Mechanics & thermodynamics
Lecture Notes

D. Clouteau

April 5, 2019

Contents

As opposed to mathematics, any theory in basic ny scientific theory

Chapter 1

Measurements

A hypothetical scientific hypothesis is turned into a scientific theory when confirmed by measurements. Without measurements, no science and in particular no physics, just speculations. But what does *measuring* actually mean?

In a nutshell, *measuring* is a process that converts characteristics or features of an object or events into numbers. Numbers that can be used in mathematical models, processed and most importantly compared. Actually, the fact that most of the laws can be reduced to numbers is one of the most distinctive features of *physics*, in particular when compared to other branches of science. Indeed in chemistry, biology, life or social sciences classification is at least equally important. One can for instance mention Mendeleiev's table, DNA sequencing, species classification.

The major challenge in performing measurements is *objectivity*, that is to say, ending up with a number, or a range of numbers, that does not change depending on who or what is involved in the measurement process ; or at least with changes that can be controlled. Short distances around us had been conveniently measured in feet but which foot is the reference: mine, yours, the right foot of the King. Long distances used to be measured in days of travel from one location to another one. But how do we travel? Walking, riding a horse, sailing ? Summer or winter time ? These questions sound silly today but had been the topic of endless disputes and countless errors. Thanks to physicists who have been able to set up procedures and tools producing *objective results* we all agree about distance, times and other quantities that are *invariant with respect to the experimenter*. Actually what has been achieved in reality is somehow weaker: it is only an agreement amongst *experimenters* about procedures that have shown to be the most *invariant* in past practice. This agreement has often required very long discussions. To this aim, old and powerful institutions have been set up to reach this consensus and define standards. Amongst these one could mention the *International Bureau of Weights and Measures*, the *International Standard Organisation* (ISO) or the NIST. This activity has become so crucial for all sciences that it has become a science by itself, the science of measurements called *Metrology*.

1.1 Data: from apparatuses to laws of physics

At first sight, the most instrumental technique to reach an agreement amongst experimenters is to use the same apparatus to produce the sought numbers. Unfortunately, one single apparatus will not meet the need of performing all required measurements. Moreover, a given apparatus may exhibit some drift over time, or even worse some failure. Hence apparatuses are duplicated and thus have to be calibrated to ensure that they all measure the same thing; that is to say give the same number under the same conditions. Moreover, any original apparatus shows some

limitations, in terms of range or precision or experimental conditions. Some new and expectedly better devices have been proposed over time giving new but presumably different numbers that have to be compared or linked with the numbers given by formerly-used ones [1]. Hence comparison, calibration, conversion are amongst the many tasks involved in the framework of *Metrology* to reach reliable and objective measurements.

Following this line of thoughts, we could consider *metrology* as a completely independent field aiming at dealing with a set of apparatuses with the purpose of providing numbers or, as we would nowadays say, data to other sciences. This is certainly not true from an historical perspective. Indeed, what has been observed over centuries is the joint progress of data acquisition techniques and the development of new theoretical ideas and methods. To restrict the discussion to the scope of the present course, the development of Newton's laws would have been impossible without the availability of precise clocks, compas and weights. On the other hand, as it will be shown in chapters on thermodynamics, it turns out that the very definition of temperature is deeply embedded in the law of ideal gases, giving it its so-remarkably-simple form.

It is also worth noticing that, for a long time, theory and data have been considered more or less independently, data being used to demonstrate the validity of theoretical developments[2]. This was somehow reconsidered with the development of quantum physics when it appeared that at very small scale any measurement would modify the characteristic of the object/event at stake. Besides, from this perspective of an interweaving between data collection and theory, one should realise that until very recently, developments in physics have occurred in the context of a scarcity of data. Theoretical laws have been built and validated trying to take the best advantage of this small amount of available data. As opposed to this old and recent past, nowadays, *big data* has become the key-word and scientists are overwhelmed by flows of data challenging former scarce-data approaches. Hence new developments will most certainly emerge to account for this profusion of data.

1.2 Physical dimension and Units

Leaving science-fiction aside, let us come back to the great achievements of physics in dealing with measurements coming from so many different devices. As already mentioned, the outcomes of two different apparatuses, that is to say two numbers, may or may not be linked. In some cases, it can be inferred that they are measuring the same characteristic with different numbers. In other cases, the numbers given by two different apparatuses do not show any dependency. Two different entities have emerged over time:

- The first one is known as *physical dimension* carrying the essence of the measured characteristic disregarding the specific scale to quantify it. This is mainly classification.
- The second one is the specific *unit* used to produce a specific number. This number being

[1] Seismology and more precisely what has been called quantitative seismology is a relatively young science that was born with one apparatus, the Richter's sismometer. Based on the measurements given by this instrument, Richter derived a scale to measure the magnitude of earthquakes known as the Richter scale aimed at replacing the former intensity scale known as the Mercalli's scale. Note that the Mercalli scale is of ordinal type with 12 levels labelled by roman numbers from I to XII characterizing the damage caused by the earthquake in a specific area. To mimic this former scale the Richer's scale uses real positive number obtained combining the logarithm of the measured acceleration with the distance from the epicenter of the earthquake. Many decades later it has been shown that this scale can be mapped onto the energy released during the earthquake. However this mapping is biased by filtering effect of the Richter apparatus. Hence even today, each time an earthquake occurs, the magnitude reported by different labs are not the same depending on applied corrections, calibrations or methods.

[2] Actually, this was the true scientific revolution introduced by *Galilei Galileo* who resorted to experiments to validate a theory. Before that time *data* were not acquired but simply given by the surrounding world and scientists would *speculate* on this. This was the time of *natural philosophy*.

nothing but the ratio with respect to the chosen *reference unit*. The *unit* can be considered as customary, cultural and practical, but yet essential in allowing to convert a physical characteristic into a number.

For instance *Length* is a physical dimension conventionally denoted by [L]. This *Length* can be expressed using different units: *meters*, *inches*, *yards* or *parsec*, depending of the ruler at hand or the country or community you belong to. The same stands for *Time*, another *physical dimension* denoted by [T] and that can be measured in *seconds*, *hours* or *whispers* for a musician.

The original idea of *physical dimension* is attributed to *Louis-Joseph Fourier* who claimed that the law of physics should be valid whichever particular system of units is used. In other words, he stated that laws of physics have to be *invariant* with respect to the particular units used in these mathematical formulae[3]. Though independent from specific units since *Fourier*, laws of physics are bound to satisfy coherency rules in the way physical dimensions are combined in mathematical expressions as it will be detailed in Section 1.3. In particular two different physical dimension, for instance *Length* and *Time* cannot be added. But they can be divided giving either *velocity* with dimension $[\mathrm{LT}^{-1}]$ or *slowness* with dimension $[\mathrm{TL}^{-1}]$.

From a practical point of view, the output of any apparatus measuring a given physical dimension, *Length* for instance, may be converted in any convenient unit valid for this physical dimension. However, the unit itself has to be explicitly mentioned together with the measured number. In this regard, any measure in physics, that is to say the number measuring the sought characteristic is a ratio with respect to a *reference unit* that is kept somewhere. Ten times the length of my 1-meter ruler is expressed as 10 *meters* or in short-hand notation 10 [m].

1.3 Dimensional Analysis

Though any number of units can be defined, actually as many as existing apparatuses, there are only five different *basic physical dimensions* for mechanics and thermodynamics all together. These *basic physical dimensions* are:

- Length [L],
- time [T],
- mass [M],
- Temperature $[\Theta]$
- Amont of substance [N]

and two additional ones to cover all physical dimensions:

- Electric charge [Q]
- Luminous intensity [J].

Hence any physical quantity P used in this course will be a combination of these basic dimensions taking the following form:

$$[\mathrm{P}] = [\mathrm{M}]^{\alpha}\ [\mathrm{L}]^{\beta}\ [\mathrm{T}]^{\gamma}\ [\Theta]^{\delta} = [\mathrm{M}^{\alpha}\mathrm{L}^{\beta}\mathrm{T}^{\gamma}\Theta^{\delta}] \tag{1.1}$$

where exponents α, β, γ and δ are signed integers[4]. For instance:

[3] From a mathematical point of view, *physical dimensions* are *equivalence classes* of units that can be one-to-one mapped with any other unit of the same class.

[4] Note that this could be extended to rational numbers for some particular quantities.

- *Velocity* already encountered being a *Length* travelled during a given period of *Time* is:

$$[\mathrm{V}] = [\mathrm{LT}^{-1}]$$

- *Acceleration* being the rate of change of velocity with respect to time is:

$$[\Gamma] = [\mathrm{LT}^{-2}]$$

- *Momentum* being the product of mass and velocity is

$$[\mathrm{P}] = [\mathrm{MLT}^{-1}]$$

- *Force* being defined by Newton's second law as equal to the rate of change of *momentum* with respect to *Time* is:

$$[\mathrm{F}] = [\mathrm{MLT}^{-2}]$$

- *Work* done being the product of a *Force* by the *Length* over which it applies is:

$$[\mathrm{W}] = [\mathrm{ML}^2\mathrm{T}^{-2}]$$

  and so is *energy*

$$[\mathrm{E}] = [\mathrm{ML}^2\mathrm{T}^{-2}]$$

This simple algebra, reducing every quantity to this very small number of *basic physical dimensions* is extremely useful in physics. It first allows to check for errors in your mathematical derivations. Indeed in pure mathematics there is no harm in writing:

$$a = b + c$$

and there is no simple way to make sure that the derivation leading to this formula is correct or not. On the contrary in physics a, b, and c must have the same *physical dimension* and even the same *unit*. Hence formulae have to satisfy *dimensional homogeneity* as a first step to their correctness. Please, always check it yourself as this will be the first thing done by others, including the person in charge of grading your work.

As a byproduct of this reasoning, it becomes obvious that in any formula involving particular functions such as *exponential*, *logarithm*, *sine*, *cosine* or any other special function, the argument of any of these should be dimensionless[5]. Once again, this is a good trick to check for some errors in your results.

Actually, this analysis using the physical dimension of various quantities has shown to be more powerful than simply checking for inconsistency in formulae. It is a tool to find the basic structure of the results - up to a multiplicative constant - without solving the governing equations.

[5] Indeed how would you obtain the exponential of the 1-meter platinum-irridium bar kept in Paris. you can divide it in any number of pieces, you can replicate it to take the square but following the well known series expansion of this function:

$$\exp x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$$

x cannot have a physical dimension as all terms in the series would have different dimensions and could not be summed up.

1.3.1 Dimensional Analysis

As a simple example, let us consider the period T_o of a pendulum as a function of its mass m length l under gravitational acceleration g. Since we are looking for a *Time* that only appears in g of physical dimension $[\mathrm{LT}^{-2}]$, we already know that the period should scale as $\frac{1}{\sqrt{g}}$. Moreover since *Length* has to be eliminated we end up with:

$$T_o = \alpha\sqrt{\frac{l}{g}}$$

We have no clue about the value of the factor α but it is likely to be of the order of 1. Actually it is equal to 2π for small amplitude vibrations of the pendulum. But what is even more striking is that this very simple dimensional analysis already shows that the period of a pendulum does not depend on its mass, but only on length and gravitational acceleration g. The same reasoning holds for free fall, leading to the result that the falling time does not depend on the mass of the object - provided that friction with air is neglected.

This dimensional analysis is condensed into the so-called *Vaschy-Buckingham* or *π-theorem* stating that:

Theorem 1 *Any physical phenomena involving n physical quantities may be reformulated as a system of equations involving $p \leq n$ dimensionless parameters $\pi_{i=1,p}$, where p in the dimension of the kernel of the $4 \times n$ dimensional matrix built with the exponents in formula 1.1 for all original variables.*

These *dimensionless parameters* are extremely useful from both the theoretical and experimental points of view. Indeed they allow to perform a reduced set of possibly smaller-scale experiments having the same dimensionless quantities as what would be required by full-scale experiments, significantly reducing the cost of an experimental campaign. From the theoretical point of view, some of the *dimensionless parameters* may appears to be small enough to simply be neglected in solving the problem giving access to simpler and almost as accurate models.

From an historical perspective, basic discoveries in mechanics and thermodynamics were done without the use of dimensional analysis which is attributed to *Joseph Fourier* and developed by *James Clerk Maxwell*. In the late XIXth century, early XXth, Lord Reynolds 1842-1912 showed the major contribution of dimensional analysis in *Fluid Mechanics* introducing the so-called Reynolds number:

$$\mathcal{R}_e = \frac{vL}{\nu}$$

where v is the velocity of the flow, L a characteristic length of the body and ν is the kinematic viscosity of the fluid. This number expresses the ratio between kinetic energy and the work of viscous forces.

1.4 Units and standards

As already mentioned, various units are used to perform data-acquisition. However, some standards have to be defined to allow a proper communication between scientists and beyond, between any user of these data. Thanks to dimensional analysis, only six basic standards have to be defined for *Length*, *Time*, *Mass*, *Temperature*, *Electric Charge* and *Luminous Intensity*.

This was the idea of the original Metric-System proposed during the French Revolution based on physical facts[6]:

[6] the idea was actually much older as in 1664 John Wilkins proposed to define the meter as the length of a pendulum beating at 2 seconds. The value would have been 0.993m.

- *Time*-unit was set based on the solar day or the solar year duration,
- *Length*-unit was linked to the size of the earth - one *meter* being defined as one 1/10000 of one quarter of an Earth's meridian,
- *kilogram* as an unit *mass* was related to the weight of one liter of water at its maximum density (at 4°C);
- *temperature* was defined from freezing and evaporation points of water at atmospheric pressure.

However over time these definitions have appeared not having the required precision. For a long time, a specimen of platinum and iridium was kept in Sèvre and was assumed to be *one-meter* long at 0°C. This specimen had been replicated in different places after careful comparison with the original one. However in 1960 the decision was made to use the speed of light, known to be constant, to define the reference unit of length.

Definition 1 *1 meter is* $\frac{1}{299,792,458}$ *the distance travelled by light in one second.*

Hence the speed of light is *by definition* fixed to:

$$c = 299,792,458 \text{ [m/s]}$$

Similarly as solar day or tropical year of year 1900 as appeared to be too imprecise and unpractical for a standard, the present definition is:

Definition 2 *One second is the time interval of 9,192,631.7700 oscillation periods of a Cesium atomic clock.*

Surprisingly the one-kilogram standard is still a piece of platinum-iridium kept in vacuum in Paris. However this might be changed in future fixing forever the value of the Planck constant h. Indeed, h being an *action* $[\mathrm{ML^2T^{-1}}]$, fixing its value would give an unequivocal definition of the mass, unit length and time standards being fixed[7].

These three Units: *meter, second, kilogram* give access to the definition of all over units commonly used:

- *Joule* [J] for energy being $[\mathrm{kg\ m^2 s^{-2}}]$
- *Newton* [N] for force being $[\mathrm{kg\ m\ s^{-2}}]$
- *Pascal* [Pa] for pressures being a force by unit surface: $[\mathrm{kg\ m^{-1} s^{-2}}]$
- *Watt* [W] for power being $[\mathrm{kg\ m^2 s^{-3}}]$

It turns out that an absolute scale for temperature is somehow more difficult as it is not so obvious that we can add two *temperatures* as we do for *Length*, *Time* or *Mass*. Actually, it will be seen in Chapter 11 that *Temperature* being an *intensive* quantity, concatenation is not allowed as it is for *Length*, *Time* or *Mass*. It will also be shown that *absolute temperature* is strongly related to the law of ideal gases scaled with two values: the absolute 0 and the triple point of water at 273.16°K. It is called the *Kelvin temperature*. It is then related to a third universal constant k known as the Boltzmann Constant, giving the rate of increase of internal energy with temperature:

$$K_B = 1.38064852 \times 10^{-23} \text{ [kg m}^2\text{s}^{-2}\text{K}^{-1}\text{]}$$

[7] $h = 6.626070040(81) \times 10^{-34}$ [kg m^2/s]

1.5 Orders of magnitude

It is somehow surprising that these Unit standards can cover so many phenomena observed in nature with so many different ranges of magnitudes. Indeed the radius of a electron, provided that it can be defined in a classical sense, is less than 10^{-18} [m] whereas the size of the universe is something of the order of 3.10^{23} [m]. As far as time is concerned, the transfer time of light over the diameter of an atomic nucleus is 10^{-25} [s] whereas the age of the universe is 5.10^{17} [s]. The mass of an electron is 9.110^{-31} [kg] and the mass of a galaxy is of the order of 10^{42} [kg]. As pointed out by the French philosopher *Blaise Pascal* in his *thoughts* we are lost between two infinities, the infinitely small and the infinitely large. However, it turns out that, whichever meaningful quantity we might choose, nothing goes beyond 10^{100}. Yet, it is impressive to have been able to derive laws covering such large ranges of scales and hundred digits is a lot to deal with in practical applications.

Hence when exposed to a new problem, we first have to figure out the orders of magnitude we are dealing with, that is to say the magnitude of the exponent of 10 we have to consider, in particular calibrating our measurement apparatus. Then we will be able to find the best unit to choose following the uniform pattern given in table 1.5 using a little bit of ancient Greek symbols.

Name	symbol	value
atto	a	10^{-18}
fermto	f	10^{-15}
pico	p	10^{-12}
nano	n	10^{-9}
micro	μ	10^{-6}
milli	m	10^{-3}
unit		1
kilo	k	10^{3}
Mega	M	10^{6}
Giga	G	10^{9}
Tera	T	10^{12}
Peta	P	10^{15}

Table 1.1: Magnitudes of physical units

However in some cases we may have to deal with measurements varying over several orders of magnitude. When trying to display them on a single graph we will find out that only a few values are visible on the graph and most of them will disappear, either cluttered together along some axis or left outside of the graph. More annoying is the following fact. If we choose to plot the period of an event in *secondes* and some other team would show the same experimental results in terms of frequency of this event measured in *Hertz* we will end up with two very different plots putting the emphasis on very different scales: large periods for us, meaning undistinguishable low frequencies for the other team and large frequencies for them leading unnoticeable tiny periods on our graph.

Two related observations are worth mentioning. From the perspective of order of magnitude, 10 is midway between 1 and 100 whereas its linear distance to 1 is only 9 and it is 90 to 100; 10 times less. This remains true when normalising the distance with respect to the original value 10. Now adding physical dimension and considering 10 [Hz] being mid way between 1 [Hz] and 100 [Hz], we may choose to measure things in periods ; effectively the inverse of the frequency. Then $10Hz$ becomes 100 [ms], 1 [Hz] is 1000 [ms] and 100 [Hz] becomes 10 [ms]. Hence, using a linear scale, even rescaled with respect to 100 [ms], now the distance from 1 [Hz]($1s$) to 10 [Hz](100 [ms]) is 10

times larger than the one from 100 [Hz](10 [ms]) to 10 [Hz](100 [ms]). Hence, as far as non-negative physical quantities are concerned, it seems that from both practical and theoretical points of view - plotting results or defining a distance that does not depend on whether we choose period of frequency - a standard linear scale is not the best suited.

And actually, most scientific publications ressort to logarithm-scales when showing the results. Though looking at first sight a bit awkward with its non equally spaced grids, these scales are not so unfamiliar in every day life. Indeed sound is commonly measured with a logarithm-scale units called dBA: Acoustic decibel. Musicians use octaves for tones : meaning doubling the frequency between two subsequent octaves. Richter magnitude for earthquakes is a scaled logarithm of the energy released along the seismic fault.

As already mentioned, taking the logarithm of some physical quantity cannot be done without first re-normalising this quantity with some reference value. For instance, sound level is measured in decibel as follows:

$$L = 20 \log_{10} \left(\frac{P}{P_o} \right) dB$$

where P is the acoustic pressure and P_o is the smallest audible change in pressure [8].

It is often argued that experimentalists plot their results on a logarithmic scale in order to obtain more or less aligned results minimising fluctuations. And indeed it can be argued that a log-scale allowing to show several orders of magnitude, tends to reduce the fluctuation amplitudes on a chart. However there is a much more fundamental reason as to why results tend to be aligned on a log-log chart. In most cases, this is nothing but the graphical image of the dimensional analysis presented in Section 1.3 (see Fig. 1.1 for an example in seismology).

1.6 Errors and uncertainties

Measurements are inherently prone to errors and uncertainties that have to be properly quantified when reporting the measured values. People have even suggested to revert the definition of measurement as "A set of observations that reduce uncertainty where the result is expressed as a quantity"[9]. Hence a major goal in physics is to adequately control and reduce these errors and uncertainties by several means:

- setting more accurate standards,
- better controlling the experimental set-up,
- resorting to skilled experimentalists,
- replicating the experiments.

Conventionally two different types of errors are considered: *systematic* and *statistical* errors. The underlying assumption is that *systematic* errors are due to weaknesses in the experimental set-up which could be corrected whereas *statistical* errors are what is left when all state-of-the-art corrections have been applied.

[8] This means that the regulatory limit of sound level of 90dB corresponds to a pressure $10^{4.5} \approx 30,000$ times the reference pressure P_o.

[9] Douglas Hubbard: "How to Measure Anything", Wiley (2007), p. 21

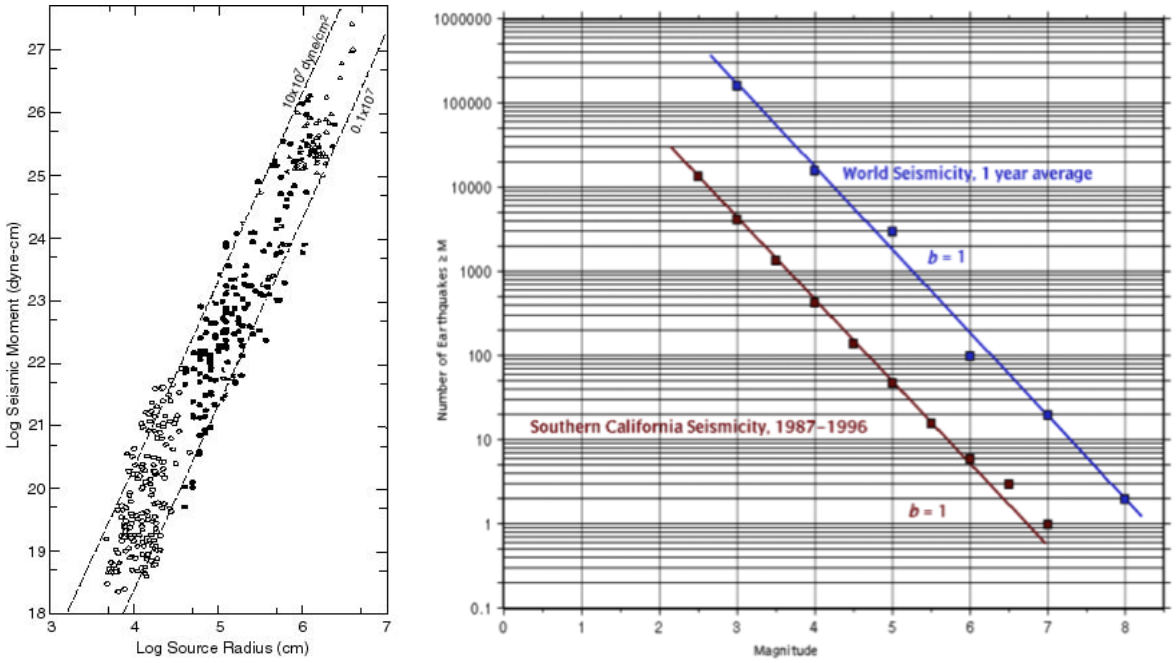


Figure 1.1: (left) Seismic moment M_o of recorded earthquakes as a function of the size L of the active fault showing a power law of the form $M_o = \Delta\sigma L^3$ where $\Delta\sigma$ is the stress-drop on the fault during the earthquake (After Hanks, 1977). (Right) *Gutenberg-Richter law* on the number of earthquakes as a function of magnitude ($M_w = 2/3 \log_{10}(M_o/M_{\text{ref}})$) for southern California and the entire world. This chart illustrates the unpredictability of earthquake triggering mechanics: the likelihood of one earthquake of magnitude 6 is equal to the likelihood of 10 earthquakes of magnitude 5.

Part I

Fundamentals in classical mechanics

Chapter 2

Kinematics

2.1 Vectors and inertial frames

Surprisingly, though classical mechanics had been developed without any use of *vectors*, most modern physics courses give a central part to this mathematical object. Actually, this is the result of a long process, or even a quest, in physics toward invariance and symmetry as pointed out by *Richard Feynman* in his Modern Physics course[1]. R. Feynman wanted to name the corresponding chapter *Symmetry* but finally opted for *Vectors* as a more conventional title. The main advantage of using *vectors* for both students or full-fledged physicists is that this account for invariance provides simpler and more generic laws, that are valid for all observers whatever the frame of reference they consider. Hence further detailed computations can be postponed to later times or even outsourced to others, engineers or computers.

In Chapter 1 the emphasis was put on measured physical quantities, being of different natures or dimensions and being measured using various scales with respect to varied references. These quantities are expressed as a set of numbers or *scalars* either positive or signed and the corresponding *units*. Since in most cases phenomena cannot be described with a single quantity, several of them have be handled at a time. This is obviously the case when some object has to be located in space. When you are in a foreign city looking for a drugstore or a train station you are not only interested in time or distance to reach the place but you actually need *directions*: 2 blocks East, 3 blocks North, 2nd floor is a valid answer when 10 [minutes] or 1 [km] will not give you all the clues to reach the place. Here is the main reason for introducing vectors: not getting lost, getting to the point. Vectors are meant to provide both *amplitude* and *direction*.

On the first hand, amplitude has already been introduced in the previous chapter as a positive number. On the other hand *direction* is something that can intuitively be understood but that is far less easy to quantify, that is to say to put numbers on it. Counting blocks in East-West/North-South direction may be a convenient way in some North-American cities, less convenient in Paris, Old Delhi or Tokyo. The GPS position given in terms of latitude and longitude angles by your smartphone may be considered a more convenient way to reach the sought-for place. Hence how many numbers are required ? what are the relevant quantities and units ? Blocks or Angles ? What is the range for each of them ? It turns out that the answer is neither unique nor straightforward, though aiming at defining the same unique location. Hence a valid-for-all generic tool is required.

Vectors have appeared to be the most convenient mathematical tool to give an unique and generic definition for all observers and postpone the precise definitions of the actual numbers or

[1] Feynman Lectures: Chapter 11.

coordinates used by each of them. In this regard it meets the invariance requirement of physicists: the invariance with respect to the observer. Readers are referred to Chapter **??** for an introduction to the mathematical framework of *vectors*. Next section focusses on the use of *vectors* to characterise the motion of an object with respect to space and time, which is the aim of *kinematics*. Before this *Frames of Reference* are defined with a special emphasis on *Inertial frames* with respect to which the *Principle of Inertia* is assumed to hold.

2.1.1 Frame of reference

Characterizing the motion of an object relies on *rulers* and *clocks* introduced in Chapter 1 and specified by procedures edicted by the *Bureau of International Standards*. To this aim, let us define a *frame of reference*.

Definition 3 *A* frame of reference *denoted by $\mathcal{R}$ is made of:*

- *3 identical straight* rulers *measuring length with a given precision Δd. These rulers are called* axis.
- *each one of them is perpendicular to the other two and they are all glued together at the given point "O" called the* origin *of the frame.*
- *Axis are labelled such that the first one - the x-axis - is transformed in the second one - y-axis - when rotating it by 90°*anti-clockwise *around the third axis - z-axis. With such an orientation, the* reference frame *is a* direct frame.
- *A clock measuring* time *is added to the frame and reset to* 0 *at the beginning of the experiment.*

Any point P in space is thus uniquely characterised by its three *coordinates* (x_P, y_P, z_P) in this *frame of reference*. The *motion* of any object M of *small size*[2] is fully characterised by the three coordinates of the point at which it is located at time t:

$$M(t) : (x_M(t), y_M(t), z_M(t))$$

x_M, y_M and z_M being real-valued functions of time t. When they take a constant value over time object M is motionless or fixed with respect to the frame of reference $\mathcal{R}$. When two of the three components are constant, the motion of the object is *one-dimensional*. This case is studied in following section before being generalised to two and three-dimensional motions in subsequent sections[3].

