J.E Clausius).*

All these formulations are equivalent and express the fact that energy cannot be created or destroyed, but it is possible to change it to another form of energy or work. It is also the law that excludes the possibility of building a machine that produces work without input energy, a 'perpetuum mobile' of the first kind.

13.8 Work of Ideal Gases

Mechanical work was defined as the scalar product of the force acting on an object and the displacement of the object due to this force.

$$W = Fs \cos \alpha \quad \text{eq. 222}$$

$$dW = F \cos \alpha \, ds$$

where F is the force acting on the object and s (ds) is the displacement of the object and α is the angle between the force and the displacement. In the thermodynamics of ideal gases, the system consists of an ideal gas, surrounded by a wall and the environment. Except for the isochoric processes, part of the wall is mobile, allowing expansion or compression of the volume of the system. In the case of reversible processes, the system is in equilibrium with the environment at every moment of the process, and the pressure of the system is equal to the pressure of the environment. The thermodynamic work relates to the volume change (Figure 87) at a certain pressure and can be derived:

$$dW = Fds = -pS \, ds = -p \, dV \quad \text{eq. 223}$$

where p is the pressure, S is the surface and dV is the infinitesimal change in the volume of the ideal gas. The sign of the changes of the thermodynamic functions (heat, work, enthalpy, entropy...) is considered from the point of view of the system according to the IUPAC (International Union of Pure and Applied Chemistry) convention, when the system receives heat or work, it is positive and when releases negative. When the volume of the system increases (expansion, $\Delta V > 0$), the work is done by the system (released from the system), so the work must be negative, when the volume decreases (compression, $\Delta V < 0$), the work is done on the system, the system received work; the work is positive.

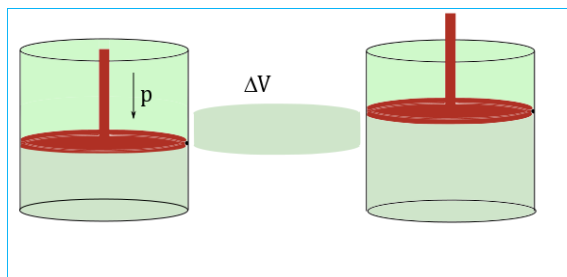


Figure 87. Volume change of gas during expansion.

In the following chapters, we inspect the internal energy, work, and entropy changes of an ideal gas-containing system when one of the parameters (volume, pressure, temperature, heat exchange) of the system is fixed. All processes are *reversible processes*, which means that the system is at every moment in equilibrium with the environment, and it is possible with a small action to change the direction of the process.

13.9 Internal Energy, First Law of Thermodynamics

Definition: In a closed system, between any two equilibrium states, the change in internal energy is equal to the difference between the heat transfer into the system and the work done by the system.

$$dU = dQ + dW \quad \text{eq. 224}$$

where U is the internal energy of the system, Q is the heat exchanged by the system and the environment, and W is the work done by/on the system. The sign of the values of the heat and work is considered according to the system, positive, when received by the system, and negative, when released by the system (not the operations of adding or subtracting in eq. 224). The system receives work and heat if:

$$dW > 0 \quad \text{and} \quad dQ > 0 \quad \text{eq. 225}$$

and work is done by the system and releases heat if:

$$dW < 0 \quad \text{and} \quad dQ < 0 \quad \text{eq. 226}$$

Internal energy is a state function, an extensive property; its value depends on the state of the system only and does not depend on the path through which the system reached the actual state. We usually calculate the change in internal energy, and it is useful to choose a reversible path. The calculated value of its change is valid for all possible paths that also have a common starting point and endpoint, even for irreversible paths.

13.10 Entropy

Entropy is a thermodynamic function; it is a state function and an extensive property.

*Definition. In a **reversible** process, an infinitesimal increase in the entropy (dS) of a system is the result of an infinitesimal transfer of heat (dQ) to a closed system divided by the common temperature (T) of the system and the surroundings which supply the heat.*

$$dS = \frac{dQ}{T} \quad \text{eq. 227}$$

In the case of an irreversible process, the change in the entropy is larger than the heat income over the temperature.

$$dS > \frac{dQ}{T} \quad \text{eq. 228}$$

The meaning of the entropy was expressed in several ways according to the field of interest (thermodynamics, informatics, astronomy, and statistical physics) and point of view of the definition.

- For a closed thermodynamic system, a quantitative measure of the amount of thermal energy not available to perform the work.
- A measure of disorder or randomness in a closed system.
- Measure of the loss of information in a transmitted message.
- The tendency of all matter and energy in the universe to evolve toward a state of inert uniformity.

From the definition formula of the entropy (eq. 227), we can substitute for dQ into the equation of the first law of thermodynamics and the definition formula for the thermodynamic work (eq. 223):

$$dU = dQ + dW = TdS - pdV \quad \text{eq. 229}$$

In this equation, we consider the interaction of the system with the environment in terms of heat exchange and the exchange of work between them.

13.11 Enthalpy, Helmholtz Energy, Gibbs Free Energy

Enthalpy is a state function and an extensive property. It is defined as the sum of the internal energy of the system and the pressure - volume product of the system.

$$H = U + pV \quad \text{eq. 230}$$

$$dH = dU + pdV + Vdp$$

Mathematical background:

The differential of a product of two functions ($f(x)$, $g(x)$) is:

$$d(f \cdot g) = g \cdot df + f \cdot dg$$

Example:

$$d(pV) = Vdp + pdV$$

The change in enthalpy is usually related to absorbing or releasing heat from or to the system, for example, chemical reactions are classified as exotherm (heat producing) or endotherm (heat absorbing) reactions according to it. We can define the change in enthalpy as a function of entropy, temperature, pressure, and volume, as in the case of the internal energy.

Substituting for dU (eq. 229):

$$\begin{aligned} dH &= TdS - pdV + pdV + Vdp \\ dH &= TdS + Vdp \end{aligned} \quad \text{eq. 231}$$

Helmholtz energy (A) is defined as the difference between the internal energy and the product of the temperature and entropy.

$$A = U - TS \quad \text{eq. 232}$$

Following a similar process as in the case of the enthalpy.

$$dA = dU - TdS - SdT = TdS - pdV - TdS - SdT = -pdV - SdT \quad \text{eq. 233}$$

Gibbs free energy is defined as:

$$\begin{aligned} G &= H - TS \\ G &= U + pV - TS \end{aligned} \quad \text{eq. 234}$$

and

$$\begin{aligned} dG &= dU + pdV + Vdp - TdS - SdT = \\ &= TdS - pdV + pdV + Vdp - TdS - SdT = Vdp - SdT \end{aligned}$$

Gibbs free energy is the key function in thermodynamics; its change decides about the direction and spontaneity of processes.

To summarise the thermodynamic potentials, we repeat the previous results and list them in one place.

All these functions are state functions; their differentials are exact differentials.

$$\begin{aligned} dU &= -pdV + TdS \\ dH &= Vdp + TdS \\ dA &= -pdV - SdT \\ dG &= Vdp - SdT \end{aligned} \quad \text{eq. 235}$$

Thermodynamic functions (U , H , A , and G) are expressed as a product of an extensive variable (V , S) and an intensive variable (T , p).

We consider in these chapters a closed system, which means that no material exchange between the system and the environment is allowed, and no chemical reaction occurs during the process. The number and chemical character of the particles do not change during the process. The total differential is the mathematical equivalence of the state function property in thermodynamics, explained in the following paragraph.

Let us have a function f (for example, Gibbs free energy) of two variables x , y (for example, p , T). The total differential of the function is defined as follows:

$$df = \left(\frac{\partial f}{\partial x} \right)_y dx + \left(\frac{\partial f}{\partial y} \right)_x dy \quad \text{eq. 236}$$

where $\left(\frac{\partial f}{\partial x} \right)_y$ is the partial derivation of f according to x ,

while y is constant and vice versa.

As a demonstration of this property, let us give an example of a phase diagram, which is an expression of the Gibbs free energy equilibrium for a material (Figure 88) as a function of pressure and temperature. On the boiling line of the material (which separates the liquid and gas phase areas) two points (Figure 88, points 1,3) are chosen, which express the change in boiling temperature, when the system pressure is changed and vice versa. The infinitesimal change in the Gibbs energy of the liquid-gas equilibrium dG (paths 1-3) is achieved, for example, by

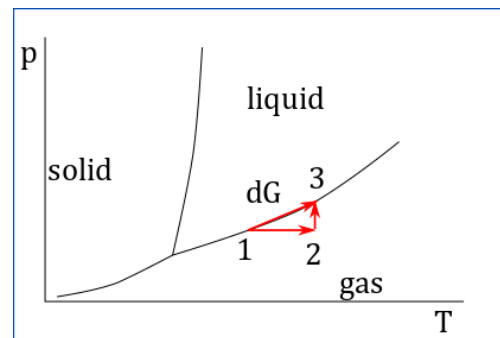


Figure 88. Transition between two states along the evaporation line curve.

increasing the system temperature at constant pressure (paths 1-2) and then increasing the pressure at constant temperature (paths 2-3). The equilibrium is achieved with new conditions; point 3:

$$\begin{aligned} G_{3l} &= G_{1l} + \left(\frac{\partial G_l}{\partial T}\right)_p dT + \left(\frac{\partial G_l}{\partial p}\right)_T dp \\ G_{3g} &= G_{1g} + \left(\frac{\partial G_g}{\partial T}\right)_p dT + \left(\frac{\partial G_g}{\partial p}\right)_T dp \end{aligned} \quad \text{eq. 237}$$

the subscripts g and l stand for the gaseous and liquid phases. G is a state function, so its changes along the two paths, 1-3 and 1-2-3, are equal.

The partial derivatives are the entropy and the volume of the system (eq. 235):

$$\left(\frac{\partial G}{\partial T}\right)_p = -S \quad \text{and} \quad \left(\frac{\partial G}{\partial p}\right)_T = V \quad \text{eq. 238}$$

The liquid and gaseous phases are in equilibrium at points 1 and 3 (not in 2), so the values of G of the two phases are equal at these points.

$$\begin{aligned} G_{1l} &= G_{1g} & G_{3l} &= G_{3g} \\ -S_l dT + V_l dp &= -S_g dT + V_g dp \end{aligned} \quad \text{eq. 239}$$

Rearranging the equation and defining the changes of entropy (ΔS) and volume (ΔV) of the material during the boiling process, we receive the Clausius-Clapeyron relation:

$$\begin{aligned} \Delta S &= S_g - S_l \\ \Delta V &= V_g - V_l \\ \frac{dp}{dT} &= \frac{\Delta S}{\Delta V} \end{aligned} \quad \text{eq. 240}$$

and substituting for ΔS at a given temperature T :

$$\begin{aligned} \Delta S &= \frac{\Delta H_{latent}}{T} \\ \frac{dp}{dT} &= \frac{\Delta H_{latent}}{T \Delta V} \end{aligned} \quad \text{eq. 241}$$

Where ΔH_{latent} is the enthalpy of the phase transition (evaporation), the so-called latent heat. The slope (derivative, dp/dT) of the evaporation lines in the phase diagram is equal to the ratio of entropy and volume changes during the phase transition, or to the ratio of evaporation heat (ΔH_{latent}) and the product of boiling temperature and volume change during the phase transition. In the case of the evaporation and sublimation, i.e., when one of the phases is the gaseous phase (at not too high pressure), the volume change is approximately equal to the volume of the gaseous phase because the volume of a gas at normal pressure is much higher than the volume of the same amount of liquid or solid material. For example, the volume of one mole of water vapour (gaseous phase) under normal conditions is ca. 22.41 L; on the other hand, in the liquid or solid phase it is around 18 ml; three orders less, it is negligible compared to 22.41 L. Substituting for the ΔV the volume of the gaseous phase and using the formula of the ideal gas law (eq. 209):

$$\begin{aligned} \Delta V &\cong V_g = \frac{nRT}{p} \\ \frac{dp}{dT} &= \frac{\Delta H_{latent} p}{nRT^2} \end{aligned} \quad \text{eq. 242}$$

The solution after separation of variables and integration is as follows:

$$\int_{p_1}^{p_2} \frac{dp}{p} = \int_{T_1}^{T_2} \frac{\Delta H_{latent}}{nRT^2} dT \quad \text{eq. 243}$$

$$\ln \frac{p_2}{p_1} = \frac{\Delta H_{latent}}{nR} \left(\frac{1}{T_1} - \frac{1}{T_2} \right)$$

Eq. 243 in integral form is **not valid** for a phase transition in which the gaseous phase is not involved; it means melting (solidification) of the solid (liquid) phase because in that case the change of the molar volume during the phase transition cannot be replaced by the volume of the gaseous phase.

Another important property of a total differential is the equivalence of the mixed derivatives, that is, the values of the mixed second partial derivatives do not depend on the order of derivation.

$$\left(\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right)_y \right)_x = \left(\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right)_x \right)_y \quad \text{eq. 244}$$

Applying this formula to the previous example, we receive:

$$\left(\frac{\partial}{\partial T} \left(\frac{\partial G}{\partial p} \right)_T \right)_p = \left(\frac{\partial}{\partial p} \left(\frac{\partial G}{\partial T} \right)_p \right)_T$$

$$\left(\frac{\partial V}{\partial T} \right)_p = - \left(\frac{\partial S}{\partial p} \right)_T$$

In an equilibrated system, the volume variation on temperature equals the entropy variation on pressure with opposite sign. This equation is one of the so-called Maxwell relations. In Table 6 are summarised the definitions, differentials, and Maxwell relations of the thermodynamic functions. The subscripts next to the parenthesis indicate the fixed variable in the process.

There are functions; thermodynamic work and heat, which are **not** state functions, but **path** functions and their differentials are not exact differentials. Contrary to state functions, the system cannot be characterised by the work or heat of the process. Their values depend on the path through which the system goes from one state to another, and the value of path functions must be found according to the given process which describes the transition between the initial and final states.

Table 6. Definition formulas of thermodynamic functions.

Thermo-dynamic function	Define-tion	differenti-al	p=	V=	T=	S=	Maxwell relation
Internal energy – $U(V,S)$	$\Delta Q+W$	$TdS-pdV$	$-\left(\frac{\partial U}{\partial V}\right)_S$	-	$\left(\frac{\partial U}{\partial S}\right)_V$	-	$-\left(\frac{\partial p}{\partial S}\right)_V = \left(\frac{\partial T}{\partial V}\right)_S$
Enthalpy – $H(p,S)$	$U+pV$	$Vdp +TdS$	-	$\left(\frac{\partial H}{\partial p}\right)_S$	$\left(\frac{\partial H}{\partial S}\right)_p$	-	$\left(\frac{\partial V}{\partial S}\right)_p = \left(\frac{\partial T}{\partial p}\right)_S$
Helmholtz energy – $A(V,T)$	$U-TS$	$-pdV-SdT$	$-\left(\frac{\partial A}{\partial V}\right)_T$	-	-	$-\left(\frac{\partial A}{\partial T}\right)_V$	$\left(\frac{\partial p}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T$
Gibbs energy – $G(p,T)$	$U+pV-TS$	$Vdp-SdT$	-	$\left(\frac{\partial G}{\partial p}\right)_T$	-	$-\left(\frac{\partial G}{\partial T}\right)_p$	$\left(\frac{\partial V}{\partial T}\right)_p = -\left(\frac{\partial S}{\partial p}\right)_T$

13.12 Thermodynamic Equilibrium

A thermodynamic equilibrium in a system is defined according to the previous state functions, for which the corresponding thermodynamic potential (Gibbs energy, Helmholtz energy, internal energy) has a minimum value. For example, in the case of a system, which is defined with the variables pressure and temperature (eq. 235), the Gibbs energy has a minimum value. The mathematical definition is as follows:

$$\begin{aligned}dG(p, T) &= 0 \\d^2G(p, T) &> 0\end{aligned}\quad \text{eq. 245}$$

A similar equilibrium condition is valid for the other thermodynamic potentials (internal energy and Helmholtz energy). On the other hand, the value of the entropy has a maximum value under the equilibrium condition, mathematically:

$$\begin{aligned}dS &= 0 \\d^2S &< 0\end{aligned}\quad \text{eq. 246}$$

The equilibrium conditions can be expressed in general with the entropy maximisation and energy minimisation principles.

Entropy Maximisation Principle. The equilibrium value of any unconstrained internal parameter is such as to maximise the entropy for the given value of the total internal energy.

Energy Minimisation Principle. The equilibrium value of any unconstrained internal parameter is such as to minimise the energy for the given value of the total entropy.

As an example, we can consider boiling water. Above 100 ° C and at normal pressure, the water evaporates to steam, gaseous phase. Under stable conditions, the system reaches the equilibrium state and persists in this state for an arbitrary long time. This is a stable steady state. In some rare cases, the water will not begin to boil at 100 ° C and normal pressure, usually if it is heated and cooled many times and contains too little of dissolved gases or other nucleation centres (impurities). After reaching the boiling temperature, the temperature increases further above 100 ° C without boiling. However, because this is superheated water, the steady state is unstable. It can persist only under undisturbed conditions and a disturbance; shaking, a piece of impurity, dust, etc. will cause a transition to a stable steady state when the water begins to boil explosively and the temperature drops rapidly to the boiling point, evolving the corresponding extra energy (difference of energy at superheated and normal boiling temperature) as an explosion. A superheated water is an unstable state; more than 100 ° C for a water at normal pressure in the liquid state is a value outside the equilibrium condition.

The equilibrium states can also be *dynamic* or *static* equilibriums. For example, let us have a big company. Under stable economic conditions, the number of employees is constant. Some of them retire during the monitored period, some of them are recruited, and the overall number is constant. This is an equilibrium, the number of employees is not changing, and it is a dynamic equilibrium, because the persons employed in the enterprise are continuously exchanged. In some cases, a small business, under stable economic conditions, may have all the positions occupied properly. In this case, the company can operate for a period with the same staff without any change. This is a static equilibrium.

In physics (chemistry, biology), dynamic equilibrium is frequent. Acetic acid is a weak acid, with a dissociation constant $K_{HA}=1.75 \text{ mol L}^{-1}$ at 25 ° C. Only a small part of the acid molecules is dissociated into hydrogen cations and acetate anions. If we were able to follow the individual molecules, we would see that the dissociated ions from time-to-time associate and form a molecule of acetic acid, and other molecules of acetic acid dissociate to ions. However, the overall number of ions and acid molecules is constant.

In the equilibrium state, the properties of the system are constant in *space* and *time*. The properties of an equilibrium system have a constant *average value* in any part of the system, and these values do not change over time. There can exist local perturbations, i.e., the measured value of a property oscillates around its steady value. In a gaseous system, when the particles are colliding, at that moment and place of the collision, the actual density is higher, and after the collision, when the particles rebound from each other, it is lower than the average value.

13.13 Isochoric Processes

An isochoric process means a process at constant volume, where the ideal gas is separated from the environment with walls that are not mobile and no (volume) work is allowed (Figure 89). The exchange of energy, heat exchange is allowed. Applying the ideal-gas law (eq. 209):

$$\frac{p}{T} = \frac{nR}{V} = \text{const}$$

The work of an isochoric process is zero:

$$dV = 0 \quad \Rightarrow \quad dW = -pdV = 0 \quad \text{eq. 247}$$

The pressure-volume plot (Figure 90) is a straight line, parallel to the pressure axis (y). Work, by definition, is the integral of the curve $p = f(V)$, the area between the curve and the volume axis (x), which is in this case zero.

The heat absorbed by the system is completely transformed into internal energy:

$$\begin{aligned} dU &= dQ = TdS \\ dQ &= nc_{mv}dT \end{aligned} \quad \text{eq. 248}$$

$$\Delta Q = nc_{mv}(T_2 - T_1)$$

where c_{mv} is the isochoric molar heat capacity and n is the number of moles of the material. The entropy change of the system is calculated from the combination of eq. 227 and eq.248.

$$\begin{aligned} dS &= nc_{mv} \frac{dT}{T} \\ \Delta S &= nc_{mv} \ln \frac{T_2}{T_1} \end{aligned} \quad \text{eq. 249}$$

The internal energy of the ideal gas is as follows:

$$U = \frac{3}{2}nRT \quad \text{eq. 250}$$

Therefore, the isochoric molar heat capacity of the ideal gas is:

$$\begin{aligned} c_{mv} &= \frac{dQ}{n dT} = \frac{dU}{n dT} = \frac{3}{2}R \\ &\cong 12.5 \text{ JK}^{-1}\text{mol}^{-1} \end{aligned} \quad \text{eq. 251}$$

The amount of heat (Q), and in this case the internal energy change, is illustrated on the temperature – entropy (T-S) diagram (Figure 91).

$$T = Ae^{\frac{S}{nc_{mv}}} \quad \text{eq. 252}$$

A is calculated from the reference value of the entropy at 1K (see later). The graph is an exponential curve, and the area between the line and the x-axis is the heat exchange (Q) between the system and the environment.

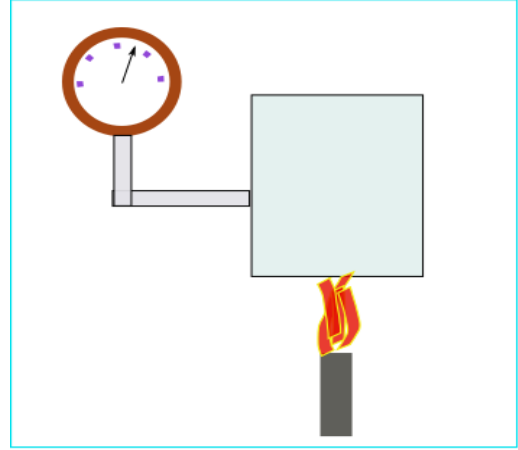


Figure 89. Isochoric system

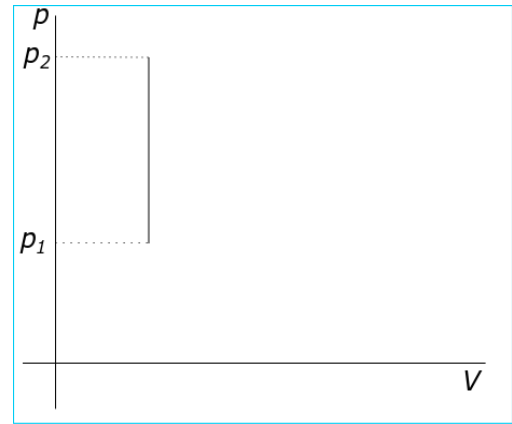


Figure 90. p-V diagram of isochoric system.