2.1.2 Principle of inertia

Opposing *Aristotle*'s mainstream views at that time, *Galileo Galilei* was the first to postulate that:

Principle 1 *An object moving with a constant velocity will keep this velocity as long as nothing acts on it.*

[2]Let us keep this loose definition of *small size* assuming that the size of the object is of the order of the measurement precision Δd. A more axiomatic view would be to consider point objects, also referred to as *point mass*.

[3]Note that the motion could still be one-dimensional with three time-varying components. What is the relation between these components in this case?

This was actually pure speculation since this *principle of inertia* somehow contradicts our common-sense. Indeed in every day life any object tends to stop after some time but Galileo claimed that it slows down and stops because the environment acts on it and not because of an internal essential feature of the object itself. *Galileo*'s view, though contradicting common sense, turned out to be more effective than *Aristotle*'s one. In particular Newton with his laws of motion studied in Chapter 3 gave a more precise and effective statement of this principle[4].

One striking consequence of this so-called *principle of inertia* is that the *velocity* of a body with respect to a given frame of reference is not the inherent quantity a physicist should be interested in as it is defined up to an additive constant depending on the chosen frame of reference. Indeed, when taking another *frame of reference* $\mathcal{R}'$ that is moving with a constant velocity with respect to the first *frame of reference* $\mathcal{R}$ the velocity of an object in uniform motion takes a different but constant value. It could even be at rest with respect to $\mathcal{R}'$ when $\mathcal{R}'$ properly chosen. Hence the laws of physics and more particularly the laws of mechanics should be invariant with respect to two *reference frames* one moving at a constant velocity with respect to the other one[5]. Hence, only the change of velocity, and in particular accelerations, can be considered as invariant over *Frames of Reference* at least for the so-called *inertial frames*.

Definition 4 *A reference frame in which* Galileo's principle of inertia *is valid is called an* inertial frame *or* Galilean frame.

As a consequence:

Property 1 *any frame moving with a constant velocity with respect to an* inertial frame *is also an* inertial frame.

2.2 Velocity and acceleration for one-dimensional Motion

We assume in this section that the motion of the body is along one single axis, the x-axis for convenience so that we can ignore the two other axes. Think for example of a car travelling on a straight road equipped with milestones and yardsticks. Knowing x_M for all time $t \geq 0$ we can plot the graph of $(t, x_M(t))$. It is expected to be a continuous function as we do not expect the car to jump instantaneously from one location to another distant one.

2.2.1 Velocity

The first question we may ask is how *fast* the body moves *with respect to the frame of reference*. Let us assume that we can measure its position in the chosen *frame of reference* at time t and then at time $t + \Delta t$ where Δt is the time delay of our apparatus measuring the distance from origin. the signed distance - positive or negative -travelled during this time delay is:

$$\Delta x = x_M(t + \Delta t) - x_M(t)$$

[4]From an epistemological point of view, *Galileo* is acknowledged as the first modern scientist, claiming that we can be misled by our perceptions of the surrounding world. *Descartes* developed his *method* based on *Galileo's ideas*. Actually, Galileo's statement turns out to be a little too restrictive. Indeed, *inertia* also refers to the mass of the body and what remains constant over time is not necessarily the velocity itself. It could be any quantity depending linearly on the velocity - and on mass - and in particular the *momentum*. Chapter 6 comes back to this.

[5]Note that even for reference frames, we cannot define an absolute velocity but only the velocity of one frame or one object with respect to another one. Note also that this invariance property is experimentally true only when the relative velocity between the two frames is much smaller than the speed of light. When it is not the case the times given by the two clocks in the two frames do not coincide as conjectured by Einstein. In the remaining part of this course it is always assumed that relative velocities are small enough to avoid this relativistic correction.

The *averaged velocity* defined as the travelled distance divided by time lapse Δt:

$$v_{\Delta t}(t) = \frac{\Delta x}{\Delta t}$$

is a first attempt to characterise how fast the body moves. This *averaged velocity* is a signed physical quantity expressed in [length-unit/time-unit]. It could be in [m/s] or [km/h] or more exotically in [mph]. It can be positive, negative or nil depending on whether the body moves towards the positive x, the negative x or is at rest with respect to the *frame of reference*. Note that the sign of the velocity is what gives the direction of a one-dimensional motion. $v_{\Delta t}$ is a function of time t - the time of the first measurement - but also depends on the time delay Δt. Taking smaller and smaller time delays, we expect to obtain some value that does not depend anymore on the time delay of the experiment. This mathematical abstraction is called the *instantaneous velocity*:

$$v_x = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} \stackrel{\text{Def}}{=} \frac{dx}{dt}$$

and mathematically speaking it is the derivative of $x(t)$ with respect to t[6].

A motion with a constant velocity v_o is defined as an *uniform motion*. The travelled distance is then obviously proportional to the travel time yielding:

$$x_M(t) = x_o + v_o t$$

where x_o is the location at the beginning of the experiment ($t = 0$).

2.2.2 Acceleration

Following *Galileo*'s reasoning, it appears that external actions on an object can be observed by recording the change in its velocity. For a single body moving along a given fixed axis, this means changes of its velocity with respect to time. More precisely the rate of change of velocity with respect to time is the *acceleration*. In particular, the *average acceleration* during time lapse $\Delta t'$ is defined as:

$$a_{\Delta t'}(t) = \frac{\Delta v}{\Delta t'} = \frac{v(t + \Delta t') - v(t)}{\Delta t'}$$

whereas the *instantaneous acceleration* is obtained taking the limit for vanishing $\Delta t'$:

$$a_x(t) = \lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t'} = \frac{dv}{dt} = \frac{d^2x}{dt^2}$$

Hence, *instantaneous acceleration* is a measure of the action applied on a body. When it vanishes, no action is applied.

From a practical point of view in order to be able to compute the *average acceleration* from experimental measurements of distances at different time instants, at least three measurements are required. Taking equal time lapses the three instants are $t - \Delta t$, t and $t + \Delta t$ leading to the so-called *Euler scheme*:

$$a_{\Delta t} = \frac{x(t + \Delta t) - 2x(t) + x(t - \Delta t)}{(\Delta t)^2}$$

Accelerometers are devices recording acceleration, thus quantifying in a sense the actions applied to them.

[6]Note that this quantity may not always be defined at all time t and may also be discontinuous with respect to time. This is the case of a bouncing ball at the very instant it reaches the ground. Velocity jumps instantaneously from a non-positive value to a non-negative one. The absolute values of these two velocities being equal in the case of an elastic rebound.

Motion with constant acceleration a_o is the simplest motion we could think of appart from an uniform motion. Integrating once over time gives access to a linearly varying velocity:

$$v_x(t) = a_o t + v_o$$

where v_o is the velocity at time $t = 0$. Integrating once more yields:

$$x(t) = \frac{1}{2} a_o t^2 + v_o t + x_o$$

that is to say a parabola with respect to time. The extremum - minimum for $a_o > 0$ and maximum for $a_o < 0$ - is reached when velocity vanishes at:

$$t_{\text{extr}} = -\frac{v_o}{a_o}$$

with extreme value:

$$x_{\text{extr}} = -\frac{v_o^2}{2a_o} + x_o$$

Motion with given acceleration is obtained integrating twice with respect to time, once to obtain the velocity:

$$v_x(t) = \int_0^t a(\tau) d\tau + v_o$$

and a second time for the motion itself:

$$x(t) = \int_0^t v(\tau') d\tau' + x_o = \int_0^t \left(\int_0^{\tau'} a_x(\tau) d\tau \right) d\tau' + v_o t + x_o$$

Nowadays, each and every smartphone has a set of accelerometers embedded inside. Hence, selecting one given direction, you have access to $a_x(t)$ and you can thus integrate to find the current position of your smartphone provided that you know the initial position x_o and the initial velocity v_o. However due to error with respect to both initial position, initial velocity and accuracy of the accelerometer a significant drift in position will be observed over time as shown in Figure 2.2.2.

2.2.3 Phase portrait and invariance

Based on *Galileo's principle of inertia* we can conclude that the motion of a body - actually an ideal point-mass - is fully characterised by the current position and velocity and future actions that will be applied on it. Hence representing position and velocity on the same graph is likely to give some insight about the applied actions. This $(x(t), v(t))$ graph is called the phase portrait that can be more or less complex depending on the phenomenon at stake.

A body at rest has a phase diagram restricted to a single point $(x_o, 0)$

A body with constant velocity has a phase diagram consisting of an horizontal line $v = v_o$

A body with constant acceleration has a parabolic phase diagram. Indeed:

$$t = \frac{v - v_o}{a_o}$$

yielding:

$$x = x_o + \frac{v^2 - v_o^2}{2a_o}$$

Hence we can detect one invariant in this motion that is:

$$\frac{v^2}{2} - a_o x = \frac{v_o^2}{2} - a_o x_o = \frac{E_o}{m} \tag{2.1}$$

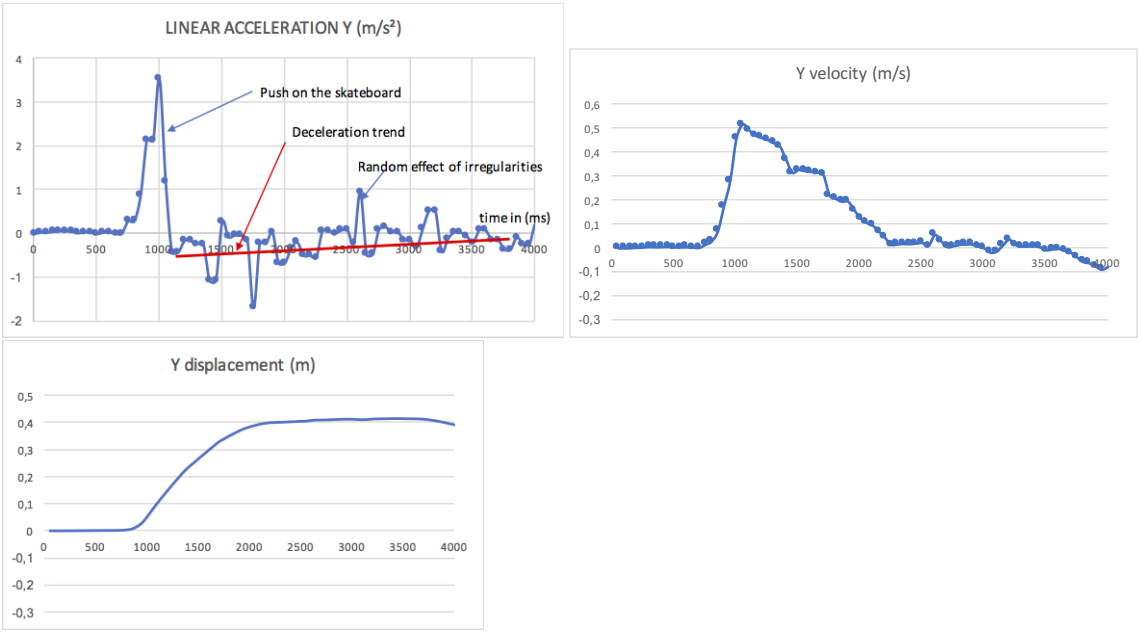


Figure 2.1: A smartphone is taped onto a skateboard. The skateboard is pushed forward and then stops after a while. (Top-left) noisy recorded acceleration. (Top-right) Velocity integrated from acceleration. (Bottom-left) Displacement integrated from recorded acceleration.

drawing a parabola along the x-axis with extremum value:

$$x_{\text{extr}} = x_o - \frac{v_o^2}{2a_o}$$

Sinusoidal motion has an elliptical phase diagram. Indeed:

$$x(t) = x_o \cos(\omega t), \qquad v(t) = -x_o \omega \sin(\omega t), \qquad a(t) = -x_o \omega^2 \cos(\omega t) = -\omega^2 x(t)$$

Hence:

$$\omega^2 x(t)^2 + v(t)^2 = \omega^2 x_o^2 = \frac{E_o}{m}$$

which draws an ellipse on the phase plane or even a circle when properly rescaled. This also reveals a time-invariant for this motion.

Exponential motion has a radial phase diagram. Indeed:

$$x(t) = x_o \exp(-\lambda t), \qquad v(t) = -x_o \lambda \exp(-\lambda t) = -\lambda x(t)$$

2.3 Two- and three-dimensional motions

Dealing with cartesian coordinates to track two- or three-dimensional motions is straight forward as the two other components of the motion $(y(t), z(t))$, of the velocity $(v_y(t), v_z(t))$ and of the acceleration $(a_y(t), a_z(t))$ can be handled separately. Actually, introducing vectors is even more appealing as it generically deals with different coordinate systems.

Let us denote by $\overrightarrow{r}(t)$ the position vector of a given object of small size with coordinates with respect to the given *reference frame* $\mathcal{R}$ being $(x(t), y(t), z(t))$.

The *Velocity-vector* is then a time-dependent vector defined as:

$$\overrightarrow{v}(t) = \frac{d\overrightarrow{r}}{dt} = \lim_{\Delta t \to 0} \frac{\overrightarrow{r}(t+\Delta t) - \overrightarrow{r}(t)}{\Delta t}$$

where the difference between two vectors is well-defined together with the multiplication by a scalar $1/\Delta t$.

Acceleration-vector is also a time-dependent vector defined as:

$$\overrightarrow{a}(t) = \frac{d\overrightarrow{v}}{dt} = \lim_{\Delta t' \to 0} \frac{\overrightarrow{v}(t+\Delta t') - \overrightarrow{v}(t)}{\Delta t'}$$

One can also write the *acceleration vector* as the double time-derivative of the position vector $\overrightarrow{r}$:

$$\overrightarrow{a}(t) = \frac{d^2\overrightarrow{r}}{dt^2}$$

Uniform motion is fully characterised by a constant velocity vector $\overrightarrow{v_o}$ and the motion is given by:

$$\overrightarrow{r} = \overrightarrow{v_o}t + \overrightarrow{r_o}$$

where $\overrightarrow{r}_o$ is the position vector at time $t = 0$ and does not depend on time. Another way to define an uniform motion is to assume that:

$$\overrightarrow{a}(t) = \overrightarrow{0}$$

meaning that all components of the *acceleration-vector* vanish.

2.3.1 Motion with constant acceleration

It can be easily verify that a motion with constant acceleration vector $\overrightarrow{a_o}$ exhibits the following linear and parabolic evolution of velocity and position vectors respectively:

$$\overrightarrow{v}(t) = \overrightarrow{a_o}\, t + \overrightarrow{v_o} \tag{2.2}$$

$$\overrightarrow{r}(t) = \frac{1}{2}\overrightarrow{a_o}\, t^2 + \overrightarrow{v_o}\, t + \overrightarrow{r_o} \tag{2.3}$$

where $\overrightarrow{v}_o$ and $\overrightarrow{r}_o$ are respectively the velocity and position vectors at time $t = 0$. It can even be noted that since $\overrightarrow{a_o}$ is constant and presumably not vanishing, this vector is associated to one particular direction in space, for instance the z-axis:

$$\overrightarrow{e_z} = \frac{\overrightarrow{a_o}}{\|\overrightarrow{a_o}\|}$$

$\overrightarrow{v_o}$ is another direction that may or may not be colinear with $\overrightarrow{a_o}$. Hence let us define:

$$v_{zo} = \overrightarrow{v_o} \cdot \overrightarrow{e_z}, \qquad , \overrightarrow{v_\perp} = \overrightarrow{v_o} - v_{zo}\overrightarrow{e_z}$$

Hence, choosing for convenience $\overrightarrow{r_o} = \overrightarrow{0}$, the motion is parabolic in time with respect to direction $\overrightarrow{e_z}$ and is linear or constant with respect to all other perpendicular directions:

$$\overrightarrow{r}(t) = \left(\frac{1}{2}a_o t^2 + v_{zo}t\right)\overrightarrow{e_z} + \overrightarrow{v_\perp}\, t$$

Assuming that $\|\overrightarrow{v_\perp}\| = v_{xo}$ does not vanish and defining:

$$\overrightarrow{e_x} = \frac{\overrightarrow{v_\perp}}{\|\overrightarrow{v_\perp}\|}$$

yields:

$$x = v_{xo}\, t \tag{2.4}$$
$$z = \frac{a_o}{2v_{xo}^2}\, x^2 + \frac{v_{zo}}{v_{xo}}\, x \tag{2.5}$$

showing that the trajectory is a parabola in the (x, z) plane. Note that $\frac{v_{zo}}{v_{xo}}$ is nothing but $\tan\alpha_o$ with α_o the angle of the initial velocity vector at $t = 0$ on the (x, z) plane.

2.3.2 Motion with non-constant acceleration

Assuming the non-constant acceleration-vector $\overrightarrow{a}(t)$ is known, the motion is obtained integrating twice the acceleration vector with respect to time and accounting for initial velocity and position vectors $\overrightarrow{v}_o$ and $\overrightarrow{r}_o$:

$$\overrightarrow{r}(t) = \int_0^t \overrightarrow{v}(\tau')d\tau' + \overrightarrow{r_o} = \int_0^t \left(\int_0^{\tau'} \overrightarrow{a}(\tau)d\tau \right) d\tau' + \overrightarrow{v_o}t + \overrightarrow{r_o}$$

2.3.3 Acceleration induced by an uniform circular motion

An uniform circular motion can easily be represented using a Cartesian frame by means of sin ans cos functions. Indeed, denoting by ω the angular velocity, the components read:

$$x(t) = R\cos(\omega t + \phi_o) \tag{2.6}$$
$$y(t) = R\sin(\omega t + \phi_o) \tag{2.7}$$
$$z(t) = z_o \tag{2.8}$$

where $R = \|\overrightarrow{r}(t)\|$ is the constant radius of the trajectory and ϕ_o is the angle at $t = 0$. Hence the components of the velocity vector are:

$$v_x(t) = -\omega R\sin(\omega t + \phi_o) \tag{2.9}$$
$$v_y(t) = \omega R\cos(\omega t + \phi_o) \tag{2.10}$$
$$v_z(t) = 0 \tag{2.11}$$

satifying:

$$\|\overrightarrow{v}\| = \sqrt{v_x^2 + v_y^2 + v_z^2} = \omega R$$

Hence the amplitude of the velocity or in other words the speed along the circular trajectory, is constant. However the acceleration does not vanish as:

$$a_x(t) = -\omega^2 R\cos(\omega t + \phi_o) \tag{2.12}$$
$$a_y(t) = -\omega^2 R\sin(\omega t + \phi_o) \tag{2.13}$$
$$a_z(t) = 0 \tag{2.14}$$

or choosing $z_o = 0$:

$$\overrightarrow{a} = -\omega^2\overrightarrow{r}$$

with constant amplitude $\|\overrightarrow{a}\| = -\omega^2 R$.

As a conclusion, a constant *speed*, that is to say a constant amplitude of the velocity vector does not lead to a vanishing acceleration. Actually when *speed* is constant, the acceleration is normal to the velocity. Indeed, taking the time derivative of:

$$\|\overrightarrow{v}\|^2 = \overrightarrow{v} \cdot \overrightarrow{v} = \text{cst}$$

leads to:

$$2\overrightarrow{a} \cdot \overrightarrow{v} = 0$$

Chapter 3

Newton's Laws of motion

The three Newton's laws are certainly one of the master piece of science. They are routinely used to analyse the solar system, the evolution of our planet and its atmosphere, including climate change, to design cars, trains, ships and aircrafts and even gene therapies. These laws and the related concept of *forces* are so well-known and have been so often used for so long time that it tends to hide the basic principles and definitions beneath. Hence, beyond reviewing these laws of motion and defining forces, this chapter and Chapter 6 on *momentum* will try to shade some light on theses important underlying principles.

3.1 The original Newton's laws

The English translation of Newton's laws originally written in Latin in *Principia Mathematica* published in 1687 reads:

Law I: Every body persists in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by force impressed.

Law II: The alteration of motion is ever proportional to the motive force impress'd; and is made in the direction of the right line in which that force is impress'd.

Law III: To every action there is always opposed an equal reaction: or the mutual actions of two bodies upon each other are always equal, and directed to contrary parts.

3.1.1 Law II: Momentum and forces

Words used by Newton, actually their translation in English, may be misleading. Referring to the original text let us first notice that *motion* is to be understood as *quantity of motion* or *momentum*. It is not pure kinematics. Mass, and more precisely *inertial mass* plays a central part here. Indeed, Newton gives the definition of the *quantity of motion* in the introduction to *principia mathematica* mentioning that it doubles when the *quantity of substance*, namely *mass*, doubles for a fixed velocity; and that it also doubles when velocity doubles for a fixed *quantity of substance*. Hence the *quantity of motion* is the product of *mass* and *velocity* and *Newton* gave a way to properly define it in mathematical terms.

Definition 5 *Since velocity is a* vector*, so is the* quantity of motion *or* momentum *defined by:*

$$\overrightarrow{p} = m\overrightarrow{v} \tag{3.1}$$

Hence in the Second Law, *Alteration of motion* has to be understood as the *rate of change of momentum* with respect to time.

Principle 2 (Second Law) *In mathematical terms this second law reads:*

$$\frac{d\overrightarrow{p}}{dt} = \overrightarrow{F} \tag{3.2}$$

From an epistemological perspective, the second law gives the first non-ambiguous definition of *Forces* though Newton still uses the more ambiguous term *Action* in the third law[1]. *Forces* can now be fully characterised as *vectors* and measured as change of *momentum* with respect to time. Beyond precisely defining *forces*, Newton's second law also gives a way to define and measure this so-far unclear *quantity of substance* called *mass* - and more precisely *inertial mass* as:

This intimate property of a body that opposes any change in its velocity*, whatever the external cause for this change.*

Indeed, considering a one-dimension motion, a force f_1 measured using the change in momentum of a given body of known acceleration a_1 and mass m_1 satisfying:

$$f_1 = m_1 a_1$$

can be applied to another body. Measuring the induced acceleration a_2 of this second body it can be inferred that:

$$f_1 = m_2 a_2$$

leading to:

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

3.1.2 Law I: Inertial frames

Surprisingly, Law I appears to be a particular case of Law II as it states that when no *force* is "impressed" on the body its *momentum* remains constant, and so does the velocity - the mass of the body is assumed to be fixed. As acknowledged by Newton himself, Law I is a restatement of the *principle of inertia* proposed by *Galileo* at the beginning of the XVIIth Century and recalled in Chapter 2. So is Law I simply a tribute paid by Newton to Galileo or is there something more fundamental in it, only implicitly stated by Newton ? Actually, Law I should be understood as the definition of a class of *frames of reference* in which the two other laws are valid. These frames are called *inertial frames*. Hence, Law I can be rephrased as:

Definition 6 *An* inertial frame of reference*, is a frame of reference in which* every body persists in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by force impressed.

3.1.3 Law III: *Actio-Reactio* or conservation of momentum

As far as the third law is concerned, let us start with an anthropological perspective. In a Western frame of mind, human beings are supposed to master *Nature*. Hence it is common-sense to say that when jumping, the Western body applies a force on Earth, instancing its power. For Native-Americans, *Earth* pushes up the foot of the jumper allowing her/him to jump. Almost two

[1] Note that today *action* has a precisely defined meaning in physics that is different from a *force*. It is the product of a *momentum* by a distance.

$M_1 = 3.4$ [kg], $M_2 = 5.2$ [kg]

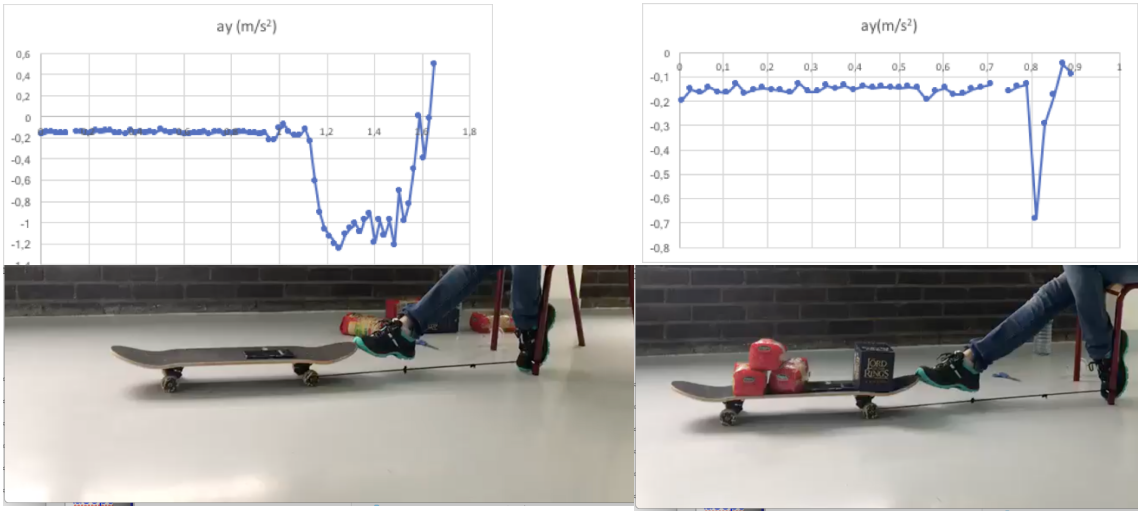


Figure 3.1: Experimental evidence of Newton's second Law done by *Hiba Belkadi and Raymond Diab*. $\frac{M_2}{M_1} = \frac{5.2}{3.4} = 1.5$, $\frac{a_1}{a_2} = \frac{1.2}{0.7} = 1.7$, $\frac{\Delta r}{r} \approx 10\%$

centuries after Christopher Colombus, Newton came to bring a perfect balance between these two perspectives stating with his third law that both are correct. Being a mathematician, he added that the two forces are equal in amplitude and opposite in direction so that the sum of the two force-vectors cancels out:

$$\overrightarrow{F}_{1\to 2} + \overrightarrow{F}_{2\to 1} = \overrightarrow{0}$$

A modern physicist would say that Newton brought back the *symmetry* in the interaction between two bodies, or the *invariance* with respect to which body is acting on the other one.

And indeed, this third law is a consistency condition for *interaction forces* in order to satisfy a more fundamental law of physics: *conservation of linear momentum* as discussed in Chapter 6.

3.1.4 Concluding remarks

As a conclusion on the original Newton's Laws of motion, let us not forget to make sure that the chosen *frame of reference* is an *inertial frame* when applying the second law. At least, up to the required precision of the sought results, since we do not know a single *true inertial frame of reference*[2].

Then the second law can be used to either compute the *applied external force* when the time-evolution of *momentum* is known or compute the change in *momentum* when the *applied forces* are known. The last and more intricated case is when *applied forces* are known as a function of the motion itself as studied in the next section.

Finally let us remark that Newton's Laws remain rather vague concerning the definition of bodies on which these laws apply. In Chapter **??**, it will be show that it applies to any body/system as long as it is restricted to the motion of the so-called *centre of mass*. Until then we will restrict its use to *point-masses*.

[2]Newton claimed in his *Principia mathematica* that the centre of the universe is such a frame of reference. Unfortunately is is nowadays conjectured that such a centre does not exist and Einstein has strongly revised Newton's laws with the theory of relativity.

3.2 Forces

Up to now, *Forces* have been introduced as *vectors* to restore the *balance of momentum* when the *momentum* happens to change in terms of either amplitude or direction. However experimental results give more substance to some *forces* being *invariant* whichever body they apply on. For instance the force applied by a spring shows such an invariance and depends only on how much the spring has been stretched or compressed. *Gravity* happens to affect bodies as a function of their mass and their location in space. Hence the concept of *force fields* is developed to define forces that will applied on a body when at a given location in space. As opposed to *gravity* and other *force fields*, *Viscous forces* happen to depend on the velocity of the body and some of its geometric features. Finally *Contact forces* are partly determined by the *balance of momentum* itself, and some constraints related to *friction*. These basic types of forces are briefly described in this section.

3.2.1 Elastic force

Applying forces on an elastic body effects reversible changes of its shape. Moreover, when these deformations are small enough their amplitude varies linearly with respect to the amplitude of the force, other parameters being kept constant. As a consequence, with proper calibration, measuring the deformations of a body is a way to quantify the amplitude of the applied forces. Hence apparatuses to measure forces have been designed along these lines instead of resorting to Newton's second law itself. However since, on the first hand, deformations of a body are of several types: compression, extension or shear, amongst others, and, on the second hand, forces have different directions, these apparatuses, called *dynamometers* are carefully designed to select only one particular type of deformation correlated to a specific direction of the force. The simplest example is a spring for which the axial force F happens in first approximation to be proportional to the change in length l of the spring. It is known as the empirical *Hooke's law*:

$$F = -k(l - l_{\text{eq}}) \tag{3.3}$$

where l_{eq} is the free-length of the spring and k is its stiffness with I.S. unit in [N/m]. The minus sign indicates that the force applied by the spring pulls back the attached body when the length of the spring is increased and pushes back when the spring is compressed. This ensures that the spring will go back to its original length l_{eq} where the applied force is released. *Forces* of this sort are also referred to as *restoring forces*.

Accounting for the direction of the force and denoting by $\overrightarrow{e_r}$ the unit directing vector of the spring at time t pointing to the end where the force is applied, the induced force vector reads:

$$\overrightarrow{F} = -k(l - l_{\text{eq}})\overrightarrow{e_r}$$

3.2.2 Weight in an uniform gravity field

Weight is the force induced on a body by gravitation. It is a well-established experimental fact that this *weight* is proportional to the *mass* of the body m and can be easily measured hanging the body to a calibrated *dynamometer* of stiffness k and measuring its change in length Δl:

$$W = k\Delta l$$

Doubling the mass the body will double the change in length, thus doubling the weight. However, weighing the same body on Earth and or the Moon will result in two different changes in length

of the spring resulting in two different *weights* for a body of a given mass m. Noticing that the *weight* is actually a force-vector $\overrightarrow{W}$ with a given amplitude and direction we can finally state that:

$$\overrightarrow{W} = m\,\overrightarrow{g} \tag{3.4}$$

where $\overrightarrow{g}$ is called the *acceleration of gravity* as it has indeed the physical dimension of an acceleration[3]. Hence an *acceleration of gravity* $\overrightarrow{g}$ induces an acceleration of $\overrightarrow{a} = \overrightarrow{g}$ of the body subjected to this only one force, all other external forces being excluded. This result holds whatever the mass of the body. This is major counter-intuitive result already suggested by Galileo that all objects fall in the same way whatever their mass when subjected to gravitation only.

It is another experimental fact that at ground level on earth, the *acceleration of gravity* is very much constant[4] and of the order of 10 [ms^{-2}]. As a result of last chapter, the trajectory of any body subjected to a constant gravity is a parabola depending only on the initial velocity and position:

$$\overrightarrow{r} = \frac{1}{2}\overrightarrow{g}t^2 + \overrightarrow{v_o}t + \overrightarrow{r_o} \tag{3.5}$$

When considering higher precision or motions of larger amplitude in the direction of gravity, its variation in space becomes significant and has to be accounted for by means of a *field of gravitational force.*

3.2.3 Fields of forces

On a sunny day, crops in a field usually have more or less same height and same direction - opposing the gravity field. Hence this crop field is a good image and even a kind of biological measure of the constant gravity field on this specific part of planet Earth. Taking a picture of this field during a storm reveals that some crops are bent in different directions by gusts of wind. Some may even be laid down. Considering each of the crops as images of vectors on a given surface, you will get a good idea of what a non-constant *force field* looks like on a surface[5]. You can also think of the hair of the student in front of you. It is another image of a force field on a sphere. Now you have to visualise how it would look like filling the entire space with arrows.

Mathematically speaking, a *force field* is a vector-valued function of space coordinate-vector $\overrightarrow{r}$ denoted by $\overrightarrow{F}(\overrightarrow{r})$. It is then a vector-valued function of a vector. In a cartesian frame, it comes down to a set of three functions (F_x, F_y, F_z) depending on three variables (x, y, z). Hence:

$$\overrightarrow{F}(\overrightarrow{r}) = (F_x(x,y,z), F_y(x,y,z), F_z(x,y,z))$$

As a consequence a body at position vector $\overrightarrow{r_1}$ will experience a force $\overrightarrow{F}(\overrightarrow{r_1})$ and a body at $\overrightarrow{r_2}$ will be subjected to a force $\overrightarrow{F}(\overrightarrow{r_2})$.

Newton has shown that the gravitational force between two massive bodies with mass m and M respectively is:

$$\overrightarrow{F}_{M\to m} = -\frac{\mathcal{G}mM}{r^3}(\overrightarrow{r_m} - \overrightarrow{r_M}), \qquad r = \|\overrightarrow{r_m} - \overrightarrow{r_M}\|$$

[3] More precisely, using the balance of momentum for a body only subjected to gravitation yields:

$$m_{\text{iner}}\,\overrightarrow{a} = \overrightarrow{W} = m_{\text{g}}\,\overrightarrow{g} \qquad \Rightarrow \overrightarrow{a} = \overrightarrow{g}$$

when assuming that the *gravitational mass* m_{g} is equal to the *inertial mass* m_{iner} which is an experimental fact with a relative precision of 10^{-10}.

[4] with actually tidy fluctuations of different length scales that are used to analyse the heterogeneities of the underground and earth interior.

[5] Note that the length of the individual crops could also vary to give a more accurate image of an non-uniform force field.

where $\mathcal{G}$ is the gravitational constant with value 6.67408 10^{-11} [$m^3kg^{-1}s^{-2}$]. Hence, taking our reference frame at the location of the body of mass M ($\overrightarrow{r_M} = \overrightarrow{0}$), this body induces in the surrounding space a gravity field:

$$\overrightarrow{g}(\overrightarrow{r}) = -\frac{GM}{r^2}\overrightarrow{e_r} \qquad \text{with} \qquad \overrightarrow{e_r} = \frac{\overrightarrow{r}}{r} \tag{3.6}$$

Since, whatever the location of mass m in space $\overrightarrow{g}$ is directed towards the the other mass M put at centre of the reference frame, it is call a *central-force field*. Gravity fields induced by extended bodies will be studied in Chapter 8, together with planetary motions.

The *Coulomb's force field* induced by an electric charge Q on an other electric charge q is also a *central force field* given by:

$$\overrightarrow{F}_{Q\to q} = \frac{1}{4\pi\epsilon_o}\frac{qQ}{r^3}(\overrightarrow{r_q} - \overrightarrow{r_Q}), \qquad r = \|\overrightarrow{r_q} - \overrightarrow{r_Q}\|$$

where ϵ_o is the dielectric constant. Note that, as compared to gravity that is always an attractive force, the *Coulomb's force* is either attracting for charges of opposite signs or repulsing for charges of same polarity. Note that Laplace's forces induced by a magnetic fields are not-central-force fields and also dependent on the velocity.

The force induced by a spring centered at point O can be modeled as a central force field since:

$$\overrightarrow{F}(\overrightarrow{r}) = -k(r - l_{\text{eq}})\overrightarrow{e_r}, \qquad r_{\min} < r < r_{\max}$$

3.2.4 Viscous forces

Viscous forces play a significant role in fluids and Newton wrote several chapters of his treaty to discuss these forces. Indeed, it has to be remembered that the *principle of inertia* was extremely controversial at that time since it was in contradiction to common-sense and experimental results: a feather does not fall at the same speed as a rock. Revealing viscous forces was a major attempt to find an agreement between the *principle of inertia* and experimental facts.

The simplest empirical model for a *viscous force* is a force opposing the motion with an amplitude that is linearly proportional to the velocity of the body. This model is valid for low velocities or more precisely low Reynolds numbers. This proportionality factor is called the *viscosity coefficient* usually denoted by c. This coefficient depends on both the fluid itself and the shape of the body.

$$\overrightarrow{F}_{\text{visc}} = -c\overrightarrow{v}$$

As opposed to elastic forces, this force - actually the work done by this force - is never restored and is dissipated in the fluid as discussed in subsequent chapters. It is then called a *dissipative force*. Note that for higher velocities - higher Reynolds numbers - this force also called the *drag force* is proportional to the square of the speed. Note also that this *drag force* is not the only force induced by the fluid flow around the body, components perpendicular to the velocity can also be induced, in particular the *lift* on an airplane or a wing.

3.2.5 Contact forces and friction

Contact forces between two bodies are very much dependent on the two surfaces in contact. Hence very few generic things can be stated about these forces. They are even partly determined by the balance of momentum itself, especially when the relative velocity between the two bodies vanishes. When sliding between the two bodies occurs, it appears that the interaction force does not depend too much on the sliding speed though it always opposes the relative motion. The strength with

which it opposes the motion happens to be proportional to the strength with which the two sliding surfaces are kept in contact. Intuitively a heavy block is more difficult to slide than a lighter one. The contact surface also plays a significant role. Sliding on ice requires a force of much smaller amplitude than the corresponding force-amplitude when sliding on sand-paper. This *empirical law* is known as the *Colomb's law* relating the normal and the tangential components of the interaction force.

For a given sliding interface with normal unit-vector $\overrightarrow{n}$ ($\|\overrightarrow{n}\| = 1$) the two components of the interaction force $\overrightarrow{R}$ are: the normal component N along $\overrightarrow{n}$ and the tangential force $\overrightarrow{T}$ being perpendicular to $\overrightarrow{n}$. When sliding occurs the direction of the tangential force $\overrightarrow{T}$ is along the opposite direction of the relative velocity vector $\overrightarrow{v}$ and is linearly proportional to the amplitude of the normal component. Since the proportionality coefficient is a dimensionless quantity it is usually taken as the tangent of the so-called *friction angle* α. This angle is nothing but the limit angle between the reaction force $\overrightarrow{R}$ and the normal unity vector of the interface $\overrightarrow{n}$. In mathematical terms the contact-friction or *Coulomb's Law* then reads:

$$\overrightarrow{R} = N\overrightarrow{n} + \overrightarrow{T} \qquad \text{and} \qquad \overrightarrow{v} \cdot \overrightarrow{n} = 0 \tag{3.7}$$

$$\text{no sliding:} \qquad \|\overrightarrow{T}\| < \tan\alpha \; N \qquad \text{and} \qquad \overrightarrow{v} = 0 \tag{3.8}$$

$$\text{sliding:} \qquad \overrightarrow{T} = -\tan\alpha \; N \frac{\overrightarrow{v}}{\|\overrightarrow{v}\|} \tag{3.9}$$

The simpler case of sliding without friction is given by setting $\alpha = 0$ leading to $\overrightarrow{T} = \overrightarrow{0}$ in all cases (sliding or non-sliding).

Chapter 4

Work and Energy

Work and *energy* are extensively-used words in day to day life, not only about basic sciences and technology but also economics, social and political sciences. Indeed terms such as *energy mix* or *work force* do not belong to physicists' parlance. Psychology and education could also be mentioned and students sometimes claim that they have had no *energy* to do their home *work*. Actually this statement has some ground from a *physics* perspective as it is shown in this chapter that no *work* can be done by a system without depletion of its *internal energy*[1].

Focussing on Mechanics, *work* takes a very practical and simple form. It basically consists in measuring the amount of weights lifted to some height or moved over some distance. When we push a stone by one meter and someone else pushes it by two meters, we can say that he has done twice more work than we have. Now if the stone is twice heavier it means that moving it by the same distance amounts to twice more work. Then if this twice heavier stone is slid on ice by my friend when mine is deeply stuck in mud I am pretty sure that my friend has not done twice more work than I have. Having carefully studied last chapter on Newton's Laws of motion and forces, we can say that twice more work is performed over the same distance only if the applied force is twice larger.

This discussion has more or less grasped the main concept of *work* that will be discussed in this chapter. *Work as the product of force over distance. Energy* is a bit more difficult-to-grasp concept when referring to daily life. Indeed the required amont of energy to do some work may depend on the *efficiency* of the system or the worker. For instance, we could spend a lot of energy trying to move a too heavy boulder, having done no work at all by the end of the day. So it seems that pushing a stone with a significant force but not being able to move it by one inch means no work has been done whatever the energy put in this task. In this chapter, only a part of the overall concept od *energy* will be discussed : *kinetic energy*, *potential energy* and their overall balance with *external work*. Actually this law of *energy conservation* will be directly deduced from *Newton's law of motion*. Hence in the particular scope of Newtonian mechanics, this cannot be considered as a new statement. The picture will change when studying thermodynamics where *thermal energy* or *heat* will be introduced to meet the *energy conservation law*. Actually, *conservation of energy* and *conservation of momentum* are two of the few arching principles in physics as they are intimately linked with two invariance principle: invariance with respect to time for energy and invariance with respect to space for momentum. Note that in this chapter we will refer only to point mass, neglecting the spatial extension of bodies under study and the associated energy, in particular the *rotational kinetic energy* that will be studied later on in Chapter **??**.

[1] Assuming no heat exchange as it will be studied in Chapter 12. Besides possible consequences on student's grades is not part of the syllabus of this physics course, though it is an experimental fact that the amont of work and grades are strongly correlated.

4.1 Work

As mentioned in the introduction, physicists define the *mechanical work* as the product of a force by a displacement. This is straight-forward when force and displacement are along the same direction. But how is the *mechanical work* defined when the force and the displacement are perpendicular? The answer is quite simple: there is no work done at all. Only the part of the displacement along the direction of the force contributes, or, equivalently, the part of the force aligned with the displacement. Fortunately our mathematical toolbox provides us with a scalar product of two vectors that exactly serves this purpose. It leads to the definition of the increment of work ΔW induced by a constant force $\overrightarrow{F}$ over an increment of displacement $\overrightarrow{\Delta r}$:

$$\Delta W = \overrightarrow{F} \cdot \overrightarrow{\Delta r}, \qquad \text{for fixed } \overrightarrow{F}$$

Note that ΔW has the physical dimension of a force times a distance, that is to say $[\mathrm{ML}^2\mathrm{T}^{-2}]$. In the System of International Units, it is thus in [Nm] or Joules [J]. Note also that when the scalar product is positive - that is to say in a loose sense when the force applied onto the body contributes to its displacement [2] - the *work done by the force onto the body* is positive. On the contrary when the scalar product is negative - that is to say in the same loose sense when the *force* applied onto the body opposes its displacement - the *work done by the force* is negative.

Yet one difficulty remains. Indeed the force may change in both amplitude and direction along a possibly irregular path followed by the body sujected to $\overrightarrow{F}$. To this aim a small increment of work δW is defined as:

$$\delta W = \overrightarrow{F} \cdot \overrightarrow{dr}$$

where $\overrightarrow{dr}$ the path increment is supposed to be small enough to be approximate by a straight segment over which the force $\overrightarrow{F}$ is assumed to be constant. The total work W along a given path $\mathcal{C}$ is then obtained summing up these small contributions of possibly varying sign resulting in a positive, negative or even vanishing total work for the overall path:

$$W = \int_{\mathcal{C}} \delta W = \int_{\mathcal{C}} \overrightarrow{F} \cdot \overrightarrow{dr} \tag{4.1}$$

Note that the path or trajectory may be a very complex one. Hence the sum over this path is a rather difficulty task to perform, especially when the force itself is path dependent. From a mathematical perspective, a curvilinear abscissa s is defined on path $\mathcal{C}$ which measures the length of the path starting from one convenient reference point. Hence the path $\mathcal{C}$ is defined as a vector-valued function $\overrightarrow{r}(s)$ of the curvilinear abscissa[3] s. With such a definition $\overrightarrow{dr}$ is nothing but ds times the unit tangent vector $\overrightarrow{t}(s)$:

$$\overrightarrow{dr} = \overrightarrow{t}\, ds, \qquad \overrightarrow{t} = \frac{d\overrightarrow{r}}{ds}(s)$$

where the tangent unit vector $\overrightarrow{t}$ can vary as a function of the curvilinear abcissa s. Hence, assuming that *force* $\overrightarrow{F}$ only depends on $\overrightarrow{r}$ yields a better defined integral over a single variable s:

$$W = \int_{s_o}^{s_1} \overrightarrow{F}\,(\overrightarrow{r}(s)) \cdot \overrightarrow{t}\,(s) ds \tag{4.2}$$

[2] at least the two vectors belong the a common half-space.

[3] Using a cartesian frame:

$$\overrightarrow{dr} = (dx, dy, dz) = (v_x, v_y, v_z)\, dt$$

and ds the elementary length of the path is equal to the norm of $\overrightarrow{dr}$

$$ds = \|dr\| = \sqrt{dx^2 + dy^2 + dz^2} = \|\overrightarrow{v}\| dt, \qquad \Rightarrow s_1 - s_o = \int_{t_o}^{t_1} \|\overrightarrow{v}\| dt$$

As a consequence, when the force only depends on location the work only depends on the path and does not depend on the speed at which this path is traveled.

Alternatively, the path and the force can be expressed as functions of time t instead of curvilinear abscissa s. Hence $d\overrightarrow{r} = \overrightarrow{v}dt$ leading to:

$$W = \int_{t_o}^{t_1} \overrightarrow{F} \cdot \overrightarrow{v} dt \tag{4.3}$$

Where $\overrightarrow{F}$ could depend on either location, velocity or time. The power of *force* $\overrightarrow{F}$ in the velocity field $\overrightarrow{v}$ is defined as the time derivative of the work:

$$\mathcal{P} = \frac{dW}{dt} = \overrightarrow{F} \cdot \overrightarrow{v} \tag{4.4}$$

The work developed by *forces* studied in Chapter 3 can easily be computed using Formulae (4.2) or (4.3):

- **Spring**: *Hooke's law* $W = \mathcal{E}_{\mathrm{p}}(l_o) - \mathcal{E}_{\mathrm{p}}(l_1)$

$$\mathcal{E}_{\mathrm{p}}(l) = \frac{k}{2}(l - l_{\mathrm{eq}})^2$$

- **Weight** $W = \mathcal{E}_{\mathrm{p}}(z_o) - \mathcal{E}_{\mathrm{p}}(z_1)$

$$\mathcal{E}_{\mathrm{p}}(z) = mgz$$

- **Viscous forces** $W \neq \mathcal{E}_{\mathrm{p}_o} - \mathcal{E}_{\mathrm{p}1}$

$$\begin{aligned} \mathcal{P}(t) &= -c_o \|\overrightarrow{v}\|^2 \leq 0 \\ W &= -c_o \int_{t_o}^{t_1} \|\overrightarrow{v}\|^2 dt \leq 0 \end{aligned}$$

- **Contact forces** $W \neq \mathcal{E}_{\mathrm{p}_o} - \mathcal{E}_{\mathrm{p}1}$

$$\begin{aligned} \mathcal{P}(t) &= -N \tan\alpha \|v\| \leq 0 \\ W &= -\tan\alpha \int_{s_o}^{s_1} N(s) ds \leq 0 \end{aligned}$$

It is worth noticing that in the elastic and gravity cases, the work does not depend on the path itself and only depends on the initial and final points. These two cases correspond to *conservative forces* deriving from a potential energy as studied in Section 4.2.

Note also that for viscosity and friction forces the work done and power are always negative. These forces fall under the category of *dissipative forces*. Finally, note that when the body is sliding without friction ($\alpha = 0$) the work done, and the related power, vanish. Indeed the normal reaction force is by definition normal to the displacement increment - or the velocity - and thus generates no *work*.

4.2 Potential energy

Definition 7 *A force field is a* conservative force field *when, for any two points* $\overrightarrow{r_o}$ *and* $\overrightarrow{r_1}$*, the work required to go from* $\overrightarrow{r_o}$ *and* $\overrightarrow{r_1}$ *does not depend on the path.*

As a consequence the work done on a body travelling along any closed loop in a conservative force field vanishes.

$$\oint \overrightarrow{F}_{\text{conservative}} \cdot \overrightarrow{dr} = 0 \tag{4.5}$$

Two particular cases of such forces have already been introduced: constant force field and elastic force. Actually any force field deriving from a so-called *potential energy* satisfy this property. This will be first studied in one dimension and than extended to higher-dimensions.

Conservative forces as the derivative of a 1D-potential energy

Restricting the analysis to one-dimensional motions, it can be easily shown that any force depending only on position x is a conservative force. Indeed for all x_o and x_1 the work done reads:

$$W = \int_{x_o}^{x_1} F(x)dx$$

Hence, defining a potential energy $\mathcal{E}_{\text{p}}(x)$ as the primitive of $-F(x)$:

$$\frac{d\mathcal{E}_{\text{p}}}{dx} = -F$$

yields:

$$W = -\int_{x_o}^{x_1} \frac{dU}{dx}dx = \mathcal{E}_{\text{p}}(x_o) - \mathcal{E}_{\text{p}}(x_1)$$

As a consequence, the work done is equal to the drop of potential energy from $\mathcal{E}_{\text{p}}(x_o)$ to $\mathcal{E}_{\text{p}}(x_1)$ and the force is equal to minus the derivative of the potential energy. This property can be extended to higher dimensions defining the gradient of a potential.

4.2.1 Potential energy of central forces

Central forces are defined as:

$$\overrightarrow{F}(\overrightarrow{r}) = -f(r)\overrightarrow{e_r}, \qquad r = \|\overrightarrow{r}\|, \qquad \overrightarrow{e_r} = \frac{\overrightarrow{r}}{r} \tag{4.6}$$

Gravitational and *Coulomb* forces both fall under this category. Hence the induced work reads:

$$W = -\int_{\mathcal{C}} f(r)(\overrightarrow{e_r} \cdot \overrightarrow{dr}) = -\int_{\mathcal{C}} f(r)dr = \mathcal{E}_{\text{p}}(r_o) - \mathcal{E}_{\text{p}}(r_1) \tag{4.7}$$

$$\frac{d\mathcal{E}_{\text{p}}}{dr} = f(r) \tag{4.8}$$

In Particular for:

- **Gravity** $\mathcal{E}_{\text{p}}(r) = -\dfrac{GMm}{r}$
- ***Coulomb's*** $\mathcal{E}_{\text{p}}(r) = \dfrac{Qq}{4\pi\epsilon_o r}$
- **Spring** $\mathcal{E}_{\text{p}}(r) = \dfrac{k}{2}(r - l_{\text{eq}})^2$ as already obtained.

4.2.2 Conservative forces as the gradient of a potential energy

Definition 8 (Gradient of a bi-dimensional potential) *(see [4] for 3D potentials)*

Given a scalar potential $U(x,y)$*, its the* gradient *is a vector-field defined in a* Cartesian frame *by:*

$$\overrightarrow{\text{grad}}\ U = \overrightarrow{\nabla} U \quad = \quad \begin{pmatrix} \frac{\partial U}{\partial x} \\ \\ \frac{\partial U}{\partial y} \end{pmatrix} \tag{4.10}$$

This gradient has a simple geometrical interpretation in terms of steepest tangent vector to the curved surface defined by $z = U(x,y)$. Note that level curves defined implicitly by $U(x,y) = U_o$ have the same altitude and are the intersection between the surface $z = U(x,y)$ and the plane defined by $z = U_o$. Hence the gradient has the following properties(See [5] for 3D potentials):

Property 2 (Properties of the gradient of a scalar bi-dimension potential)

- *The gradient is normal to level curves defined by* $U(x,y) = U_o$*.*
- *The gradient is tangent to the surface defined by* $z = U(x,y)$*.*
- *The gradient at each point is such a vector that its direction indicates the steepest slope at that point and its amplitude is the slope in this direction.*

Differential vector calculus and in particular the *stokes theorem* tells us that any *conservative force field* can be expressed as the opposite of the gradient of a *potential energy*. Note that the minus sign is used to make sure that the work is positive when there is a drop in potential energy.

Theorem 2 (Potential energy of a conservative force field) *Any force field deriving from a* potential $-\mathcal{E}_\text{p}$*, where* $\mathcal{E}_\text{p}$ *is called the* potential energy*, is a* conservative force field*. Indeed:*

$$W = \int_\mathcal{C} \overrightarrow{F} \cdot \overrightarrow{dr} = -\int_\mathcal{C} \left(\frac{\partial \mathcal{E}_\text{p}}{\partial x} dx + \frac{\partial \mathcal{E}_\text{p}}{\partial y} dy \right) = -\int_{E_o}^{E_1} d\mathcal{E}_\text{p} = E_o - E_1 \tag{4.11}$$

Where $E_o = \mathcal{E}_\text{p}(x_o, y_o)$ *and* $E_1 = \mathcal{E}_\text{p}(x_1, y_1)$ *are the values of the potential energy at the two ends of contour* $\mathcal{C}$ *with respective coordinates* (x_o, y_o) *and* (x_1, y_1)*. Moreover for any* conservative force field*, there exist a* potential energy $\mathcal{E}_\text{p}$ *defined up to an additive constant such that:*

$$\overrightarrow{F}_{\text{conservative}} = -\overrightarrow{\text{grad}}\ \mathcal{E}_\text{p} \tag{4.12}$$

[4] **Gradient of a three-dimensional potential**: Given a scalar *potential* $U(x,y,z)$, its the *gradient* of vector-field defined in a *Cartesian frame* by:

$$\overrightarrow{\text{grad}}\ U = \overrightarrow{\nabla} U \quad = \quad \begin{pmatrix} \frac{\partial U}{\partial x} \\ \\ \frac{\partial U}{\partial y} \\ \\ \frac{\partial U}{\partial z} \end{pmatrix} \tag{4.9}$$

[5] **Properties of the gradient of a scalar three-dimension potential**

- The gradient is normal to the equipotential surfaces defined by $U(x,y,z) = U_o$.
- The direction of the gradient is the direction along which the potential $U(x,y,z)$ varies more rapidly at a given point. It is called the steepest ascent.