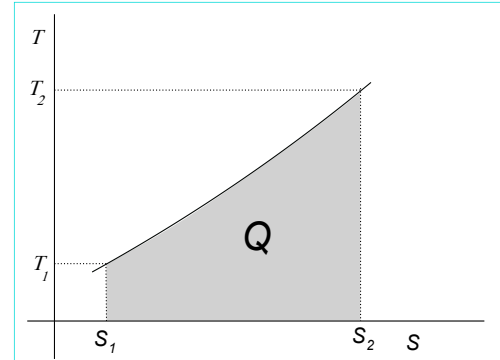


Figure 91. T-S diagram of isochoric system, the exchanged heat is the grey area.

Mathematical background.

$$S = \int \frac{dQ}{T} = \int nc_{mv} \frac{dT}{T} = nc_{mv} \ln T + \text{const}$$

If $T=1 \Rightarrow \ln T=0$

$$\begin{aligned} \text{const} = S_{1K} &\Rightarrow S = nc_{mv} \ln T + S_{1K} \\ T = e^{\frac{S-S_{1K}}{nc_{mv}}} &= e^{\frac{-S_{1K}}{nc_{mv}}} e^{\frac{S}{nc_{mv}}} = Ae^{\frac{S}{nc_{mv}}} \end{aligned}$$

13.14 Isobaric Processes

The isobaric condition means that the pressure is constant and equal to the pressure of the environment. Common physical processes are carried out usually under isobaric conditions.

$$p = \text{const}$$

From the ideal gas law, we have:

$$\frac{V}{T} = \frac{nR}{p} = \text{const} \quad \text{eq. 253}$$

$$dW = -pdV$$

$$W = - \int_{V_1}^{V_2} p dV = -p(V_2 - V_1) \quad \text{eq. 254}$$

Let us have a system of ideal gas in a container that has a piston that moves without any friction (Figure 92).

The pressure in the system is equal to the pressure in the surroundings, and it is constant. Heating the system increases the temperature and volume of the system. The heat added to the system is transformed into the internal energy of the system and partially into the work done by the system.

The p-V diagram is a line parallel with the volume (x) axis, and the work is the area below it, the area of a rectangle (Figure 93). The minus sign is, since pressure is always a positive number, the volume change is positive in the case of expansion when the system is doing the work, i.e., the work is negative. The volume change is negative, in the case of compression, when the work is performed on the system and is positive ($V_2 - V_1 > 0$, $W < 0$ or $V_2 - V_1 < 0$, $W > 0$).

The transferred heat to the system is calculated:

$$dQ = nc_{mp} dT \quad \text{eq. 255}$$

where c_{mp} is the isobaric molar heat capacity. The isobaric and isochoric molar heat capacities of the system are different (see later).

The entropy definition (and graph) is like the isochoric system graph but with the isobaric molar heat capacity.

$$dQ = TdS = nc_{mp} dT$$

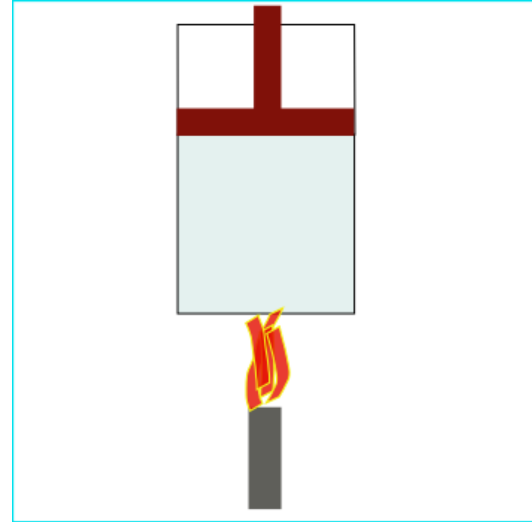


Figure 92. Isobaric system

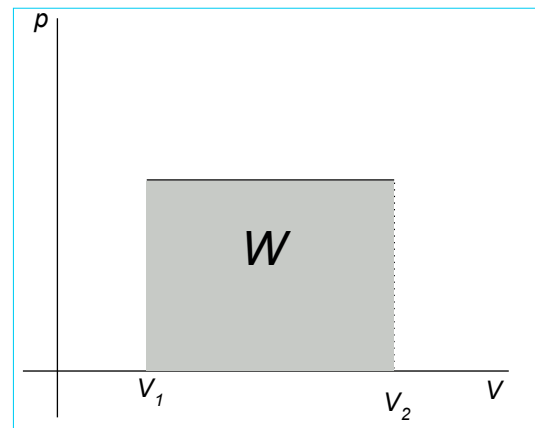


Figure 93. p-V diagram of isobaric system. Work is the grey area.

$$dS = nc_{mp} \frac{dT}{T}$$

$$\Delta S = nc_{mp} \ln \frac{T_2}{T_1}$$
eq. 256

And

$$T = Ae^{\frac{S}{nc_{mp}}}$$

Using the definition of internal energy, we can find the difference between the isobaric and isochoric heat capacities.

$$dU = dQ + dW$$

$$nc_{mv}dT = nc_{mp}dT - pdV = nc_{mp}dT - nRT \frac{dV}{V}$$

$$(c_{mp} - c_{mv}) \frac{dT}{T} = R \frac{dV}{V}$$

After integration in limits T_1 , T_2 , and V_1 , V_2 :

$$(c_{mp} - c_{mv}) \ln \frac{T_2}{T_1} = R \ln \frac{V_2}{V_1}$$

For an isobaric system, we can write:

$$\frac{V_1}{T_1} = \frac{V_2}{T_2} \quad \Rightarrow \quad \ln \frac{T_2}{T_1} = \ln \frac{V_2}{V_1}$$

and so

$$c_{mp} - c_{mv} = R$$
eq. 257

Eq.257 is known as the Mayer relation, which tells that the isobaric molar heat capacity of the ideal gas is greater than the isochoric molar heat capacity, and the difference is equal to the universal gas constant. Heat capacity is the amount of heat necessary to increase the temperature of the system (of a given mass or amount of moles) by one degree. In the case of an isochoric process, the consumed heat contributes only to the increase of the internal energy, which for the ideal gas only relates to an increase of the temperature. However, in the case of isobaric processes, the consumed heat partially contributes to the increase in internal energy (temperature) and partially is used to perform a work. Because of this, more energy is required to increase the temperature in the same measure in isobaric conditions than in isochoric conditions.

Isobaric processes are often characterised by the enthalpy of the system. Taking the first law of thermodynamics at constant pressure, the change of the internal energy is as follows:

$$\Delta U = Q - p\Delta V$$

$$Q = U_2 - U_1 + p(V_2 - V_1) = (U_2 + pV_2) - (U_1 + pV_1) = H_2 - H_1$$

The received heat changed the value of the enthalpy of the system:

$$H = U + pV$$
eq. 258

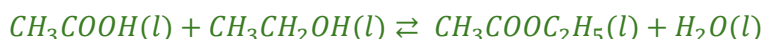
Chemical reactions usually run under isobaric conditions, so the process can be characterised with a change in the enthalpy. If heat has evolved and released to the environment, the reaction is exothermic (the system loses heat), and the enthalpy change is *negative*. If the system consumes heat (receiving heat) during the reaction (endothermic reaction), the enthalpy change is *positive*.

$$\Delta H < 0 \quad \text{exothermic reaction}$$

$$\Delta H > 0 \quad \text{endothermic reaction}$$

Because enthalpy is a state function, the reaction heat is not affected with the reaction path, only with the states of the system before and after the reaction, so when the enthalpy of a reaction path is unknown, it is possible to calculate it using a different reaction path with known enthalpy.

Problem. Standard molar enthalpy (in standard reaction conditions) of the esterification reaction of acetic acid and ethanol cannot be measured directly because the reaction is too slow.



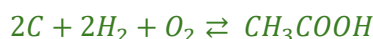
Find the standard reaction enthalpy if the standard enthalpies of formation are as follows:

$$\begin{aligned}\Delta H^0(\text{CH}_3\text{COOH}) &= -484.2 \text{ kJ mol}^{-1} \\ \Delta H^0(\text{CH}_3\text{CH}_2\text{OH}) &= -277.0 \text{ kJ mol}^{-1} \\ \Delta H^0(\text{CH}_3\text{COOC}_2\text{H}_5) &= -486.6 \text{ kJ mol}^{-1} \\ \Delta H^0(\text{H}_2\text{O}) &= -285.7 \text{ kJ mol}^{-1}\end{aligned}$$

Solution: The model reaction path, which can be used in a theoretical reaction, is the decomposition of the reactants into pure substances (molecules of one chemical element) and the formation of the products from them. The standard enthalpies of formation of elements (molecules which contain one type of atom) are zero!



Standard enthalpies of reactions are the differences between products and reactants enthalpies. For example, the $\Delta H^0(\text{CH}_3\text{COOH})$ is a standard enthalpy change of the reaction:



The enthalpy change of esterification is calculated in the following way:

$$\begin{aligned}\Delta H^0(\text{reaction}) &= -\Delta H^0(\text{CH}_3\text{COOH}) \\ &\quad - \Delta H^0(\text{CH}_3\text{CH}_2\text{OH}) \\ &\quad + \Delta H^0(\text{CH}_3\text{COOC}_2\text{H}_5) + \Delta H^0(\text{H}_2\text{O}) \\ \Delta H^0(\text{reaction}) &= -(-484.2 - 277.0) - 486.6 \\ &\quad - 285.7 = -11.1 \text{ kJ mol}^{-1}\end{aligned}$$

The enthalpy change is negative, the reaction is exothermic, heat has evolved and released into the environment during the reaction. The opposite reaction (decomposition of the ester) has the same enthalpy change with opposite sign (+11.1 kJ mol⁻¹). Heat is consumed by the system during the reaction, it is an endothermic reaction.

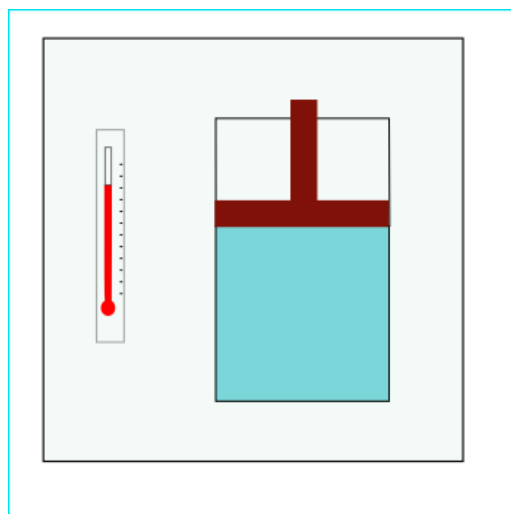


Figure 94. Isothermal system.

13.15 Isothermal (Isothermic) Processes

The next type of process is the isothermal process, which is defined with a constant temperature ($T=\text{const}$). It is the case when the system is in thermal equilibrium with its environment of constant temperature (Figure 94). Because the internal energy of an ideal gas depends on the temperature and the number of particles, in a closed isothermal system the internal energy is *constant*. Heat consumed/released by the system from/to the environment is transformed into equivalent volume work.

$$dU = 0 \quad dQ = -dW$$

The p-V diagram is a hyperbola (Figure 95):

$$p = \frac{nRT}{V} = \frac{\text{const}}{V}$$

The integral below the $p=f(V)$ curve is the work:

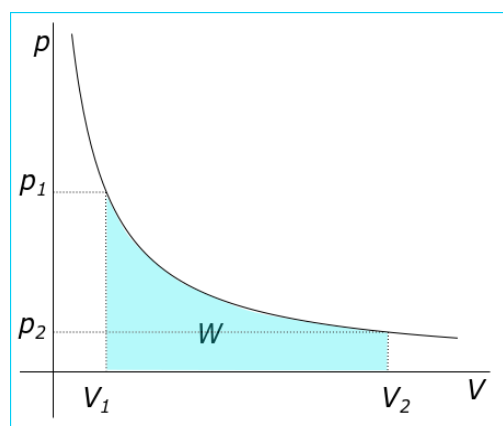


Figure 95. p-V diagram of isothermal process with the reversible work – cyan coloured area.

$$W = - \int p dV = -nRT \int \frac{dV}{V} \quad \text{eq. 259}$$

$$W = nRT \ln \frac{V_1}{V_2} \quad \text{eq. 260}$$

$$Q = nRT \ln \frac{V_2}{V_1}$$

The heat is determined directly from the work with opposite sign (the numerator and denominator of the logarithm are exchanged).

13.16 Irreversible processes

What is the difference if the process is reversible or irreversible? Work is a path function, not a state function, so its value depends on the path along which the system proceeds from one state to another. In the case of a reversible process, the paths of direct and reversed processes continue along the same path in opposite directions, but in the case of irreversible processes, the two paths are different. Let us consider the previous example of isothermal expansion/compression. The path of the irreversible process is not along the isotherm (Figure 95), but in the case of the expansion (Figure 96, from state 1 to 2) proceeds along the green line, it means that the pressure drops from the high p_1 value to the low p_2 at volume V_1 , and the volume expansion (from V_1 to V_2) follows it latter. The work is the area between the green line and the volume axis.

$$W_{1-2} = -p_2(V_2 - V_1) \quad \text{eq. 261}$$

On the other hand, the irreversible compression is done along the red line, the pressure is rapidly increased from p_2 to p_1 at the V_2 volume, and the volume compression follows that.

$$W_{2-1} = -p_1(V_1 - V_2) \quad \text{eq. 262}$$

The difference of the work in absolute values for this case is:

$$|W_{1-2}| - |W_{2-1}| = (p_1 - p_2)(V_2 - V_1)$$

As the two areas differ significantly, so do the two works, too. The work done on the system (compression) is bigger than the work done by the system (expansion) (in absolute values). To increase the work received from the system and reduce the work done on the system, it is possible to achieve this by changing the pressure in more steps (like stairs). In the limit case, when the pressure changes in very small steps, the process becomes reversible, and the work done by the system and on the system will be the same. **This means that reversible work is the maximum work a system can do during isothermal compression, and reversible work is the minimum work necessary to perform on the system to transfer it between states.**

The entropy of an isothermal process (Figure 97) is calculated from the definition formula and eq. 260:

$$dQ = TdS \quad Q = T(S_2 - S_1) \quad \text{eq. 263}$$

$$\text{eq. 264}$$

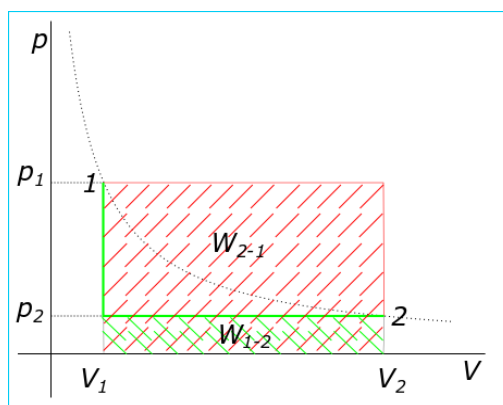


Figure 96. Isothermal process and irreversible work of isothermal expansion – green rectangle and irreversible compression – red rectangle.

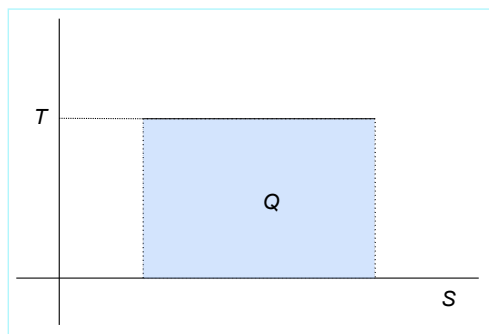


Figure 97. T-S diagram of isothermal process, the light blue area is the heat exchanged during the process.

$$\Delta S = nR \ln \frac{V_2}{V_1}$$

In the case of isothermal processes, the product of volume and pressure is constant:

$$p_1 V_1 = p_2 V_2 \quad \Rightarrow \quad \frac{V_2}{V_1} = \frac{p_1}{p_2}$$

So, the work, heat, and entropy are:

$$\begin{aligned} W &= nRT \ln \frac{p_2}{p_1} \\ Q &= nRT \ln \frac{p_1}{p_2} \\ \Delta S &= nR \ln \frac{p_1}{p_2} \end{aligned} \quad \text{eq. 265}$$

Heat capacities (specific, molar, etc.) are not defined (infinite) for the isothermal process because despite the heat input/output to/from the system, there is no temperature change.

$$dQ = c dT \quad dQ \neq 0 \quad \wedge \quad dT = 0 \quad \Rightarrow \quad c = \infty \quad \text{eq. 266}$$

13.17 Adiabatic Processes

When a system is perfectly insulated or the process is too fast, no energy exchange is possible between the system and the environment; the process is running under adiabatic conditions. This means that the heat exchange between the system and the surrounding environment is zero ($Q = 0$), and the change of internal energy is equal to the work done on or by the system. When the work is done by the system (expansion), its internal energy is reduced, and when it is done on the system (compression), it increases.

$$\begin{aligned} dU &= dW = -pdV \\ nc_{mv}dT &= -\frac{nRT}{V}dV \end{aligned}$$

After rearranging and integrating,

$$\begin{aligned} c_{mv} \frac{dT}{T} &= -R \frac{dV}{V} \quad / \int_1^2 \\ c_{mv} \ln \frac{T_2}{T_1} &= -R \ln \frac{V_2}{V_1} \\ \left(\frac{T_2}{T_1} \right)^{c_{mv}} &= \left(\frac{V_1}{V_2} \right)^R \end{aligned}$$

From the Gay-Lussac laws:

$$\begin{aligned} \frac{T_2}{T_1} &= \frac{p_2 V_2}{p_1 V_1} \\ \left(\frac{p_2 V_2}{p_1 V_1} \right)^{c_{mv}} &= \left(\frac{V_1}{V_2} \right)^R \\ (p_1 V_1)^{c_{mv}} V_1^R &= (p_2 V_2)^{c_{mv}} V_2^R \\ R &= c_{mp} - c_{mv} \\ p_1^{c_{mv}} V_1^{c_{mp}} &= p_2^{c_{mv}} V_2^{c_{mp}} \quad / \quad \sqrt[c_{mv}]{} \\ pV^\gamma &= \text{const} \quad \text{where} \quad \gamma = \frac{c_{mp}}{c_{mv}} \end{aligned} \quad \text{eq. 267}$$

γ is the ratio of the isobaric ($5/2 R$) and isochoric ($3/2 R$) molar heat capacity. For the ideal gas γ is $5/3 (=1.667)$.

The adiabatic work is possible to calculate from the basic definition formulas:

$$\begin{aligned}dW &= -pdV \\ pV^\gamma &= K \\ pV &= nRT\end{aligned}$$

$$\begin{aligned}W &= - \int_{V_1}^{V_2} p dV = - \int_{V_1}^{V_2} \frac{K}{V^\gamma} dV = -K \left[\frac{V^{1-\gamma}}{1-\gamma} \right]_{V_1}^{V_2} = \frac{1}{\gamma-1} [KV_2^{1-\gamma} - KV_1^{1-\gamma}] \\ W &= \frac{1}{\gamma-1} [p_2 V_2^\gamma V_2^{1-\gamma} - p_1 V_1^\gamma V_1^{1-\gamma}] = \frac{1}{\gamma-1} [p_2 V_2 - p_1 V_1] = \frac{nR}{\gamma-1} [T_2 - T_1]\end{aligned}$$

Substituting for R and γ :

$$W = \frac{n(c_{mp} - c_{mv})}{\frac{c_{mp}}{c_{mv}} - 1} [T_2 - T_1] = nc_{mv}[T_2 - T_1] = \Delta U \quad \text{eq. 268}$$

We see that really; the work equals the change of the internal energy. This is not surprising because heat exchange between the system and the environment is impossible; therefore, the work must be done only on account of the internal energy only. An adiabatic expansion also means a decrease in internal energy and a decrease in the system temperature. Adiabatic compression means an increase in the internal energy, and so the temperature of the ideal gas. The change in entropy is zero, because $dQ=0$. The TS plot (Figure 98) is a line parallel to the temperature axis, and the integral below it is zero, as well as the change in entropy of an adiabatic process.

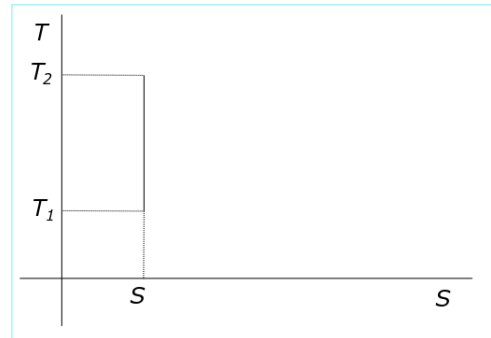


Figure 98. T-S diagram of adiabatic process.

13.18 Efficacy of the Steam Engine and the Second Law of Thermodynamics

Machines have been known for a long time, but their use was limited due to restrictions arising from the energy sources. These were usually based on human, animal, or on available natural phenomena, such as wind, water, etc. In the 16th and 17th centuries, the first usages of steam engines were recorded, and in the 19th century the utilising of machines on an industrial scale. The changes provoked by the construction of the steam machines caused an industrial revolution, with a huge increase of the goods production, making it available to a wider range of population. Today we see enormous environmental problems, which have resulted (also) from this industrial revolution; however, in the 19th century it was seen only as a progress and scientists dealt with the steam engine from a theoretical point of view, also. One of them was Sadi Nicolas Leonard Carnot (1796-1832), who calculated the efficacy of a theoretical steam engine.

An engine of practical importance must work in a cyclical way. The Carnot cycle (Figure 99) consists of isothermal expansion, adiabatic expansion, isothermal compression, and adiabatic compression of an ideal gas. The engine receives energy from the heat source with temperature T_1 , which is followed by isothermal expansion (steps 1-2).

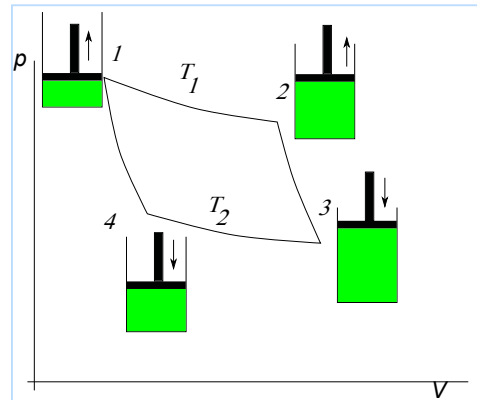


Figure 99. p-V diagram of Carnot's cycle.

In the next step, the expansion continues as an adiabatic process (2-3), without any additional heat income, so it is done on account of the internal energy of the system, resulting in a decrease in the temperature to T_2 . The next step is isothermal compression (3-4), when the work done on the system is removed in the form of heat, which is dispersed to the environment. The last step is adiabatic compression (4-1) when the system is compressed to its initial volume and its temperature increases to the initial T_1 .

In the case of the Carnot cycle, the energy input is the heating during the isothermal expansion (steps 1-2, red area on the T-S plot, Figure 100) and the work input is during the adiabatic and isothermal compressions (steps 3-4, 4-1, p-V plot, Figure 99), respectively. Work outputs are during the isothermal expansion (1-2) and adiabatic expansion (2-3) and heat output during isothermal compression (3-4). The output work of the adiabatic expansion is of the same magnitude as the work of the adiabatic compression with opposite sign, so they annul each other. The heat output during isothermal compression is useless because this heat is dispersed in the environment. The net useful output work is the difference between the isothermal expansion and compression. Changes in state functions (internal energy and entropy) are zero during the entire process because of its cyclic character. The efficacy is the ratio of the sum of works over the heat input with negative sign (due to the sign convention of the work done on or by the system), dispersed heat is not included:

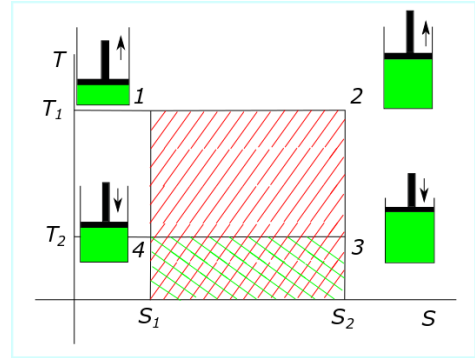


Figure 100. T-S diagram of the Carnot's cycle. Red rectangle is the input heat, green is the released (dispersed) by the system.

The net useful output work is the difference between the isothermal expansion and compression. Changes in state functions (internal energy and entropy) are zero during the entire process because of its cyclic character. The efficacy is the ratio of the sum of works over the heat input with negative sign (due to the sign convention of the work done on or by the system), dispersed heat is not included:

$$\eta = \frac{-\sum_{i=1}^4 W_i}{Q_{1-2}}$$

$$\eta = - \frac{nRT_1 \ln \frac{V_2}{V_1} - nc_{mv}(T_2 - T_1) + nRT_2 \ln \frac{V_4}{V_3} - nc_{mv}(T_1 - T_2)}{nRT_1 \ln \frac{V_2}{V_1}}$$

$$\eta = - \frac{T_1 \ln \frac{V_2}{V_1} + T_2 \ln \frac{V_4}{V_3}}{T_1 \ln \frac{V_2}{V_1}}$$

From the condition that the entropy change (state function) in a cyclic process is zero (Table 7):

$$\sum_{i=1}^4 S_i = 0 = nR \ln \frac{V_2}{V_1} + nR \ln \frac{V_4}{V_3}$$

$$\ln \frac{V_4}{V_3} = -\ln \frac{V_2}{V_1}$$

$$\eta = - \frac{T_1 \ln \frac{V_2}{V_1} - T_2 \ln \frac{V_2}{V_1}}{T_1 \ln \frac{V_2}{V_1}} = 1 - \frac{T_2}{T_1} \quad \text{eq. 269}$$

Temperature T_1 is higher than T_2 , so the efficacy is between 0 and 1 (in percentage 0-100%). The higher the difference between the temperature of the heater and the cooler, the higher the efficiency. It is very important that the final formula does not contain any material characteristics of the engine, and

due to that it has general validity and importance. The whole cycle consists of reversible processes, and, as before we showed in the case of an isothermal process, the reversible process provides the maximum work gained from the system and the minimum work, if consumed. The efficiency of the Carnot cycle is the maximum efficiency of the heat engine. This maximum efficacy cannot be 1 (or 100%), because T_2/T_1 is never zero. Due to that, it is impossible to create a heat engine that works with 100% effectivity.

Table 7. Processes and thermodynamic functions of the Carnot cycle.

Path	Type	$\Delta U=$	$\Delta W=$	$\Delta Q=$	$\Delta S=$
1-2	Isothermal expansion	0	$-nRT_1 \ln \frac{V_2}{V_1}$	$nRT_1 \ln \frac{V_2}{V_1}$	$nR \ln \frac{V_2}{V_1}$
2-3	Adiabatic expansion	$nc_{mv}(T_2 - T_1)$	$nc_{mv}(T_2 - T_1)$	0	0
3-4	Isothermal compression	0	$-nRT_2 \ln \frac{V_4}{V_3}$	$nRT_2 \ln \frac{V_4}{V_3}$ dispersed	$nR \ln \frac{V_4}{V_3}$
4-1	Adiabatic compression	$nc_{mv}(T_1 - T_2)$	$nc_{mv}(T_1 - T_2)$	0	0
Σ	Cyclic process	0	$\neq 0$	$\neq 0$	0

This finding arising from the Carnot cycle is the basis of the second law of thermodynamics, which was formulated in several ways and by several scientists; however, they are equivalent, and the meanings are the same:

*It is impossible to create a **periodic** working heat engine that extracts heat and converts it **all** into useful work. Part of the input energy is transferred to the cold reservoir, and it is a waste of energy.* (W. Thompson-lord Kelvin, M. Planck)

*Heat will not flow **spontaneously** from a cold object to a hot object.* (R. Clausius)

The first law of thermodynamics talks about the equivalence of different types of energy (work, heat, etc.) and does not exclude the construction of a machine with hundred percent effectivity; it is the second law of thermodynamics which forbade it.

The **entropy** of a physical system is proportional to *the amount of energy no longer available to do physical work* (the energy dissipated to the environment during isothermal compression). Entropy plays a central role in the second law of thermodynamics, which states that in an isolated system, any activity increases the entropy.

The second law of thermodynamics is expressed with the change of entropy as well.

*In an **isolated** system, natural processes are **spontaneous** when they lead to an increase in entropy.*

$$\Delta S \geq 0 \quad dS \geq 0 \quad \text{eq. 270}$$

$$dS \geq \frac{dQ}{T}$$

The equal sign represents the reversible processes, whereas the greater sign represents the irreversible (real) processes. In the equilibrium state of the system, the entropy has a maximum.

$$S = S_{max} \quad dS = 0 \quad \text{eq. 271}$$

Problem. Find the change of entropy of an ideal gas of temperature 20 ° C pressure 0.1 MPa and volume 2 litres if it expands reversible (ΔS_{rev}) and irreversible (ΔS_{irev}) isothermally to twice its initial volume!

Solution. The process is isothermal and from the First Law of Thermodynamic (eq. 224) follows:

$$\begin{aligned}
dQ &= -dW && \text{because} && dU = 0 \\
dS_{rev} &= \frac{dQ}{T} = \frac{-dW}{T} = p \frac{dV}{T} = \frac{nRT}{V} \frac{dV}{T} \\
&&& nRT = pV_1 \\
\Delta S_{rev} &= \frac{pV_1}{T} \ln \frac{V_2}{V_1} = \frac{0.1 \cdot 10^6 \cdot 2 \cdot 10^{-3}}{293.15} \ln 2 \doteq 0.473 \text{ J K}^{-1} \\
Q_{rev} &= T \Delta S_{rev} \doteq 138.63 \text{ J}
\end{aligned}$$

The entropy change of the reversible isothermal expansion was 0.473 J K^{-1} and the heat received by the system was 138.63 J . In the case of expansion of gases (work done by the system), their temperature generally decreases. To keep the temperature constant, heat must be supplied to the system. In the case of irreversible expansion, we can choose a path, when the pressure suddenly drops to half of the initial pressure,

$$p_2 = \frac{V_1}{V_2} p_1 = \frac{1}{2} p_1$$

and subsequently doubles the volume as shown on Figure 96 and the heat (work) is calculated according to eq. 261.

$$Q_{irev} = -W = p_2(V_2 - V_1) = 0.05 \cdot 10^6 \cdot (4 \cdot 10^{-3} - 2 \cdot 10^{-3}) = 100 \text{ J}$$

The entropy change for the irreversible process is (eq. 270):

$$\Delta S_{irev} > \frac{Q_{irev}}{T} > \frac{100}{293.15} > 0.341 \text{ J K}^{-1}$$

The entropy change of the irreversible expansion is greater than 0.341 J K^{-1} . However, what is the actual value of the entropy change? We must consider that entropy is a state function, so its change does not depend on the path of the process; the entropy changes of the irreversible and reversible path **equal**, the entropy change of the irreversible expansion is also 0.473 J K^{-1} , and this value is, in fact, bigger than 0.341 J K^{-1} .

Let us reverse the direction of the process to isothermal compression. Compression of gases (the work done on the system) produces heat, which must be subtracted from the system to keep its temperature constant. In the case of the reversible process, the entropy change is equal to that of the reversible compression with opposite sign; -0.473 J K^{-1} . The irreversible heat release, when the pressure increases suddenly from p_2 to p_1 and the volume changes from V_2 to V_1 at p_1 , is:

$$\begin{aligned}
Q_{irev} &= -W = p_1(V_1 - V_2) = 0.1 \cdot 10^6 \cdot (2 \cdot 10^{-3} - 4 \cdot 10^{-3}) = -200 \text{ J} \\
\Delta S_{irev} &> \frac{Q_{irev}}{T} > \frac{-200}{293.15} > -0.682 \text{ J K}^{-1}
\end{aligned}$$

This value is again the lower limit of this process, and the actual value is that calculated from the reversible one; -0.473 J K^{-1} is bigger than -0.682 J K^{-1} .

13.19 Third Law of Thermodynamics

In most cases, we do not calculate the absolute value of the entropy but its change, the difference of its values of two different states. It is possible even in the case that the absolute values are not known in any of the considered states. If we know the absolute value of the entropy of a system at an arbitrary point, then it is possible to find it (theoretically) in any other state of the system. The reference point can be a perfect (ideal) crystal state of a material at 0 K. This state is a pure theoretical state; it is impossible to reach. The perfect crystal state means that the crystal is free from defects, and all atoms are perfectly at their sites. The 0 K temperature means that all motions are stopped. There is only one possible configuration of that system; there is only one microstate to provide that macrostate.

Ludwig Boltzmann (1844-1906) defined the entropy using microstates.

$$S = k_B \ln w \quad \text{eq. 272}$$

k_B is the Boltzmann constant ($=1.380649 \cdot 10^{-23} \text{ J K}^{-1}$), w is the number of microstates. Substituting $w = 1$, we get the value zero.

Third Law of Thermodynamics: The entropy of a perfect crystal of a pure compound at absolute zero (0K) is zero.

It is impossible to reach 0 K with a finite number of steps.

$$\lim_{T \rightarrow 0} S = 0 \quad \text{eq. 273}$$

It is impossible to reach 0 K, but many attempts have been made to approach this temperature as closely as possible. Certain properties of materials become different and interesting in comparison to high temperatures. For example, electrical resistance of certain materials vanishes, and magnetic flux fields are expelled, i.e., become superconductive. Superconductive materials are important for scientific, technological, and economic reasons, because, for example, electrical resistance causes significant waste worldwide in energy transport.

13.20 Spontaneity of Processes

When we drop a few crystals of salt (NaCl) into water, it dissolves immediately and will never spontaneously crystallise. To crystallize them, we must evaporate the solvent, i.e. to add a significant amount of heat to the system. When two containers, filled with different gases, are connected, they do mix and will not separate from each other. Everyone knows the fairy tale about Cinderella and her bad stepmother. When she wanted to go to the ball, her stepmother mixed the lentils with the ash and ordered her to separate them. Mixing the lentils with the ash only took a little effort and a while. However, separating the lentils from the ash was a task that should hinder Cinderella for hours to arrive on time. Why is it possible to mix things with so little effort and why is it so difficult and sometimes even impossible to separate them? What is the driving force of these processes? In all cases, the entropy was increasing. Separated crystalline salt and solvent, separated gases, lentils and ash have lower entropy compared to solution, mixed gases, and mixed lentils and ash. Spontaneous processes proceed in the direction of increasing entropy, and in equilibrium, the entropy has a maximum. Entropy is not the only function that affects the direction of the process flow. The next part of the equilibrium definition tells us that the Gibbs free energy (Helmholtz energy, internal energy) has a minimum value. When a surface-active material (surfactant or soap) is dissolved, the process is more complex, as in the case of NaCl. This material contains a hydrophilic group that is well soluble in water and a hydrophobic group (long aliphatic chain) that is soluble in oil but is little soluble in water. At low concentrations (below the so-called critical micelle concentration, CMC), the surfactant is dispersed uniformly in water and forms a true solution. Above the critical micelle concentration, the surfactant molecules begin to form aggregates and separate from the water on a microscopic scale. This separation is a spontaneous process but means a decrease in entropy. The decrease in entropy must be compensated for with a larger benefit of the system. It is the Gibbs free energy:

$$G = H - TS \quad \text{eq. 274}$$

$$\Delta G = \Delta H - T\Delta S < 0 \quad \text{eq. 275}$$

Therefore, it is evident that the decrease in the entropy must be compensated with a decrease in the enthalpy to have a decrease in the Gibbs energy and a spontaneous process. Let us look at the processes that are present when solving a surfactant in water. The dissolution of the molecules results in a more uniform system, i.e., an increase in entropy. There are interactions between the hydrophilic group of the surfactant and water molecules, an energetically favourable process, and an exothermic

process. These two processes support the dissolution. There is the hydration of the hydrophobic group; this is an energetically unfavourable endothermic process that acts against the solution. Solving or not solving is the question of proportions of these processes. When the concentration exceeds the critical micelle concentration, the increase in gain from the entropy and the hydrophilic group hydration cannot compensate for the energy need of the hydrophobic groups' hydration, the result is the formation of aggregates, where the hydrophobic parts of the molecules are oriented inside the micelles, and the hydrophilic parts are orientated toward the water molecules.

From this and from the definition of the Gibbs free energy, it is clear which parameters of the system, and the properties of the molecules affect the critical micelle concentration. Increasing the temperature increases the importance of the entropy part of the Gibbs energy and increases the CMC. A longer alkyl chain in the surfactant molecule means less favourable hydration of the sole surfactant molecule, i.e., decrease in CMC.

Spontaneous processes are running in the direction of decreasing of Gibbs free energy. A process can be exothermic or endothermic from the point of view of the enthalpy and with increasing or decreasing entropy.

Table 8. Spontaneity of processes depending on the changes of enthalpy and entropy.

Enthalpy change	Entropy change	The change of Gibbs free energy	The process is spontaneous
+	+	depends on T, may be + or -	if the temperature is high enough $T\Delta S > \Delta H$
-	+	is always negative	always
-	-	depends on T, may be + or -	if the temperature is low enough $ \Delta H > T\Delta S $
+	-	is always positive	never

In Table 8, the possible combinations of enthalpy and entropy changes of the processes, which decide whether a process is always or never spontaneous, and processes whose spontaneity is possible to control with temperature, which are spontaneous at low or high temperatures, are listed. In the case of processes, where the spontaneity depends on temperature, the enthalpy term (at low temperatures, exotherm, enthalpy-governed process) prevails over the entropy term, or reversed (at high temperatures, endotherm, entropy-governed process) (Figure 101).

Problem. The ice melts at 273.15 K (0 °C) The melting molar heat is 6 kJ/mol. Calculate the change in the Gibbs free energy of melting at 265 K and 280 K.

Solution:



$$\Delta H_m = H_m(l) - H_m(s) = 6 \text{ kJ/mol}$$

In equilibrium, T = 273.15K and $\Delta G=0$; the change in the molar melting entropy is as follows:

$$\Delta S_m = \frac{\Delta H_m}{T} = \frac{6 \cdot 10^3}{273.15} = 22 \text{ J/molK}$$

The molar Gibbs energy change of melting at T=265 K is:

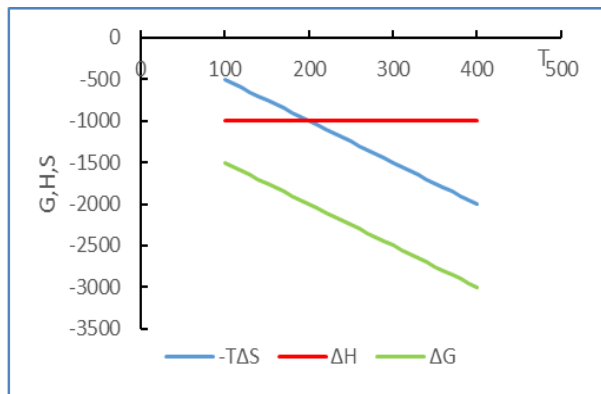
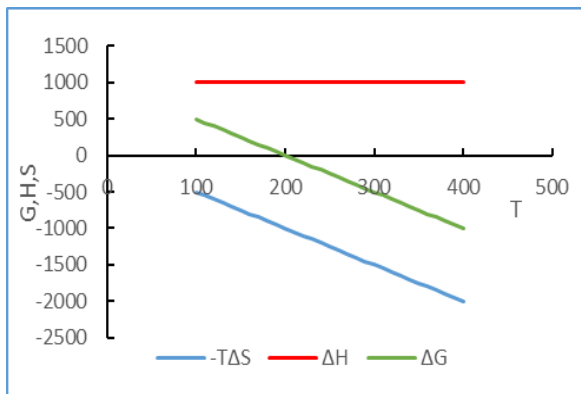
$$\Delta G_m = \Delta H_m - T\Delta S_m = 6 \cdot 10^3 - 265 \cdot 22 = +170 \text{ J/mol}$$

The value is positive, so the melting of ice at 265 K (-8.15 °C) is not a spontaneous process.

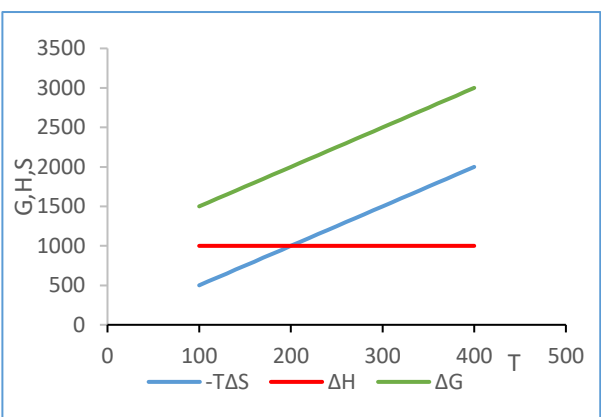
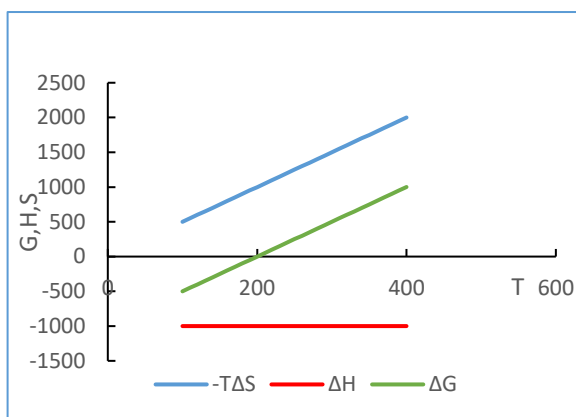
At T=280 K:

$$\Delta G_m = \Delta H_m - T\Delta S_m = 6 \cdot 10^3 - 280 \cdot 22 = -160 \text{ J/mol}$$

The molar Gibbs energy change is negative, that is, the melting of ice is spontaneous, of course. The process is endothermic (heat is consumed), but the increase in entropy prevails at 280 K, the process is controlled by entropy, while at 265 K, the enthalpy prevails, and the enthalpy controls the process.



$\Delta H > 0$ $\Delta S > 0$ (+ +) spontaneous at high T $H < 0$ $\Delta S > 0$ (– +) always spontaneous.



$\Delta H < 0$ $\Delta S < 0$ (– –) spontaneous at low T, $\Delta H < 0$ $\Delta S > 0$ (– +) never spontaneous

Figure 101. Theoretical graphs of the changes in Gibbs free energy, enthalpy, and entropy (in arbitrary units). A process is spontaneous when ΔG (green line) is negative.

13.21 Heat Transfer and Diffusion

Spontaneous heat transfer always proceeds from a place with higher temperature in a direction of the temperature decrease to eliminate the temperature difference. For example, when we put a cup of hot coffee on a table, after a certain time, the coffee cools down to room temperature and the room temperature will increase slightly. This is the case when the system (room and coffee together) behaves as an adiabatic system, and we neglect the temperature exchange between the room and the surrounding area. The heat exchange between the coffee and the room is the most intensive at the beginning of the process, when the difference between their temperature is the largest and the least and stops in the moment when their temperatures become equal. We can find another example where the heat transfer is constant in time, and the heat transfer continues practically for ever. This is the case of a planet that is orbiting around its star having one side toward the star (gravitational locking, spin-orbit locking). In this case, the temperatures of the two objects have been constant for an extremely long time and the heat transfer is constant for this time. The temperatures of the star and the planet will not be equal, because the star compensates for its heat lost with a thermonuclear reaction; on the other hand, the planet dissipates to Space the same quantity of heat as it receives from the star. The previous case (coffee – room) was an example of a static equilibrium, when after a certain time no

process occurs in the system; the second one was an example of a dynamic equilibrium, when the star releases the same amount of heat as it is producing, the planet is dissipating the same amount of heat as it is receiving from its sun, resulting in constant temperatures of the Sun and the planet and constant heat flux between them.

Heat transfer is done by three mechanisms, conduction, convection, and radiation. In a real process, depending on the conditions, one, two, or all three types are present.

Conduction is provided with microscopic motion of atoms of the material. The heat is transferred with rapid motion, vibrations from the source through a conductor. In this type of mechanism, an object is required to facilitate heat transfer. The heat flow (heat transfer rate) is proportional to the heat transfer area (S) and temperature gradient ($\frac{dT}{dx}$).

$$\frac{dQ}{dt} = -\lambda S \frac{dT}{dx} \quad [W, Js^{-1}] \quad \text{eq. 276}$$

Where λ is the heat transfer coefficient [$Wm^{-1}K^{-1}$]. The heat flux (q , units, Wm^{-2}) is defined with Fourier's law, as the heat flow through a unit area:

$$q = \frac{1}{S} \frac{dQ}{dt} = -\lambda \frac{dT}{dx} \quad \text{eq. 277}$$

The negative sign on the right-hand side is because the heat flux (flow) proceeds in the direction against the temperature gradient (decrease of the temperature along the direction x).

Diffusion. Diffusion is the transport of fluids due to atomic motion. When we put a spoon of sugar in our hot coffee, the sugar dissolves but mostly remains on the bottom of the cup in the first moment. We used to mix it to have a uniform sweet taste in the whole volume of coffee, but after a certain time the sugar will also be dispersed uniformly without any mixing because of diffusion. Diffusion also relates to heat transport indirectly, as much as the transported material takes its thermal state. The reason for discussing the diffusion here is that the mathematical formalism of the diffusion is the same as the conduction. We can define diffusive flux (J , units $mol m^{-2}s^{-1}$), Fick's first law, similarly, to heat flux:

$$J = \frac{1}{S} \frac{dn}{dt} = -D \frac{dc}{dx} \quad \text{eq. 278}$$

$$n = -DS\Delta t \frac{\Delta c}{\Delta x}$$

t is time, n is the number of moles of the transported material, D is the temperature-dependent diffusion coefficient, and c is the concentration at a certain place of the solution. The equation is valid (similarly to Fourier's law of heat flux) for the case of steady diffusion, when the conditions are stable (temperature, concentration gradient, etc.).