For a bi-dimensional potential one can easily deduce an equivalent condition for a *force-field* to be *conservative* (See [6] for 3D potentials). It simply reads:

$$\frac{\partial F_x}{\partial y} = \frac{\partial F_y}{\partial x}$$

Indeed, when a force field $\overrightarrow{F}$ is conservative it reads $\overrightarrow{F} = -\overrightarrow{\text{grad}}\ \mathcal{E}_\text{p}$ leading to:

$$F_x = -\frac{\partial \mathcal{E}_\text{p}}{\partial x}, \qquad F_y = -\frac{\partial \mathcal{E}_\text{p}}{\partial y}$$

As a consequence:

$$\frac{\partial F_x}{\partial y} = \frac{\partial}{\partial y}\left(-\frac{\partial \mathcal{E}_\text{p}}{\partial x}\right) = -\frac{\partial^2 \mathcal{E}_\text{p}}{\partial y \partial x} = -\frac{\partial^2 \mathcal{E}_\text{p}}{\partial x \partial y} = \frac{\partial}{\partial x}\left(-\frac{\partial \mathcal{E}_\text{p}}{\partial y}\right) = \frac{\partial F_y}{\partial x}$$

4.3 Kinetic energy and energy conservation

The *work* done by a force has been defined in the previous section. In addition it has been shown in Chapter 3 that the sum of the external forces is equal to the rate of change of the *momentum*. Hence, for a single particle of mass m, $\overrightarrow{F}_\text{ext}$ denoting the sum of the forces applied on a body, W_ext being the work done by all these forces can be expressed as:

$$W_\text{ext} = \int_{t_o}^{t_1} \overrightarrow{F}_\text{ext} \cdot \overrightarrow{v}\, dt = \int_{t_o}^{t_1} m\frac{d\overrightarrow{v}}{dt} \cdot \overrightarrow{v}\, dt = \int_{t_o}^{t_1} m\frac{d}{dt}\left(\frac{\overrightarrow{v}\cdot\overrightarrow{v}}{2}\right) dt = \mathcal{E}_\text{kin}(t_1) - \mathcal{E}_\text{kin}(t_o)$$

Where the kinetic energy $\mathcal{E}_\text{kin}$ is defined as:

Definition 9 (Kinetic energy)

$$\mathcal{E}_\text{kin}(t) = \frac{1}{2} m \|\overrightarrow{v}\|^2 \tag{4.13}$$

Hence we have proven the conservation of energy in the case of a point-mass with fixed mass m:

Theorem 3 *The change in kinetic energy of a point-mass between two instant t_o and t_1 is equal to the work done by external force along the trajectory $\mathcal{C}$ of the particle between these two time instants:*

$$\Delta\mathcal{E}_\text{kin} = \mathcal{E}_\text{kin}(t_1) - \mathcal{E}_\text{kin}(t_o) = W_\text{ext} = \int_\mathcal{C} \overrightarrow{F}_\text{ext} \cdot \overrightarrow{dr} \tag{4.14}$$

Taking the derivative of this formula with respect to time gives:

[6] For 3D potentials the two additional conditions are:

$$\frac{\partial F_x}{\partial z} = \frac{\partial F_z}{\partial x}, \qquad \frac{\partial F_z}{\partial y} = \frac{\partial F_y}{\partial z}$$

These three conditions are equivalent to:

$$\overrightarrow{\text{curl}}\ \overrightarrow{F} = \overrightarrow{0}$$

with:

$$\overrightarrow{\text{curl}}\ \overrightarrow{F} = \overrightarrow{\nabla} \times \overrightarrow{F} = \begin{pmatrix} \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \\ \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \\ \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \end{pmatrix}$$

Theorem 4 (Theorem of Kinetic Energy) *The rate of change of the* kinetic energy *of a point-mass is equal to the power of external forces:*

$$\frac{d\mathcal{E}_{\text{kin}}}{dt} = \frac{dW_{\text{ext}}}{dt} = \mathcal{P}_{\text{ext}} = \overrightarrow{F}_{\text{ext}} \cdot \overrightarrow{v} \tag{4.15}$$

Note that any force $\overrightarrow{F}$ applied to a given point-mass can be split into a conservative force deriving from a potential energy $\mathcal{E}_{\text{p}}$ and a non-conservative force $\overrightarrow{F}_{\text{nc}}$:

$$\overrightarrow{F} = -\overrightarrow{\text{grad}}\ \mathcal{E}_{\text{p}} + \overrightarrow{F}_{\text{nc}} \tag{4.16}$$

Since:

$$\begin{aligned}
\int_{\mathcal{C}} (-\overrightarrow{\text{grad}}\ \mathcal{E}_{\text{p}}) \cdot \overrightarrow{dr} &= -\int_{t_o}^{t_1} \overrightarrow{\text{grad}}\ \mathcal{E}_{\text{p}} \cdot \overrightarrow{v}\, dt \\
&= -\int_{t_o}^{t_1} \left(\frac{\partial \mathcal{E}_{\text{p}}}{\partial x}\frac{dx}{dt} + \frac{\partial \mathcal{E}_{\text{p}}}{\partial y}\frac{dy}{dt} \right) dt \\
&= -\int_{t_o}^{t_1} \frac{d\mathcal{E}_{\text{p}}(\overrightarrow{r}(t))}{dt} dt = \mathcal{E}_{\text{p}}(\overrightarrow{r}(t_o)) - \mathcal{E}_{\text{p}}(\overrightarrow{r}(t_1))
\end{aligned}$$

the *Theorem of Kinetic Energy* (4.15) becomes:

$$\Delta(\mathcal{E}_{\text{kin}} + \mathcal{E}_{\text{p}}) = W_{\text{nc}} = \int_{\mathcal{C}} \overrightarrow{F}_{\text{nc}} \cdot \overrightarrow{dr} \tag{4.17}$$

As a consequence, when the *non-conservative forces* are dissipative forces ($W_{\text{nc}} \leq 0$) the total energy defined as $\mathcal{E}_{\text{kin}} + \mathcal{E}_{\text{p}}$ always decreases:

$$\Delta(\mathcal{E}_{\text{kin}} + \mathcal{E}_{\text{p}}) = W_{\text{nc}} \leq 0$$

Moreover when the *non-conservative forces* generates no work ($W_{\text{nc}} = 0$), as for instance in the case of interaction forces without friction or rolling without sliding, the total energy remains constant:

$$\mathcal{E}_{\text{kin}} + \mathcal{E}_{\text{p}} = E_o$$

with a balance over time between kinetic and potential energy[7].

[7] It is worth noticing that kinetic, potential and total energies are defined up to an additive constant. Indeed, the kinetic energy in an inertial frame with constant velocity $\overrightarrow{v_o}$ with respect to an other inertial frame a $1/2\ mv_o^2$ term will be added. Note that the other additional term $m\overrightarrow{v} \cdot \overrightarrow{v_o}$ will cancel the change in the external work. Indeed the incremental travelled path $\overrightarrow{dr}$ has changed by $\overrightarrow{v_o}dt$.

Chapter 5

Applications of energy conservation

As mentioned in previous chapter, the theorem of kinetic energy is a direct consequence of Newton's laws of motion. Hence all analyses performed with the conservation of energy could also be solved directly with Newton's laws of motion. However, solution methods taking advantage of the conservation of energy may be simpler and faster. This is the case in particular when the motion is governed by a single parameter. Indeed one single equation is required in this case instead of the set of Newton's laws.

This chapter first considers a set of analyses that can be conducted using this technique. Then a qualitative discussion of classes of solutions will be introduced.

5.1 Point-mass under constant gravity

Weight is a conservative force with potential energy given by $\mathcal{E}_{\mathrm{p}} = mg(h - h_o)$ where h_o is the initial height. Hence the total energy is constant with value:

$$E_o = \frac{1}{2}mv_o^2$$

where v_o is the amplitude of the initial velocity. Thanks to Newton's law of motion, we know that only the vertical velocity v_z will change over time. Hence:

$$\frac{1}{2}m(v_z^2(+v_{xo}^2 + v_{yo}^2)) + mg(h - h_o) = \frac{1}{2}m(v_{zo}^2(+v_{xo}^2 + v_{yo}^2))$$

leading to:

$$v_z^2 - v_{zo}^2 = -2g(h - h_o)$$

Since the vertical velocity vanishes when the body reaches it highest altitude we can directly obtain this highest altitude not even knowing the trajectory:

$$h_{\max} = h_o + \frac{v_{zo}^2}{2g}, \qquad v_{zo} > 0$$

Furthermore, the trajectory can be obtained integrating the vertical velocity noticing that $\frac{dh}{dt}$:

$$\frac{dh}{dt} = v_z \qquad \text{with} \qquad |v_z| = \sqrt{v_{zo}^2 - 2g(h - h_o)}$$

Assuming that $v_{zo} > 0$ we can assume that v_z remains positive until time t_h when the highest altitude is reached. Hence we have:

$$t = \int_{h_o}^{h} \frac{dh'}{\sqrt{v_{zo}^2 - 2g(h' - h_o)}} = \frac{1}{-g}\left[\sqrt{v_{zo}^2 - 2g(h' - h_o)}\right]_{h_o}^{h} = \frac{1}{-g}\left(\sqrt{v_{zo}^2 - 2g(h - h_o)} - |v_{zo}|\right)$$

for $h < h_{\max}$ and hence for $t < v_{zo}/g$. This yields to:

$$v_{zo} - gt = \sqrt{v_{zo}^2 - 2g(h - h_o)} \qquad \text{for} \qquad t < v_{zo}/g$$

and finally:

$$h = -\frac{1}{2}gt^2 + tv_{zo} + h_o \qquad \text{for} \qquad t < v_{zo}/g$$

For larger time t the vertical velocity will become negative and the same procedure can be followed starting from $t_o = v_{zo}/g$ leading to the classical results obtained directly from Newton's Law:

$$h = -\frac{1}{2}gt^2 + tv_{zo} + h_o \qquad \text{for all} \qquad t$$

5.2 Point-mass under gravity and normal reaction force

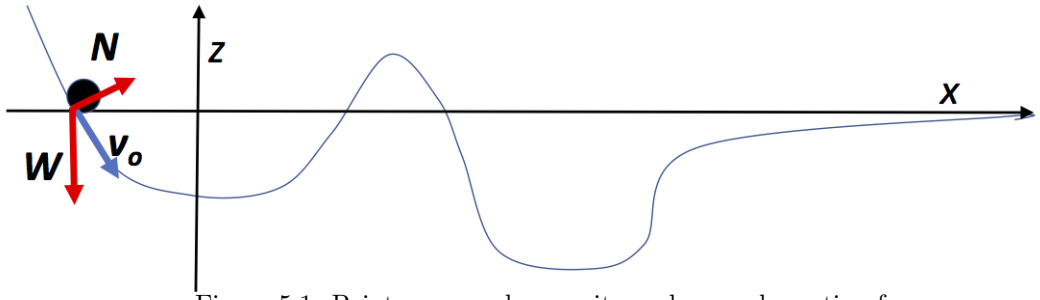


Figure 5.1: Point-mass under gravity and normal reaction force

When a point-mass subjected to its own weight is constrained to follow a given trajectory with a reaction force $\overrightarrow{R}$ that is perpendicular to the trajectory the work done by this reaction force vanishes and the total energy remains constant all along the trajectory. This is the case of a rollercoaster or any body constrained by a string under tension such as a pendulum. As opposed to the free fall studied previously, the horizontal components of the velocity are not constant since the horizontal components of the reaction force vary with the trajectory. But the direction of the velocity is known as it is along the tangent of the given trajectory. Hence the only unknown is the amplitude of the velocity vector, namely the speed v. This speed, actually its square, can be computed at any location of the mass using the conservation of mechanical energy. In two dimensions the trajectory is given by the profil $z(x)$ and the potential energy as a function of x reads:

$$\mathcal{E}_{\mathrm{p}}(x) = mgz(x)$$

and the total energy is:

$$E_o = \frac{1}{2}m\|\overrightarrow{v_o}\|^2 + mgh_o$$

or could be equivalently written as $E_o = mgh_m$ where h_m is the highest altitude the body could reach. For the sake of simplicity the vertical axis could be set such that $h_m = 0$. Hence:

$$v = \sqrt{-2gz(x)} = \frac{ds}{dt}$$

which can be integrated with respect to time noticing that $ds = \sqrt{1 + (dz/dx)^2}dx$:

$$t = \int_0^{x(t)} \sqrt{\frac{1 + (dz/dx)^2}{-2gz(x)}}dx$$

Note that this integration can be performed only up to $z(x) = 0$. Further computation can be performed accounting for a backwards travel path along the $z(x)$ curve.

5.3 Pendulum

In the particular case of a pendulum of length L with initial angle θ_o will have a vanishing initial potential energy with $h_o = L(1 - \cos\alpha_o)$ and will reach its maximum velocity $v_{\max}$ at $\theta = 0$ with value:

$$v_{\max} = \sqrt{2gL\cos\theta_o}$$

The time evolution can further be obtained noticing that:

$$L\frac{d\theta}{dt} = \|v\| = \sqrt{\frac{2(E_o - \mathcal{E}_{\mathrm{p}}(\theta))}{m}}, \qquad \frac{d\theta}{dt} \geq 0$$

with $\mathcal{E}_{\mathrm{p}}(\theta) = mgz(\cos\theta - \cos\theta_o)$ (see Figure 5.2). Hence integrating with respect to θ between 0 and θ_o given a quarter an oscillation period T of the pendulum:

$$\frac{T}{4} = \sqrt{\frac{L}{2g}} \int_0^{\theta_o} \frac{d\theta}{\sqrt{\cos\theta - \cos\theta_o}}$$

When the total energy is strictly positive with $\theta_o = \pi$ the velocity at $\theta = \pi$ does not vanish and the pendulum keeps on rotating. The phase diagram is given in Figure 5.2.

5.4 Gravitational potential and escape velocity

The gravitational force field, as a central force field is a conservative force deriving from a isotropic potential energy $\mathcal{E}_{\mathrm{p}}(r)$ that is only dependent on the distance r with respect to the center of the attracting body assumed to be fixed. Hence[1]:

$$\mathcal{E}_{\mathrm{p}}(r) = -\frac{GMm}{r} \tag{5.1}$$

[1] A complete derivation is given in this footnote:

$$-\overrightarrow{\mathrm{grad}}\ \mathcal{E}_{\mathrm{p}}(r) = -\mathcal{E}_{\mathrm{p}}{}'(r)\overrightarrow{\mathrm{grad}}\ r$$

Since $r = \sqrt{x^2 + y^2 + z^2}$ one obtains:

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2 + y^2 + z^2}} = \frac{x}{r}$$

Performing similar calculations for the other components gives:

$$\overrightarrow{\mathrm{grad}}\ r = \overrightarrow{e_r}$$

and finally using the definition of the gravitation force:

$$\mathcal{E}_{\mathrm{p}}{}'(r) = \frac{GMm}{r^2}, \qquad \Rightarrow \mathcal{E}_{\mathrm{p}}(r) = \mathcal{E}_{\mathrm{p}_o} - \frac{GMm}{r}$$

Constant $\mathcal{E}_{\mathrm{p}_o}$ is set to zero such that the potential energy vanishes when the two attracting bodies are far away from each other ($r = +\infty$). Hence:

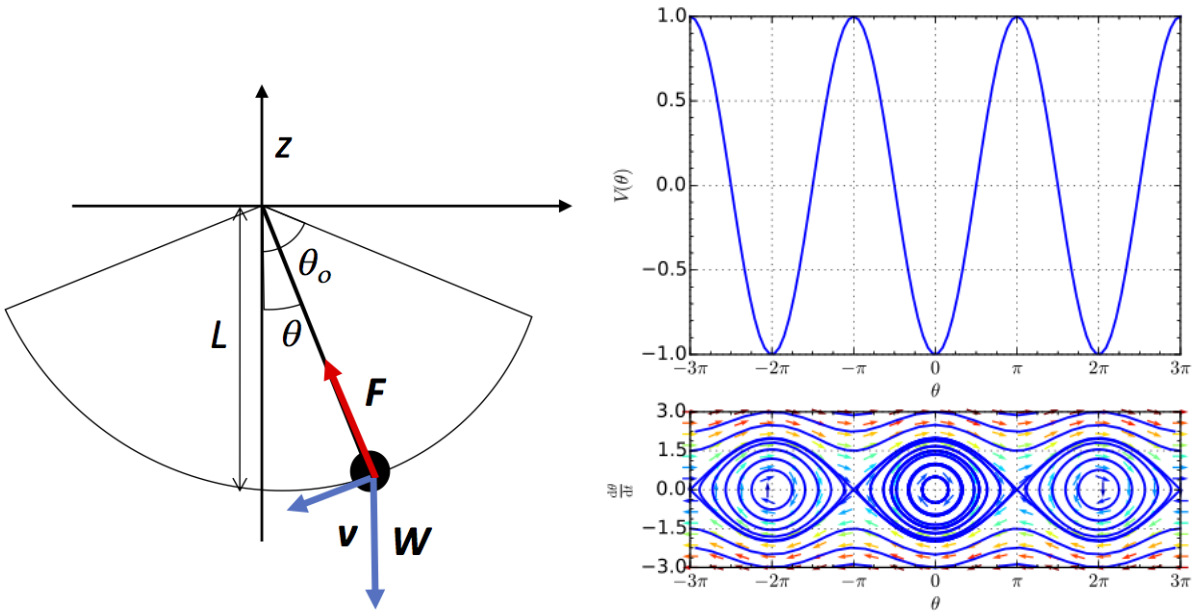


Figure 5.2: Potential energy and phase diagram of a simple pendulum

It is then possible to compute the required initial velocity v_l for a rocket to leave the attraction of the earth assuming it reaches infinity with a vanishing velocity:

$$v_\mathrm{l} = \sqrt{2\frac{GM}{R_o}} = \sqrt{2gR_o}$$

Where R_o is the Earth's radius and g is the gravity acceleration at Earth's surface. Hence:

$$v_\mathrm{l} \approx \sqrt{2 \times 9.8 \times 6.4}\ [\mathrm{km/s}] \approx 11\ [\mathrm{km/s}]$$

5.5 Qualitative study of equilibrium for one-parameter systems

The qualitative analysis of the response of systems depending on only one parameter can be done using a graph of the potential energy as the one given in Figures 5.1 and 5.2. When the potential energy has no upper bound on both sides, the trajectory is bounded in space for a bounded total mechanical energy. This is the case of the elastic potential ($\mathcal{E}_\mathrm{p}(x) = kx^2/2$). When the potential energy has a finite limit on at least one side, there are unbounded trajectories. This is the case for the gravitational potential with and escape total energy ($E_o = 0$) and an escape velocity from any initial radius R_o.

Any local minimum in the potential energy ($\frac{d\mathcal{E}_\mathrm{p}}{dx}(x_\mathrm{min}) = 0$ and $k = \frac{d^2\mathcal{E}_\mathrm{p}}{dx^2}(x_\mathrm{min}) > 0$) is a stable equilibrium position and small vibration can take place in its neighbourhood. In this neighbourhood, small vibrations can be modelled using the an approximate quadratic potential energy:

$$\mathcal{E}_\mathrm{p}(x) \approx \frac{k}{2}(x - x_\mathrm{min})^2$$

where $k = \frac{d^2\mathcal{E}_\mathrm{p}}{dx^2}(x_\mathrm{min})$ is the tangent stiffness.

Any local maximum in the potential energy ($\frac{d\mathcal{E}_\mathrm{p}}{dx} = 0$ and $\frac{d^2\mathcal{E}_\mathrm{p}}{dx^2} \leq 0$) is an unstable equilibrium position.

5.6 Solution in Phase-plane

For a system depending on a single parameter x, the total mechanical energy $\mathcal{E}_\mathrm{kin}(v)+\mathcal{E}_\mathrm{p}(x)$ can be represented on the (x, v) plane, called the *Phase-plane*, by means of its iso-energy curves implicitly defined as:

$$\mathcal{E}_\mathrm{kin}(v) + \mathcal{E}_\mathrm{p}(x) = E$$

Each energy level defining a curve as shown of Figure 5.2 for the pendulum. These curves never cross for two different energy levels. These curves are closed curves when for this energy level E the system is locally trapped in a potential well but can have some open branches. The intersection of these curves with $x-$axis are the values x_i for which:

$$\mathcal{E}_\mathrm{p}(x_i) = E$$

When these points are not extremum values of the potential energy ($\frac{d\mathcal{E}_\mathrm{p}}{dx}(x_i) \neq 0$), the iso-energy curve is tangent to the $v-$axis. Indeed:

$$\frac{1}{2}mv^2 + \frac{d\mathcal{E}_\mathrm{p}}{dx}(x_i)(x - x_i) = 0 \qquad \Rightarrow |v| = \sqrt{\frac{2\left|\frac{d\mathcal{E}_\mathrm{p}}{dx}(x_i)(x - x_i)\right|}{m}}$$

When one of these points is a local minimum of the potential energy, the iso-energy curve reduces to a single point. When it is a local maximum of the potential energy ($\frac{d^2\mathcal{E}_\mathrm{p}}{dx^2}(x_i) < 0$), two branches cross at that point with tangent equations given by:

$$v = \pm\sqrt{\frac{1}{m}\frac{d^2\mathcal{E}_\mathrm{p}}{dx^2}(x_i)}(x - x_i)$$

When the potential reaches a bounded limit $\mathcal{E}_\mathrm{p}(\pm\infty)$ at infinity the iso-energy curves have an horizontal asymptot with value:

$$v = \pm\frac{2(E - \mathcal{E}_\mathrm{p}(\pm\infty))}{m}$$

Chapter 6

Dynamics of two-point-mass systems

Up to now, analysis has been restricted to a single body and even a point-mass. In this chapter, we extend the analysis to systems made of several point-masses. Hence, when momentum was unambiguous for a single particle we have now to precisely define what would be the momentum of a set of particle each one having its own momentum. We will then refer to the *linear momentum* of the system as the sum of the momenta of all particles in that system. Besides *momentum* and *linear momentum* of a single body or a particle coincide.

To this aim, we first restrict ourselves to the case of a system made of two point-masses and show that conservation of linear momentum is fulfilled thanks to third Newton's Law. This is then extended to conservation of linear momentum of a set of point-masses and its restriction to the center of mass is considered. Afterwards, the general solution for the collision between two particles is discussed solely based on conservation of linear momentum and energy. Finally conservation of *linear momentum* is used to analyse bodies with varying mass considering sets of auxiliary two-body systems with constant mass.

6.1 Linear momentum of a two-body system

Let us consider a system made of two or more sub-bodies and let us define its linear momentum as the sum of the momentum of its parts. Hence for two parts:

$$\overrightarrow{p} = \overrightarrow{p}_1 + \overrightarrow{p}_2 \tag{6.1}$$

Assuming that this system is only subjected to internal forces, we will show that this momentum remains constant in an *inertial frame of reference*, that is:

$$\frac{d\overrightarrow{p}}{dt} = 0 \tag{6.2}$$

Indeed let us consider each part of this system, labelled 1 and 2 for convenience, each one having a non-vanishing mass. Pulling appart this original body into two distinct pieces would require some forces to be applied. Hence the force $\overrightarrow{F}_{1\to 2}$ applied by part 1 on the other part 2 is unlikely to vanish. So is force $\overrightarrow{F}_{2\to 1}$ applied by part 2 on part 1. At the same time each part being considered in the same *inertial frame of reference*, Newton's Law II applies for each with no additional external

forces to be considered appart from the interaction forces. Hence:

$$\frac{d\overrightarrow{p_1}}{dt} = \overrightarrow{F}_{2\to 1}, \qquad \frac{d\overrightarrow{p_2}}{dt} = \overrightarrow{F}_{1\to 2}$$

Taking the sum of the two equations yields:

$$\overrightarrow{F}_{2\to 1} + \overrightarrow{F}_{1\to 2} = \frac{d\overrightarrow{p_1}}{dt} + \frac{d\overrightarrow{p_2}}{dt} = \frac{d\overrightarrow{p}}{dt}$$

Hence Newton's Law III is equivalent to the conservation of the linear momentum of the system made of the two parts in the absence of external forces. As a matter of fact, conservation of momentum appears to be much more general that Newton's Laws and is nowadays one of the pilar of modern physics.

Let us finally remark that defining the center of mass C of the two particles through its position vector $\overrightarrow{r}_C$ in any reference frame as a function of the position vectors $\overrightarrow{r}_1$ and $\overrightarrow{r}_2$ of each of the two masses by:

$$(m_1 + m_2)\overrightarrow{r}_C = m_1\overrightarrow{r}_1 + m_2\overrightarrow{r}_2$$

one can remark that :

$$\overrightarrow{p} = m_1\overrightarrow{v}_1 + m_2\overrightarrow{v}_2 = \frac{d}{dt}(m_1\overrightarrow{r}_1 + m_2\overrightarrow{r}_2) = (m_1 + m_2)\frac{d\overrightarrow{r}_C}{dt}$$

Hence the *linear momentum* of the system of two particles equals the *mometum* of a single fictitious particle of mass $m_1 + m_2$ located at the center of mass and travelling at $\overrightarrow{v}_C$.

These properties are generalized in next section for multiple-particles systems and for extended systems in Chapter **??** giving some rationals to the classical assumption consisting of modeling a extended body by a point-mass. We now know that the point-mass is located at the center of mass and that the conservation of linear momentum is preserved.

6.2 Linear momentum conservation of a multiple-particle system

In this section conservation of *linear momentum* is extended to an arbitrary set of particles together with the related center of mass of this system. In particular, it is shown that the conservation of linear momentum of the system of particles coincides with the second Newton's law of a fictitious particles carrying the total mass of the system and located at the center of mass. Hence, Newton's laws are shown to be scale invariant as a set of particles can be replaced by an equivalent point-mass as long as we are interested neither in the detailed distribution of the particles in the set nor in the spatial orientation of this set - Angular momentum will be introduced in Chapter 7 to account for the overall spatial orientation of the system.

Let us thus consider an assembly of N interacting bodies labelled from 1 to N with $\overrightarrow{F}_{i\to j}$ the interaction force applied by body i on body j. These could be either contact or distant forces but all of them satisfying the action-reaction law:

$$\overrightarrow{F}_{i\to j} = -\overrightarrow{F}_{j\to i}$$

In addition to these interaction forces let us call $\overrightarrow{F}_{o\to i}$ the forces applied by the outside of the system of N bodies on body i. Let then $\overrightarrow{p}_i$ be the momentum of body i defined as:

$$\overrightarrow{p}_i = m_i\overrightarrow{v}_i = m_i\frac{d\overrightarrow{r}_i}{dt}$$

with m_i the mass of body i. For each body the balance of momentum reads:

$$\frac{d\overrightarrow{p}_i}{dt} = \overrightarrow{F}_{o\to i} + \sum_{j\neq i} \overrightarrow{F}_{j\to i}$$

Summing these equations for all bodies gives:

$$\underbrace{\sum_i \frac{d\overrightarrow{p}_i}{dt}}_{\frac{d\overrightarrow{p}}{dt}} = \underbrace{\sum_i \overrightarrow{F}_{o\to i}}_{\overrightarrow{F}} + \sum_i \sum_{j>i} \underbrace{\left(\overrightarrow{F}_{j\to i} + \overrightarrow{F}_{i\to j}\right)}_{\overrightarrow{0}}$$

where $\overrightarrow{p}$, the sum of the momentum of all bodies, is the linear momentum of the system:

$$\overrightarrow{p} = \sum_i \overrightarrow{p}_i = \sum_i m_i \frac{d\overrightarrow{r}_i}{dt} = \frac{d}{dt}\left(\sum_i m_i \overrightarrow{r}_i\right)$$

Hence $\overrightarrow{p}$ can be seen as the momentum of a point-mass of mass $M = \sum_i m_i$ located at the *center of mass* C of the N bodies defined as:

Definition 10 *The center of mass C of N bodies of respective mass m_i and located at $\overrightarrow{r_i}$ is defined as:*

$$\overrightarrow{r_C}(t) = \sum_i \frac{m_i}{M}\ \overrightarrow{r}_i(t), \qquad M = \sum_i m_i \tag{6.3}$$

As a consequence, the balance of linear momentum of the system of N bodies reads:

$$M\frac{d\overrightarrow{v_C}}{dt} = \overrightarrow{F} \tag{6.4}$$

when $\overrightarrow{v_C}$ is the velocity of the center of mass and $\overrightarrow{F}$ is the sum of all external forces applied on all the bodies of the system:

$$\overrightarrow{F} = \sum_i \overrightarrow{F}_{o\to i} \tag{6.5}$$

Hence, whatever the spatial extension of a body or a system made of several interacting bodies, the *balance of momentum* applies at the *center of mass* of the system for the total mass M and external forces applied at the very same point. However, this global equation does not tell us anything about the local motion of the individual parts of the system, including the spatial orientation of the overall system. This will be studied in Chapter 7 introducing the *angular momentum* of the system of particles and in Chapter **??** for rigid bodies.

6.3 Collision between two particles

The time during which particles collide is very short and a detailed description of the interactions forces acting during collision is difficult, if not even possible, to access experimentally. Newton's second and third laws of motion are then of little help. Conservation laws for momentum and energy, expressing the underlying invariance of the law of physics with respect to space and time happen to be much more adequate to analyse data acquired from a collider experiment.

In this section we consider two interacting particles with respective masses m_1 and m_2 and velocities $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ long before collision and $\overrightarrow{v_1}'$ and $\overrightarrow{v_2}'$ long after collision in any convenient

inertial frame of reference. Respective momenta are denoted by $\overrightarrow{p_1} = m_1\overrightarrow{v_1}$ and $\overrightarrow{p_2} = m_2\overrightarrow{v_2}$ before collision and $\overrightarrow{p_1}' = m_1\overrightarrow{v_1}'$ and $\overrightarrow{p_2}' = m_2\overrightarrow{v_2}'$ after collision.

Let us first choose the inertial frame that moves with m_2 before collision being also called the *lab frame* in a collider. Hence $\overrightarrow{p_2} = \overrightarrow{0}$ in that frame and conservation of momentum yields:

$$\overrightarrow{p_1} = \overrightarrow{p_1}' + \overrightarrow{p_2}' \tag{6.6}$$

In this very same *lab frame*, energy conservation reads:

$$\frac{p_1^2}{2m_1} = \frac{p_1'^2}{2m_1} + \frac{p_2'^2}{2m_2} + \Delta U \tag{6.7}$$

where p_1, p_1' and p_2' are the amplitudes of the individual momentum-vectors and ΔU is the internal energy created after collision. When $\Delta U = 0$ the collision is defined as *elastic*, *anelastic* otherwise.