The case when the concentrations are not constant in time is described with the Fick's second law:

$$\frac{\partial c(x, t)}{\partial t} = -\frac{\partial J}{\partial x} = D \frac{\partial^2 c(x, t)}{\partial x^2} \quad \text{eq. 279}$$

The equation has the same form as the equation for the conduction of heat:

$$\frac{\partial T}{\partial t} = a \frac{\partial^2 T}{\partial x^2} \quad \text{eq. 280}$$

The coefficient a has a different meaning; of course, it is defined as

$$a = \frac{\lambda}{c_p \rho} \quad \text{eq. 281}$$

where c_p is the specific heat capacity and ρ is the density.

Both equations are second-order partial differential equations, whose solutions depend on the initial and boundary conditions. An example of a solution of both diffusion and heat conduction is:

$$c(x, t) = \frac{\alpha}{\sqrt{t}} e^{-\frac{x^2}{4Dt}} \quad T(x, t) = \frac{\alpha}{\sqrt{t}} e^{-\frac{x^2}{4at}} \quad \text{eq. 282}$$

where α is a constant, which depends on the conditions.

Convection is a type of heat transfer in which a flowing material is required, at a different temperature as the environment, to transfer heat. The heat transfer rate depends on the mass transfer rate, so hydrodynamic equations are solved to find the amount of the transported heat. A simple example of convection is a hairdryer, which generates a hot air stream. In this case, the heat-carrying medium is air.

Radiation is the third type of heat transfer mechanism. In this type, no carrying medium is necessary for heat transport. A typical example is the sunshine. An essential role in providing appropriate conditions for life on the planet Earth plays the Sun and the energy coming from it. Between Earth and Sun there is the density of gases and dust so low that we can consider it practically as a vacuum, so no conduction or convection is possible. The only way to transport heat is by radiation. Radiation is basically done by electromagnetic waves, which are fluxes of photons. Their speed is the speed of light ca. 300 000 km/s, so the approximate Sun-Earth distance of 150 million km is done in more than eight minutes.

The heat flow density (q) of a radiation is expressed with the *Stefan-Boltzmann equation*, which describes the heat flow density radiated from a blackbody in dependence on its temperature. A blackbody is an idealised object that absorbs or emits all frequencies of the spectrum. The radiation power is proportional to the fourth power of the temperature.

$$q = \frac{1}{S} \frac{dQ}{dt} = \varepsilon \sigma T^4 \quad \text{eq. 283}$$

Where ε is the emissivity of the object, a perfect blackbody has value 1, and real objects have emissivity between 0 and 1. $\sigma = 5.67 \times 10^{-8} \text{ W/m}^2\text{K}^4$ is the Stefan-Boltzmann constant.

The maximum intensity is at wavelength $\lambda_{\max}$ (in m), which is called the Wien law:

$$\lambda_{\max} = \frac{2.9 \times 10^{-3}}{T} \quad \text{eq. 284}$$

Every object continuously emits radiation. Substituting common temperatures, e.g., room temperature, or normal temperature of a human body, the peaks of the emission spectra are in infrared area (approx. 9-10 μm) invisible for human eyes, visible for certain animals that have receptors for infrared radiation (snakes). Using an infrared camera, the radiation is detectable, and the temperature is measurable even, when the object is in a certain distance. Infrared cameras are exploited by soldiers to identify targets, by rescue teams to find lost persons, a quick check if a person has elevated temperature, to reveal glowing fire remnants after extinction of hazardous fires by firefighters, etc. In medicine, thermography is used to visualise the part of the organism with tissues that are usually higher in temperature than in the surroundings to reveal possible anomalies and diseases.

14 Electricity

Nowadays, electricity is the most important source and medium of energy. It is valid entirely for households, and partially in industry and economy, also. Its advantages are the relatively simple technological requirements and usage. The serious disadvantage is the difficulty of storing electricity, it is necessary to use it immediately after production.

Sources of electricity are the power plants, where in general other types of energy are transformed with certain efficiency into electrical energy. When we look for the original source of the energy, usually

we end up at the Sun, whose energy is transformed in several steps into electricity. For example, coal, crude oil, or natural gas comes from living organisms (vegetation, plankton), which were stored for many million years beneath the surface and converted to coal, crude oil, or natural gas. When these materials are used as fuels in power plants, the energy conserved in them is converted into electricity. Unfortunately, the efficacy of these processes is rather low, and the polluting of the environment is high.

There exist relatively novel methods, which exploit the energy of the Sun and immediately transfer it to electrical energy using photovoltaic panels. Wind farms, where the mechanical energy of the wind is transformed into electricity, use the energy of the Sun, also, because the wind is generated due to the pressure variation, i.e., the temperature differences of the atmosphere, which has its origin in the Sun, also. Another very important method is to build water dams, where the potential energy of the water (evaporated by the Sun after falling as rain or snow) is used to produce electricity. Unfortunately, the construction of a dam always means an encroachment on nature and a threat to the local ecosystem.

There exist a few sources and methods that are independent of solar energy, such as nuclear energy, geothermal energy, where the first one uses the energy of splitting of radioactive elements (uranium, plutonium) when after a controlled chain reaction, new elements are created, and huge amount of energy is released. Nuclear energy has its advantages and disadvantages. The release of carbon dioxide is low, but nuclear power plants produce radioactive waste, which is dangerous to the environment. It is necessary to store it for a very long time under special conditions, and there is also a risk of nuclear accident when the chain reaction gets out of control and radioactive material is released to the environment or even a nuclear blast occurs. The biggest nuclear accidents were in Fukushima in 2011 and in Chernobyl in April 1986, when, following the explosion of the reactor, huge areas surrounding the plant were contaminated for a long time by radioactive materials and became uninhabitable for humans.

Although the first evidence of electricity was known in ancient cultures, widespread utilisation of electricity occurred only in the twentieth century. People have known electric fish for a long time, lightning occasionally caused serious accidents or death, and it was also known that an amber bar scrubbed with fur attracts another object (due to static electricity). There exist ancient artefacts which might have connection with certain knowledge of electricity, one of them is the so-called Baghdad Battery or Parthian Battery found in Iraq. Its purpose is not clear and under discussion; however, it could serve as a primitive battery. It consists of three pieces; from a ceramic pot, a cylinder from rolled copper plate, and an iron rod. Filling the pot with an appropriate electrolyte and placing the copper cylinder and the iron rod inside represents a battery.

For a long time, electricity remained rather a curiosity, but the interest of physicists becomes more intense in the 18th century, when several scientists, *Luigi Galvani*, *Alessandro Volta*, *Benjamin Franklin*, *Andre-Marie Ampere*, *Michael Faraday*, *Georg Ohm*, etc. revealed the basic features and relations in electricity and magnetism, which lead to the possibility of enormous progress in industry, economy, science and society, also.

The key to electricity (and magnetism) lies in the structure of atoms, where the nucleus carries positive and the surrounding electrons negative charge. Basically, the net charge of the atoms is zero because the quantities of the positive and negative charges equal and compensate each other. The electrons are bound with different strength to the nucleus in different elements, the material can lose some electrons or acquire additional electron; for example, when the amber rod is rubbed with fur. The contact of the two materials can cause an imbalance in the positive and negative charges on them;

however, the total positive and negative charges remain the same. Such an imbalance is the key to the exploiting of electricity because it enables (controlled) flow of the electrons, which is the electric current capable of driving machines. During these processes no electron is formed or destroyed; they are only replaced from one place, one material, to the other. The conservation law of the charge is valid:

The net electric charge (algebraic sum of charges) of an isolated system does not change, no matter what interaction occurs within the system.

Under certain conditions, when enough energy is concentrated into one point (synchrotrons), the energy can be changed to an equivalent amount of mass, for example, electron-positron pair (or other type of particles, according to the energy). The charge of the energy beam before the process was zero, and the sum of the charges of the positron and electron is zero, also. Although the mass of the beam before the process was zero and the mass of the electron-positron pair is ca. $2 \times 9.31 \cdot 10^{-31} \text{ kg}$, but this mass was formed from equivalent amount of energy, so the energy-mass conservation law is not violated at all.

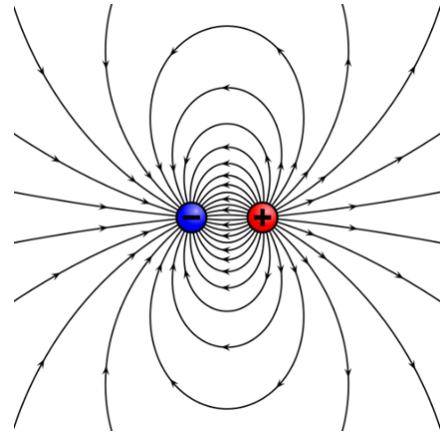


Figure 102. Electric field-lines of a positive-negative charge pair.¹¹

Electrons are carriers of the negative charge and protons of the positive charge in the atom; besides, there exist other elementary particles (muons, gluons, bosons, etc.) and some of them have charges. The antiparticles of the electrons and protons are called positrons and antiprotons, respectively. Positrons and antiprotons have the same properties as their antipairs but opposite charge. In the nuclei of the atoms, in addition to protons, neutrons are also present. Neutrons are neutral, have no charge, and their mass is slightly larger than that of protons.

The electric charge is quantised, which means that it can acquire only discrete values, which is related to the charge of the single proton e ($q_e = 1.602 \cdot 10^{-19} \text{ C}$, elementary charge), it is always its integer products; $n \cdot q_e$, n is a whole number. The charge of the electron has the same magnitude as the proton, but opposite sign; $-1 \cdot q_e$. The SI units of the electric charge is the **coulomb (C)**, expressed in basic SI units **1 C = 1 A s**.

14.1 Electric Field, Intensity of the Electric Field

Every particle with an electric charge generates an electric field (electrostatic field) and interacts with other particles with charge by this electric field. The interaction is attractive when it interacts with a particle of an unlike (opposite) charge, and it is repulsive when with a like charged particle. The force of action of the electric field on a charge is:

$$\vec{F} = q \cdot \vec{E} \quad \text{eq. 285}$$

$$[E] = 1 \text{ NC}^{-1} = 1 \text{ V m}^{-1}$$

Here, E is the intensity of the electric field and q is the charge interacting with the field. The electric field is a vector quantity, the direction of the electric field is graphically represented by field lines, and the orientation is **from positive charge to negative charge** (Figure 102). The density of field lines

¹¹ Image source https://cs.wikipedia.org/wiki/Elektrick%C3%BD_dip%C3%B3l?oldid=10718986

represents the intensity of the electric field. A single, point-like electric charge (the charge is concentrated to one point) forms an electric field of spherical symmetry.

Gauss law. The net electric flux (Φ_E) across any hypothetical closed surface is equal to $1/\epsilon$ times the net electric charge enclosed by the surface.

Electric flux (Φ_E) is the measure of the distribution of the electric field across a given surface, although an electric field cannot flow. It is a way of describing the electric field strength at any distance from the charge that causes a field.

$$\Phi_E = \frac{q}{\epsilon} = \vec{E} \cdot \vec{S} = ES \cos \theta \quad \text{eq. 286}$$

where q is the charge inside the volume, which generates the electric field, ϵ is permittivity of the nonconducting environment (material constant) surrounding the charge q , E intensity of the electric field, S the surface enclosing the charge q . θ is the angle between the electric field line and the normal vector of the surface. The surface is defined here as a vector, the magnitude is the area of the surface, the orientation is the normal vector, which is perpendicular to the tangent to the surface at the given point (Figure 103). In the case of a point-like charge (spherical symmetry), the area is the surface of the sphere ($4\pi r^2$) around the point-like charge:

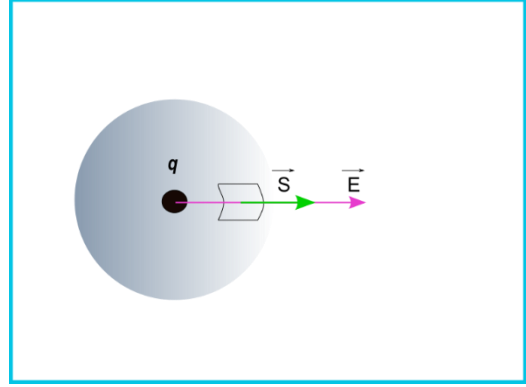


Figure 103. Point-like charge and a spherical electrostatic field around it.

$$\frac{q}{\epsilon} = E 4\pi r^2 \cos 0$$

$$E = \frac{1}{4\pi\epsilon} \frac{q}{r^2} \quad \text{eq. 287}$$

The force that exerts the electric field (*Coulomb's law*) on a charge (q') is:

$$\vec{F} = q' \vec{E} = \frac{1}{4\pi\epsilon} \frac{qq'}{r^2} \frac{\vec{r}}{r} \quad \text{eq. 288}$$

The last fraction ($\frac{\vec{r}}{r}$) has absolute value one and converts the right-hand side to a vector. The magnitude of the force is:

$$F = \frac{1}{4\pi\epsilon} \frac{qq'}{r^2} \quad \text{eq. 289}$$

The charges have positive or negative signs, so in the case of unlike charges and attractive force between them the force is negative, while in the case of like charges (both positive or both negative charges), the force is repulsive and has positive sign. Eq.288 and eq. 289 represent Coulomb's law, which is the simplest and most general case of electrostatic force between point-like charges.

Other types of fields with a different geometry might have a different equation (plate capacitor, cylindrical capacitor). For example, let us have a rod with charge q and around a cylinder with charge q' , and calculate the force between the rod and the cylinder with different, cylindrical geometry. The electric field intensity derived from Gauss's law:

$$E = \frac{q}{\varepsilon S} = \frac{q}{\varepsilon 2\pi r l}$$

And the force is:

$$F = \frac{1}{2\pi\varepsilon} \frac{qq'}{rl} \quad \text{eq. 290}$$

where r is the radius of the cylinder and l is the length of the cylinder.

14.1.1 Shell Theorem

A charge is usually not concentrated at one point but dispersed on a surface, for example, on the surface of a sphere. In this case, it is rather a complicated problem to calculate the intensity of the electric field at a certain point, where contributions from every point of the spherical surface should be considered. To simplify the problem, we use the so-called shell theorem (also valid for gravitational force).

Theorem 1. *A uniform spherical shell of charge behaves for external points, as if all its charge were concentrated at its centre.*

$$E = \frac{1}{4\pi\varepsilon} \frac{q}{r^2} \quad \text{for } r > r_{\text{sphere}} \quad \text{eq. 291}$$

Theorem 2. *A uniform spherical shell of charge exerts no force on a charged particle placed inside the shell.*

$$E = 0 \quad \text{for } r < r_{\text{sphere}} \quad \text{eq. 292}$$

Inside the sphere, the intensity of the electric field is zero and outside the sphere is the same value as if the entire charge would be concentrated in the centre of the sphere.

14.1.2 Electric Dipole, Dipole Moment

Compounds basically have no charge, molecules, and atoms are neutral. The situation is different when we dissolve a salt in a solvent (electrolyte). Instead of neutral particles (atoms, molecules), ions (cations and anions) occur, with positive and negative charges. The total charge of the solution remains zero, but inside the solution charge-carrying particles are present and provide a random motion in the solution.

Electronegativity is the measure of the ability of an element to attract electrons. The higher the electronegativity of an element, the more it attracts electrons and easily forms anions. A typical example is fluorine, which has the highest electronegativity of all elements. The lower the electronegativity, the easier the element loses electrons and forms cations, which is characteristic for metals. In chemical compounds of different elements with significantly different electronegativity, the bonding electrons are differently bound to the atoms, belong to the connected atoms, and the average time spent in the vicinity of each connected atom is different. The bond is polar; the atom with higher electronegativity has a partial negative charge, and the atom with lower electronegativity has a partial positive charge. In the limit case, when the whole bond belongs to one atom (group of atoms), anions and cations are formed which are held together mainly with electrostatic interaction, and this is called an ionic bond. The other type of chemical bond, where the bonding electrons are shared and belong to both connected atoms, is the covalent bond. There exists no pure ionic bond, and ionic bonds have a covalent character to some extent when the bonding electrons belong to both connected atoms.

In the case of polar bonds, the whole molecule is electroneutral, but the molecule has a negative and a positive pole. The measure of the polarity of the molecule is the electric dipole moment of the molecule. In general, the electric dipole moment is formed when the positive and negative charge is in a certain distance from each other (Figure 104). The dipole moment is a vector defined as the charge (q) multiplied by the distance (r) between them. The orientation of the vector is from the negative charge to the positive charge (probably for historical reasons).

$$\vec{\mu} = q \cdot \vec{r} \quad [\vec{\mu}] = 1 \text{ Cm} \quad \text{eq. 293}$$

In the case of a system that consists of several positive and negative charges, the total dipole moment is the vector sum (not a simple arithmetic sum) of the partial dipole moments:

$$\vec{\mu} = \sum_i q_i \vec{r}_i \quad \text{eq. 294}$$

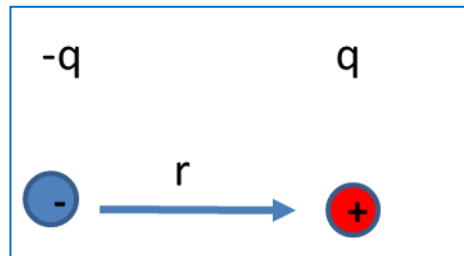


Figure 104. Electric dipole

The dipole moment in molecules also depends on the geometry of the molecule.

Problem. Find the dipole moment of the water! A water molecule consists of two hydrogen and one oxygen atom; both hydrogens are bound with a simple bond to oxygen. The angle between the bond is 104.45° and the length of the bond is $0.9584 \cdot 10^{-10} \text{ m}$. For convenience of calculation, we can place the oxygen at the origin of the coordinate system, one hydrogen on the x axis and the second hydrogen according to the bond angle 104.45° (Figure 105, the blue vectors correspond to the dipole moments of the bonds, not to the length of the bond). The partial charge on oxygen is $-0.671e$ (e is the elementary charge, $1.602 \cdot 10^{-19} \text{ C}$) and on both hydrogens $0.3355e$. The charges, coordinates, and dipole moments are listed in the following table:

Table 9. Partial charges and dipole moment of the water molecule.

	$q \text{ (C)}$	$x_i \text{ (m)}$	$q_i x_i \text{ (Cm)}$	$y_i \text{ (m)}$	$q_i y_i \text{ (Cm)}$
O	$-10.7494 \cdot 10^{-20}$	0	0	0	0
H1	$5.3747 \cdot 10^{-20}$	$0.9584 \cdot 10^{-10}$	$5.1511 \cdot 10^{-30}$	0	0
H2	$5.3747 \cdot 10^{-20}$	$-0.2392 \cdot 10^{-10}$	$-1.2856 \cdot 10^{-30}$	$0.9281 \cdot 10^{-10}$	$4.9883 \cdot 10^{-30}$
$\Sigma \mu_i$			$3.8655 \cdot 10^{-30}$		$4.9883 \cdot 10^{-30}$

The vector of the dipole moment is defined with its μ_x and μ_y components:

$$\vec{\mu}(\mu_x, \mu_y) = \vec{\mu} \left(\sum_i q_i x_i, \sum_i q_i y_i \right) = \vec{\mu}(3.8655 \cdot 10^{-30}, 4.9883 \cdot 10^{-30})$$

The magnitude of the vector is as follows:

$$|\vec{\mu}| = \sqrt{\mu_x^2 + \mu_y^2} = 6.31 \cdot 10^{-30} \text{ Cm}$$

The experimental value is $6.17 \cdot 10^{-30} \text{ Cm}$. To find the angle (θ) between the dipole moment's vector and the bond between the first (or second) hydrogen is, in our case, very simple, because the bond O-H₁ lays on the x-axis.

$$\text{tg } \theta = \frac{\mu_y}{\mu_x} = 1.29047 \quad \theta = 52.227^\circ$$

The calculated angle is half of the bond angle; this means that the dipole moment lies in the symmetry axis of the water molecule, of course.

The partial charges and the dipole moment of the water molecule play an important role in the physics and chemistry

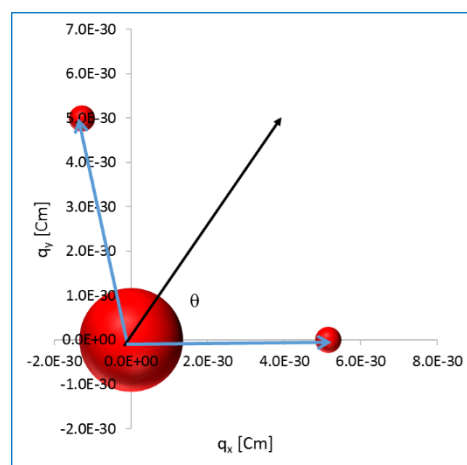


Figure 105. Dipole moment of water molecule.

of water. Water is also the most important solvent, widely used, and the basis of life on Earth. It is a polar solvent, dissolves polar molecules, and does not dissolve nonpolar molecules, for example alkanes, alkenes, etc. As a counterexample, we can consider carbon dioxide, which is a molecule with zero dipole moment; though, a partial charge should exist on the carbon and the oxygen due to their different electronegativity (electronegativity of O: 3.44, C: 2.55, H: 2.1). However, in a carbon dioxide molecule, all atoms lay inline, the molecule is symmetric, and because of this, the dipole moments of the two C-O bonds compensate each other and the overall dipole moment of the molecule is zero. In Table 10, there are the boiling points, molar masses, and dipole moments of water, several alcohols, and carbon dioxide. The dipole moments are in *debye (D)*, which is a non-SI unit, but the dipole moments in debyes are simple numbers and debyes are still used in physics and chemistry.

$$1D \approx 3.33564 \cdot 10^{-30} Cm$$

Table 10 demonstrates two very interesting things; the first is the variation of the boiling point of the homologous series of alcohols; $H-(CH_2)_n-OH$, where n is an integer and n is increasing from zero (water) to three (1-propanol). To evaporate a liquid, we must increase the kinetic energy of the molecules to overcome the attractive forces between the molecules inside the liquid. The higher molecular mass means a higher inertia of the molecule and a higher energy request to speed up the molecule to the escape speed. A higher dipole moment means stronger electrostatic interactions between the polar molecules.

Table 10. Boiling points, molar masses and gas phase dipole moments of alcohols, water, and carbon dioxide.

Liquid	Boiling point [°C]	Molar mass [g/mol]	Dipole moments ¹² [D]
water	~100	18	1.85
methanol	~64	32	1.70
ethanol	~78	46	1.69
1-propanol	~97	60	1.68
Carbon dioxide (sublimation)	-78.5	44	0

The boiling point dependence on the molar mass of the series provides an approximately linear relationship, except of water, that is, the addition of a CH_2 group increases the boiling point by about 14-19°C. It is mainly a result of the increase in the molar weight of the compound. Accordingly, the boiling point of the water should be less than 50 °C instead of 100 °C. In the case of water, the prevailing property is the dipole moment and, therefore, the hydrogen bonds present in water. The situation is very different for carbon dioxide, which has a boiling point of -78.5 °C. Its molar weight is comparable to that of ethanol, but the lack of polarity of the molecule caused this low temperature of sublimation (solid-gaseous phase transition) of dioxide.

¹² Nelson, R., Lide, D. and Maryott, A. (1967), *Selected values of electric dipole moments for molecules in the gas phase*, National Institute of Standards and Technology, Gaithersburg, MD, [online], <https://doi.org/10.6028/NBS.NSRDS.10>

14.2 Work and Potential Energy in an Electric Field

If a force (F_{act}) moves an object in gravitational field against the gravitational force ($F_{act} = -F_g$, higher), its potential energy will increase. When a charge is moved by a force (F_{act}) in an electric field (E), against the electric forcefield ($F_{act} = -F_{el}$), the situation is not so obvious; its potential energy can increase or decrease. It depends on the sign of the charge and the orientation of the electric field. Let us have a homogeneous electric field between two plates of a capacitor (Figure 106). The upper plate is charged to $Q_1 (>0)$ and the lower plate to $-Q_1 (<0)$. The motion of a positive charge $q > 0$ from the positive plate (case 1) to the negative plate, is spontaneous, producing work ($W < 0$), and the potential energy of the charge will decrease. In the case of a negative charge $q < 0$, to move it from the positive plate to the negative plate (case 2), it is necessary to add work ($W > 0$), and the potential energy of the charge increases. The work is expressed as follows:

$$W = \int_{x_1}^{x_2} \vec{F}_{act} \cdot d\vec{x} = - \int_{x_1}^{x_2} \vec{F}_{el} \cdot d\vec{x} = - \int_{x_1}^{x_2} q\vec{E} \cdot d\vec{x} = - \int_{x_1}^{x_2} qE \cos\theta dx \quad eq. 295$$

Where F_{el} is the force between the charge q and the electrostatic field E , which must be compensated (Figure 106, green vector), x is the path of the charge's motion, $\theta (=0)$ is the angle between the motion vector and the electric field vector between the capacitor plates. The direction of the electric field vector (vector line orientation) is from positive plate to negative. The direction of the charge motion in Figure 106 (defined by us) is the same, i.e., the angle between them (θ) is zero and $\cos\theta$ is 1. Therefore, the sign of the work, done by external force, is negative for positive charge ($q > 0$), and positive for negative charge ($q < 0$). It corresponds to the conclusion made intuitively above, considering the repulsing force between like charges and the attractive force between unlike charges. The energy/work conservation law is valid also; the work done on the system is transformed to the increased potential energy, and the received work from the system causes a decrease of the potential energy. The convention for the sign of the work is identical to that in the case of thermodynamics, i.e., the work done by the system is negative and done on the system is positive.

However, if we consider the work done by the electric field (W_{el} , electric work), the sign of the work is opposite.

The electric field intensity in general is given by the Gauss' Law (Φ_E is the electric flux, see Gauss' Law, eq. 286, ϵ permittivity of the environment):

$$\Phi_E = \frac{Q}{\epsilon} = \vec{E} \cdot \vec{S} = ES \cos \theta \quad eq. 296$$

Where θ is the angle between the normal vector of the plane enclosing the charge (Q) and the electric field lines:

$$E = \frac{Q}{\epsilon S \cos \theta} \quad eq. 297$$

And substituting for E into eq. 295:

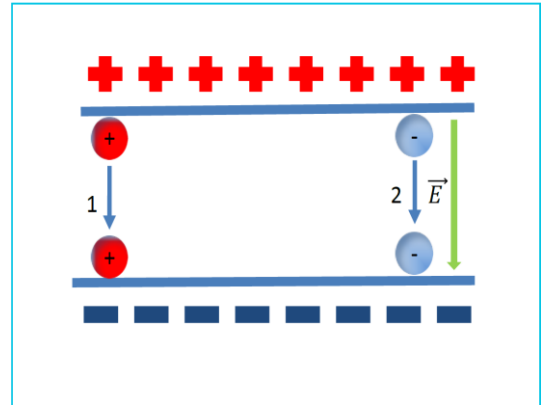


Figure 106. Transport of charges between charged plates.

eq. 298

$$W = - \int_r q \frac{Q}{\epsilon S \cos \theta} dr$$

In the case of a plate capacitor (Figure 106, previous example), where S is the area of one plate, d is the distance between the plates, Q is the charge on the plates, the work necessary to move a charge q from one plate to the other, depending on the sign of the charge and direction of the motion (for example $\theta=0$, $\cos\theta=1$):

$$W = -q \frac{Qd}{\epsilon S} = -q \frac{Q}{C} \quad \text{eq. 299}$$

$$C = \frac{\epsilon S}{d} \quad \text{unit F (farad)} \quad \text{eq. 300}$$

where d is the distance between the capacitor plates, S is the area of the plates, and ϵ the permittivity of the material between the plates. C stands for the fraction, which is usually constant, and it is called the capacitance of the capacitor. The capacitance depends on the geometrical characteristics of the capacitor (S, d) and the permittivity of insulating material between the plates. To increase the capacitance of the given capacitor (with the given area and plate distance), the permittivity of the medium is usually increased; for example, the space between the plates is filled with special oil, with higher permittivity, instead of air.

Definition. An **electric potential (φ)** is the amount of work needed to move a unit positive charge from a reference point (infinite distance, with zero potential) to a specific point inside the electric field without producing any acceleration.

$$\varphi_A = \int_{\infty}^A d\varphi = \int_{\infty}^A \frac{dW}{q} \quad \text{eq. 301}$$

The zero acceleration ensures constant kinetic energy of the charged particle, so the added work is converted to increase its potential energy. The potential difference between two points is the voltage (U). Let us consider a case where the electric field is generated by a point like charge (Q) and q is the charge moved in the field (Figure 107) from point a to point b . The angle between the path ($r_b - r_a$) and the electric field vector line of the electric field is 0° , $\cos 0^\circ = 1$. Using eq. 287 - eq. 289 and eq. 295

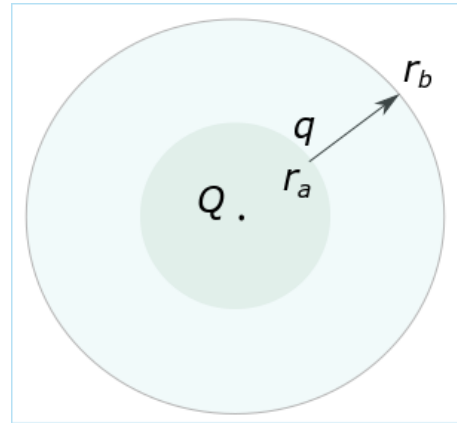


Figure 107. Transport of a charge in spherical electrostatic field.

$$\begin{aligned} \Delta\varphi = U &= -\frac{Q}{4\pi\epsilon} \int_{r_a}^{r_b} \frac{1}{r^2} dr = -\frac{Q}{4\pi\epsilon} \left[-\frac{1}{r} \right]_{r_a}^{r_b} \\ &= \frac{Q}{4\pi\epsilon} \left[\frac{1}{r_b} - \frac{1}{r_a} \right] = \varphi_b - \varphi_a \\ 1V &= 1J \, C^{-1} = 1kg \, m^2 s^{-3} A^{-1} \end{aligned} \quad \text{eq. 302}$$

The unit of the difference in electric potential and the electric potential is *volt*, abbreviation V . The work done moving the charge q from the eq. 301 is:

$$\begin{aligned} dW &= q \, d\varphi \\ W &= q \int_{r_a}^{r_b} d\varphi = q(\varphi_b - \varphi_a) = qU \end{aligned} \quad \text{eq. 303}$$

In physics, the energy (work) of a particle is sometimes expressed in electronvolts (eV , *not SI unit*). $1eV$ is the amount of kinetic energy gained by a particle with unit charge that is accelerated from rest through an electric potential difference of one volt in vacuum.

The joule - electronvolt transfer is according to the electrical energy (work) definition eq. 303:

$$1eV = 1.602 \cdot 10^{-19} C \cdot 1 V = 1.602 \cdot 10^{-19} J$$

where the elementary charge is multiplied with the unit voltage.

The power (rate of work) is defined as:

$$P = \frac{W}{t} = \frac{qU}{t} = IU \quad \text{eq. 304}$$

$$I = \frac{q}{t}$$

Where I is the electric current, which is the rate of the flow of electric charge. The unit of the electric current is *ampere* (A), one of the *SI* basic units.

14.3 Electrical Resistance, Ohm's Law

The relation between the electric potential, current, and resistance is expressed by **Ohm's law**.

Ohm's law: the current through a conductor (see in below) between two points is directly proportional to the voltage across the two points.

$$I = \frac{U}{R} \quad \text{eq. 305}$$

$$1\Omega = 1 VA^{-1} = 1 kgm^2 s^{-3} A^{-2}$$

The unit of electrical resistance is the *ohm* (Ω). The resistance R of an object is defined as the ratio of voltage U on the ends of the object, that induces current I .

Electrical resistance is the measure of the resistance against the flowing of electric current. It depends on the characteristic of the conductor material and its geometrical dimensions. In the case of a metal conductor, the resistance is proportional to its length (l) and inversely proportional to the cross-sectional area (S):

$$R = \rho \frac{l}{S} \quad \text{eq. 306}$$

ρ - is the resistivity of the material, a material constant. The units of the resistivity are Ωm . From the Ohm's law is clear, that when current flows through a component of a circuit, the voltage-drop on this component is proportional to the resistance of the component.

$$U_{comp} = IR_{comp} \quad \text{eq. 307}$$

Materials are divided according to their conductivity to conductors, semiconductors, and insulators. Metals (Ag, Cu, Fe, etc.) are very good conductors, with low resistivity, where the electrical conduction is provided by the motion of electrons, which are relatively free in metal lattice. On the other hand, insulators have high resistance to electric current. They are nonmetals (glass, paper, plastic), in which electrons cannot move freely. The third type of materials are the semiconductors, which are positioned with their conductivity between the previously mentioned types of materials. Their utilisation in electronics and industry caused enormous changes, and all electronic devices (computers, mobile phones, etc.) are based on semiconductor technology.

The resistance is temperature dependent, but the relation is different for conductors and semiconductors. The resistance of conductors increases linearly with temperature:

$$R = R_0(1 + \alpha\Delta T) \quad \text{eq. 308}$$

In the case of semiconductors, the resistance decreases exponentially with increasing temperature:

$$R = A e \frac{B}{T} \quad \text{eq. 309}$$

The reversed value of the resistance is the electrical conductance (G , unit S , siemens).

$$G = \frac{1}{R} = \kappa \frac{S}{l} \quad \text{eq. 310}$$

κ is the conductivity (reversed value of the resistivity, units $S m^{-1}$).

14.4 Materials in Electric Field

14.4.1 Conductors, Semiconductors, Insulators

Electric current is a synchronised motion of charged particles controlled by an external electric field, oppositely to the random motion of charged particles. Conductors, semiconductors, and insulators behave differently in an external electric field.

In the case of conductors, and in the presence of an external electric field, the charged particles (electrons) move against the external field to compensate for it. This motion continues until the external field is fully compensated with an induced electric field in the material (Figure 108, top) and the intensity of the electric field inside the material is zero.

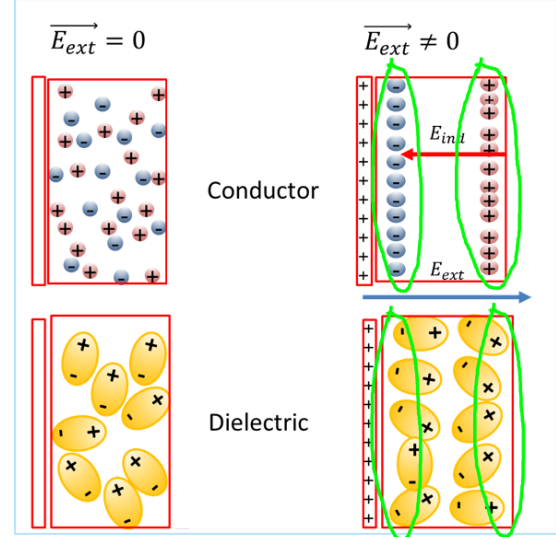


Figure 108. Charge separation in conductor (top) and in a dielectric material (bottom).

$$E_{ext} + E_{ind} = 0$$

From the Gauss law:

$$E_{ext} = \frac{q}{S \epsilon_0} = \frac{\sigma}{\epsilon_0} \quad \text{eq. 311}$$

where σ is surface charge density of the electric field.

The electric field has the same intensity behind the material as before it, although inside the material has zero intensity. Charges in a dielectric material, an electric insulator (Figure 108, bottom), are not as mobile as in the case of a conductor. The charges (dipoles) are partially orientated against the external field (electrical polarisation), but they are not capable to compensate for the external electric field. The electric field inside (E) the material is weaker than the external field, but it is not zero. The electric field on the other side of the material is weaker than the initial field; the material behaves as an insulator and, in the case of a perfect insulator, it is zero. In the material, electric polarisation occurs (P), which is proportional to the intensity of the electric field in the material:

$$P = \chi \epsilon_0 E \quad \text{eq. 312}$$

χ is the electric susceptibility of the material, ϵ_0 permittivity of vacuum. Polarisation acts against the external field and decreases the electric field intensity inside the studied volume:

$$E = E_{ext} - \frac{P}{\epsilon_0} = E_{ext} - \chi E \quad \text{eq. 313}$$

$$E = \frac{E_{ext}}{(1 + \chi)} = \frac{E_{ext}}{\epsilon_r} = \frac{q}{4\pi\epsilon_0\epsilon_r r^2}$$

The permittivity of the material is the product of the permittivity of vacuum (ϵ_0) and the relative permittivity ($\epsilon_r=1+\chi$). Relative permittivity is a dimensionless number, which expresses how much decreases the intensity of the electric field passing through a material. The permittivity of the vacuum ϵ_0 is a basic physical constant, its value is:

$$\epsilon_0 = 8.8542 \cdot 10^{-12} F \cdot m^{-1} (farad \text{ per meter}) \text{ or } C \cdot V^{-1} \cdot m^{-1}$$

Polarisation is the response of the material to the external electric field. Three types of polarisations can be observed:

1. Electronic polarisation is due to the displacement of the centres of the positive and negative charges in the **atoms**; the nucleus and the electron cloud of atoms.
2. Atomic polarisation is a **deformation** of **molecules**, where the centres of the partially positive and negative charges of the molecules are shifted away from each other and the dipole moments of the molecules are increased.
3. Orientation polarisation occurs when the initially random orientation of polar molecules becomes less random, and the **molecules orientate** their positive and negative poles according to the external electric field.

14.4.2 First and Second Type Electrical Conductors – Metals and Electrolytes

Metals belong to the first-type electrical conductor group, where electrons are the charge carriers, while the positive (usually) metal cations are bound to the lattice positions. Second-type conductors are electrolytes (solutions of ions) in which the electric current is provided by the flow of ions. When electrodes are immersed in electrolytes, the initial random motion of the ions becomes orientated, the cations flow to the negative electrode (cathode), and the anions flow to the positive electrode (anode). In that case, the electric current is provided by both types of charge carriers, positive and negative, while in the case of the conductors of the first type, only the electrons are mobile, and the metal cations are immobile. The ions are substantially larger than electrons; their motion in a liquid is hindered by permanent collisions with the molecules of the solvent. Similarly to sedimentation speed, conductivity depends on the viscosity of the solution. Increasing temperature increases the particle speed in a solution and decreases its viscosity, so the resistance of the solutions decreases with increasing temperature, contrary to the temperature dependence of the metal resistance. The conductance of materials (G , unit S - siemens) is the reversed value of the electric resistance:

$$G = \frac{1}{R} = \frac{\kappa}{C} \quad \text{eq. 314}$$

Where κ is conductivity of the solution and C is the cell constant, (unit m^{-1}) characterises the geometry of the electrode. Unit of conductivity is:

$$[\kappa] = S \cdot m^{-1}$$

The conductivity is characteristic for the given solution; the electric conduction of a solution depends on the type and number of ions, so conductivity depends on the concentration (more precisely on the activity) of the solution.

Molar conductivity (λ) is the conductivity of the given solution of unit concentration (activity).

$$\lambda = \frac{\kappa}{c} \quad \text{eq. 315}$$

c is the concentration of the solution.

The molar conductivity of the solution is concentration dependent, and at low concentration is characterised by the Kohlrausch's law for strong electrolytes:

$$\lambda = \lambda_{\infty} - a\sqrt{c} \quad \text{eq. 316}$$

Where λ_{∞} is limiting molar conductivity at infinite dilution, that is, the molar conductivity of the solution extrapolated to zero concentration. The limiting molar conductivity is an additive quantity, it is the sum of the limiting molar conductivities of the cations and anions, expressed by the Kohlrausch's law of independent ion migration.

$$\lambda_{\infty} = \lambda_{\infty \text{cation}} + \lambda_{\infty \text{anion}} \quad \text{eq. 317}$$

14.4.3 Semiconductors

Semiconductors are materials whose ability to conduct electrical current differs from that of conductors and insulators. The conductivity of metals is of order 10^5 - 10^7 S/m, insulators' 10^{-11} - 10^{-25} S/m, the value for semiconductors is located between these values ca. 10^{-3} - 100 S/m (Table 11). Semiconductors in their native state are poor conductors; they do not have free electrons in their crystalline structure as metals. Semiconductors are based on silicon and germanium (elements of the fourth column of the periodic system), which have half of the electrons to fully occupy their valence orbitals. However, doped with elements positioned next to them in the periodic system (As, P, Ga, etc.), we have p-type or n-type semiconductors.

Table 11. Examples of electrical resistivity and conductivity of conductors, semiconductors, and insulators.

Material	Electrical resistivity - ρ (Ωm)	Electrical conductivity, κ (Sm^{-1}) $= 1/\rho$	Type of material
Silver ¹³	$1.59 \cdot 10^{-8}$	$6.3 \cdot 10^7$	metal, conductor
Copper ¹³	$1.72 \cdot 10^{-8}$	$5.8 \cdot 10^7$	metal, conductor
Gold ¹³	$2.44 \cdot 10^{-8}$	$4.2 \cdot 10^7$	metal, conductor
Aluminium ¹³	$2.82 \cdot 10^{-8}$	$3.4 \cdot 10^7$	metal, conductor
Germanium ¹³	0.46	2.2	semiconductor
Silicon ¹³	$3.5 \cdot 10^{-8}$	$4.3 \cdot 10^{-4}$	semiconductor
Polyethylene ¹³	$>10^{13}$	$<10^{-13}$	insulator
Diamond ¹⁴	$\sim 10^{16}$	10^{-11} - 10^{-18}	Insulator

The n-type semiconductor is formed when, for example, silicon is doped with an element from the fifth group of periodic systems, e.g., phosphorus, which has five electrons on its valence shell. Four phosphorus electrons are involved in the bonding with neighbouring silicon atoms, but the fifth electron is free to conduct electric current (Figure 109, left).

In the case of elements of the third group, for example aluminium, it has three electrons in the valence shell, all three involved in the bonds with neighbouring silicones, but with the fourth silicon the bond is incomplete, there is a 'hole' in place of the missing electron. In an electric field, the electrons of neighbouring atoms of silicon can fill the hole, leaving their original position as a new hole (Figure 109, right). In this way, the hole can 'travel' along the crystal and conduct electricity.

¹³ CRC Handbook of Chemistry and Physics (n7)

¹⁴ P.W. May, "CVD Diamond - A New Technology for the Future?", *Endeavour*, **19**(3) (1995), 101-106.

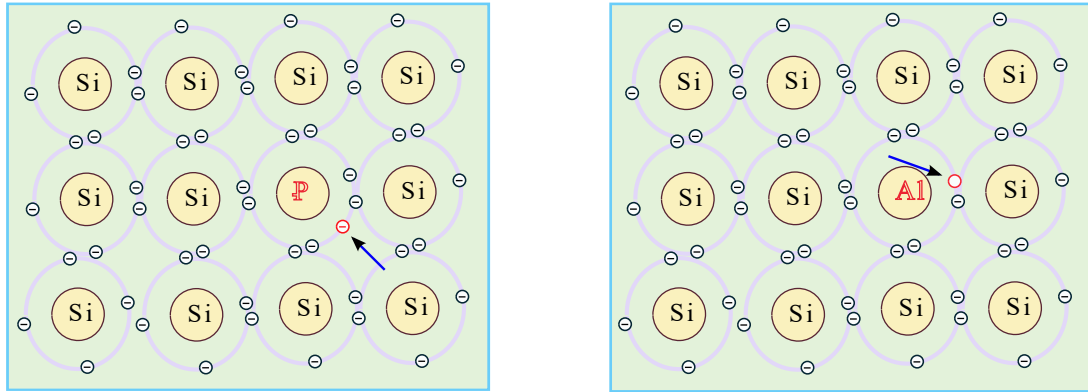


Figure 109. n-type (left) and p-type (right) semiconductor.

According to band theory, the valence electrons occupy the valence band in the material (Figure 110, red rectangles), and the conduction of electricity is performed in the conduction band (green rectangles). The valence and conducting band of a conductor are, in the basic energy state, overlapped. In the case of insulators, the energy difference between the valence and conduction bands is big (usually $>3\text{eV}$), it is difficult to transfer an electron from the valence band to the conduction band. In the case of semiconductors, the band gap is small, and it is relatively easy to excite electrons from the valence band to the conduction band and to conduct electricity. This excitation can be done either by supplying energy to the material, e.g., as the temperature increases, a few electrons are moved to the conduction band (intrinsic semiconductors).

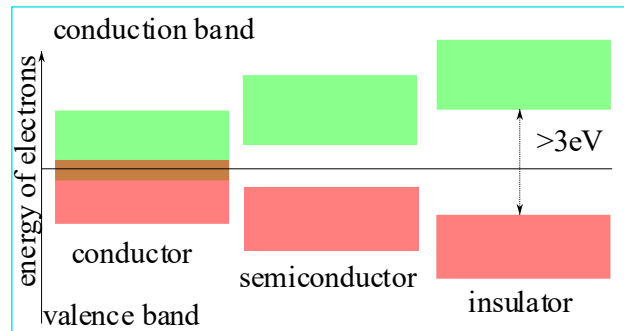


Figure 110. Valence and conduction bands of conductors, semiconductors, and insulators.