6.3.1 Center of mass and reduced mass

$\overrightarrow{v_C}$ the velocity vector of the center of mass of the two particles read:

$$M\overrightarrow{v_C} = \overrightarrow{p_1} + \overrightarrow{p_2} = \overrightarrow{p_1}' + \overrightarrow{p_2}' \tag{6.8}$$

with $M = m_1 + m_2$ and remains constant during the experiment assuming no external force for both *elastic* or*anelastic* collisions. We then denote by $\overrightarrow{p_1}_r'$ and $\overrightarrow{p_2}_r'$ the relative momenta defined as:

$$\overrightarrow{p_1}_r' = \overrightarrow{p_1}' - m_1\overrightarrow{v_C}, \qquad \overrightarrow{p_2}_r' = \overrightarrow{p_2}' - m_2\overrightarrow{v_C}$$

Hence $\overrightarrow{p_1}_r'$ and $\overrightarrow{p_2}_r'$ are the momenta of each particle in the inertial frame attached to the center of mass. Conservation of momentum yields:

$$\overrightarrow{p_1}_r' = -\overrightarrow{p_2}_r'$$

and the same property for the initial relative momenta:

$$-\overrightarrow{p_{2r}} = \overrightarrow{p_{1r}} = m_1(\overrightarrow{v_1} - \overrightarrow{v_C}) = \mu\overrightarrow{v_1}$$

where the reduced mass μ is defined as:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{6.9}$$

Hence, the initial energy reads:

$$E_{\text{ini}} = \frac{p_1^2}{2m_1} = \frac{p_C^2}{2M} + \frac{p_{1r}^2}{2\mu}$$

and simularly:

$$\frac{p_1'^2}{2m_1} + \frac{p_2'^2}{2m_2} = \frac{p_C^2}{2M} + \frac{p_{2r}'^2}{2\mu}$$

As a consequence in the inertial frame attached to the center of mass of the two particles:

$$p_r^2 = p_r'^2 + 2\mu\Delta U$$

Hence for *elastic collisions* ($\Delta U = 0$) the relative momenta have the same amplitude and opposite directions for the two particles both before and after collision. The only free parameter is the angle between the initial and final momenta of each particle. For inelastic collisions, the change in the amplitude is a measure of the energy transfer from kinetic to internal energy.

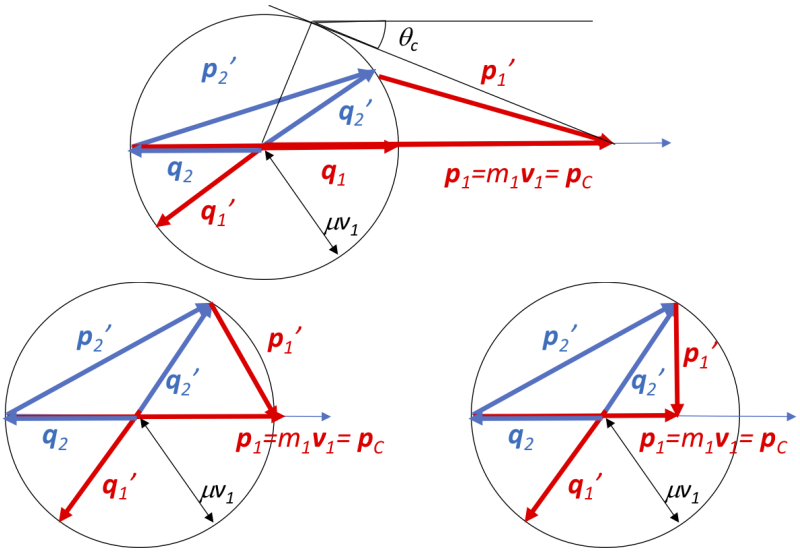


Figure 6.1: Momentum for a elastic collision (top) $m_1 > m_2$, (bottom left) $m_1 = m_2$, (bottom right) $m_1 < m_2$

6.3.2 Scattering pattern

When coming back to the inertial frame of the lab with particle m_2 initially at rest ($\overrightarrow{p_2} = 0$), momenta $\overrightarrow{p_1}$, $\overrightarrow{p_C}$, $\overrightarrow{p_{1r}}$ and $\overrightarrow{p_{2r}}$ are aligned. Moreover $\overrightarrow{p_1}$ and $\overrightarrow{p_C}$ are equal ($\overrightarrow{p_C} = \overrightarrow{p_1}$) and:

$$\overrightarrow{p_1} = \left(1 + \frac{m_1}{m_2}\right) \overrightarrow{p_{1r}}$$

Hence they are all along the same positive axis, chosen for convenience as the horizontal one. Moreover:

$$p_1 \geq p_{1r}$$

For $m_1 < m_2$, $\overrightarrow{p_1}$ is inside the cercle centered at $(\mu v_1, 0)$ with radius μv_1, outside for $m_1 > m_2$ and on the cercle for $m_1 = m_2 = m$ (see Figure 6.3.2). Moreover:

$$\overrightarrow{p_2}' = \overrightarrow{p_{2r}}' + \frac{m_2}{M}\overrightarrow{p_C} = \overrightarrow{p_{2r}}' + \mu\overrightarrow{v_1}$$

Hence the tip of vector $\overrightarrow{p_2}'$ originating at O is on the cercle centered at $(\mu v_1, 0)$ with radius μv_1. Vector $\overrightarrow{p_1}'$ is then built with its origin at the tip of $\overrightarrow{p_2}'$ and ending at the tip of $\overrightarrow{p_1}$ since:

$$\overrightarrow{p_2}' + \overrightarrow{p_1}' = \overrightarrow{p_C} = \overrightarrow{p_1}$$

As a consequence:

- when $m_1 = m_2 = m$, $\overrightarrow{p_2}'$ and $\overrightarrow{p_1}'$ are orthogonal vectors. The two particles are scattered at right angle.
- When $m_1 > m_2$ the diffraction angle of m_1 is less than a limiting angle $\alpha_{\max}$ such that:

$$\sin \alpha_{\max} = \frac{\mu v_1}{(m_1 - \mu)v_1} = \frac{m_2}{m_1}$$

- When $m_1 > m_2$, mass m_1 is forward-scattered ($\overrightarrow{p_1}' \cdot \overrightarrow{p_1} > 0$).

These aforementioned properties are derived from the conservation of momentum and energy. Some additional information on the system is required to fully determine the scattering of the two particles after collision and in particular the scattering angles. Conservation of angular momentum studied in Chapter 7 brings some additional elements but the precise knowledge of the interaction potential or interaction forces is required for this full description. Indeed the change oi momentum of each particle is given by:

$$\Delta\overrightarrow{p_1} = \overrightarrow{p_1}' - \overrightarrow{p_1} = \int_{-\infty}^{+\infty} \overrightarrow{F}_{2\to 1} dt$$

6.4 Bodies with varying mass

A system with a varying mass such as a rocket or a trolley being filled when rolling can be analysed as a two-part system with fixed total mass during a short period a time Δt. Let us denote by $m(t)$ the mass of the main system of interest at time t and δm the amount of mass lost between t and $t + \Delta t$. Mass conservation implies that:

$$m(t + \Delta t) + \delta m = m(t) \qquad \text{and} \qquad \lim_{\Delta t \to 0} \frac{\delta m}{\Delta t} = \lim_{\Delta t \to 0} \frac{m(t) - m(t + \Delta t)}{\Delta t} = -\frac{dm}{dt}$$

$\overrightarrow{v}(t)$ denotes the velocity of the main part m at time t and $\overrightarrow{v_r}$ is the relative velocity of the lost part δm with respect to the main part. Hence the linear momentum at time t and $t + \Delta t$ of the entire system read:

$$\begin{aligned} \overrightarrow{p}(t) &= m(t)\overrightarrow{v}(t) \\ \overrightarrow{p}(t + \Delta t) &= (m(t) - \delta m)\overrightarrow{v}(t + \Delta t) + \delta m(\overrightarrow{v_r} + \overrightarrow{v}(t + \Delta t)) \\ &= m(t)\overrightarrow{v}(t + \Delta t) + \Delta m\overrightarrow{v_r} = \overrightarrow{p}(t) + m(t)\overrightarrow{\Delta v} + \delta m\overrightarrow{v_r} \end{aligned}$$

As a consequence the rate of change of momentum of the entire system reads:

$$\lim_{\Delta t \to 0} \frac{\Delta \overrightarrow{p}}{\Delta t} = m(t)\frac{d\overrightarrow{v}}{dt} - \frac{dm}{dt}\overrightarrow{v_r}$$

The first term on the right-hand side is the classical rate of momentum of a body with a fixed mass $m(t)$ whereas the second term represents the change in momentum generated by the rate of mass being lost - note that $\frac{dm}{dt}$ is negative when mass is lost. Obviously, this last term vanishes when the relative velocity tends to zero. When analysing a rocket or an airplane, this last term in the rate of momentum (6.10) is put on the right hand side of Newton's Law II as an additional force called the *thrust* of the engin. This fictitious force $\overrightarrow{T} = \frac{dm}{dt}\overrightarrow{v_r}$ is directed towards the opposite of the relative velocity when $\frac{dm}{dt}$ is non-positive. For a rocket $q_m = -\frac{dm}{dt}$ is the flux of mass outflowing the jet engin and $\overrightarrow{v_r}$ is the ejection-velocity of gases. Hence to increase thrust we can either increase the mass flux or the ejection-velocity. The former option is always better for a rocket as increasing the flux would increase the total initial mass of the rocket [1].

[1] Note that for an airplane, the mass remains constant as the mass of air outflowing from the jet engins is balanced by the mass of inflowing air. Hence $-\frac{dm}{dt}$ should be replaced by the the mass flux through the engins q_m . Moreover the momentum of the inflow in the jet engin has to be accounted for as a drag force. Indeed its relative velocity is the opposite of the aircraft velocity $\overrightarrow{v}$ since the air is at rest with respect to a fixed reference frame. Hence the *thrust* $\overrightarrow{T}$ reads:

$$\overrightarrow{T} = -q_m(\overrightarrow{v_r} - \overrightarrow{v})$$

Hence to increase thrust increasing the mass flux is a good option here as the total mass remains constant. And increasing the mass flux at the inflow simply consists in increasing the cross-section of the intake. This is why jet-engins have become larger and larger over years.

Chapter 7

Angular Momentum

From the original Newton's laws of motion studied in Chapter 3 two essential principles have been derived: *conservation of linear momentum* linked to the invariance with respect to translation in space; *conservation of energy* related to invariance with respect to time.
In this chapter, invariance with respect to rotation and the related *conservation of angular momentum* is discussed. To this aim several concepts have to be introduced: *angular momentum*, *torque* or *moment of a force*. Related concepts of *rotational kinetic energy* and *moment of inertia* will be studied in the subsequent chapter **??**. Before doing so, an appropriate tool from our vector-analysis toolbox has to be selected to account for rotations. The vector product appears a good candidate as suggested by the following remarks.

Remark 1 (*Vector products* as rotation generator) *Using a Cartesian frame we know that:*

$$(\overrightarrow{e_z}\times)\overrightarrow{e_x} = \overrightarrow{e_y}, \qquad \textit{and} \qquad (\overrightarrow{e_z}\times)\overrightarrow{e_y} = -\overrightarrow{e_x} \tag{7.1}$$

Hence, applying $(\overrightarrow{e_z}\times)$ *on a vector in the* $(\overrightarrow{e_x},\overrightarrow{e_y})$*-plane rotates this vector by* $\pi/2$ [rad] *around the* $\overrightarrow{e_z}$ *axis.*

Indeed any vector $\overrightarrow{a}$ in the $(\overrightarrow{e_x},\overrightarrow{e_y})$-plane reads:

$$\overrightarrow{a} = a_x\overrightarrow{e_x} + a_y\overrightarrow{e_y}$$

Hence $\overrightarrow{b} = (\overrightarrow{e_z}\times)\overrightarrow{a}$ reads:

$$\overrightarrow{b} = \overrightarrow{e_z}\times(a_x\overrightarrow{e_x} + a_y\overrightarrow{e_y}) = a_x\overrightarrow{e_z}\times\overrightarrow{e_x} + a_y\overrightarrow{e_z}\times\overrightarrow{e_y} = -a_y\overrightarrow{e_x} + a_x\overrightarrow{e_y}$$

Obviously vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ are vectors of the $(\overrightarrow{e_x},\overrightarrow{e_y})$-plane having the same amplitude $\sqrt{a_x^2+y_y^2}$ and being perpendicular since:

$$\overrightarrow{a}\cdot\overrightarrow{b} = 0$$

Hence $(\overrightarrow{e_z}\times)$ transforms any vector of the $(\overrightarrow{e_x},\overrightarrow{e_y})$-plane into a perpendicular vector of same amplitude, leading to a rotation of $\pi/2$ [rad]. As a consequence in cylindrical coordinates around the $\overrightarrow{e_z}$-axis, we also have[1]:

$$(\overrightarrow{e_z}\times)\overrightarrow{e_r} = \overrightarrow{e_\theta}, \qquad \text{and} \qquad (\overrightarrow{e_z}\times)\overrightarrow{e_\theta} = -\overrightarrow{e_r}$$

[1] Take $\overrightarrow{a} = \overrightarrow{e_r} = \cos\theta\overrightarrow{e_x} + \sin\theta\overrightarrow{e_y}$ to prove it.

7.1 Angular velocity vector

Incremental motion induced by a rotation around the $\overrightarrow{e_z}$ axis can easily be characterised defining the position vector of a point on a circle of radius R in the $(\overrightarrow{e_x}, \overrightarrow{e_y})$-plane as:

$$\overrightarrow{r} = R\overrightarrow{e_r}(\theta)$$

where $\overrightarrow{e_r}(\theta)$ is the unit vector with angle θ with respect to $\overrightarrow{e_x}$ ($\overrightarrow{e_r} = \cos\theta\overrightarrow{e_x} + \sin\theta\overrightarrow{e_y}$). A small change $\Delta\theta$ in the angle, induced a change $\overrightarrow{\Delta r}$ in the position vector defined as:

$$\overrightarrow{\Delta r} = \overrightarrow{r}(\theta + \Delta\theta) - \overrightarrow{r}(\theta) = R(\overrightarrow{e_r}(\theta + \Delta\theta) - \overrightarrow{e_r}(\theta))$$

Under the small-change-in-angle hypothesis, vector $\overrightarrow{\Delta r}$ is tangent to the circle of radius R and thus collinear to $\overrightarrow{e_\theta}$ with amplitude $R\Delta\theta$:

$$\overrightarrow{\Delta r} \approx R\Delta\theta\ \overrightarrow{e_\theta} = (\Delta\theta\ \overrightarrow{e_z}) \times \overrightarrow{r}$$

Assuming that this change in position vector has occurred during a time interval Δt, the average velocity vector reads:

$$\overrightarrow{v}_{\Delta t} = \frac{\overrightarrow{\Delta r}}{\Delta t} \approx \left(\frac{\Delta\theta}{\Delta t}\overrightarrow{e_z}\right) \times \overrightarrow{r}$$

Taking the limit for vanishing time intervals yields:

$$\overrightarrow{v} = \frac{d\overrightarrow{r}}{dt} = \underbrace{\left(\frac{d\theta}{dt}\overrightarrow{e_z}\right)}_{\overrightarrow{\omega}} \times \overrightarrow{r}$$

The first vector on the right-hand side denoted by $\overrightarrow{\omega}$ is defined as the *angular velocity vector*:

Definition 11 (Angular velocity vector) *The* angular velocity vector *around a fixed axis* $(O, \overrightarrow{e_z})$ *(with S.I. physical unit* [rad/s]*) is defined as:*

$$\overrightarrow{\omega} = \frac{d\theta}{dt}\overrightarrow{e_z} \tag{7.2}$$

with $\theta(t)$ *the angle as a function of time in a cylindrical coordinate system around axis* $(O, \overrightarrow{e_z})$*. Its amplitude* $\omega = \frac{d\theta}{dt}$ *is the* angular velocity*.*

From velocity vector to the angular velocity vector

Hence the velocity of a point moving on a circle of fixed radius R is nothing but the vector product between the *angular velocity vector* $\overrightarrow{\omega}$ and the position vector $\overrightarrow{r}$:

$$\overrightarrow{v} = \overrightarrow{\omega} \times \overrightarrow{r}$$

However, when the radius is not fixed, a radial component has to be added leading for a planar motion normal to the $\overrightarrow{e_z}$-axis to:

$$\overrightarrow{v} = v_r\overrightarrow{e_r} + \overrightarrow{\omega} \times \overrightarrow{r} \tag{7.3}$$

where $v_r = dr/dt$. Applying the vector product by $\overrightarrow{r}$ to this equation the *angular velocity vector* can be expressed as a function of the velocity. Indeed, remembering that:

$$\overrightarrow{a} \times (\overrightarrow{b} \times \overrightarrow{c}) = \overrightarrow{b}(\overrightarrow{a} \cdot \overrightarrow{c}) - \overrightarrow{c}(\overrightarrow{a} \cdot \overrightarrow{b})$$

it reads:

$$\overrightarrow{r} \times \overrightarrow{v} = rv_r(\overrightarrow{e_r} \times \overrightarrow{e_r}) + \overrightarrow{\omega} r^2 - \overrightarrow{r}(\overrightarrow{\omega} \cdot \overrightarrow{r})$$

and since both $\overrightarrow{e_r} \times \overrightarrow{e_r}$ and $\overrightarrow{\omega} \cdot \overrightarrow{r}$ vanish:

$$r^2 \overrightarrow{\omega} = \overrightarrow{r} \times \overrightarrow{v} \tag{7.4}$$

It is worth noticing that the scaling factor r^2 is required by the misfit in physical dimension between the *angular velocity* $\overrightarrow{\omega}$ in $[\mathrm{rad/T}]$ and $\overrightarrow{r} \times \overrightarrow{v}$ in $[\mathrm{L^2T^{-1}}]$. Actually the amplitude of this vector is twice the area covered by the position vector $\overrightarrow{r}$ over any fixed time unit[2].

It is important to notice that the angular velocity vector, expressed as a function of both the position vector and the velocity vector is strongly frame-dependent. Considering for instance a body with constant velocity v_o along the x-axis with a constant offset a in the y-direction. Hence:

$$\overrightarrow{r} = v_o t \overrightarrow{e_x} + a\overrightarrow{e_y} \qquad \Rightarrow \overrightarrow{r} \times \overrightarrow{v} = -av_o \overrightarrow{e_z}, \qquad \overrightarrow{\omega} = -\frac{-a}{a^2 + (v_o t)^2} \overrightarrow{e_z}$$

Shifting the origin of the frame by $a\overrightarrow{e_y}$ would cause the vector product to vanish together with the angular velocity.

7.2 Angular momentum of a point-mass

Since the *angular velocity vector* is obtained, up to a scaling factor, applying the vector product with $\overrightarrow{r}$ onto velocity $\overrightarrow{v}$, the *angular momentum* will be defined applying the same vector product with $\overrightarrow{r}$ on the *momentum* itself, though without any rescaling.

Definition 12 *The* angular momentum *of a point-mass of momentum* $\overrightarrow{p} = m\overrightarrow{v}$ *with respect to the origin of the frame is defined as:*

$$\overrightarrow{L} = \overrightarrow{r} \times \overrightarrow{p} \tag{7.5}$$

As a consequence, for a single point-mass the *angular momentum* with respect to the centre of the reference frame is collinear with the *angular velocity vector* and reads:

$$\overrightarrow{L} = mr^2 \overrightarrow{\omega}$$

Hence, the amplitude of the *angular momentum* is proportional to the *angular velocity* $\omega = d\theta/dt$ with a scaling factor mr^2. The physical dimension of the angular momentum is $[\mathrm{ML^2T^{-1}}]$. Note that ω is the instantaneous angular velocity with respect to an instantaneous rotation axis that is normal to the plan defined by both the position vector $\overrightarrow{r}$ and the velocity vector $\overrightarrow{v}$. Note that this instantaneous axis is likely to change over time together with the angular velocity ω.

It is worth noticing again that, as the *angular velocity*, the *angular momentum* is strongly dependent on the frame of reference and more particularly on the reference point with respect to which the angular momentum is defined.

---

[2]This general result is better known in astronomy as the second Kepler's law. Indeed we know that $\|\overrightarrow{a} \times \overrightarrow{b}\|$ is the area of the parallelogram generated by the two vectors $\overrightarrow{a}$ and $\overrightarrow{b}$. Hence the area covered during Δt by the position vector is:

$$\Delta A = \frac{1}{2} \left\| (\overrightarrow{r} + \overrightarrow{\Delta r}) \times \overrightarrow{r} \right\|$$

dividing by Δt gives:

$$\frac{\Delta A}{\Delta t} = \frac{1}{2} \left\| \overrightarrow{0} + \frac{\overrightarrow{\Delta r}}{\Delta t} \times \overrightarrow{r} \right\| = \frac{1}{2} \|\overrightarrow{v} \times \overrightarrow{r}\|$$

Taking now the time derivative of the *angular momentum* gives:

$$\frac{d\overrightarrow{L}}{dt} = \overbrace{\underbrace{\frac{d\overrightarrow{r}}{dt}}_{\overrightarrow{v}} \times \overrightarrow{p}}^{\overrightarrow{0}} + \overrightarrow{r} \times \frac{d\overrightarrow{p}}{dt}$$

where the first term on the right-hand side vanishes since $\overrightarrow{p} = m\overrightarrow{v}$ is collinear with $\frac{d\overrightarrow{r}}{dt} = \overrightarrow{v}$. Hence accounting for Newton's second law:

$$\frac{d\overrightarrow{L}}{dt} = \overrightarrow{r} \times \overrightarrow{F} \tag{7.6}$$

where the right hand side $\overrightarrow{r} \times \overrightarrow{F}$ defines the *torque* or *moment of force* $\overrightarrow{F}$ with respect to O. *Torque* will be studied within details in section 7.3. Before doing so, it is worth studying what happens when this external *torque* vanishes.

7.2.1 Constant-angular-momentum motion

The angular momentum of a single particle vanishes when the right-hand side of Equation (7.6) vanishes which occurs in any of the following cases:

- the external force vanishes ($\overrightarrow{F} = 0$) ;
- the external force is applied at the reference of the frame ($\overrightarrow{r} = 0$);
- the two vectors $\overrightarrow{r}$ and $\overrightarrow{F}$ are collinear. This is in particular the case of central force fields provided that the centre of the central force field is on the $\overrightarrow{e_r}$ axis.

In all these cases the *angular momentum vector* $\overrightarrow{L}$ is constant, both in direction $\overrightarrow{e_z} = \overrightarrow{L_o}/\|\overrightarrow{L_o}\|$ and amplitude. Note that when the *angular momentum* itself vanishes $\overrightarrow{r}$ and $\overrightarrow{v}$ are collinear and the motion is thus one-dimensional.

When the angular momentum is constant and does not vanishes, the motion remains in a plane that is normal to $\overrightarrow{L}_o$. Indeed the body cannot leave this plane since both the position vector and the velocity are perpendicular to $\overrightarrow{L}_o$. One sticking consequence is that planets subjected to the central gravitational field of the Sun have their trajectory restricted to the so-called *ecliptic plane*. Moreover the amplitude of the *angular momentum vector* with value $mr^2 d\theta/dt$ has to remain constant too. When, the radius is fixed the angular velocity is constant and the motion is uniform along the corresponding cercle. When the radius is not fixed it corresponds to the already mentioned Kepler's Law of equal areas covered by the position vector over equal time:

$$\int_{\Delta\theta} r^2 d\theta = \frac{L_o}{m}\Delta t$$

It will be shown in Chapter 8 that the trajectory is either an ellipse or an hyperbola depending on the total energy of the system.

7.3 Torque

The *torque* or *moment of a force* with respect to the origin of the *reference frame* has been introduced in the conservation of *angular momentum* for a single particle (7.6). A more general definition is the following.

Definition 13 *The* torque *with respect to point A of a force $\overrightarrow{F}$ applied at point M is defined as:*

$$\overrightarrow{\tau_A} = \overrightarrow{AM} \times \overrightarrow{F} \tag{7.7}$$

Note that when the *torque* is defined with respect to the origin of the reference frame O the subscript is often omitted. Once known with respect a given point A the torque with respect to any other point B can easily be computed as:

$$\overrightarrow{\tau_B} = \overrightarrow{\tau_A} + \overrightarrow{BA} \times \overrightarrow{F} \tag{7.8}$$

It is also worth noticing that when $\overrightarrow{AM}$ and $\overrightarrow{F}$ are collinear the related torque vanishes. Hence, all points on a line including the location at which the force is applied and that is directed by this very same force vector experience no torque from this force. Similarly two points belonging to a line that is parallel to the force experience the same torque for this particular force. Hence the torque only varies on a plane that is perpendicular to the direction of the force and is invariant along that direction.

From a more physical point of view Formula (7.7) shows that *torque* has the same physical dimension $[\mathrm{ML}^2\mathrm{T}^{-2}]$ as a *work* or an *energy*. Yet it is different in nature and in SI units Torques are expressed in [Nm] and not in Joules. Actually it is the *torque* times the angle of rotation that has to be compared with *energy* or a *work*. Indeed, let us consider a force of fixed amplitude F along direction $\overrightarrow{e_\theta}$ applied on a point-mass located at $\overrightarrow{r} = R\overrightarrow{e_r}$, radius R being fixed. It then generates a torque $\overrightarrow{\tau}$ with respect to point O given by:

$$\overrightarrow{\tau} = (R\overrightarrow{e_r}) \times (F\overrightarrow{e_\theta}) = RF\ \overrightarrow{e_z}$$

with constant amplitude $\tau = RF$. The work done by the force when it rotates by any angle $\Delta\theta$, that is to say integrating along an elementary path $Rd\theta\overrightarrow{e_\theta}$ in cylindrical coordinates, reads:

$$W = \int_{\theta_o}^{\theta_o+\Delta\theta} \overrightarrow{F} \cdot (Rd\theta\overrightarrow{e_\theta}) = RF\Delta\theta = \tau\Delta\theta$$

This result can be summarised as follows:

Property 3 (Incremental work induced by a torque) *The increment of work ΔW induced by a constant torque $\tau\overrightarrow{e_z}$ in a rotation of angle $\Delta\theta$ around axis $(O, \overrightarrow{e_z})$ is:*

$$\Delta W = \tau\Delta\theta \tag{7.9}$$

This property can be generalised to a *torque* of varying direction and amplitude introducing the *angular velocity*.

Property 4 (Work done by a torque) *Work induced between time t_1 and t_2 by a torque $\overrightarrow{\tau}_A$ with respect to any point A is:*

$$\Delta W = \int_{t_1}^{t_2} \overrightarrow{\tau}_A \cdot \overrightarrow{\omega}_A dt \tag{7.10}$$

where $\overrightarrow{\omega}_A$ is the angular velocity vector *with respect to the same point A. The instantaneous power is then given by:*

$$\mathcal{P} = \frac{dW}{dt} = \overrightarrow{\tau}_A \cdot \overrightarrow{\omega}_A \tag{7.11}$$

From a practical point of view it is worth noticing that *power* and *torque* are the two main characteristics of a car engine. A high torque gives a good speed-up at low angular velocity whereas maximum power is usually reached at high angular velocity.

7.3.1 Torque on a system of point-masses

Up to now, forces have been applied on bodies considered as point-masses. Hence little attention has been paid to the exact location at which these forces were applied. *Torque* being now defined, the role played by this particular location at which this force is applied can be further investigated. Considering a set of forces $\overrightarrow{F}_i$ applied at location A_i with coordinate vector $\overrightarrow{r}_i$ in a given frame of reference, each one of then generating a torque $\overrightarrow{\tau}_i = \overrightarrow{r}_i \times \overrightarrow{F}_i$ with respect to the origin of the frame. The total or *net force* $\overrightarrow{F}$ and the *net torque* $\overrightarrow{\tau}$ are then equal to the sum of all individual terms:

$$\overrightarrow{F} = \sum_i \overrightarrow{F}_i, \qquad \overrightarrow{\tau} = \sum_i \overrightarrow{r}_i \times \overrightarrow{F}_i \tag{7.12}$$

The *net torque* can also be computed with respect to any point A as:

$$\overrightarrow{\tau}_A = \sum_i (\overrightarrow{r}_i - \overrightarrow{r}_A) \times \overrightarrow{F}_i = \overrightarrow{\tau} - \overrightarrow{r}_A \times \overrightarrow{F} \tag{7.13}$$

Hence, the detailed distribution of forces is not required to transfer the *net torque* from one point to another. Only the knowledge of the *net force* is required. When the *net force* $\overrightarrow{F}$ vanishes, the *net torque* with respect to any point is constant[3].