When doped with other elements, new energetic levels occur in the gap between the valence and conducting band, making the gap between the bands smaller. When a p- and n-type semiconductors are joined, a diode is prepared. A diode will operate as a switch, open, or closed, according to the polarity of the source in the circuit. When three layers of semiconductor (two p-type, one n-type, or reverse) are joined, a transistor is prepared. All electronic devices are constructed based on semiconductor technology.

14.5 Electromotive Force (EMF)

Definition. Electromotive force (ξ) is the energy (E) transferred per unit charge (Q) when a certain type of energy is converted into electrical energy.

$$\xi = \frac{E}{Q} \quad [\xi] = \frac{1 \text{ J}}{1 \text{ C}} = 1 \text{ V} \quad \text{eq. 318}$$

The voltage difference between the electrodes of an electric cell is the electromotive force (emf) of the cell in the case of zero current. The electromotive force of an ideal source when its internal resistance (R_i) is zero is

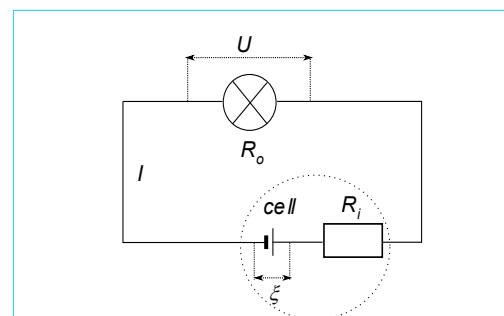


Figure 111. Electrical circuit with a lamp (circle with cross inside) and a cell with internal resistance.

independent of the current. The internal resistances of real cells are not zero; the voltage (U) provided by a source of electromotive force depends on the current that flows through the circuit.

$$U = \xi - IR_i \quad \text{eq. 319}$$

Part of the produced voltage is unusable; it is consumed inside the cell because of its internal resistance (dissipated into the environment as heat). The efficacy of the cell (η) is calculated as the ratio of the useful energy (E_u) in the circuit over the total energy of the circuit (E_t) (assuming ideal conductors, with zero resistances).

$$\eta = \frac{E_u}{E_t} = \frac{ItU}{It\xi} = \frac{U}{\xi} = \frac{IR_o}{I(R_o + R_i)} = \frac{R_o}{R_o + R_i} \quad \text{eq. 320}$$

The efficacy of the cell is defined by the ratio of the outer resistance (on the equipment) to the sum of the resistances on the equipment and the internal resistance of the cell (Figure 111).

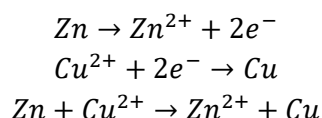
14.6 Galvanic Cells

An electric current is a flux of charged particles, usually electrons, but sometimes ions or other particles. The unit of electric current is *ampere* – A, one of the basic SI units. The direction of an electric current is, by convention, the direction in which a **positive charge would move**. If the charge carriers are electrons, the direction of their flow is opposite to the direction of the current flow. Two types of current are basically distinguished: direct current and alternating current. In the case of a direct current, the charge carriers flow in one direction; on the other hand, in the case of an alternating current, the direction of flow is frequently and periodically reverted.

The electricity used in households (in Europe) is alternating current with frequency 50 Hz, which means that the voltage and current have a sinusoidal shape with frequency 50 s^{-1} , and it changes its direction 100 times (two times in one cycle) in one second.

Small electronic devices often use batteries, galvanic cells, as a source of electricity. Galvanic cells produce direct currents. Many types of galvanic cells exist, but most of them are based on chemical reactions that produce electricity.

The principle can be explained by the example of the *Daniell cell*, which produces at 25°C 1.1018 V. It consists of a zinc and copper electrode, each of them immersed in their $1 \text{ mol} \cdot \text{L}^{-1}$ sulphate solution. Two metal electrodes are connected to an outer circuit and the solutions are connected with a salt bridge (Figure 112). The zinc electrode dissolves, producing Zn^{2+} cations (released to the solution) and two electrons for every cation. The copper cations of the solution take up two electrons and precipitate on the electrode. Electrochemical reactions are as follows:



The two electrons released by zinc in the first reaction are transported through the wire connection of the electrodes to the copper electrode and captured with the copper ion. The flow of these electrons

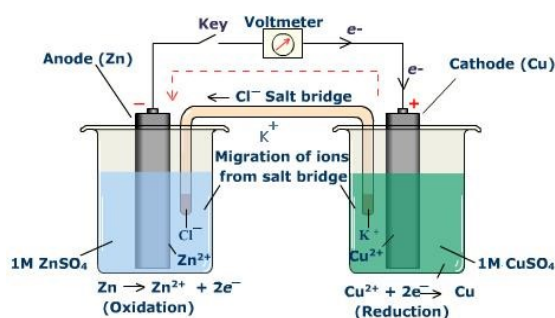


Figure 112. Daniell cell scheme.¹⁵

¹⁵ Image source https://chem.libretexts.org/@api/deki/files/147212/Fig_6-18.PNG?revision=1.

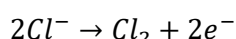
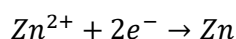
represents the electric current, and the electric potential difference between the electrodes is the voltage of the cell.

The electric potential of the galvanic cell is the difference between the potentials of its electrodes. Potentials of the electrodes depend on the concentration of the ions in the solutions, and having hydrogen cations in the solution, the measured cell potential depends on the pH of the solution. The measurement of pH is based on the measurement of electrode potentials.

14.7 Electrolysis

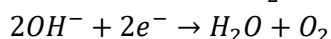
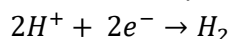
Electrolysis is a process reversed to the processes of galvanic cells. An outer electric source is connected to the electrodes, which determines the value and polarity of the electrode potentials, and so the chemical processes running on the electrodes. Electrolysis involves electrochemical reactions, which occur in solutions (or melts) because of an electric current. It is used in industry (aluminium production, hydrogen-oxygen production, galvanisation, etc.).

When electrodes of a battery are submerged in a solution of ions, the ions migrate to the oppositely charged electrode and react with it, producing neutral substances. The further processes depend on the chemical character of the produced neutral substances, sometimes forming a surface covering layer on the electrode from the dissolved material in the solution (metals). For example, iron or steel parts can be protected by coating with zinc against corrosion (galvanisation). When they are used as cathodes in electrolysis, their surface becomes covered with an appropriate protecting layer of zinc. The electrode reactions of $ZnCl_2$ are as follows:

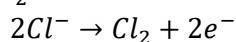
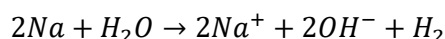
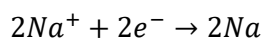


In this case, the neutral chlorine forms a gas that is produced on the anode and released from the solution.

Electrolysis of the sodium chloride solution is surprisingly more complicated. There are two theories for the mechanism of the process, the first (preferred by the authors of the most textbooks, e.g.¹⁶) is the direct reduction of hydrogen cations and oxidation hydroxide anions, releasing H_2 and O_2 gases.

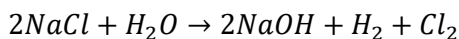


The Na^{+} and Cl^{-} ions are not involved in the electrochemical reaction and provide only the electric charge transport to the electrodes. This theory is supported with the lower (in absolute values) redox potential of H^{+} and OH^{-} ions compared to the Na^{+} and Cl^{-} ions, and the Arrhenius theory of water dissociation, in which hydrogen and hydroxide ions are present in neutral solutions and pure water, with concentrations of 10^{-7} mol/L. The problem with this theory is the low concentration of the H^{+} and OH^{-} ions in pure water; therefore, for noteworthy water electrolysis pure water is inappropriate; ions, salts, acids, or basis must be added. The second theory (Daniell-Miller view¹⁷) supposes reduction of sodium cations to neutral sodium with an immediate reaction with water to produce sodium hydroxide solution and gaseous hydrogen on the cathode and oxidation of chloride anions to chlorine gas.



¹⁶ Oxtoby D., Gillis H. and Butler L., (2016), *Principles of Modern Chemistry*, 8th edn, Boston, MA: Cengage.

¹⁷ Chang H., Duncan K., Kim K., Paik S.H.: *Chem. Educ. Res. Pract.*, 2020, **21**, 806-822.



The reduction of sodium cations to neutral sodium atoms during the electrolysis of solution is often questioned because the reduction potential Na^+ in a 0.1M NaCl solution is -2.71V and that of H^+ (at normal pressure) is only -0.414V, and the oxidation potential of OH^- 0.815V, while that of Cl^- is 1.42V (at normal conditions). It should be easier to reduce hydrogen than sodium and oxidise hydroxide than chloride anions. However, if a mercury electrode is used in electrolysis, part of the formed neutral sodium is dissolved in mercury, which hinders the reaction with water, and the products are sodium amalgam, NaOH and chlorine gas¹⁸, which supports the mechanism of the reduction of sodium and oxidation of chloride ions. The question of the mechanism is still open, the survey and history of the scientific debate can be found in paper¹⁷.

Production of hydrogen gas and its use in automobile industry as alternative fuel for gasoline is one of the most prospective ambitions if the electrolysis is performed with green technology.

The amount of precipitated metal on the electrode is calculated using Faraday's laws of electrolysis. Faraday's first law of electrolysis expresses the mass of the material produced with the electrochemical reaction.

The mass (m) of elements deposited on an electrode in grams is directly proportional to the charge (q) in coulombs:

$$m = A q = A I t \quad \text{eq. 321}$$

Where A is the electrochemical equivalent of the given material, I is the current that flows through the circuit, t is the time of electrolysis.

Faraday's second law of electrolysis defines the value of the electrochemical equivalent:

The mass of substances deposited or dissolved by the same amount of electricity is directly proportional to their equivalent masses (A) or equivalent weights.

$$m = nM = \frac{q}{|z|F} M = \frac{M}{|z|F} I t \quad \text{eq. 322}$$

So

$$A = \frac{M}{|z|F}$$

Where M is the molar mass of the material, z is the number of electrons in the electrochemical reaction and F is the Faraday constant; the charge of one mole of the elementary charge ($96\,486.7\text{ C mol}^{-1}$).

15 Magnetism

A magnetic field is an invisible field that exert a vector force, characterized by both strength and direction, and is produced by magnetic objects or varying electric fields.

The simplest magnetic structure that can exist in nature is the magnetic dipole, with a north and a south magnetic pole. Cutting a bar magnet to two pieces, the result will be not one north- and one south-pole magnet, but two pieces with their own new north and south poles (Figure 113). There **exists no** separate north or south pole of magnets. This is an essential difference between electricity and magnetism, because in the case of electricity, it is possible to separate positive and negative electric charges.

¹⁸ K. H. Büchel, H.H. Moretto, P. Woditsch: *Industrial Inorganic Chemistry*, Wiley-VCH, Weinheim, Germany, 2000.

Magnetism and electricity are interlinked phenomena that do not exist separately. The motion of a charged particle (electric current) produces a magnetic field. Around a wire, where an electric current flows, a magnetic field is generated (macroscopic current) as well as inside the atoms and molecules from the motion of electrons and protons (microscopic currents). According to that, all materials should produce magnetic fields; however, it is not so.

The reason is that usually the elementary magnetic fields of electrons and protons are in opposed couples, and the produced fields eliminate mutually, resulting in a zero magnetic field. Bulk alignment of the minor magnetic fields produces a non-zero magnetic field, as is the case in bar magnets. The alignment of these minor magnets inside the material is possible to eliminate by heating them above the so-called *Curie temperature (point)* and changing the magnetic character of the object. The *Curie temperature* is the temperature above which the materials lose their permanent magnetic properties. The process works in an opposite way, also, heating a piece of appropriate mineral above its Curie point in an external magnetic field, the previously randomly orientated elementary magnets will be aligned according to the external magnetic field producing *magnetic domains*, and cooling the object below the Curie point will keep their orientation producing a bar magnet. This is used in geology to determine the flip of the magnetic poles of the Earth. It is the exchange of the magnetic poles of our planet, which occurs randomly every ca. 450 000 years. Certain minerals melt due to volcanic activities and preserve the actual magnetic pole orientation after cooling. By dating the age of the melting, the age of the flipping of the Earth magnetic poles is possible to determine.

Permanent magnets are certain metals or metal alloys; for example, *alnico* is an alloy of aluminium, nickel, and cobalt with some minor moieties. *Ferrite* is an alloy of Fe_2O_3 with barium, manganese, nickel and zinc, which retain their magnetic property without an external magnetic field. Other types of materials can be magnetised to a certain degree, but they lose their magnetic feature beyond an external magnetic field.

Placing an object into a magnetic field, it will behave in a different way, according to its magnetic properties. According to this, materials are divided into several groups: *diamagnetic*, *paramagnetic*, and *ferromagnetic materials*. Magnetic properties can be characterised by the *magnetic permeability* of the material - μ (similarly to the electric permittivity in the case of an electric field).

Magnetic permeability involves the vacuum permeability - μ_0 and relative permeability - μ_r .

$$\mu = \mu_r \mu_0 \quad \text{eq. 323}$$

The relative permeabilities of the

- diamagnetic materials are less than one $\mu_r < 1$
- paramagnetic materials more than one $\mu_r > 1$
- ferromagnetic materials are much more than one $\mu_r \gg 1$

The value of the permeability of the vacuum belongs to basic physical constants, and its value is:

$$\mu_0 = 4 \cdot \pi \cdot 10^{-7} \text{ NA}^{-2} \text{ (or } \text{Hm}^{-1}, \text{ henry per metre)}$$

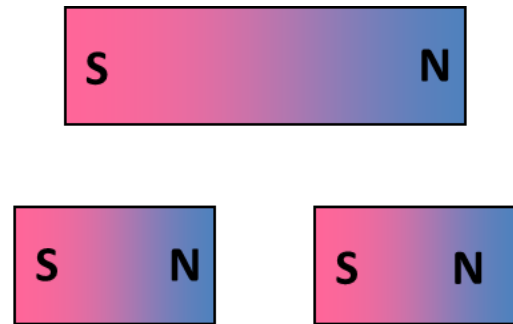


Figure 113. Bar magnet and the situation after its cutting.

15.1 Magnetic Flux Density (Magnetic Induction, B), Magnetic Field Strength(H)

The magnetic field is a vector field, determined by its strength, direction, and orientation, and defines the interaction (attraction or repulsion) of the field with magnetic objects (bar magnets, moving electric charge, etc.). The magnetic properties are described with two distinct, but related quantities, *magnetic flux density (magnetic induction)* and *magnetic field strength*.

The unit of magnetic induction (B) is *tesla (T)*:

$$[B] = 1 \text{ T} = 1 \text{ kg s}^{-2} \text{ A}^{-1}$$

and for the magnetic field strength (H) is henry (read [anri]) (H), unit Am^{-1}

$$B = \mu H = \mu_0 \mu_r H = \mu_0 (1 + \chi) H \quad \text{eq. 324}$$

where χ is the magnetic susceptibility. Its relation to μ_r is clear from eq. 323 and eq. 324. Magnetic induction is a linear function of the magnetic field strength for diamagnetic, paramagnetic, and non-magnetic ($\mu_0=1$) materials, but for ferromagnetic materials, the relationship is more complex.

15.2 Lorentz Law

The force exerted by an electromagnetic field of electric field intensity E and magnetic induction B to a moving electric charge (q) is expressed with the Lorentz law (Hendrik Antoon Lorentz, 1853 -1928):

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \quad \text{eq. 325}$$

The first part is the electrostatic interaction, and the second part is the magnetic interaction. Magnetic force is a vector product of the velocity and magnetic induction multiplied by the electric charge (scalar) of the particle.

$$\vec{F} = q \cdot \vec{v} \times \vec{B} \quad \text{eq. 326}$$

The magnitude of the force is the following.

$$F = qvB \sin \alpha$$

Where α is the angle between the velocity vector and the magnetic field vector lines. It follows from the rules of the vector product that the direction of the resulting force is perpendicular to the plane of the charged particle's velocity and the magnetic field lines. The direction of the force can be determined with the right-hand rule of the vector product (Figure 12). The orientation is flipped, in the case of negative charge (Figure 114). The force has a maximum value when the particle enters the magnetic field perpendicularly to the magnetic field lines ($\sin \alpha=1$). The resulting force bends the initial linear trajectory of the particle. The radius of the trajectory can be found from the

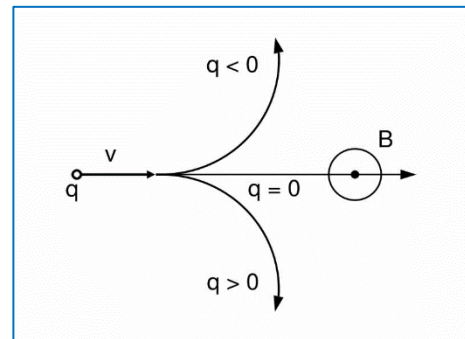


Figure 114. Trajectory bending of a charged particle in a magnetic field.¹⁹

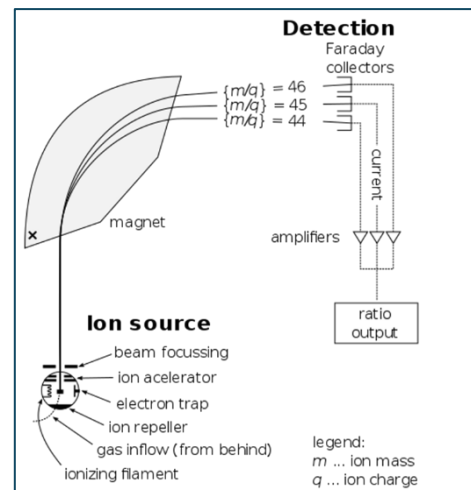


Figure 115. Scheme of a mass spectrometer.²⁰

¹⁹ Image source https://en.m.wikipedia.org/wiki/File:Lorentz_force.svg.

²⁰ Image source https://simple.m.wikipedia.org/wiki/File:Mass_spectrometer_schematics.png.

formula of the centripetal acceleration and force (eq. 85, eq. 86) and from the Lorentz law (eq. 326):

$$F = m a_c = m \frac{v^2}{r} = qvB$$

$$r = \frac{m v}{q B}$$
eq. 327

Mass spectroscopy (Figure 115), an analytical method used to identify the chemical or isotope composition of unknown samples, is based on this concept; the investigated sample is fragmented, ionised, and shot in a magnetic field. The trajectory of the fragments is bent, and the radius of bending is a function of the ratio of their mass and charge (eq. 327). From the position of the signals of the fragments, the original molecule composition is identified. The speed of the particles is possible to find from kinetic energy of the fragments. The fragments are accelerated by an electric field of voltage (U), the electric energy (work, eq. 303) is transformed to the kinetic energy of the particle:

$$\frac{1}{2} m v^2 = qU$$

$$v = \sqrt{\frac{2qU}{m}}$$

The radius of the trajectory is the function of the fragment's mass and charge:

$$r = \frac{m}{qB} \sqrt{\frac{2qU}{m}} = \sqrt{\frac{m}{q}} \sqrt{\frac{2U}{B^2}}$$
eq. 328

The expression $\sqrt{\frac{2U}{B^2}}$ contains the parameters of the mass spectrometer, which are known.

15.3 Magnetic Field Along a Straight Wire

A moving electric charge induces a magnetic field, so a magnetic field is induced around wires that conduct electric current (a flow of electric charges).

If we consider a current in a straight wire (Figure 116), the induced magnetic field is circular symmetry, with the wire in its centre.

Ampere's law: The magnetic field on the perimeter of a region is proportional to the current that passes through the region.

$$\sum_{\text{closed path}} B_{\parallel} \Delta L = \mu_0 I_{\text{enclosed}}$$

$$\oint_C B dl = \mu_0 I_{\text{enclosed}}$$

$$B 2\pi r = \mu_0 I_{\text{enclosed}}$$

$$B = \frac{\mu_0 I_{\text{enclosed}}}{2\pi r}$$

$$\mu_0 = 4\pi \cdot 10^{-7} \text{ Tm/A}$$

eq. 329

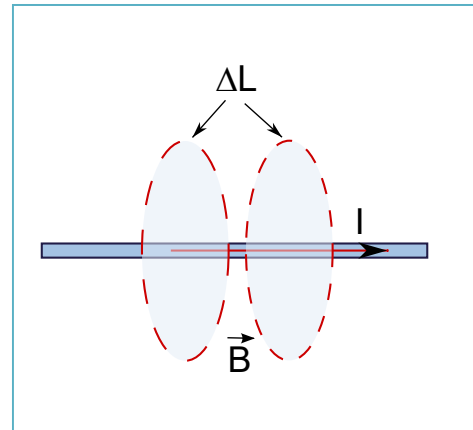


Figure 116. Magnetic field around a wire with electric current.

where B is the magnetic induction around the wire at distance r . At that distance the circumference is ΔL , I_{enclosed} is the electric current in the wire. The orientation of the magnetic field lines can be found

according to the right-hand rule; if the thumb on the right-hand points in the direction of the current, the other fingers point in the direction of the magnetic field lines.

15.3.1 Magnetic field of electric current in two parallel wires

Let us have two parallel wires with a unit current flowing in the same direction. The interaction between them (F_1 , F_2) is possible to describe according to the magnetic part of the Lorentz law (eq. 326):

$$\vec{F}_1 = q_1 \vec{v}_1 \times \vec{B}_2$$

$$\vec{F}_2 = q_2 \vec{v}_2 \times \vec{B}_1$$

The current of the first wire is in the magnetic field of the second wire, and vice versa.

$$\begin{aligned} |\vec{F}_1| &= |\vec{F}_2| = |q_1 \vec{v}_1 \times \vec{B}_2| = \\ &= I_1 t v_1 B_2 \sin 90^\circ = I_1 L_1 B_2 = I_2 L_2 B_1 \end{aligned}$$

using

$$q = I t \quad L = vt \quad B = \frac{\mu_0 I}{2\pi r} \quad \text{eq. 330}$$

$$F = \frac{\mu_0 I_1 I_2 L_1}{2\pi r}$$

L_1 is the length of the wires and r is their distance.

In the case of currents 1 A, the force of attraction (repulsion) between the wires placed at 1 m, on a length of 1 m, and $\mu_0 = 4\pi \cdot 10^{-7} \text{ Tm/A}$, between the two wires is (see the 'old' definition of ampere):

$$F = \frac{4\pi \cdot 10^{-7} \cdot 1 \cdot 1 \cdot 1}{2\pi \cdot 1} = 2 \cdot 10^{-7} \text{ N}$$

The attractive or repulsive character of the force depends on the mutual orientation of the currents in the wires. In the case of the same orientation of the current, the force is attractive. In the case of the opposite direction of currents, the force between the wires is repulsive.