7.4 Conservation of Angular Momentum

Conservation of *angular momentum* has already been derived from Newton's law in the particular case of a single particle (see Equation 7.6). Besides, introducing the *centre of mass* (see Equation 6.3) of a set of point-masses, the *conservation of linear momentum* has been derived for this assembly of point-masses accounting for the Newton's third *actio-reactio* law. The purpose of this section is to generalise the conservation of angular to sets of point-masses. To this aim Newton's third *actio-reactio* law has to be extended to torque. Hence $\overrightarrow{F}_{j\to i}$ being the *interaction force* applied by body j on body i and $\overrightarrow{\tau}_{j\to i}$ being the *interaction torque* with respect to the origin of the frame O that is applied by body j on body i one assumes that:

$$\begin{aligned} \overrightarrow{F}_{j\to i} + \overrightarrow{F}_{i\to j} &= \overrightarrow{0} \\ \overrightarrow{\tau}_{j\to i} + \overrightarrow{\tau}_{i\to j} &= \overrightarrow{0} \end{aligned} \tag{7.14}$$

Let us now consider a set of N bodies with interaction forces $\overrightarrow{F}_{j\to i}$ of body j applied at the centre of mass C_i with location vector $\overrightarrow{r}_i$ of body i and interaction moments $\overrightarrow{\tau}_{j\to i}$ with respect

[3]It is also worth noticing that when $\overrightarrow{F}$ does not vanish, the torque remains constant when transported along the direction of this net force $\overrightarrow{F}$. As a consequence, there exist a line $(A, \overrightarrow{F})$ along which the torque vanishes. This particular point A is given by:

$$\overrightarrow{r}_A = \frac{\overrightarrow{\tau} \times \overrightarrow{F}}{\|\overrightarrow{F}\|^2}$$

to O and external forces $\overrightarrow{F}_{o\to i}$ and moments $\overrightarrow{\tau}_{o\to i}$ with respect to O. Each of the bodies satisfies the conservation of momentum and angular momentum:

$$\begin{aligned}\frac{d\overrightarrow{p_i}}{dt} &= \sum_{j\neq i}\overrightarrow{F}_{j\to i}+\overrightarrow{F}_{o\to i}\\ \frac{d\overrightarrow{L_i}}{dt} &= \sum_{j\neq i}\overrightarrow{\tau}_{j\to i}+\overrightarrow{\tau}_{o\to i}\end{aligned}$$

Summing up these two sets of equations with respect to all bodies gives the conservation of net momentum and net angular momentum:

$$\frac{d\overrightarrow{p}}{dt} = \overrightarrow{F}_{\text{ext}} \tag{7.15}$$

$$\frac{d\overrightarrow{L}}{dt} = \overrightarrow{\tau}_{\text{ext}} \tag{7.16}$$

With:

$$\begin{aligned}\overrightarrow{p} &= \sum_i \overrightarrow{p_i} = M\overrightarrow{v}_C\\ \overrightarrow{L} &= \sum_i \overrightarrow{L_i} = \overrightarrow{r}_C\times\overrightarrow{p}+\overrightarrow{L}_C, \qquad \overrightarrow{L}_C=\sum_i \overrightarrow{CC_i}\times(\overrightarrow{p}_i-\overrightarrow{p})\\ \overrightarrow{F}_{\text{ext}} &= \sum_i \overrightarrow{F}_{o\to i}\\ \overrightarrow{\tau}_{\text{ext}} &= \sum_i \overrightarrow{\tau}_{o\to i}\end{aligned}$$

and where C is the centre of mass of the system of particles.

It is worth noticing that the torque *actio-reactio* law (7.14) is not straightforward since the interaction forces could be applied at distant locations. When the two bodies are in close contact, the interaction forces apply at the same point. Hence this law states that local interaction torques at the contact point have equal amplitude and opposite directions. When the two bodies are at some distance, with only interaction forces applying on each of the two bodies at C_i and C_j respectively and no local interaction moments Equation (7.14) reads:

$$\overrightarrow{C_jC_i}\times\overrightarrow{F}_{j\to i}=\overrightarrow{0}$$

Hence the interaction force has to be aligned with the relative position vector. Central distant forces such as gravity or electrostatic forces satisfy this property. When this alignment condition is not satisfied a local distant torque has to be accounted for to fulfil the torque *actio-reactio* law.

Conservation of *angular momentum* , first derived from Newton's second law of motion for a point-mass, has been generalised to extended bodies. When *conservation of linear momentum* and *conservation of energy* are respective consequences of the invariance with respect to translation in space and time, *conservation of angular momentum* is a consequence of the invariance with respect to rotation in space.

7.4.1 Conservation of angular momentum at the centre of mass

Conservation of angular momentum in Equation (7.16) is expressed with respect to the origin of the frame O but can also be written with respect to any other fixed point A. Indeed, using the transport of both *angular momentum and torque* yields:

$$\begin{aligned} \frac{d}{dt}\left(\overrightarrow{L_A} + \overrightarrow{OA} \times \overrightarrow{p}\right) &= \overrightarrow{\tau_A} + \overrightarrow{OA} \times \overrightarrow{F} \\ \frac{d\overrightarrow{L_A}}{dt} + \overrightarrow{v_A} \times \overrightarrow{p} &= \overrightarrow{\tau_A} + \overrightarrow{OA} \times \left(-\frac{d\overrightarrow{p}}{dt} + \overrightarrow{F}\right) \end{aligned}$$

Where the last term of the left-hand side vanishes since A is fixed and the last term on the right-hand side vanishes when accounting for the balance of momentum. Note that the last term on the left-hand side also vanishes when A is the centre of mass of the body ($A = C$), even though this point is not fixed. Hence the conservation of angular momentum with respect to the moving centre of mass C reads:

$$\frac{d\overrightarrow{L_C}}{dt} = \overrightarrow{\tau_C} \tag{7.17}$$

Note that torque $\overrightarrow{\tau_C}$ applied at the centre of mass of the body does not depend on forces $\overrightarrow{F}$ applied at this very point. Hence Equation (7.17) known as the *Euler Equation* fully governs the rotation of the overall body independently from the motion of the centre of mass. When the torque at the centre of mass of the body vanishes the angular momentum with respect to the centre of mass is constant.

Constant angular momentum with respect to the centre of mass is satisfied when net torque at the centre of mass $\overrightarrow{\tau_C}$ vanishes. Since gravity does not induce any external moment at the centre of mass this occurs in many practical situations: Earth's rotation, a gymnast doing a somersault or an ice-skater spinning at various angular velocity. Mastering this constant angular momentum while spinning, somersaulting, flipping, hence controlling the angular velocity is the art of the gymnast. This is only achieved in changing the shape of the body.

The picture is simpler when the shape of the body does not change. We then refer to *rigid bodies*. Still the shape of the body plays a significant role on the angular velocity as obviously shown by an ice-skater whose angular velocity increases when her arms are closer to her body. Hence the way angular velocity scales with the angular momentum depends on a shape-related factor called the *moment of inertia*. Chapter **??** will introduce this important concept an will restrict the analysis to the rotation of rigid bodies with respect to a fixed axis.

Chapter 8

Central Force Motion

Centrale force motion and more particularly motion of planets have played a very significant role in the history of science and its developments. Though the heliocentric model had originally been suggested by Aristarchus of Samos 3rd century BCE, the geocentric Ptolemy's model had prevailed for many centuries for both its practicality and theological interpretations. Actually it is only with Kepler's law that the heliocentric model became more accurate and with Newton's laws of motion that it became simpler, though not closing philosophical arguments about the location of the centre of the universe.

All theoretical building blocks having been presented in previous chapters, this chapter will concentrate on the solving the the two-body gravitational problem, retrieving the three Kepler's Laws.

8.1 Planetary motion

In this section, the motion of two bodies of mass M_1 and M_2 are assumed to be under gravitational interaction only in an inertial frame. $\overrightarrow{r}_1$ and $\overrightarrow{r}_2$ are the position vectors of the centre of mass of these two bodies. As long as the two bodies do not collide the interaction force is:

$$\overrightarrow{F}_{1\to 2} = -\overrightarrow{F}_{1\to 2} = -\frac{\mathcal{G}M_1M_2}{r_{12}^2}\overrightarrow{e}_{12}, \qquad \overrightarrow{e}_{12} = \frac{\overrightarrow{r}_2 - \overrightarrow{r}_1}{r_{12}}, \qquad r_{12} = \|\overrightarrow{r}_2 - \overrightarrow{r}_1\|$$

The position vector of the centre of mass $\overrightarrow{r}_C$ of the two particles reads:

$$\overrightarrow{r}_C = \frac{M_1\overrightarrow{r}_1 + M_2\overrightarrow{r}_2}{M_1 + M_2}$$

Applying the conservation of momentum on the two bodies yields:

$$M_1\overrightarrow{v}_1 + M_2\overrightarrow{v}_2 = (M_1 + M_2)\overrightarrow{v}_C = \overrightarrow{\text{cst}} \tag{8.1}$$

Hence a frame of reference attached to the centre of mass is an inertial frame. In the followings, this frame is chosen to proceed with the analysis. In particular, without any lack of generality it is assumed that:

$$\overrightarrow{r}_C = \overrightarrow{0}, \qquad \overrightarrow{v}_C = \overrightarrow{0} \tag{8.2}$$

As a consequence, the system is fully characterise by the time-evolution of the relative position vector $\overrightarrow{r} = \overrightarrow{r}_2 - \overrightarrow{r}_1$. Dividing each of the second Newton's laws for each of the body by the mass

of each of them and substracting the two yields:

$$\frac{d^2}{dt^2}(\overrightarrow{r}_2 - \overrightarrow{r}_1) = \frac{\overrightarrow{F}_{1\to 2}}{M_2} - \frac{\overrightarrow{F}_{2\to 1}}{M_1} = \frac{M_1 + M_2}{M_2 M_1}\overrightarrow{F}_{1\to 2}$$

The reduced mass being defined as

$$\mu = \frac{M_2 M_1}{M_1 + M_2} \tag{8.3}$$

yields:

$$\mu \frac{d^2\overrightarrow{r}}{dt^2} = \overrightarrow{F}_{1\to 2} \tag{8.4}$$

Hence, $\overrightarrow{r}$ satisfies the second Newton's law for a fictitious body of mass μ and a gravitational force centered at the origin of the frame - actually the centre of mass of the system. Note that for this gravitational force to be the one applied on a mass μ the mass at the centre is the sum of the masses of the two objects $M = M_1 + M_2$.

8.1.1 Solution of the reduced gravitational problem

The reduced gravitational problem

$$\mu \frac{d^2\overrightarrow{r}}{dt^2} = -\frac{\mathcal{G}\mu M}{r^2}\overrightarrow{e_r} \tag{8.5}$$

is characterised by a constant angular momentum:

$$\overrightarrow{L}_o = \mu r^2 \overrightarrow{\omega} \tag{8.6}$$

Hence the trajectory is restricted to a plane normal to $\overrightarrow{e_z} = \frac{\overrightarrow{L}_o}{\|\overrightarrow{L}_o\|}$ and expressed in polar coordinates $(r(t), \theta(t))$ with:

$$r^2 \frac{d\theta}{dt} = \frac{L_o}{\mu} \tag{8.7}$$

which is nothing but the already introduced *second Kepler's Law*. Moreover it is characterised by a constant total energy E_o:

$$E_o = \mu \left(\frac{1}{2} \left(\left(\frac{dr}{dt} \right)^2 + \left(r \frac{d\theta}{dt} \right)^2 \right) - \frac{\mathcal{G}M}{r} \right) \tag{8.8}$$

The angular velocity in the second term on the right-hand side of the conservation of energy can be substituted using Equation (8.7)

$$E_o = \left(\frac{\mu}{2} \frac{dr}{dt} \right)^2 + \frac{L_o^2}{2\mu r^2} - \frac{\mathcal{G}\mu M}{r}$$

Hence, the analysis reduces to the case of one particle of mass μ moving along one direction r and subjected to a so-called *effective potential*:

$$V_{\text{eff}}(r) = \frac{L_o^2}{2r^2} - \frac{\mathcal{G}M}{r} = \frac{\mathcal{G}M}{r^2}(r_o/2 - r) \tag{8.9}$$

with r_o a characteristic distance defined as:

$$r_o = \frac{L_o^2}{\mu^2 \mathcal{G} M}$$

This effective potential consists in the classical attractive gravitational potential $V(r) = -\frac{\mathcal{G}M}{r}$ that dominates at large distances and a repulsive or centrifugal potential $\frac{L_o^2}{2r^2}$ that dominates at small radius. The stationary solution with fixed radius r is reached when:

$$V'_{\text{eff}} = 0$$

Hence for:

$$r_{\text{eq}} = r_o, \qquad \frac{E_{\min}}{\mu} = V_{\text{eff}}(r_o) = -\frac{GM}{2r_o}$$

The radius being fixed, the conservation of angular momentum $L_o = \mu r_{\text{eq}}^2 \omega$ shows that the angular velocity ω is fixed too, yielding:

$$r_{\text{eq}}^3 = \frac{\mathcal{G}M}{\omega^2}$$

Noticing that the period reads $T = 2\pi/\omega$ a particular instance of Kepler's second law stating that *the square of the period is proportional to the cube of the semi-axis* is retrieved. The general proof is given in next section through a qualitative analysis of the trajectory.

8.1.2 Qualitative analysis of the trajectory and third Kepler's Law

A first analysis of the solution can be conducted as a function of the total energy E_o Indeed, since:

$$\lim_{r\to 0} V_{\text{eff}} = +\infty, \qquad \lim_{r\to+\infty} V_{\text{eff}} = 0$$

when $-\frac{G\mu M}{2r_o} < E_o < 0$ the particle is trapped in the potential between the two roots with $r_{\min}$ and $r_{\max}$ of:

$$V_{\text{eff}}(r) = E_o$$

that reads:

$$r^2 - 2r_1 r + r_o r_1 = 0, \qquad \text{with} \qquad r_1 = \frac{r_{\max} + r_{\min}}{2} = -\frac{\mathcal{G}\mu M}{2E_o} \geq r_o$$

Hence:

$$r_{\min} = r_1\left(1 - \sqrt{1 - \frac{r_o}{r_1}}\right), \qquad r_{\max} = r_1\left(1 + \sqrt{1 - \frac{r_o}{r_1}}\right)$$

and the period can be computed as:

$$\begin{aligned} T/2 &= \int_{r_{\min}}^{r_{\max}} \frac{dr}{\sqrt{2}\sqrt{E_o/\mu - V_{\text{eff}}(r)}} \\ &= \frac{1}{\sqrt{-2E_o/\mu}} \int_{r_{\min}}^{r_{\max}} \frac{r dr}{\sqrt{-r^2 + 2r_1 r - r_o r_1}} \end{aligned}$$

Noticing that:

$$-r^2 + 2r_1 r - r_o r_1 = (r - r_{\min})(r_{\max} - r)$$

leads to the change of variable:

$$r = r_1 + \frac{r_{\max} - r_{\min}}{2}\cos\phi$$

yielding:

$$\begin{aligned} T/2 &= \frac{r_{\max}-r_{\min}}{2\sqrt{-2E_o/\mu}}\int_0^\pi \left(\cos\phi + \frac{r_{\max}+r_{\min}}{r_{\max}-r_{\min}}\right)d\phi \\ T &= 2\pi\frac{r_{\max}+r_{\min}}{2\sqrt{-2E_o/\mu}} \end{aligned} \tag{8.10}$$

Finally, since $r_1 = \frac{r_{\max}+r_{\min}}{2} = -\frac{\mathcal{G}\mu M}{2E_o}$ is half of the larger axis of the trajectory, we finally obtain the *third Kepler's Law* stating that for any two planets in the solar system, the ratio of the square of their periods is equal to the cube of ratios of their large axes:

$$\frac{T^2}{r_1^3} = \frac{2\pi}{GM}$$

When $E_o \geq 0$ the fictitious particle is not anymore trapped by the potential. The fictitious particle will reach infinity at large times meaning that the distance between the two original bodies is not bounded.

8.1.3 Trajectories and First Kepler's Law

The generic trajectory can be obtained using the change of variable $u(\theta) = 1/r$. Hence:

$$\dot{r} = -u'\dot{\theta}/u^2 = -u'L_o/\mu$$

with $u' = \frac{du}{d\theta}$ and $\dot{\theta} = \omega = \frac{d\theta}{dt}$. When used in the energy conservation equation this change of variable leads to:

$$\frac{1}{2}u'^2 + \frac{1}{2}\left(u - \frac{1}{r_o}\right)^2 = \frac{-1}{2r_1 r_o} + \frac{1}{2r_o^2} = \frac{1}{2}\frac{e^2}{r_o^2}$$

with $e^2 = \frac{2E_o r_o}{GM\mu} + 1$ which also reads $e^2 = \frac{r_1 - r_o}{r_1}$ when E_o is non-positive.

$$u - \frac{1}{r_o} = \frac{e}{r_o}\cos\theta$$

being a straightforward solution, we finally obtain the generic trajectory in polar coordinates:

$$r = \frac{r_o}{e\cos\theta + 1}$$

This gives an elliptic trajectory when $e < 1$ (circular for $e = 0$) which is nothing but the *first Kepler's Law*. It is worth noticing that a parabolic trajectory is obtained for $e = 1$ and an hyperbolic one for $e > 1$.

8.2 Gravitational fields of extended bodies (not covered in class)

The elementary gravitational field induced at point M by a element of mass $dm' = \rho dV'$ located at M' reads:

$$d\overrightarrow{g} = -\frac{\mathcal{G}}{\|\overrightarrow{MM'}\|^3}\overrightarrow{MM'}dm' \tag{8.11}$$

Hence the cumulative gravity field at M is obtained summing over the entire body Ω:

$$\overrightarrow{g} = -\int_{\Omega} \frac{\mathcal{G}}{\|\overrightarrow{MM'}\|^3} \overrightarrow{MM'} dm' \tag{8.12}$$

At first sight, the integral is rather difficult to compute especially for complex shapes and density distributions ρ over volume Ω. Keeping in mind applications to celestial bodies having in first approximation a spherical symmetry the gravitational field is also expected to have this spherical symmetry, at least as a first approximation. Let us now make a very simple observation on the gravitation field $\overrightarrow{dg}$ induced by a point-mass dm'. The sum of its radial component $\overrightarrow{dg} \cdot \overrightarrow{e_r}$ other any sphere S_o of radius R centered on this mass is equal to $4\pi\mathcal{G}dm'$. Indeed, this radial component is uniform over the sphere of surface $4\pi R^2$ with value $\mathcal{G}dm'/R^2$ Hence:

$$\int_{S_o} \overrightarrow{g} \cdot \overrightarrow{n}\, dS = -4\pi R^2 \, \frac{\mathcal{G}dm'}{R^2} = -4\pi\mathcal{G}dm'$$

Let us then notice that this property remains valid for every surface S enclosing the point mass:

$$\int_{S} \overrightarrow{g} \cdot \overrightarrow{n}\, dS = -4\pi\mathcal{G}dm' \qquad \text{when the point-mass } dm' \text{ is inside } S \tag{8.13}$$

To prove this result, we only need to prove that:

$$\int_{S} \overrightarrow{g} \cdot \overrightarrow{n}\, dS = 0 \qquad \text{when the point-mass } m' \text{ is outside } S \tag{8.14}$$

which is a direct consequence of the so-called divergence theorem as shown in this note[1]. Then Equation (8.13) can be proven in choosing a sphere S_o inside S and noticing that $S \cup S_o$ is a closed surface with no mass inside. Hence:

$$\int_{S} d\overrightarrow{g} \cdot \overrightarrow{n}\, dS = \int_{S \cup S_o} d\overrightarrow{g} \cdot \overrightarrow{n}\, dS + \int_{S_o} d\overrightarrow{g} \cdot \overrightarrow{n}\, dS = 0 - 4\pi\mathcal{G}dm' \tag{8.15}$$

We can then sum up all the elementary gravitational fields induced by all masses in volume Ω_{int} inside S to obtain:

$$\int_{S} \overrightarrow{g} \cdot \overrightarrow{n}\, dS = -4\pi\mathcal{G}M_{\text{int}} \qquad \text{with} \qquad M_{\text{int}} = \int_{\Omega_{\text{int}}} dm' \tag{8.16}$$

[1] The divergence theorem states that for any volume Ω with smooth boundary S and any smooth vector field $\overrightarrow{g}$:

$$\int_{\Omega} \text{div}\overrightarrow{g}\, dV = \int_{S} \overrightarrow{g} \cdot \overrightarrow{n}\, dS$$

where $\overrightarrow{n}$ is the outer unit normal vector of the boundary and div is the divergence operator defined in cartesian coordinates as:

$$\text{div}\overrightarrow{g} = \frac{\partial g_x}{\partial x} + \frac{\partial g_y}{\partial y} + \frac{\partial g_z}{\partial z}$$

Remembering that the gravity field $\overrightarrow{g}$ is the gradient of the gravity potential V:

$$g_x = \frac{\partial V}{\partial x}, \qquad g_y = \frac{\partial V}{\partial y}, \qquad g_z = \frac{\partial V}{\partial z}$$

yields:

$$\text{div}\overrightarrow{g} = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = \Delta V$$

where Δ is the so-called Laplacian operator. Since $V = \mathcal{G}m'/r$ is only a function of the radius, its Laplacian takes a very simple form when $r \neq 0$:

$$\Delta V = \frac{d}{dr}\left(r^2 \frac{dV}{dr}\right) = -\frac{d}{dr}\left(r^2\, \mathcal{G}m' \frac{1}{r^2}\right) = \frac{d}{dr}\left(\mathcal{G}m'\right) = 0$$

which shows the expected property.

Hence, the average normal component of the gravity field on a sphere of radius R is given by:

$$g_{\text{avg}} = \frac{1}{4\pi R^2} \int_S \overrightarrow{g} \cdot \overrightarrow{n}\, dS = -\frac{\mathcal{G}M_{\text{int}}}{R^2}$$

which turns out to be the same normal field as the one induced by a point-mass M_{int} located at the center of the sphere. Obviously, when the gravitational object has a spherical symmetry - Ω is a sphere and the mass density only depends on the radius - so is the induced gravitational field. Thus the gravity field has only a radial component at each point the amplitude of which is uniform on the sphere and equal to the average value:

Theorem 5 *The gravity field induced at a distance R from the center from a spherical body with a mass density depending on the radius only is identical to the gravity field induced by a point-mass located at the center of the sphere with a total mass M_{int} equal to the mass of the body located below R:*

$$\overrightarrow{g} = -\frac{\mathcal{G}M_{\text{int}}}{R^2}\overrightarrow{e_r} \tag{8.17}$$

As a consequence the gravity field induced by a spherical homogeneous sphere of mass M and radius R is:

$$\overrightarrow{g}(r) = -\frac{r\mathcal{G}M}{R^3}\overrightarrow{e_r}, \qquad \text{when} \qquad r \leq R \tag{8.18}$$

$$\overrightarrow{g}(r) = -\frac{\mathcal{G}M}{r^2}\overrightarrow{e_r}, \qquad \text{when} \qquad r \geq R \tag{8.19}$$

The related gravitational potential V is given by:

$$V(r) = \frac{(3R^2 - r^2)\mathcal{G}M}{2R^3}, \qquad \text{when} \qquad r \leq R \tag{8.20}$$

$$V(r) = \frac{\mathcal{G}M}{r}, \qquad \text{when} \qquad r \leq R \tag{8.21}$$

Note that this result shows that measuring of the gravity field at free surface does not give any clue about the way mass is distributed as a function of depth. However buoyancy principles, deduced from Newton's law of motion concludes that materials with higher density will be drawn deeper by gravity. Indeed, let us consider a certain volume V with mass M subjected to a given gravity field $\overrightarrow{g}$ and being at rest with respect to its surrounding environment. We can then conclude from Newton's second law that the environment applies a force $\overrightarrow{F} = -M\overrightarrow{g}$ on V. Keeping the same volume V but adding an change in mass ΔM, the force applied by the environment will not change while the weight will increase when $\Delta M > 0$ or decrease when $\Delta M < 0$. Hence the net force on V becomes $\Delta M\overrightarrow{g}$ inducing following Newton's law of motion an acceleration:

$$\overrightarrow{a} = \frac{\Delta M}{M + \Delta M}\overrightarrow{g}$$

Hence volume V will move towards $\overrightarrow{g}$ when $\Delta M > 0$, that is to say downwards, or upyards when $\Delta M > 0$ segregating thus material with larger density ($\rho = V/M$) at the bottom and lighter density on top. Actually when Earth was formed about 4.5 Billion years ago, with a few hundreds of millions of years, the heavier elements such as iron sunk down to the core. Lighter elements moved up in the atmosphere. Continental plates made of materials lighter than the mantel - mainly aluminium and silicates - and the oceanic plate float on top of it.

8.2.1 Non-uniform gravity fields

Apart from this spherical stratification of density induced by buoyancy, earth shows lateral heterogenities in density that induces small variation of the gravitational fields that can be detected by several means, including the change in the frequency of a pendulum. This technique is extensively used to detect cavities in the soil from surface observations. Indeed, the fluctuation of the gravity field induced by lack of mass $-\Delta M = -4\pi/3 d^3 \rho$, with d the radius and ρ the mass density, at depth h scales as:

$$\Delta g = \mathcal{G}\frac{\Delta M}{h^2}$$

Hence the relative change in gravity field is given by:

$$\frac{\Delta g}{g_o} = \frac{\Delta M}{M_E}\frac{R_E^2}{h^2} = \frac{\rho}{\rho_E}\frac{d^2}{h^2}\frac{d}{R_E}$$

With R_E the radius of Earth and ρ_E its average density being twice to three times the density of the crust. Hence assuming that d and h are of the same order of magnitude, the required precision is of the order of $\frac{d}{R_E}$, hence of the order of 10^{-6}.

Lateral variations of the gravity fields induced at the free-surface of the oceans by the topography of the sea bed is nowadays currently measured by satellites. Indeed, let g_o be a the gravity field at the surface of the ocean with a flat seabed 4km deep. An underwater volcano of conical shape, 3km high with a base radius of 20 km has a volume of $\pi R^2 h/3 \approx 1.2\ 10^{12} m^3$ and a differential mass as compared to water of $\Delta M = 1.2\ 10^{15} kg$ with its center of mass at depth $d = 3km$. The additional gravity potential ΔV at the original free-surface given by:

$$\Delta V = \frac{\mathcal{G}\Delta M}{d} \approx \frac{7.\ 10^{-11} 1.2\ 10^{15}}{3\ 10^3}\ [\mathrm{m}^2/\mathrm{s}^2] \approx 30\ [\mathrm{m}^2/\mathrm{s}^2]$$

Since the actual free-surface is defined by a constant gravity potential ($V(h) = \mathrm{cst}$). hence the modified height Δh with the perturbation satisfies:

$$V(z_o) = (V + \Delta V)(z_o + \Delta h) \approx V(z_o) + \Delta V + \underbrace{\frac{dV}{dz}} = -g_o \Delta h$$

where g_o is the gravity at ground surface. Hence, simplifying on both sides:

$$\Delta h = \frac{\Delta V}{g_o} \approx 3\ [\mathrm{m}] \tag{8.22}$$

Present precision level of altimetric data by satellites is ± 1 [cm], hence bathymetry of oceans can be done with a precision of the order of ± 30 [m]. However raw data have to be corrected from many factors including: earth rotation[2], ocean and continental tides[3], deeper density heterogeneities in the Earth's mantel. Figure 8.1 shows the large wavelength anomalies of the Earth surface and the estimated bathymetry.