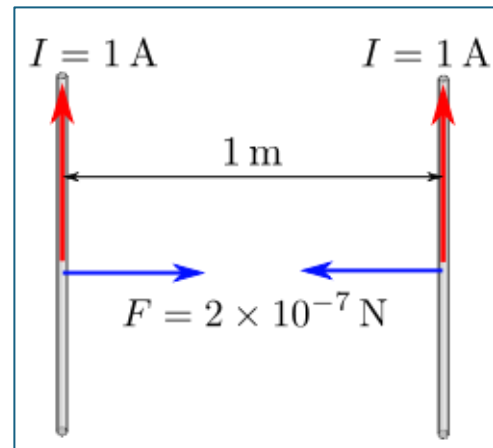
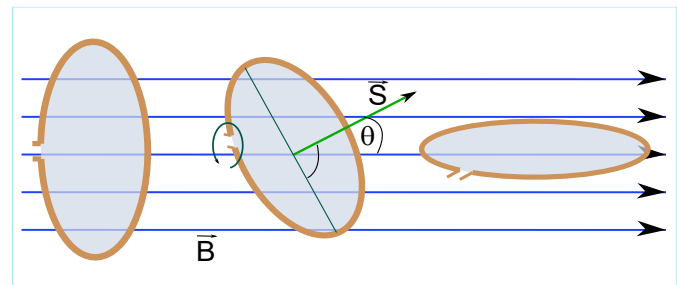


Figure 117. Attractive force between two wires with electric current.²¹

15.4 Magnetic Flux – Electromagnetic Induction, Faraday Law, Lenz Law

An electric current generates a magnetic field. A varying magnetic field also generates an electromotive force. This is the basis of energy production in most powerplants, where a conductor spins in a magnetic field. The rotation is provided by water flow from a hydroelectric dam, steam in the case of coal, oil-based power plants, etc. The variation of the magnetic field is usually achieved by varying the angle of the magnetic field lines



$$\Phi_B = BS$$

$$\Phi_B = BS \cos \theta$$

$$\Phi_B = 0$$

Figure 118. A wires' plane in magnetic field. The angle between the field-lines and wires' plane normal vector (green line) is left 0° , middle $0^\circ < \theta < 90^\circ$, right 90° .

²¹ Image source <https://en.wikipedia.org/wiki/File:Ampere-def-en.svg>.

and the position of the wires in which the electricity is produced (Figure 118).

*Faraday law of electromagnetic induction: the magnitude of the induced electromotive force (emf) equals the rate of the **change** of the magnetic flux.*

Magnetic flux is defined as a scalar product of the magnetic induction and the involved area:

$$\Phi_B = \vec{B} \cdot \vec{S} = BS \cos \theta \quad \text{eq. 331}$$

Angle θ is the angle between the magnetic field line and the normal vector of the wire plane (Figure 118). The voltage induced in the wire (electromotive force) is the time derivative of the magnetic flux.

$$\varepsilon_{\text{induced}}(U) = -\frac{d\Phi}{dt} = -\frac{dBS \cos \theta}{dt} \quad \text{eq. 332}$$

Magnetic induction B and the wires plane S in the case of an electric dynamo (power generator) are usually constant, the angle (θ) between the plane and the magnetic field changes due to the wires plane's rotation, so the induced electromotive force (voltage) is a periodic function of time:

$$\varepsilon_{\text{induced}}(U) = BS \sin \theta = U_{\text{max}} \sin \theta \quad \text{eq. 333}$$

The minus sign in the *Faraday Law* (eq. 332) means that the generated magnetic field is orientated in opposite direction as the original (external) magnetic field. This conclusion is so important that it is expressed in a separate rule.

Lenz Law: The magnetic field produced by an induced current in a closed circuit always opposes any changes in the magnetic flux.

$$U_{\text{ind}} = -\frac{d\Phi_B}{dt} \quad \text{eq. 334}$$

Let us have a bar magnet that approaches a solenoid connected to a galvanometer. The induced electromotive force in the solenoid has polarity, that generates a magnetic field with poles opposite to the bar magnet poles. The generated magnetic field exerts a repulsive force against the approaching bar magnet (Figure 119, top). When the bar magnet is taken away from the solenoid, the orientation of the electromotive force is reversed, and the generated magnetic field attracts the bar magnet (Figure 119, bottom).

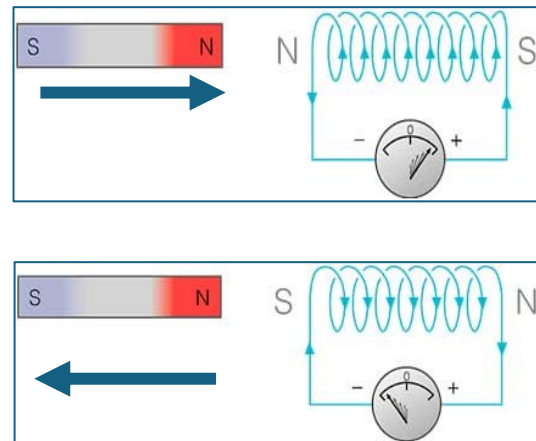


Figure 119. Electric voltage and current induced by change of magnetic field.²²

16 Electromagnetic Radiation

Electromagnetic radiation is a collection of electromagnetic waves with different energy, wavelength, and frequency. On the other hand, it is a flux of elementary particles, photons, which travel with constant speed in the given medium; speed of light ($c=299\,792\,458\text{ m/s}$ in vacuum). The speed of light in vacuum is a fundamental physical constant; in SI, the unit for the length, metre, is defined according to the speed of light in vacuum. Electromagnetic radiation does not require any medium for propagation.

The dual nature of the electromagnetic radiation means that it is considered in the same moment as:

²² Image source <https://riverglennapts.com/sk/law/513-lenz8217s-law-of-electromagnetic-induction.html>.

- a) waves characterised by frequency, wavelength, etc.
- b) particles – photons with momentum

Waves means a periodically oscillating motion; in the case of the electromagnetic waves, the vectors of the intensities of the electric and magnetic fields are oscillating perpendicularly to the direction of the beam (Figure 120), that is, electromagnetic waves are transverse waves. The vectors of the electric and magnetic fields are also perpendicular to each other. The energy of electromagnetic radiation is proportional to its frequency (ν) and inversely proportional to its wavelength (λ).

$$E = h\nu = \frac{hc}{\lambda} \quad [J] \quad \text{eq. 335}$$

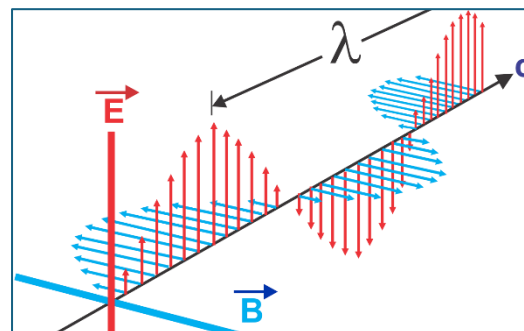


Figure 120. Propagation of an electromagnetic ray and the field-lines of its electric and magnetic fields.²³

The constant h , Planck's constant (Max Planck, 1858-1947), is a fundamental constant in physics, used for the definition of kilogram in SI.

$$h = 6.626\,070\,15 \times 10^{-34} \text{ Js}$$

Formulas to calculate the energy (E), wavelength (λ), frequency (ν), wave number ($\tilde{\nu}$) are:

Table 12. Relations between the parameters of electromagnetic radiation.

	E	λ	ν	$\tilde{\nu}$	Units
$E =$	-	$\frac{hc}{\lambda}$	$h\nu$	$hc\tilde{\nu}$	J (eV)
$\lambda =$	$\frac{hc}{E}$	-	$\frac{c}{\nu}$	$\frac{1}{\tilde{\nu}}$	m (nm)
$\nu =$	$\frac{E}{h}$	$\frac{c}{\lambda}$	-	$c\tilde{\nu}$	s^{-1} , Hz
$\tilde{\nu} =$	$\frac{E}{hc}$	$\frac{1}{\lambda}$	$\frac{\nu}{c}$	-	m^{-1} (cm^{-1})

The electromagnetic spectrum is divided into several regions according to wavelength (frequency). These are gamma rays, X-rays, ultraviolet, visible, infrared, and radio waves. The range of light, visible to the human eye is only a small portion of the entire spectrum, from about 380 to 750 nm. In this region, a different wavelength (frequency) means a different colour to us. The human eye is the most sensitive to radiation of near 555 nm, which is green in colour.

As the energy of the electromagnetic radiation depends on the frequency (wavelength), its impact on a living organism depends on the frequency (wavelength) of the radiation, too. The activity of enzymes and proteins are sensitive to the radiation of visible light and the infrared region, but basically the radiation with long wavelength (low frequency and energy) is harmless, with decreasing wavelength and increasing energy, it becomes more destructive and dangerous, even fatal to biological systems. The border between harmless and harmful electromagnetic radiation is in the ultraviolet region,

²³ Image source <https://www.cleanpng.com/png-electromagnetic-radiation-electromagnetism-wavelength-2463258/>

sunbathing without protective creams causes skin problems and even skin cancer. The bond energies for C-C and C-H bonds are ca. 346 kJ/mol ($5.75 \cdot 10^{-19}$ J/bond)²⁴ and ca. 411 kJ/mol ($6.83 \cdot 10^{-19}$ J/bond)²⁴, respectively. It means that if a tissue absorbs radiation with higher energy, chemical changes in the biological materials can occur. The fission of the bonds often results in a pair of very reactive free radicals that must be captured and transformed to less reactive ones with radical scavengers (vitamins C, E, D3).

The wavelength of the above-mentioned radiation is:

$$\lambda = \frac{hc}{E} = \frac{6.626 \cdot 10^{-34} \cdot 3 \cdot 10^8}{5.75 \cdot 10^{-19}} = 3.46 \cdot 10^{-7} \text{ m} = 346 \text{ nm} \quad \text{eq. 336}$$

The wavelength of 346 nm (and 292 nm for 411 kJ/mol) lies in the ultraviolet region. X-rays and gamma rays, with shorter wavelengths and higher energy, are even more dangerous to a living organism.

Electromagnetic radiation is not the only dangerous radiation. It consists of photons, which are particles with zero rest mass. The ability to penetrate to material is big, but it loses its energy along a long path. In the case of other types of radiations; α (${}^4_2\text{He}$; helium nuclei), β (e^- ; electron or e^+ positron), proton (p^+), neutron (n), the rays consist of particles with non-zero rest masses and very often non-zero charges. The interaction of these types of radiation with (biological) material is stronger than in the case of electromagnetic radiation (γ), although they penetrate less deeply into the material, but the energy loss of the radiation is done along a short(er) path, causing greater damage in the target.

Radiation is expressed with exposure, absorbed dose, dose rate, and biological effect.

The exposure (irradiance, X) is a measure of the ionisation of the air due to ionising radiation of photons, the amount of charge (q) formed in 1 kg of the air due to the radiation.

$$X = \frac{q}{m} \quad \text{eq. 337}$$

where m is the mass of the air. Units are C kg^{-1} (coulomb per kilogram). The absorbed dose (D) is the energy absorbed by 1 kg of the material:

$$D = \frac{E}{m} \quad \text{eq. 338}$$

Units are gray (Gy , $\text{J} \cdot \text{kg}^{-1}$). Dose rate (J_D) is defined as the amount of dose absorbed during one second. Its units are Gy s^{-1} .

$$J_D = \frac{D}{t} \quad \text{eq. 339}$$

Equal absorbed doses do not necessarily produce equal biological effects. To compare them, dose equivalent (H , units sievert Sv) is stated, which is expressed by the product of the dose and a weighting factor (w_R).

$$H = D w_R \quad \text{eq. 340}$$

In the following table are the weighting factors of the different types of radiations. The weighting factor is increasing with increasing mass, energy, and the charge of the radiation, which corresponds well with the increasing interaction with biological material. For example, the biological effect of alpha particles is twenty times bigger than the gamma rays of the same energy. The radiation dose, absorbed by living organisms, is not zero under common conditions. Radiation from the Sun, from natural radioactive minerals, and from human activities is present everywhere. This is the so-called background radiation, which is slightly different depending on the geographical location, altitude, presence of (nuclear) powerplants, content of natural and artificial radioactive isotopes in the environment, etc. Certain

²⁴Petkowski, J. J., Bains, W., Seager, S.: On the Potential of Silicon as a Building Block for Life. *Life* 2020, **10** (6), 84. Doi:10.3390/life10060084

activities elevate the radiation dose, work with radioactive isotopes (in medicine, nuclear power plants), frequent journeys with aeroplanes (as a member of an aeroplane crew), etc.

Table 13. Radiations and their biological weighting factors²⁵.

Radiation	Energy (E)	W _R
Photons, all energies		1
Electrons and muons, all energies		1
Neutrons	< 10 keV	5
	>10 keV - 100 keV	10
	> 100 keV to 2 MeV	20
	> 2 MeV to 20MeV	10
	> 20MeV	5
Protons, other than recoil protons, energy > 2 MeV		5
Alpha particles, nuclear fission products, heavy nuclei		20

16.1 Refraction and Reflection of Electromagnetic Radiation

No medium is required for the propagation of electromagnetic radiation. The speed of electromagnetic radiation in vacuum (c) is constant and independent on the motion of the source. It is a fundamental physical constant; the SI unit of length, *metre*, is defined according to the speed of the light in vacuum. Its value $c=299\,792\,458\text{ m/s}$. It is involved in the probably most famous Einstein equation, which expresses the energy-mass equivalence:

$$E = m c^2 \quad \text{eq. 341}$$

In annihilation processes, when the material interacts with its antimatter pair and is transformed to an equivalent amount of energy, the released energy is defined according to the previous equation. In nuclear scission and fusion processes (including nuclear bombs), the ‘missing’ material (the difference between the mass of the products and reactants) is transformed to equivalent amount of energy.

The finite value of the speed of light is not so obvious; it seemed infinite due to its large value. In 1676 *Ole Romer* was following the disappearance and appearance of Jupiter’s moon *Io*, and found a discrepancy with the calculated time, which was explained by the motion of the Earth on his trajectory around the Sun and with the time difference of travelling the light from *Io* to Earth between its closest and farthest point of this trajectory. In 1849 *Armand Fizeau* measured the speed of light with a rotating toothed wheel. The light passed through a gap between the teeth of the wheel, was reflected from a mirror in the distance, and was spotted with the observer. At a certain frequency of rotation of the wheel, the reflected ray met the next tooth on its route from the mirror and no light was seen. From the distance of the mirror and the frequency of the toothed wheel, the speed of light was calculated.

²⁵ The 2007 Recommendations of the International Commission on Radiological Protection. ICRP publication 103. Ann ICRP. 2007;**37(2-4)**:1-332. doi: 10.1016/j.icrp.2007.10.003. PMID: 18082557.

The speed of light in vacuum is the upper limit of the speed of an object, which is impossible to achieve with an object with non-zero mass. In materials (not vacuum), its value (v) is smaller, and the ratio of the speed of light in vacuum and in the given material is the refractive index (n) of the material.

The speed of light in a material depends on the electric and magnetic properties of the substance.

$$n = \frac{c}{v} \quad \text{eq. 342}$$

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$$

$$v = \frac{1}{\sqrt{\epsilon_r \epsilon_0 \mu_r \mu_0}} = \frac{c}{\sqrt{\epsilon_r \mu_r}} = \frac{c}{n} \quad \text{eq. 343}$$

$$n = \sqrt{\epsilon_r \mu_r}$$

In the case of paramagnetic and diamagnetic materials, when the relative permeability differs little from one, the refractive index is close to the square root of the relative permittivity of the given material.

$$n \approx \sqrt{\epsilon_r}$$

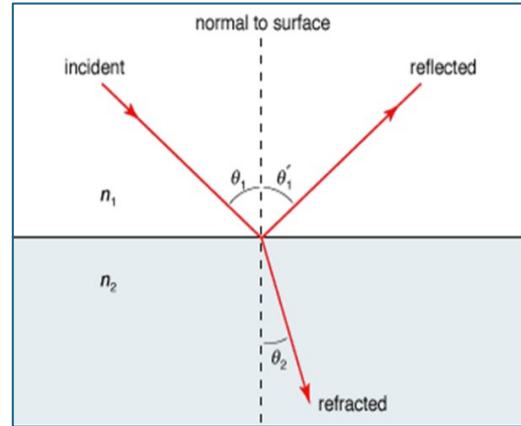


Figure 121. Reflection and refraction of a beam on the border of two phases.²⁶

eq. 344

The refractive index and the relative permittivity are frequency dependent. The previous equations are valid when the dielectric constant and the refractive index are measured at similar frequencies.

On the border of two phases, with different refractive indices, the incident beam is partially reflected to the first phase and partially refracted to the second phase. The originally straight beam trajectory is broken on the phase border (Figure 121). The relation between the refractive indices (n_1, n_2) and the incident and refracted angles (θ_1, θ_2) is the Snell law:

The product of the sine of the angle formed between the incident and refracted rays of light with the normal straight line, and the refractive index of the media (n_1 and n_2) is constant.

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad \text{eq. 345}$$

Angles θ_1 and θ_2 are measured according to the normal plane, that is, perpendicular to the surface. The refraction angle is defined with the ratio of the refractive indices of both phases and the impact angle of the ray:

$$\sin \theta_2 = \frac{n_1}{n_2} \sin \theta_1 \quad \text{eq. 346}$$

The sine is a monotone increasing function of the angle (argument) in the region between 0° and 90° , that is, a bigger angle means bigger sine value, an increasing impact angle means increasing refraction angle.

- The incident and reflected beams lay in one plane.
- The angles of the impact and reflected beams, measured according to the normal plane of the border, are the same.

$$\theta_1 = \theta_1'$$

- When the beam passes from an optically less dense phase to an optically denser phase ($n_1 < n_2$), the beam is refracted towards the normal plane, in the case of equal refractive

²⁶ Image source <https://www.britannica.com/science/Snells-law/images-videos>

indices, passes the border without changing direction ($n_1 = n_2$) and is refracted from the normal in the case that $n_1 > n_2$.

- When the incident beam lays in the normal plane (right angle impact, $\theta_1 = 0^\circ$), it crosses the border without refraction ($\theta_2 = 0^\circ$).

When the refraction is away from the normal plane ($\theta_1 < \theta_2$), exists an impact angle, when the refraction angle is 90° , the beam does not enter the second phase, but continues along the border of the two phases. This is the limit angle, called the critical angle. Higher impact angles cause total beam reflection. Total reflection is used in optical fibres (used in medicine in colonofiberscope, in information technology for data transmission, etc.). Optical fibres are produced from a core material with a high refractive index and coated with a material of low refractive index. The beam is almost always reflected from the border between the two materials and travels practically lossless along the fibre.

In the case of a polychromatic light (contains different frequencies, wavelength, colours), the refractive indices in a material are different for the different colours (wavelength, frequency). *The higher the frequency of light, the higher is the refractive index.* For example, blue light is refracted more than red light. That is the origin of the rainbow, when the white colour enters a raindrop, the different colours are refracted differently, and after a reflection on the backside of the drop and a refraction on the drop-air border, they leave the drop separately and allow the observer to see the components of the original white light.

16.1.1 Measurement of the Refractive Index

The refractive index can be measured with the Abbe's refractometer (Figure 122). The sample is placed in the cavity of the apparatus, between two prisms. The light is scattered by an upper illuminating prism and passes through the sample at different angles. The light is refracted in the sample and after that in the refracting prism, also, and lightens part of the shield. The border position between the lightened and dark parts of the shield determines the beam, which passes the sample in the cavity along the diagonal, and it is the beam, which reaches the farthest position on the shield. The angle of impact of the beam on the sample refracting prism border (θ_i) is determined by the geometry of the apparatus (the geometry of the cavity). The position of the dark and illuminated area border is measured and determines the angle of refraction angle (θ_r) in the refracting prism. The refractive index (n_r) of the refracting prism is known. The refractive index of the sample (n_i) is calculated from the Snell law (eq. 345):

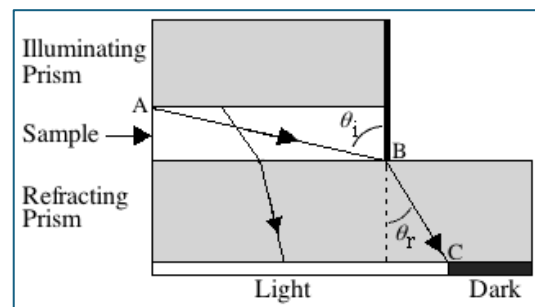


Figure 122. Schema of Abbe's refractometer.²⁷

16.1.2 Polarisation of Electromagnetic Radiation

Electromagnetic radiation is a transverse wave; the oscillation of the electric and magnetic vector of the radiation is perpendicular to the trajectory vector of the wave (Figure 123, left). The electric and magnetic vectors are mutually perpendicular as well. Electric (and magnetic) vectors are pointing in all possible directions around the axis of beam's

²⁷ Image source Manohar Reddy N., Kothandan D., Chandra Lingam S., Ahmad A.: *International Journal of Innovative Technology and Creative Engineering*, **2 (8)**, 23-25, 2012.

trajectory. This type of radiation is *unpolarized*. Under certain circumstances, one direction of electric (and magnetic) vector is kept, while all others are filtered.

The radiation becomes *polarised*. Unpolarized light can be changed to polarised:

- by birefringent through an anisotropic crystal.
- by a polarising filter;
- by reflection.

In a special case, when the refracted and reflected beams are perpendicular (90°) during the reflection on the border of two phases, the reflected beam becomes polarised, and the refracted beam is

slightly (not totally) polarized. This special impact angle is called the *Brewster angle*. Polarised light is also produced reflecting the sunshine by flat and smooth surfaces such as glass, snow, or water surfaces. To avoid becoming dazzled by the reflected light, sunglasses with polarisation filters are used. The polarisation filter offers a shelter against the polarised light (the reflected light from the surface) because the polarization plane of the beam differs from the polarization plane of the filter. It does not protect against direct sunlight, because direct sunshine contains all possible planes of the radiation, including the plane of the polarisation filter.

Optically active materials rotate the plane of polarisation of the polarized light. Optical activity of the materials relates to either the crystalline structure (temporary) of the materials, so the change in crystallinity causes change or diminishing of this feature or it relates to the molecular structure (permanent) of the compounds. These materials contain so-called chirality centres, carbon atoms with four different substituents.

The angle of rotation is proportional to the concentration of the optically active material and the optical path of the beam, i.e., the thickness of the liquid sample (measuring cell).

$$\alpha = [\alpha]_{\lambda}^t c d \quad \text{eq. 347}$$

Where $[\alpha]_{\lambda}^t$ is a material constant, c is molarity, d is the optical path. The molar concentration c is possible to replace with other types of concentrations (mass fraction w , W , in percent):

$$\alpha = [\alpha]_{\lambda}^t \frac{m}{MV_{\odot}} d = [\alpha]_{\lambda}^t \frac{m}{M\rho_{\odot}m_{\odot}} d = \frac{[\alpha]_{\lambda}^t}{M} w \frac{d}{\rho_{\odot}} = [\alpha']_{\lambda}^t w \frac{d}{\rho_{\odot}} = [\alpha']_{\lambda}^t W \frac{d}{100\rho_{\odot}} \quad \text{eq. 348}$$

$[\alpha]_{\lambda}^t$ and $[\alpha']_{\lambda}^t$ are the corresponding proportionality constants, the so-called molar or specific rotation constants, depending on the wavelength used and on the temperature. Using the previous equations, the concentration of the optically active substance can be calculated. The angle of the rotation of a polarized ray by the sample is measured with a polarimeter.

The light is emitted from the source, proceeding through a polarising filter, the polarising filter (polariser) produces polarised light which enters the sample tube, and the plane of the polarised light is turned proportionally to the amount of the optically active compounds. After that, the light enters another polarising filter (analyser), but the observer can see any light passing through the analyser only in the case where the analyser's and the polariser's mutual angle is set to the value of α , the angle of

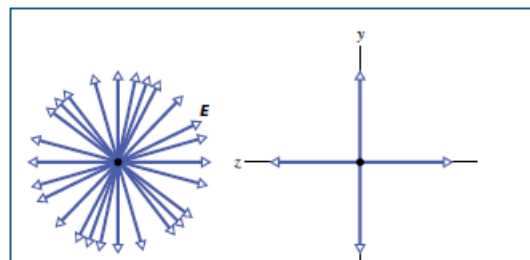


Figure 123. Vector lines of unpolarized (left) and polarized (right) beam.²⁸

²⁸ Image source <https://dilipbyella.files.wordpress.com/2020/10/combinepdf.pdf>

rotation of the plane of the polarised light by the optically active substance. The polarimeter scheme is on the Figure 124.

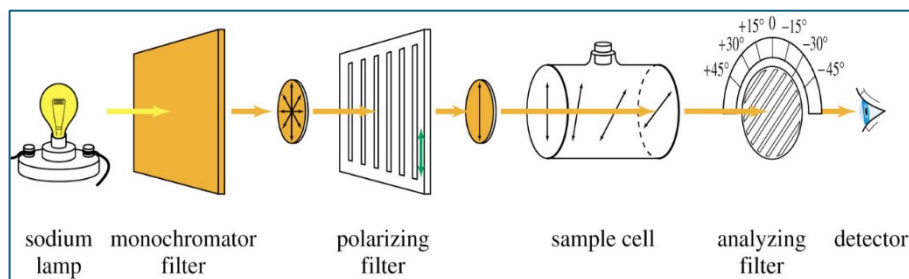


Figure 124. Basic components of a polarimeter.²⁹

16.2 Absorption of Ultraviolet and Visible Light –UV-VIS Spectroscopy

In the previous chapters, the behaviour of electromagnetic radiation on the border of two phases with different refractive indices causing reflection and refraction of the beam was discussed. Electromagnetic radiation also interacts with substances in the sense of absorption. The energy states of atoms and molecules are quantised, i.e., atoms and molecules cannot persist in an arbitrary energy state, and the differences between the available energy states are characteristic for the given material. Atoms and molecules absorb characteristic energy quanta and transfer to an excited energy state. Relaxation to the lower-energy state is deexcitation. The excitation (or deexcitation) energy provides information about the physicochemical character of the material and the intensity of the absorbed or released radiation about the amount of the material.

In the case of ultraviolet and visible spectroscopy, the samples are usually solutions, which absorb photons with energy in the visible or ultraviolet region (190-800 nm). This energy excites electrons from the bonding orbitals (σ - single bond, π - double bond) or non-bonding (n) to the anti-bonding orbitals (σ^* , π^*) of the molecules (Figure 125). The energy of excitation of these electrons is always significantly higher than that of the excitation of the vibration and rotation energy states, so the vibration and rotation excitations of the molecules follow the electron excitations. The function (curve) of the absorbance on the wavelength of the absorbed radiation is the *spectrum*. The UV-VIS spectrum of the molecules contains maxima and minima. It is a band spectrum, contrary to the electron spectra of atoms, which consist of separate lines, which correspond to the energy differences of basic and excited electron states of atoms. The atoms in the gaseous phase perform translational motions only, and there are no vibrational or rotational motions. In the case of molecular spectra, to every electron transition belong many vibrational transitions and many rotational transitions, whose energies slightly modify the electron transition energy and sum up to a band spectrum.

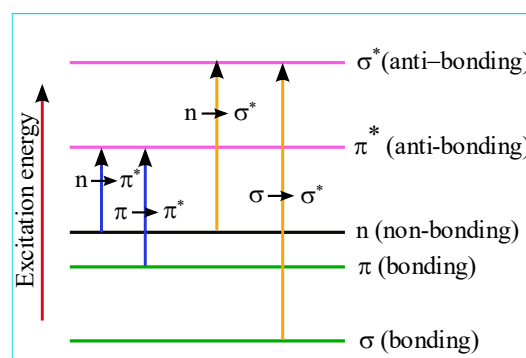


Figure 125. Types of molecular orbitals and excitation of electrons between them.

The position of the maximum of the spectrum is characteristic for the material studied under given experimental conditions (temperature, solvent, pH, etc.). The environmental conditions can shift the

²⁹ Image source <https://readchemistry.com/2022/09/28/optical-activity-in-organic-compounds>

position of the absorption maximum, decrease, or increase the excitation energy, when the basic and excited energy states are preferred in different measures in the modified conditions. The shift toward lower wavelength is called *bathochromic* (red) shift and toward higher energy, *hypsochromic* (blue) shift. For example, the change of solvent from nonpolar to polar in the case of a compound with double bond (π) and free electrons (non-bonding, n) affects the transition energy. The π and n electrons are located outside the connection between the bound atoms, contrary to the single (σ) bonds and are more affected by the change in the polarity of the solvent. The measure of sensitivity of the electrons to polarity is as follows.

$$n > \pi^* > \pi > \sigma$$

The intensity of the absorbance is proportional to the amount of material studied. UV-VIS spectroscopy is suitable also for quantitative analysis, to determine the concentration of the studied material. The absorbance (A) is defined according to the *Lambert-Beer law*, as the measure of decrease in intensity of the beam along the optical path, i.e., thickness of the sample (d or l).

$$A = -\log \frac{I}{I_0} = \epsilon cd \quad \text{eq. 349}$$

$$I = I_0 \cdot 10^{-A} = I_0 \cdot 10^{-\epsilon cd}$$

where I and I_0 are the intensity of the exiting and entering beams, respectively, and c is the concentration of the material studied. ϵ is the molar absorption coefficient (absorptivity), a constant characteristic of the given material at the given wavelength. The increase or decrease in absorbance at maximum is called a *hyperchromic* and, *hypochromic* effect and as follows from the Lambert-Beer law, it relates to the change of ϵ (units: $\text{m}^2\text{mol}^{-1}$).

The ratio of the intensities of the exiting and entering beams is the *transmittance*:

$$T = \frac{I}{I_0} = 10^{-A}$$

Sometimes expressed in percent,

$$T(\%) = \frac{I}{I_0} \cdot 100\% = 10^{-A} \cdot 100\% \quad \text{eq. 350}$$

16.3 Geometrical Optics

The main source of our information about the surrounding world comes from visual impulses. The image formed on a human retina is the interaction of beams reflected from or pathed through an object with the eye. The reliability of the information depends on the reliability of the formed image. In the case of certain distortion of the image, correction is necessary, which is usually done with glasses or contact lenses for sight corrections. Sometimes distortion is used to hinder or mask objects or make certain optical illusions. For the small size of the object or its large distance, we need to magnify the image to see it. Mirrors are used to visualise objects reflecting light. The mirrors used in households are flat-plane mirrors, where the purpose of their usage is to provide an image as like the object as possible (Figure 126). Mirrors with more complicated geometry and construction are used, e.g., to concentrate the rays to one point (parabola mirrors, torch, lighthouse) or divide a beam for interference

apparatus, etc. The reflection of light from the mirrors obeys the same rules as the reflection on the border of two phases with different refractive indices (see before) the impact beam and the reflected beams lays in the same plane, and the impact angle and the reflected angle are equal. The mirrors are constructed from non-transparent materials, which reflect the beam, without (significant) refraction or absorption, and create the image of the object with reflection.

Lenses are transparent materials that create images with refraction and obey the rules of refraction on the border of two phases with different refractive indices. The rays of light are straight-line spreading beams without significant scatter, and the object size should be larger than the wavelength of the wave. Simple rules of refraction are valid if the lenses used are thin.

Two types of lenses are basically used: *converging* and *diverging* lenses. Converging lenses have a *convex* surface and diverging lenses have a *concave* surface.

There are three characteristic rays, and an arbitrary two of them determine the image of the object. These three rays are: rays passing through the centre of the lens, rays parallel with the axis of the lens, and rays passing through the focal point of the lens. Important parameters that are required to find the position, orientation and size of the image are the focal length (f), object distance (o), image distance (i) and magnification (m) of the image. A real image is possible to project to a screen, to a retina; the image is behind the lens. In the case of an imaginary image, it is not possible to project it; it is before the lens.

The *lens equation* provides the relationship between the previously mentioned parameters:

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f} \quad \text{eq. 351}$$

The *magnification* is calculated from:

$$m = \frac{-i}{o} \quad \text{eq. 352}$$

$$|m| = \frac{\text{image height}}{\text{object height}}$$

The *power of the lens* is the reversed value of the focal length (measured in *metres*) expressed in units of *dioptre* (D, m^{-1}) and is used as a measure of the sight correction of glasses or contact lenses.

$$D = \frac{1}{f} \quad \text{eq. 353}$$

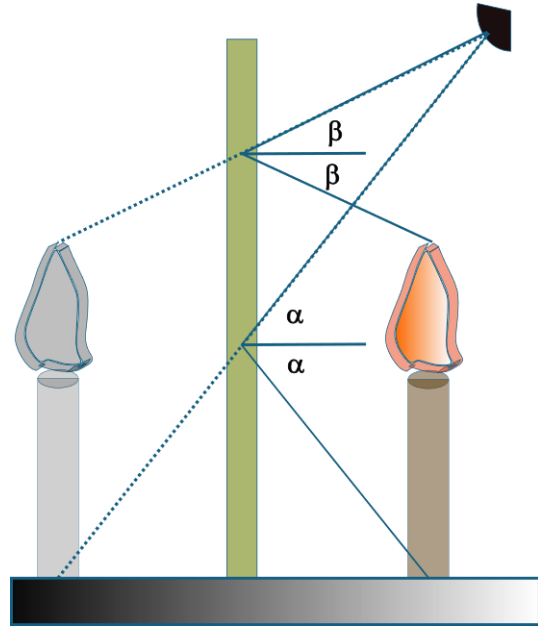


Figure 126. An image reflected by a mirror.

Converging lenses (Figure 127) have a positive focal length, so they have *positive* dioptries, while diverging lenses (Figure 128) have a negative focal length and dioptries.

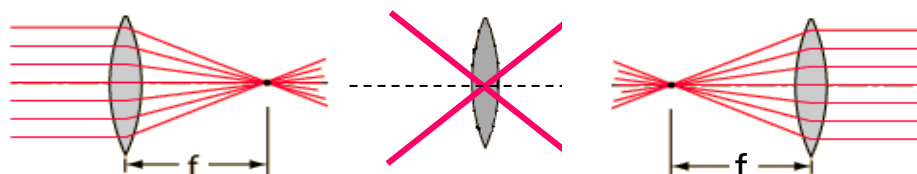


Figure 127. Parallel rays passing through a converging lens (left), rays passing through the centre of lens (middle), rays coming from the focal point of lens (right).³⁰

Converging lenses are lenses in which rays that enter it parallel to its axis, cross one another at a single point on the opposite side of the converging lens. The point at which the rays cross is called the focal point F of the lens. The distance from the centre of the lens to its focal point is called the focal length f . Rays passing through the centre of the lenses are transmitted without reflection (along a straight line). Rays passing through the focal point continue in the opposite side of the lens, parallel to the axis of the lens (Figure 127).

The focal length depends on the lens material (refractive index) and the geometry (radius of curvature) of the lens.

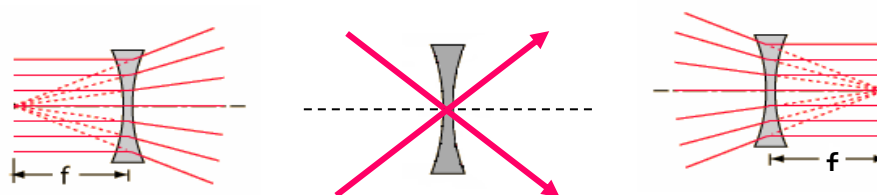


Figure 128. Parallel rays passing through a diverging lens (left), rays passing through the centre of lens (middle), rays pointing to the focal point of lens (right).³⁰

Divergent lenses (Figure 128) are lenses, in which rays that enter it parallel to its axis, diverge from each other on the opposite side of the lens. The paths of these rays cross each other in the focal point before the lens. Rays that pass through the centre of the lens continue without changing direction. The rays, which point to the focal point behind the lenses, continue their path parallel to the axis (Figure 128).

Conventions concerning the lenses.

1. We always assume that light travels through the lens from left to right, so the object is always located to the left of the lens.
2. The object distance is *positive* when the object is to the *left* of the lens. With our assumption in convention 1, it is always positive.
3. The image distance is *positive* if the image is on the *right* of the lens and *negative* if it is on the *left* side of the lens.
4. The focal length f of a converging lens is *positive*, whereas it is *negative* for a diverging lens.
5. The height of the object is positive if the object extends above the horizontal axis and negative if the object extends below.

³⁰ Image source <http://hyperphysics.phy-astr.gsu.edu/hbase/geoopt/foclen.html>.

6. The height of the image is positive if the image extends above the axis and negative if the image extends below.

From the fourth point comes the conclusion that the diopetre of a converging lens is positive and that of a diverging lens is negative.

Let us have an object (Figure 129, Figure 130, red rod) and a converging lens. The object distance from the lens (o) is larger than the focal length in the first case and less than the focal length in the second case. In the following part, we will find the image of the top point (blue dot), all other points are possible to find in a similar way. Due to the special shape of our object, the image is drawn as a perpendicular line from the top point to the x-axis. The image (top point, green points) is constructed using a ray passing through the top point and the centre of the lens (blue points) and the other ray is a parallel ray with the x-axis, passing through the focal point and crossing the lens at the height of the object (red points). In the first case (Figure 129), the image appears behind the lens (right-hand side), it is a real image. In the second case (Figure 130), the image appears before the lens, it is a virtual image. This image is above the x-axis, an non-inverse image, in the previous image is below the x-axis, it is a inverse image. Both images are bigger than the object, they are magnified.

To find the position (i) of the image, its magnification (m), and its orientation (inversed or not), it is possible to solve the analytical equations of lines (a slope, b intercept).

$$y = ax + b \quad \text{eq. 354}$$

The lens is placed on the y axis, its centre point coordinates are $[0,0]$. The coordinates of the top point of the object (red bar, blue point) are $[-o, h]$, the coordinates of the top point of the image (pink bar, green point) are $[i, h']$, the focal point coordinates are $[f, 0]$. The bottom of the object is on the x-axis as well as the bottom of the image. The coordinates of these points are $[-o, 0]$ and $[i, 0]$. The minus sign of the object position is due to the convention that the object distance is positive when the object is before the lens (left-hand side of the graph).

Let us find the equations of the lines that pass through the red and blue points and their intercept, which is the coordinate of the height of the image (green points).

Substituting the coordinates of the red points ($[0, h]$, $[f, 0]$) into the linear equation (eq. 354):

$$h = a \cdot 0 + b$$

$$0 = a \cdot f + b$$

so

$$b = h, \quad a = -\frac{h}{f}$$

and the equation is:

$$y = -\frac{h}{f}x + h \quad \text{eq. 355}$$

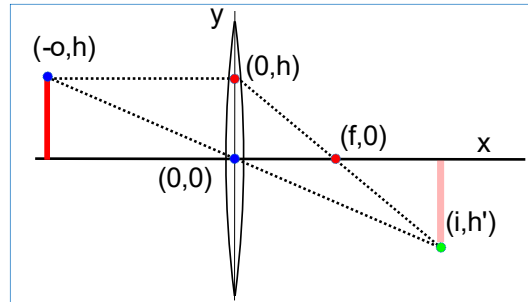


Figure 129. Construction virtual images of converging lens.

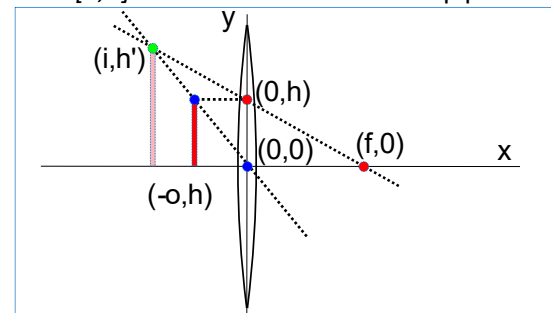


Figure 130. Construction of real images of a converging lens.

The equation of the second line comes from the blue points $([0,0], [-o, h])$:

$$\begin{aligned} 0 &= a \cdot 0 + b \\ h &= a \cdot (-o) + b \end{aligned}$$

so

$$b = 0, \quad a = -\frac{h}{o}$$

and the equation is:

$$y = -\frac{h}{o}x \quad \text{eq. 356}$$

For the coordinates of the intercept (green point), both equations (eq. 355, eq. 356) are valid; solving them as a system of two equations, we find the image coordinates (top point) $[i, h']$ from the parameters of the system.

$$\begin{aligned} -\frac{h}{f}i + h &= -\frac{h}{o}i \\ i &= \frac{1}{\frac{1}{f} - \frac{1}{o}} = \frac{fo}{o - f} \end{aligned} \quad \text{eq. 357}$$

Rearranging the equation, we receive:

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}$$

which is the previously introduced lens equation (eq. 351).

The height of the image h' is calculated from an arbitrary one of the eq. 355 or eq. 356 of the lines substituting i for x :

$$y = h' = -\frac{h}{o}i$$

The magnification of the image is as follows:

$$m = -\frac{h'}{h} = \frac{i}{o} \quad \text{eq. 358}$$

From eq. 357 follows that if the object's distance is greater than the focal length; $o > f$, the image distance $i > 0$ and vice versa (by definition, f and o are positive numbers). In the case of positive i , the position of the image is behind the lens (to the right of the lens, Figure 129), that is, the image is real and, in the case of negative i , before the lens (left-hand side), a virtual image (Figure 130).

From the equation:

$$h' = -h \cdot \frac{i}{o}$$

we can find out whether the image is inverted ($h' < 0$) or non-inverted ($h' > 0$). For a positive i (real image), h' is negative, that is, inversed, and for a negative i , h' is positive, the image is non-inverted.

The image is magnified when the magnification (m) is greater than 1 and reduced if it is less than 1:

$$\begin{aligned} |m| &= \left| \frac{h'}{h} \right| = \left| -\frac{i}{o} \right| = \left| -\frac{f}{o - f} \right| = \left| \frac{f}{f - o} \right| \\ |m| &= \left| \frac{f}{f - o} \right| \geq 1 \end{aligned} \quad \text{eq. 359}$$

Solving the previous inequality (eq. 359), the image is magnified when the object position o is from the interval $(0, 2f)$. The image is reduced when the distance from the lens is greater than $2f$. The magnification is *one* (the image size is the same as the object) if $o=0$ (the lens position) or $o=2f$ (twice the focal length). The case $o=f$, the magnification is infinite, and no image is created because the rays passing through the centre of the lens and the focal point (through blue and red points) are parallel. Substituting f for o into eq. 355 and eq. 356:

$$y = -\frac{h}{f}x + h$$

$$y = -\frac{h}{o}x = -\frac{h}{f}x$$

The slopes of both equations are identical, so they are parallel.

Divergent lens. In the case of a thin divergent lens, the focal distance is negative ($f < 0$), the focal point is to the left of the lens (Figure 131).

The lens is placed along the y -axis, with the centre in the $(0, 0)$ point. The procedure is analogous as in the case of converging lens; the two equations from the blue and the red points are the same as for the converging lens.

Analysing eq. 357, with conditions $f < 0$ and $o > 0$, i is always negative (negative numerator and positive denominator), the image is always *virtual*.

The height h' of the image is:

$$h' = -h \frac{f}{o - f} \quad \text{eq. 360}$$

The height of the image h' is always positive ($h > 0$, $o > 0$, $f < 0$), the image is always non-inverted. Magnification m is defined as follows:

$$m = \frac{h'}{h} = \left| \frac{i}{-o} \right| = \left| \frac{f}{f - o} \right| \quad \text{eq. 361}$$

The magnification is always $m < 1$, i.e., the image is reduced, because $f < 0$, $o > 0$, the denominator is greater than the numerator. The image created by the thin diverging lens is always non-inverted, virtual, and reduced.

In the case of sight correction with glasses or compact lenses, the near-sightedness is corrected with diverging lenses and the far-sightedness with converging lenses (Figure 132). In both cases, the image position is moved to the retina with these lenses instead of a spot behind or before the retina, respectively.

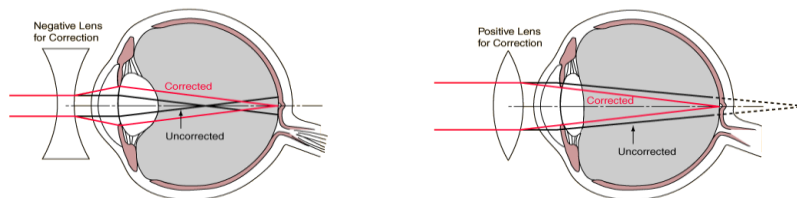


Figure 132. Sight correction with diverging and converging lenses.³¹

³¹ Image source <http://hyperphysics.phy-astr.gsu.edu/hbase/vision/eyedef.html>.

Alexander Búcsi

Pharmaceutical Physics

Published by Comenius University Bratislava in Comenius University Press

Edited by author

Manuscript 134 pages, 11,18 AH, 1st edition,

Published as an electronic publication

ISBN 978-80-223-6073-9 (online)



ISBN 978-80-223-6073-9 (online)