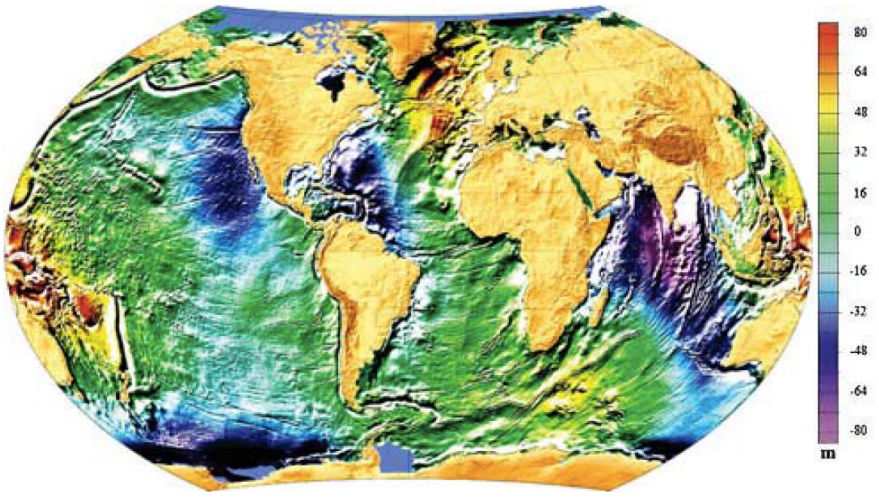
---

[2]

$$\Delta g_\omega = \omega^2 R = R\frac{\Delta g_\omega}{2} = \left(\frac{2\pi}{24 \cdot 3600\ [\mathrm{s}]}\right)^2\ (6.410^6)\ [\mathrm{m}] \approx 0.03\ [\mathrm{m/s}^2]$$

That is 0.3 % change in apparent gravity g_o. $\Delta V = \omega^2 R^2/2$ and hence:

$$\Delta h = \frac{R}{2}\frac{\Delta g_\omega}{g_o} \approx 10\ [\mathrm{km}]$$



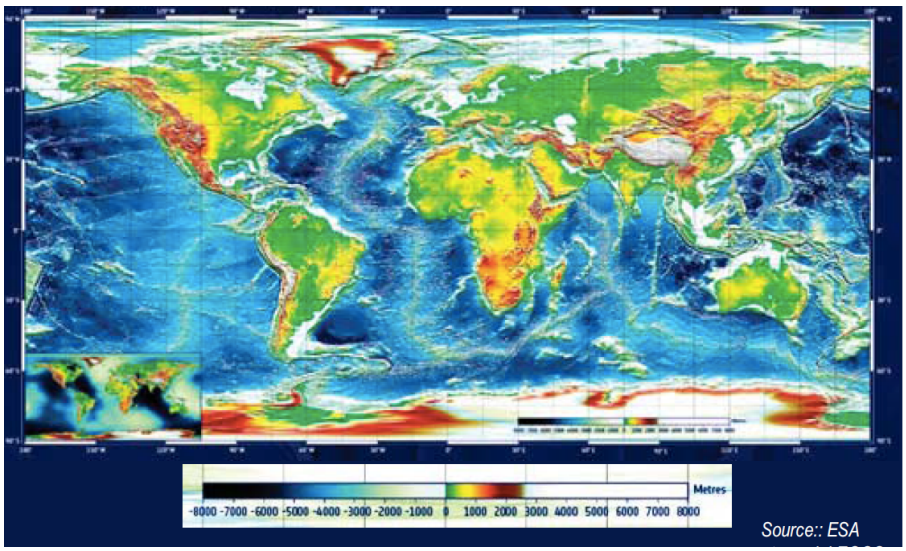


Figure 8.1: (top) Deviation of Earth's geoid from the reference ellipsoid shape of a rotating spherical symmetric Model, (bottom) induced bathymetry.

[3]

$$\Delta V \approx \Delta g_{\text{moon}} R_{\text{earth}} = \frac{2R_{\text{earth}}^2}{d} g_{\text{moon}}$$

Hence:

$$\Delta h = \frac{2R_{\text{earth}}^2}{d} \frac{g_{\text{moon}}}{g_o} \approx 0.4 \text{ [m]}$$

where $R = 6.4\ 10^6$ [m] is the Earth's radius, $d = 3.810^8$ [m] is the distance between Earth and Moon and $M_{\text{moon}} = 7.3\ 10^{22}$ [kg] is the mass of the Moon. Note that this value is much less than what is observed. In particular the highest tide reaches an amplitude of 13 [m] at Mont Saint-Michel. This amplification is due to a resonance phenomenon when high waters are confined along the coastline. Actually this value is close to the observed continental tide od about $30cm$. Note also that this change in height does not occur instantaneously when the Moon is above the observation point due to some viscosity in water, crust and mantel causing some dissipation. This phenomena caused this alignement of the rotation of the moon around its axis with its rotation around Earth. This also causes a slowing down of the rotation of the Earth and as a consequence an increase of the Earth-Moon distance, preserving then the angular momentum of the Earth-Moon system.

Chapter 9

Harmonic oscillator

The harmonic oscillator or harmonic resonator is a simple and abstract model being at the heart of theoretical physics and in particular quantum physics. It is also at the basis of many practical applications both in mechanical or electrical engineering. It is used to model: building resistance under wind or earthquake loadings ; ship stability and comfort when subjected to waves ; vibration-isolation systems, music instruments, car-suspension systems, sloshing of your cup of coffee.

In this chapter we first consider a simple model consisting of a one dimensional mass-spring-damper system subjected to external forces as depicted in Figure 9.1 to obtain the corresponding second order differential equation governing its dynamic behavior. In addition we will show how similar equations can be derived by means of a perturbation approach applied to a generic energy conservation law depending on a single parameter.

Afterwards the analysis of the response of such system is conducted in two steps : first the free-vibration response in the case of vanishing applied loads and subsequently the case of harmonic - or sinusoidal - input loads. Extension to transient applied loads is finally introduced.

9.1 The Single-Degree of Freedom oscillator

The one dimensional mass-spring-damper system shown on Figure 9.1 is a model system of a Single-Degree-of Freedom (SDOF) oscillator. Mass M is allows to move without friction along the horizontal x-axis. x_{eq} is coordinate of the equilibrium position and $x(t)$ the actual coordinate at time t. $x_o = x(0)$ is the initial position. $u(t) = x(t) - x_{eq}$ denotes the spread from the equilibrium position inducing a horizontal reaction force from the spring $F_k(t) = -Ku(t)$, with k the stiffness of the spring. Horizontal velocity relative is $v(t) = \frac{dx}{dt} = \frac{du}{dt}$ and horizontal acceleration

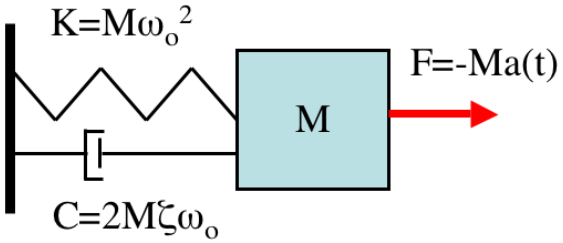


Figure 9.1: Model of an harmonic oscillator as a one dimensional mass-spring-damper system

is $\frac{dv}{dt}(t) = \frac{d^2u}{dt^2}$. The force induced by the damper is proportional to the velocity with opposite direction $F_d(t) = -cv(t)$ with c the damping coefficient. Initial velocity is $\frac{du}{dt}(0) = v_o$. An external horizontal force $f(t)$ is assumed to be applied on mass M. Hence the horizontal component of Newton's second law reads:

$$F_k(t) + F_d(t) + f(t) = M\frac{dv}{dt}(t)$$

Taking

$$\omega_o = \sqrt{\frac{K}{M}} \quad , \quad 2M\omega_o\zeta = c \tag{9.1}$$

finally yields to the model second-order differential equation:

$$\begin{aligned} \omega_o^2 u + 2\zeta\omega_o\dot{u} + \ddot{u} &= \frac{f(t)}{M} && (9.2)\\ u(0) &= u_o && (9.3)\\ \dot{u}(0) &= v_o && (9.4) \end{aligned}$$

It is worth noticing that when the anchoring point of the spring and the damper is moving with respect to an inertia frame of reference with a given function $x_b(t)$, relative displacement u and velocity v remain unchanged. However total acceleration of the mass becomes $\frac{d^2x}{dt^2} = \frac{d^2x}{dt^2} + \frac{d^2x_b}{dt^2}$. Hence, in absence of any other external force, Equation (9.2) is still valid for this system taking $\frac{f(t)}{M} = -\frac{d^2x_b}{dt^2}$.

9.1.1 Linearisation

We now consider a dynamic system depending on a single parameter x with potential energy $\mathcal{E}_{\mathrm{p}}(x)$ and kinetic energy $\mathcal{E}_{\mathrm{kin}}$ that reads:

$$\mathcal{E}_{\mathrm{kin}}(x, \dot{x}) = \frac{M(x)}{2}\dot{x}^2$$

with $\dot{x} = \frac{dx}{dt}$ and $M(x)$ the generalised mass. Conservation of energy reads:

$$\frac{d}{dt}(\mathcal{E}_{\mathrm{kin}}(x, \dot{x}) + \mathcal{E}_{\mathrm{p}}(x)) = \mathcal{P}_{\mathrm{ext}}(x, \dot{x}) + \mathcal{P}_{\mathrm{dis}}(x, \dot{x})$$

where $\mathcal{P}_{\mathrm{ext}}(x, \dot{x})$ is the power of external forces and $\mathcal{P}_{\mathrm{dis}}(x, \dot{x})$ the power of dissipative ones. The system is assumed to have a stable equilibrium position at x_{eq}. Hence :

$$\frac{d\mathcal{E}_{\mathrm{p}}}{dx}(x_{eq}) = 0 \quad , \quad \frac{d^2\mathcal{E}_{\mathrm{p}}}{dx^2}(x_{eq}) = K > 0 \quad \text{and} \quad \frac{d^2\mathcal{E}_{\mathrm{kin}}}{d\dot{x}^2}(x_{eq}, 0) = M(x_{eq}) = M > 0$$

and the analysis is restricted to small perturbations $u(t)$ of the parameter x around x_{eq}:

$$u(t) - x(t) - x_{eq} \quad , \quad \dot{u}(t) = \dot{x}(t)$$

Hence:

$$\begin{aligned} \mathcal{E}_{\mathrm{p}}(x) &\approx \mathcal{E}_{\mathrm{p}}(x_{eq}) - \frac{K}{2}u^2\\ \mathcal{E}_{\mathrm{kin}}(x, \dot{x}) &\approx \frac{M}{2}\dot{u}^2\\ \mathcal{P}_{\mathrm{ext}}(x, \dot{x}) &\approx f(t)\ \dot{u}\\ \mathcal{P}_{\mathrm{dis}}(x, \dot{x}) &\approx -c\ \dot{u}^2 \leq 0 \end{aligned}$$

As a consequence the conservation of energy takes the following form:

$$M\ddot{u}\dot{u} + Ku\dot{u} = f(t)\ \dot{u} - c\ \dot{u}^2$$

Simplifying this equation by $\dot{u}$ gives Equation (9.2). Hence even though the perturbation of parameter x around the equilibrium position x_{eq} is not a displacement of a well-identified mass M the system will show the same time evolution as the one of the model one-dimensional mechanical oscillator. This time evolution is studied in the next two sections, but first it is worth to highlight some of its properties.

9.1.2 Properties

It is important to notice that when referring to the time evolution of a linear Single-Degree-of-Freedom oscillator the three equations (9.2), (9.3) and (9.4) are to be considered. Equation (9.2) is the second order differential equation governing the evolution with time when Equations (9.3) and (9.4) are the initial conditions. Hence, the parameters governing the time evolution are:

- the constitutive parameters ω_o and ζ known respectively as the *natural circular frequency* and the *damping ratio*. Other parameters could be defined such as the *natural period* $T_o = 2\pi/\omega_o$ or the quality factor $Q = \frac{1}{2\zeta}$.
- the applied loads: u_o, v_o and $f(t)/M$, that can be further decomposed into:
  - Initial conditions: initial displacement u_o and initial velocity v_o,
  - and applied equivalent acceleration $f(t)/M$.

Linearity of the time evolution u with respect to the applied loads is quite obvious. Indeed if $u(t)$ satisfies the three equations (9.2), (9.3) and (9.4), then for any α, $\alpha u(t)$ satisfies the same three equations with applied loads αu_o, αv_o and $\alpha f(t)/M$. Similarly, when $u_1(t)$ and $u_2(t)$ are respective solutions of equations (9.2), (9.3) and (9.4) with respective applied loads $(u_{1o}, v_{1o}, f_1(t)/M)$ and $(u_{2o}, v_{2o}, f_2(t)/M)$, then $u_1(t) + u_2(t)$ is a the solution of equations (9.2), (9.3) and (9.4) with applied loads $(u_{1o} + u_{2o}, v_{1o} + v_{2o}, (f_1(t) + f_2(t))/M)$.

Hence to find the general solution, it is convenient to first choose $(u_{1o} = u_o, v_{1o} = v_o, f_1(t)/M = 0)$ and then $(u_{2o} = 0, v_{2o} = 0, f_2(t)/M = f(t)/M)$. u_1 is called the *Free oscillation response* as no forces are applied over time and u_2 is called *the forced vibration* response as vanishing initial conditions are imposed.

9.2 Free vibration of a damped SDOF

From a linearity argument we already know that the free-vibration solution defined by equations (9.2), (9.3) and (9.4) with $f(t) = 0$ simply reads:

$$u(t) = u_o g_1(t) + \frac{v_o}{\omega_o} g_2(t)$$

where $g_1(t)$ and $g_2(t)$ are functions satisfying the homogeneous second order differential equation:

$$\omega_o^2 g + 2\zeta\omega_o \dot{g} + \ddot{g} = 0$$

When ζ vanishes, the oscillator is said undamped and the solution is simply $g_1(t) = \cos(\omega_o t)$ and $g_2(t) = \sin(\omega_o t)$. In other case, the response of the Free-Vibration response of a SDOF oscillator is:

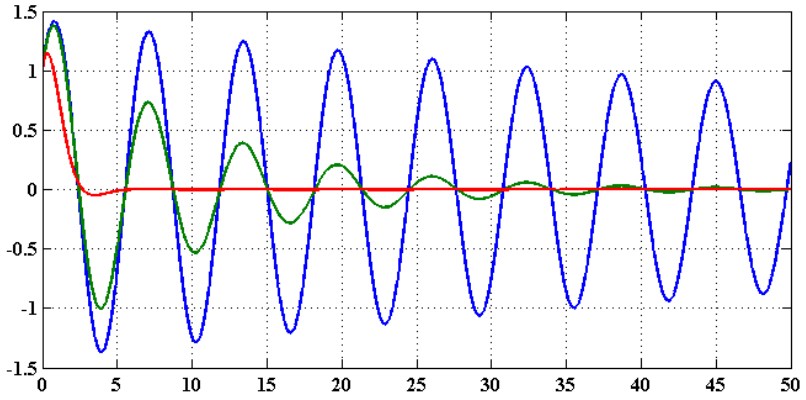


Figure 9.2: The response of a SDOF oscillator $\omega_o = 1, \zeta = 0.01, 0.1, 1$ to initial conditions $u_o = 1, v_o = 1$

Underdamped case: for $0 \leq \zeta < 1$, $\omega_d = \sqrt{1-\zeta^2}\omega_o$:

$$u(t) = e^{-\zeta\omega_o t}\left(u_o \cos(\omega_d t) + \frac{v_o + \zeta\omega_o u_o}{\omega_d}\sin(\omega_d t)\right) \tag{9.5}$$

Overdamped case: for $\zeta > 1$, $\omega'_d = \sqrt{\zeta^2-1}\omega_o$:

$$u(t) = e^{-\zeta\omega_o t}\left(u_o \cosh(\omega'_d t) + \frac{v_o + \zeta\omega_o u_o}{\omega'_d}\sinh(\omega'_d t)\right) \tag{9.6}$$

Critically damped case: for $\zeta = 1$:

$$u(t) = e^{-\omega_o t}\left(u_o + (v_o + \omega_o u_o)t\right) \tag{9.7}$$

These solutions are plotted on Figure 9.2 for the underdamped and critically damped cases where it can clearly be seen that the exponential decay of the solution is governed by the relaxation time $\tau = 1/(\zeta\omega_o$ whereas the oscillating feature is controlled by the modified circular frequency $\omega_d = \sqrt{1-\zeta^2}\omega_o$. It is worth noticing that when ζ is small $\omega_d \approx \omega_o$.

Actually, the solution can be further simplified taking:

$$\begin{aligned} u(t) &= (v_o + 2\zeta\omega_o u_o)H(t) + \dot{H}(t)u_o \\ H(t) &= e^{-\zeta\omega_o t}\frac{\sin(\omega_d t)}{\omega_d}, \qquad \omega_d = \omega_o\sqrt{1-\zeta^2} \end{aligned}$$

where in the over-damped case $\omega_d = j\omega'_d$ with $j = \sqrt{-1}$ yielding:

$$\frac{\sin(\omega_d t)}{\omega_d} = \frac{\sinh(\omega'_d t)}{\omega'_d}$$

In the critically damped case $\omega_d \to 0$ yielding:

$$\frac{\sin(\omega_d t)}{\omega_d} \to t$$

As explained in a subsequent section $H(t)$ is the *impulse response function* of the SDOF linear oscillator.

9.3 Harmonic response of a damped SDOF

Section 9.2 has shown that the free-vibrating response of a linear SDOF oscillator has an exponential decay over time provided that some damping is present. In this section, we will analyse the long-term response of an oscillator subjected to harmonic loads. Hence, we do not pay attention to initial conditions and only consider an applied force $f(t) = F_o \cos \omega t$ where ω is the circular frequency of the load, not to be confused with ω_o the natural circular frequency of the oscillator.

It can be shown that the long term response of such an linear oscillator has to be an oscillating function of time with the same circular frequency ω, an unknown amplitude $|\hat{u}|$ and a so-called phase shift ϕ:

$$u(t) = |\hat{u}| \cos(\omega t + \phi)$$

Since the complex exponential function is defined as:

$$e^{j\theta} = \frac{\cos\theta + j \sin\theta}{2}$$

One can write:

$$f(t) = \Re(F_o e^{j\omega t})$$

where $\Re$ stands for the real part of a complex number, and:

$$u(t) = \Re(\underbrace{|\hat{u}|e^{j\phi}}_{\hat{u}(\omega)} \; e^{j\omega t})$$

where $\hat{u}$ is the complex amplitude of the response bearing at the same time the amplitude $|\hat{u}|$ and the phase $\phi = \frac{\Im(\hat{u})}{\Re(\hat{u})}$ ($\Im$ being the imaginary part of a complex number). Thanks to linearity of the real part one can easily notice that:

$$\dot{u}(t) = \frac{d}{dt}\Re\left(\hat{u}e^{j\omega t}\right) = \Re\left(\hat{u}\;\frac{d}{dt}\left(e^{j\omega t}\right)\right) = \Re\left(j\omega\hat{u}\;e^{j\omega t}\right)$$

and:

$$\ddot{u}(t) = \frac{d^2}{dt^2}\Re\left(\hat{u}e^{j\omega t}\right) = \Re\left(-\omega^2\hat{u}\;e^{j\omega t}\right) = -\omega^2 u(t)$$

Noticing now that all coefficients in Equation (9.2) are real number yields:

$$\Re\left(\left(\omega_o^2 + 2j\zeta\omega_o\omega - \omega^2\right)\hat{u}\;e^{j\omega t}\right) = \Re\left(\frac{\widetilde{F}}{m}e^{j\omega t}\right)$$

Hence, taking:

$$\hat{u} = \hat{H}(\omega)\frac{F_o}{M\omega_o^2} \qquad \text{with} \qquad H(\omega) = \frac{\omega_o^2}{\omega_o^2 + 2j\zeta\omega_o\omega - \omega^2} \tag{9.8}$$

Equation (9.2) is satisfied taking $u(t) = \Re(\hat{u}\; e^{j\omega t})$. This solution is fully governed by the complex-valued function of the frequency $\hat{H}(\omega)$ known as the *Frequency Response Function* of the SDOF oscillator. The amplitude and the phase of the function is plotted on Figure 9.3 using a Log-Log scale for different values of the damping ratio ζ. Three different regimes can easily be identified on this plot:

- **Low frequency regime** with $\omega \ll \omega_o$ for which the amplitude is almost equal to unity and the phase vanishes. Hence almost nothing happens at low frequency and the oscillator acts as a low-pass filter.

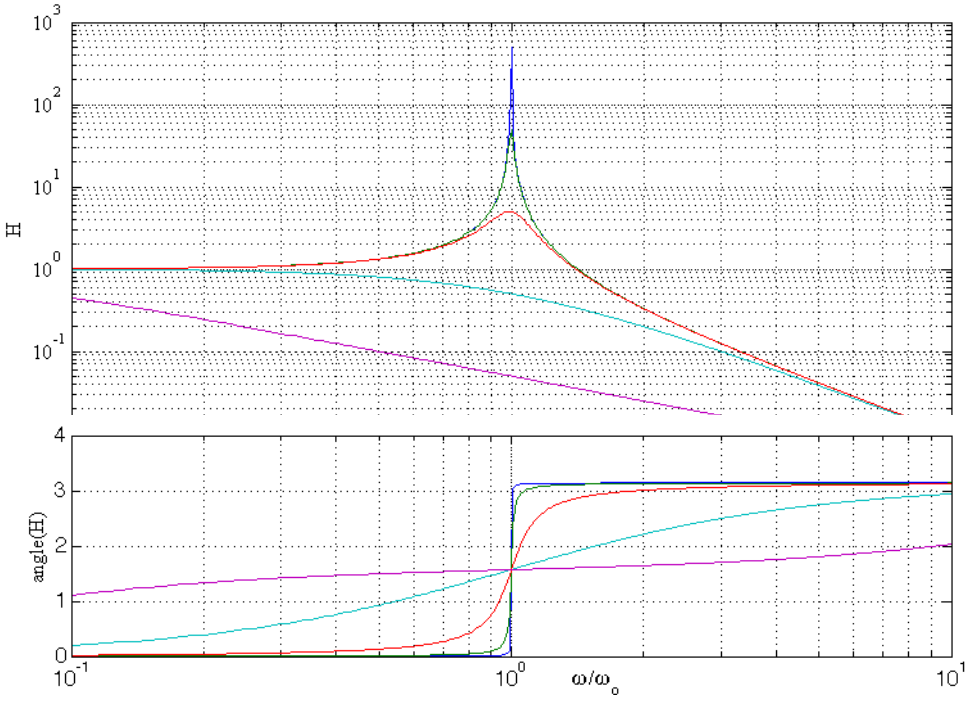


Figure 9.3: Amplitude and phase of the Frequency-Response-Function, FRF, for $\zeta = .001, 0.01, 0.1, 1, 10$

- **Resonance regime** with $\omega \approx \omega_o$ for which the amplitude is strongly amplified in the underdamped case with an amplitude inversely proportional to the damping ratio:

$$|H(\omega_o)| = \frac{1}{2\zeta\omega_o^2} = \frac{Q}{\omega_o^2}$$

  and a phase shifting rapidly with a value of $-\pi/2$ at ω_o.

- **High frequency regime** with $\omega_o \ll \omega$ for which the amplitude decreases as ω^{-2} and the phase approaches $-\pi$. Hence for input frequencies that are higher than the natural frequency of the oscillator the amplitude of the response is reduced. This phenomenon is essential to design vibration reduction systems.

This solution is only valid in the longterm but can be used to compute a transient solution with an harmonic force applied[1].

[1] Indeed, displacement and velocity at $t = 0$ are given by:

$$u(0) = \Re(\hat{H}(\omega))\frac{F_o}{M} \qquad , \qquad \dot{u}(0) = \Re(j\omega\hat{H}(\omega))\frac{F_o}{M} = -\omega\Im(\hat{H}(\omega))\frac{F_o}{M}$$

Hence the transient solution with applied force $f(t) = F_o\cos(\omega t)$ starting at time $t = 0$ and vanishing initial conditions is given by:

$$u(t) = \frac{F_o}{M}\left((\omega\Im(\hat{H}(\omega)) - 2\zeta\omega_o\Re(\hat{H}(\omega)))H(t) - \dot{H}(t)\Re(H(\omega)) + \Re(\hat{H}e^{j\omega t})\right)$$

Chapter 10

Forced vibration of a SDOF oscillator

The response of a Single Degree of Freedom oscillator to any input force $f(t)$ and for given initial conditions u_o and v_o can be derived following two approaches, both taking advantage of the linearity of the response. Either the input signal $f(t)$ is decomposed into a set harmonic signals with varying frequencies or $f(t)$ is decomposed into a set of impulses distributed over time. The latter approach, also known as the convolution approach is first studied in this chapter together with some practical application in earthquake engineering. Then it is used to derive the harmonic approach which bring addition insight on the filtering induced by a SDOF oscillator. Application to vibration reduction techniques is briefly discussed.

10.1 Convolution approach

From Newton's Law, we know that applying a force $F(t')$ at time t' over a short period of time Δt on a particle of mass M will change its velocity by:

$$\Delta v(t') = \frac{f(t')\Delta t}{M}$$

Hence, the change $\Delta u(t;t')$ in the response of the oscillator for all time $t > t'$ due to the force impulse applied at t' is the same as the change induced by applying an initial velocity $v_o = \Delta v(t')$ at time t' to the oscillator with a vanishing initial displacement $u_o = 0$. Hence, based on Section 9.2, we know that:

$$\Delta u(t,t') = \Delta v(t')H(t-t')$$

where the argument of the *Impulse Response Function* H is $t-t'$ since the fictitious initial condition is applied at t' and not at $t = 0$. By linearity the final displacement at t is the superposition of all displacements induced by all the impulses that have occurred between time 0 and time t assuming vanishing initial conditions at $t = 0$.

$$u(t) = \int_0^t \Delta u(t,t') = \int_0^t H(t-t')\Delta v = \int_0^t H(t-t')\frac{f(t')}{M}dt'$$

with for $t > 0$

Underdamped case: for $0 \leq \zeta < 1$:

$$H(t) = e^{-\zeta\omega_o t}\frac{\sin(\omega_d t)}{\omega_d} \tag{10.1}$$

Overdamped case: for $\zeta > 1$:

$$H(t) = e^{-\zeta\omega_o t}\frac{\sinh(\omega_d t)}{\omega_d} \tag{10.2}$$

Critically damped case: for $\zeta = 1$:

$$H(t) = e^{-\omega_o t}\, t \tag{10.3}$$

Hence the response is the convolution between the applied load and the *Impulse Response Function* of the oscillator. However, the result of this convolution is not easy to decipher as $H(t)$ is both oscillating and decaying. This will be studied in greater details in the following sub-section. Finally when adding the contribution of non-vanishing initial conditions at $t = 0$ the SDOF general response reads:

$$u(t) = (v_o + 2\zeta\omega_o u_o)H(t) + \dot{H}(t)u_o + \int_0^t H(t-t')\frac{f(t')}{M}dt' \tag{10.4}$$

10.1.1 Application in earthquake engineering

The transient response of an SDOF oscillator plays a central in assessing the safety of buildings and civil-works under seismic loads. Indeed under restrictions defined by all modern building codes and Eurocode 8 in particular, the seismic response of common building can be modelled by the response of a Single Degree of Freedom oscillator with mass M consisting of 85% of the total mass of the building and a natural period T_o being proportional to the number of storeys[1]. Damping ratio is given by the regulation depending on the type of material (Steel frame or Reinforced concrete) and the level of damage. This damping ratio ranges from 3 to 7%. Displacement $u(t)$ represents the horizontal drift between basement and a point located at 7/10 of the total height of the building. It is used to compute forces and torques in the structural elements of the building. The applied load is the input horizontal acceleration $a(t) = -f(t)/M$ at the basement of the building. Neither the detailed evolution of the input acceleration nor of the output drift u are of interest from an engineering point of view. Therefore only maximum values for the output drift are of interest as a function of the building characteristics represented by the natural period T_o or the natural circular frequency ω_o. Hence engineers and regulators use to plot the so-called *response spectrum*, being the maximum drift as a function of T_o for a given input acceleration and a given damping ratio ζ:

$$S_d(T_o) = \max_t |u(t; T_o)|$$

Figure 10.1 details the construction of the *Response spectrum* for a given input acceleration. It can be observed that for low natural period, the output signal is very similar to the input one appart from a scaling factor equal to $T_o^2/4\pi^2$. Actually the oscillator is very stiff and the differential displacement is very small. With increasing natural period, the high frequency part of the input signal is gradually filtered out. Maximum amplitude finally reaches a peak around a natural period of 10 seconds. For longer periods the output response converges to minus the input displacement.

[1] $T_o[s] = \frac{N_{\text{Storey}}}{10}$ is a well accepted rule of thumb.

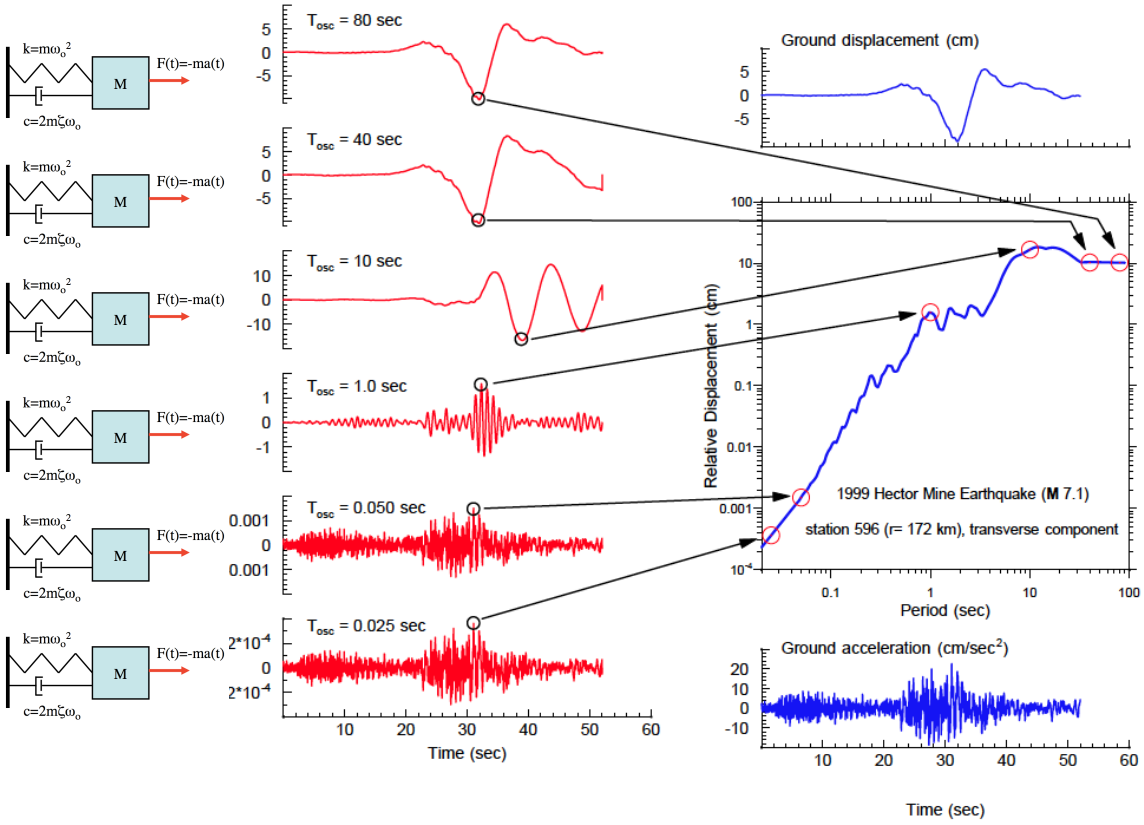


Figure 10.1: Response spectrum of acceleration recorded during Hector Mine seismic event of Magnitude M=7.1. The input acceleration (bottom-right chart) is filtered by a set of SDOF oscillators with varying natural period (Left column of charts) and maximum absolute value of the response is reported on the response spectrum as a function of the natural period T_o in log-log scale (Right chart). The opposite of the input displacement is given in the top-right chart.

Indeed, the oscillator is extremely soft and the mass is in fact decoupled from its support. It hardly moves and the differential displacement is nothing but the opposite of the displacement of the support.

Hence, these results show that the oscillator selects the frequencies in the input signal that are close to its own natural frequency, filtering the others. The property, hidden in the convolution formula will be better revealed using an expansion of the input signal over the frequency domain.

10.2 Harmonic approach

When the applied force is harmonic ($f(t) = F_o \cos(\omega t)$) with circular frequency ω, the longterm response can be obtained by pushing the initial time to $-\infty$ and setting the upper bound to $+\infty$ assuming that $H(t) = 0$ for $t < 0$. Hence:

$$u(t) = \lim_{t_o \to -\infty} \int_{t_o}^{t} H(t-t') \Re\left(\frac{F_o}{M} e^{-j\omega t'}\right) dt' = \Re\left(\int_{-\infty}^{+\infty} H(t-t') \frac{F_o}{M} e^{-j\omega t'} dt'\right)$$

and using the change of variable $\tau = t - t'$:

$$u(t) = \Re\left(\int_{-\infty}^{+\infty} H(\tau) e^{-j\omega\tau} d\tau \frac{F_o}{M} e^{j\omega t}\right) \tag{10.5}$$

Since the harmonic response is given by Equation (9.8) ($u(t) = \Re\left(\hat{H} \frac{F_o}{M} e^{j\omega t}\right)$), we finally obtain the following relation between the *Frequency Response Function* $\hat{H}(\omega)$ and *Impulse Response Function* $H(t)$:

$$\hat{H}(\omega) = \int_{-\infty}^{+\infty} H(\tau) e^{-j\omega\tau} d\tau \tag{10.6}$$

In mathematical terms, the *Frequency Response Function* $\hat{H}(\omega)$ is the Fourier Transform of the *Impulse Response Function* $H(t)$. Conversely one can show that:

$$H(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{H}(\omega) e^{j\omega t} d\omega \tag{10.7}$$

The Fourier transform can be used to decompose any input force $f(t)$ in terms of a sum of harmonic inputs:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega) e^{j\omega t} d\omega \qquad \text{with} \qquad \hat{f}(\omega) = \int_{-\infty}^{+\infty} H(\tau) e^{-j\omega\tau} d\tau$$

Since the total response is the superposition of individual responses to harmonic loads, we finally have, for vanishing initial conditions:

$$u(t) = \frac{1}{2\pi\ M} \int_{-\infty}^{+\infty} \hat{H}(\omega) \hat{f}(\omega) e^{j\omega t} d\omega \tag{10.8}$$

As compared to Formula (10.4), Equation (10.8) better illustrates how the oscillator filters the input load. Indeed, the frequency content of the input signal at ω is given by $|\hat{f}(\omega)|$ whereas the frequency content of the output is given by $|\hat{H}(\omega)\hat{f}(\omega)/M|$. Hence, $|\hat{H}|$, the amplitude of the *Frequency Response Function* plotted on figure 9.3 fully characterises the amplification/de-amplification induced by the oscillator. This feature is a the heart of vibration reduction techniques.

10.3 Vibration reduction

Vibration reduction techniques aim at either reducing vibration transmitted to the environment or preventing ambiant vibrations to reach a particular sensitive piece of equipment. It appears that both cases reduce to the so-called *transmissivity function*.

As we have noticed that any input signal can be decomposed into a set harmonic signals, the analysis is conducted here in terms of harmonic input only. Considering an input force $f(t) = \Re(F_o \cos(\omega t))$, the force applied on the support $f_s(t)$ is the sum of the force applied by the spring and the one induced by the damper:

$$f_s(t) = k\ u(t) + c\ u(t) = M\ \Re\left((\omega_o^2 + 2j\zeta\omega\omega_o)\hat{H}F_o e^{j\omega t}\right)$$

Hence *Transmissivity* is defined as the modulus of the ratio between the complex amplitude of the output force and the input force:

$$\hat{T}(\omega) \quad = \quad \left|\frac{K\hat{u} + i\omega C\hat{u}}{F_o}\right| = \left|\frac{1 + 2i\Omega\zeta}{1 + 2i\Omega\zeta - \Omega^2}\right| \qquad \text{with} \qquad \Omega = \frac{\omega}{\omega_o} \tag{10.9}$$

Let us now consider the case of an oscillator subjected to ambiant vibrations, that is to say, subjected to an input harmonic acceleration $a_o(t) = A_o \cos(\omega t)$ of its support. The total acceleration of the mass is:

$$a_T(t) = a_o(t) + \ddot{u} = \Re\left(\left(A_o - \omega^2\hat{H}\frac{-MA_o}{M}\right)e^{j\omega t}\right)$$

Hence the ratio between the complex amplitude of the output and input acceleration reads:

$$\left|1 + \omega^2\hat{H}\right| = \left|\frac{1 + 2i\Omega\zeta}{1 + 2i\Omega\zeta - \Omega^2}\right| = \hat{T}(\omega)$$

Transmissivity as a function of the input frequency is plotted on Figure 10.2 for different values of the damping coefficient. It shows that all frequency above $\sqrt{2}\omega_o$ are filtered out. Surprisingly the smaller the damping the better the amplitude reduction in this range. However the price to pay is that the lower the damping the larger the amplification around the resonance frequency ω_o. Hence

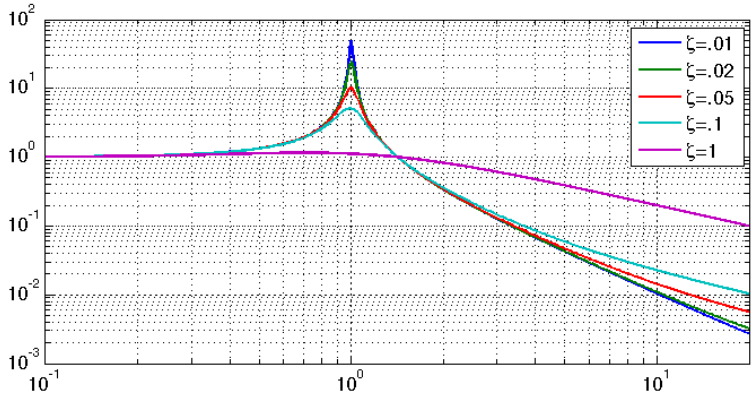


Figure 10.2: Transmissivity of a SDOF system for $\zeta = .001, 0.01, 0.1, 1, 10$

in terms of design, the key parameter is the natural frequency of the oscillator that has to be set

as small as possible to filter all frequencies above $\sqrt{2}\omega_o$. Damping has then to be adjusted to find a trade-off between filtering of high frequencies and amplification around the natural frequency of the oscillator. Therefore designing a vibration reduction system simply consists in designing an oscillator with a natural frequency that is below the frequency content of the input signal. This is achieved either reducing the stiffness of the oscillator or increasing the total mass. Air-cushion is a common technique to achieve low stiffness, yet with some limitations. To go beyond, some "negative stiffness" has to be added in the system to further reduce the natural frequency. It then consist in designing systems that are close to instability, such as an inverted pendulum.

Part II

Fundamentals in thermodynamics

Chapter 11

Kinetics of ideal gases

The renown physicist *Richard Feyman* once said that engineers' contribution to physics is restricted to one single but major discovery known by the exotic name of *Entropy* - a term meaning "the power of transformation". Feynman attributed this contribution to two different engineers being distant both in time, space and field of research: a French mechanical Engineer of the early XIXth century named Saadi Carnot, graduate from Ecole Polytechnique, and an American telecommunication engineer of the mid XXth century named *Claude Shannon*, from MIT and Bell labs.

It is rather difficult to figure out which practical, and at the same time fundamental concept, may relate the efficiency of an engine and the coding of a message. Actually an easy connection can only be fully established in one particular case: the physics of ideal gases, the topic of this chapter. However before being able to show this connection, basic concepts of, pressure, temperature and heat have to introduced.

Hence, the aim of this first chapter on thermodynamics is to show how macroscopic quantities such as, *pressure*, *temperature* are linked with microscopic quantities such as the velocity distribution of individual molecules in the gas, and how the originally empirical *law of ideal gases* can be derived from the microscopic behaviour. To this aim let us start with a brief account of the discovery of this law.

11.1 Historical development and hydrostatic pressure

At the beginning of the XIXth century, Newton's Laws were well understood together with the notions of *Work*, *Energy* though restricted to *Mechanical Energy*, either kinetic and potential. *Pressure* was understood as a force $[N]$ per unit-area $[m^2]$ and measured in Pascal [Pa] as developed in Section 11.2. Indeed Blaise Pascal studied pressure in liquids and showed through striking experiments that the pressure excess Δp from atmospheric pressure is equal to the height h of liquid above the measured level multiplied by the weight per unit volume of the liquid ρg, where g the acceleration of gravity and ρ the mass density or mass per unit volume:

$$\Delta p = \rho g h$$

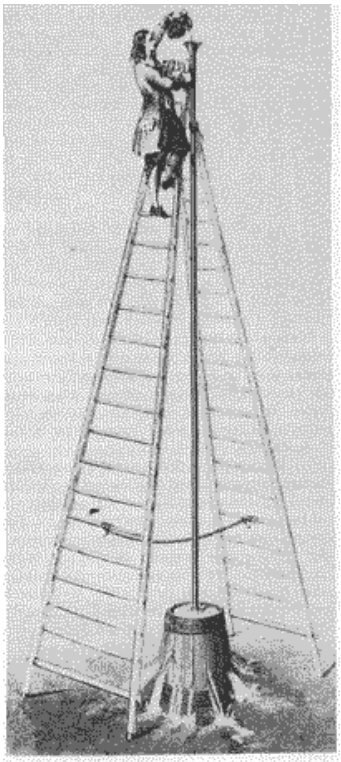


Fig. 2. Pascal's cask-bursting experiment, from Guillemin.

Though not surprising when the free-surface is above the measured point, this result is less obvious when the free-surface is far away or is of a very limited extend as in the figure on the left. This allowed the development of apparatuses to measure pressures using an heavy liquid at ambiant temperature: mercury. Hence pressure was measured in terms of height of mercury. In particular Torricelli had invented his barometer back in the XVIIth century showing that the air pressure fluctuates around 76 cm of mercury.
In 1662, *Robert Boyle* and *Edme Mariotte* independently discovered that at constant temperature the pressure of gases in a closed container is inversely proportional to the volume of the container which is known as the *Boyle-Mariotte law*:

$$PV = \text{cst}$$

It was only around 1787, when steam engines started to be developed that the influence of temperature was demonstrated by *Jacques Charles* who showed that at constant pressure P, volume V of a given amount of gas is proportional to temperature T. This experimental fact latter allowed *Joseph Louis Gay-Lussac* to devise the existence of an *absolute zero temperature*, simply extrapolating Charles's Law. *Avogadro* then showed that at constant pressure and temperature, the volume V is proportional to the *amount of substance* n independently from its mass. These three laws were put together under the name of *law of ideal gas*:

$$PV = nRT \tag{11.1}$$

where n is measured in [mol], $R = 8.314$ [m^3 Pa mol^{-1}K^{-1}] is the *universal gas constant* and T is the *absolute temperature* attributed to *Lord Kelvin*. This law being universal for ideal gases it gave a very convenient way to measure and calibrate *temperature* on a wide range of values.

The striking element here is not to combine three empirical laws in a single one, it is the fact that for different gases the same constant R is valid calling for a more fundamental explanation. *Newton* had tried to explain *Boyle's law* using repulsion forces between particles of gas or liquid but it was not satisfactory. However the idea of referring to a collective action of particles was correct. It was stated again by *Bernoulli* who failed to convince the scientific community. Finally the *kinetic theory of gases* was introduced by *Clausius* and then developed by *Maxwell* and *Boltzmann* in late XIXth century merely based on the atomic conjecture though without any experimental proof of it. The proof was given by *Einstein* in its analysis of the Brownian Motion in 1906 only.

In the meantime the analogy between *temperature* and *pressure* and the mere fact that *pressure* is used to mesure *temperature* led scientists on a wrong path for over a century. Indeed many thought that *heat* would flow as a fluid from a high temperature reservoir to lower temperature one in the same way a fluid flows from a high pressure reservoir to a lower pressure one - or similarly from a high water level reservoir to a lower water level one. Doing so they missed the important principle of energy conservation we will use in next chapter to define *heat* and the concept of *efficiency* or *entropy* introduced by Carnot.

Actually, taking a microscopic approach of a gas or a liquid as a collection of particles has finally appeared a much more fruitful approach to understand the macroscopic quantities that are *pressure*, *temperature* or *internal energy* and finally *heat*.

11.2 Hydrostatics

Liquids and gases at rest - at a macroscopic level - are fully characterised from the mechanical point of view by only two quantities, pressure P and mass density ρ. Indeed, fluids have no ability to sustain shear on the long term and then flow until stopped by the wall of the container. This is true as long as surface tension phenomena including wetting can be neglected. The consequence is that the elementary force $d\overrightarrow{F}$ applied by the fluid on any elementary surface dA with normal unit vector $\overrightarrow{n}$ - conventionally pointing outwards to the fluid - is along this outward normal direction and proportional to the contact area:

$$d\overrightarrow{F}_{\text{fluid}\to\text{wall}} = P\overrightarrow{n}\,dA$$

where P is the pressure field at this location. Hence the amplitude of the force does not depend on the orientation of the normal vector in space. Pressure P is positive when the force pushes the wall outwards, tending to expend the volume occupied by the fluid, negative otherwise. Assuming the wall is at rest, Newton's *actio-reactio law* requires the elementary surface to exert an opposite force $d\overrightarrow{F}_{\text{ext}}$ onto the fluid:

$$d\overrightarrow{F}_{\text{ext}} = -P\overrightarrow{n}\,dA \tag{11.2}$$

Pressure can vary smoothly inside the fluid - mathematical speaking the pressure field is continuous and differentiable. Considering the balance of linear momentum of a small cube of fluid of size $dV = dx \times dy \times dz$ and summing the elementary forces applied on the six edges of respective areas $dx \times dy$, $dx \times dz$ and $dy \times dz$ and respective normal vectors $\pm\overrightarrow{e_z}$, $\pm\overrightarrow{e_y}$ and $\pm\overrightarrow{e_z}$ yields:

$$\begin{aligned} -\left(\left(p(x+dx,y,z)-p(x,y,z)\right)\,\overrightarrow{e_x}\;dy\;dz\right) &+ \\ \left(p(x,y+dy,z)-p(x,y,z)\right)\,\overrightarrow{e_y}\;dx\;dz &+ \\ \left(p(x,y,z+dz)-p(x,y,z)\right)\,\overrightarrow{e_z}\;dx\;dy\right) &+ \overrightarrow{f_v}dV = 0 \end{aligned} \tag{11.3}$$

where $\overrightarrow{f_v}dV$ is the sum of all external bulk forces applying on volume dV excluding surface forces accounted for by the pressure terms. For instance, considering the weight of the cube of fluid gives:

$$\overrightarrow{f_v}dV = \overrightarrow{g}\,dm$$

where dm is the elementary mass of fluid contained in dV. Defining the mass density as:

$$\rho = \lim_{dV\to 0}\frac{dm}{dV} \tag{11.4}$$

finally yields:

$$\overrightarrow{f_v} = \rho\overrightarrow{g}$$

Dividing the conservation of linear momentum (11.3) by dV and taking the limit for both dx, dy and dz approaching 0 gives:

$$-\left(\frac{\partial p}{\partial x}\overrightarrow{e_x}+\frac{\partial p}{\partial y}\overrightarrow{e_y}+\frac{\partial p}{\partial z}\overrightarrow{e_z}\right)+\overrightarrow{f_v}=0$$

or using the *gradient*:

$$-\overrightarrow{\text{grad}}\;P+\overrightarrow{f_v}=0 \tag{11.5}$$

In the case where only gravity applies with $\overrightarrow{g} = -g\overrightarrow{e_z}$ one can easily deduce that:

$$\frac{\partial p}{\partial x}=0, \qquad \text{and} \qquad \frac{\partial p}{\partial y}=0$$

showing that the pressure field only depends on z with:

$$\frac{\partial p}{\partial z} = -\rho g \tag{11.6}$$

where gravity is toward the negative $z-$axis.

11.2.1 Incompressible liquids under gravity

As opposed to gazes, liquids can hardy be compressed when pressure is increased. Therefore, in a first approximation, the density can be assumed not to depend on the applied pressure ($\rho = \rho_o$). All other parameters being fixed, Equation (11.6) can easily be integrated:

$$p(z) = p(z_{\text{ref}}) + \rho_o g(z_{\text{ref}} - z)$$

where g is also assumed to be constant over z and z_{ref} is some reference level at which pressure is assumed to be known. This reference level is usually set at the free surface z_{surf} where the pressure is equal to the atmospheric pressure $p_{\text{atm}} \approx 10^5$ [Pa]. Note that with $g \approx 10$ [ms^{-2}] and $\rho_o \approx 1000$ [kg/m^3] for water, the pressure increases by one atmosphere (10^5 [Pa]) every 10 meters in depth.

11.2.2 Compressible gases under gravity

As opposed to fluids, gases are highly compressible. In the case of an ideal gas, dividing the empirical Law (11.1) by the total mass M and noticing that $\rho = M/V$ yields:

$$\frac{P}{\rho} = \frac{nRT}{M} = \frac{RT}{m_l} = c_o^2$$

where m_l is the molar mass of the gas and c_o has the physical dimension of a velocity. Hence for a constant temperature T:

$$\frac{dP}{dz} = \frac{dP}{d\rho}\frac{d\rho}{dz} = c_o^2 \frac{d\rho}{dz}$$

Hence the hydrostatic law (11.6) becomes:

$$\frac{d\rho}{dz} = -\frac{\rho}{l}, \qquad l = \frac{c_o^2}{g}$$

the solution of which gives an exponential decay of the mass density with altitude:

$$\rho(z) = \rho_o e^{-\frac{z}{l}} = \rho_o e^{-\frac{m_l g z}{RT}}$$

leading to the same exponential decay for the pressure:

$$P(z) = P_o e^{-\frac{z}{l}} = P_o e^{-\frac{m_l g z}{RT}}$$

11.3 Kinetic theory of ideal gases

The kinetic theory of ideal gases simply consists in applying Newton's law at a microscopic level and compute average quantities. It assumes that a gas is made of a large collection of N particles with non-relativistic velocity $\overrightarrow{v}$ and constant mass m. Interaction forces between particles are neglected and collisions between two or more particles are assumed to preserve *momentum*. Moreover, when

a given particle with velocity $-v$ hits the surface of area ΔA with a normal incidence - we assume that the normal vector of the surface is in the x-direction - it bounces back in the opposite direction with velocity v, corresponding to an *elastic* collision. Hence the change in linear momentum of the particle during this collision is:

$$\Delta \overrightarrow{p} = -2mv_x \overrightarrow{e_x}$$

During the time of the impact Δt, this change in linear momentum must result in a force applied on the wall:

$$\Delta \overrightarrow{F} = \frac{2mv_x}{\Delta t} \overrightarrow{e_x}$$

so that the second Newton's law is satisfied. Unfortunately the collision time is so small that this force is almost impossible to measure. Let us then try to take the sum of all these forces over area ΔA and for a larger period of time Δt and denoted it by $\langle \overrightarrow{F} \rangle$. Hence:

$$\langle \overrightarrow{F} \rangle = \sum_N \frac{2mv_x}{\Delta t} \overrightarrow{e_x}$$

where N is the number of particles hitting ΔA during the time interval Δt. Actually the same result holds for particles with non-normal incidence provided that only the x-component of the velocity is kept ($v_x > 0$) since the other components remain unchanged after hitting the wall in an elastic collision and assuming that now tangential force is applied on the wall. For a given normal velocity v_x only particles in volume $= \Delta A \cdot v_x \Delta t$ reach the surface. Denoting $\rho'(v_x)dv_x$ the mass density of particles with velocity in $[v_x, v_x + dv_x]$ in the $\overrightarrow{e_x}$ direction, the mass of such particles reaching area ΔA is:

$$mN(v_x)dv_x = \rho(v_x)\Delta A \cdot v_x \Delta t \; dv_x$$

Hence the force is obtained summing over all velocities from 0 to $+\infty$:

$$\langle \overrightarrow{F} \rangle = \int_0^{+\infty} \rho'(v_x) \frac{2v_x}{\Delta t} \overrightarrow{e_x} dv_x = \int_0^{+\infty} \rho'(v_x) 2v_x^2 dv_x \Delta A \overrightarrow{e_x}$$

The expected symmetry of the velocity distribution yields:

$$P = \frac{\langle \overrightarrow{F} \rangle \cdot \overrightarrow{e_x}}{\Delta A} = \int_{-\infty}^{+\infty} \rho'(v_x) v_x^2 dv_x = \rho \langle v_x^2 \rangle$$

where $\langle v_x^2 \rangle$ is the mean square of the velocity in the x direction and where ρ is the total mass density accounting for particles with all possible velocities:

$$\rho = \int_{-\infty}^{+\infty} \rho'(v_x) dv_x$$

Let us now consider the mean translational kinetic energy of all particles in a given volume V with N particles:

$$\langle \mathcal{E}_{\text{kin}} \rangle = \sum_N \frac{1}{2} m(v_x^2 + v_y^2 + v_z^2)$$

Since the velocity distribution is isotropic, the three components must have the same contribution leading to:

$$\langle \mathcal{E}_{\text{kin}} \rangle = \frac{3}{2} \sum_N m v_x^2 = \frac{3}{2} V \cdot \int_{-\infty}^{+\infty} \rho'(v_x) m v_x^2 dv_x = \frac{3}{2} PV$$

or :

$$PV = \frac{2}{3}\langle \mathcal{E}_{\text{kin}} \rangle$$

Hence we have shown that the constant factor in the *Boyle-Mariotte* law is proportional to the mean kinetic energy of the gas particles in the container. Defining now the *absolute temperature* T as the mean translational kinetic energy for each translational degree of freedom of one particle in the gas scaled by a factor $2/k_B$ with k_B the *Boltzmann constant*:

$$T = \frac{2}{k_B}\frac{\langle \mathcal{E}_{\text{kin}} \rangle}{3N}$$

one finally obtains the Law of ideal gases:

$$PV = Nk_BT$$

or:

$$PV = nRT$$

where n is the number of mols, $R = N_A k_B$ is the *ideal gas constant* for mono-atomic gases and $N_A = 6.022140857(74) \diamond 10^{23}$ [mol^{-1}] is the Avogadro's number defined as the number of particles per mol. Note that for gases made of diatomic molecules other degrees of freedom than the translational ones contribute to the kinetic energy. Indeed any particle made of two atoms has at least two additional degrees of freedom consisting of the two rotations with respect to the two axis perpendicular to the symmetry axis of the molecule. Vibration degrees of freedom may also be activated[1]. It is then assumed that each of these degrees of freedom, when activated, carries the same mean energy as the translational degree of freedom v_x -and v_y and v_z. Hence, a more general definition of *temperature* of an ideal gas in terms of internal kinetic energy reads:

$$T = \frac{(\gamma - 1)}{k_B}\frac{\langle E \rangle}{N}$$

where, at macroscopic level, the dimensionless parameter γ will be shown to be the ratio between heat capacities at constant pressure and constant volume respectively. At a microscopic level it can be expressed as a function of the number of activated degrees of freedom n_{dof} as:

$$\gamma = \frac{2}{n_{\text{dof}}} + 1$$

Note that many implicit or explicit hypotheses have been made to obtain this seemingly very simple and universal result. In applying the Newton's law, all interaction forces between particles hitting the wall of the container have been ignored. This is likely to be correct only for low density of particles and/or negligible cross-section of these particles. Elastic collisions, between molecules and with the wall of the container, have also been assumed and no potential energy between molecules has been accounted for.

Hence the energy "stored" in the gas is solely made of the kinetic energy. As a consequence, the *internal energy* of the gas denoted by U is nothing but the *average kinetic energy*[2]:

$$U = \langle \mathcal{E}_{\text{kin}} \rangle = \frac{n_{\text{dof}}}{2} nRT = \frac{nRT}{\gamma - 1} \tag{11.7}$$

[1] Actually, the fact that for some molecules these degrees of freedom are activated or not had been a pending shortfall of classical mechanics and was only overcome by quantum physics.

[2] Since the gas involves a very large number of particles, fluctuations on the *internal energy* of the gas are expected to be extremely small, thanks to the *Law of large numbers*.

It could be argued that the *kinetic approach* originally developed by *Kelvin* has allowed to retrieve the already known expression of the *internal energy* of an ideal gas, and in particular its linear variation with respect to *temperature* T on the first hand and *amont of substance* n on the second hand. However it could, and should be looked at the other way round, that is to say as a *mere definition* of *the absolute temperature* as anticipated by *Gay-Lussac*.

As a conclusion, this section has unveiled the link between macroscopic quantities such as pressure P, temperature T, internal energy U and microscopic ones: essentially the velocity distribution of individual particles, leading to two main results in thermodynamics of ideal gases: the *ideal gas law* and the *internal energy*. Actually it even provides a way to defined, and to measure *temperature*[3]. Chapter 12 will explore the consequences of these results in terms of standard thermodynamics and will show that some elements are still missing to close the theory. These missing elements will be studied in chapter 13 going a little bit further in terms of statistical analysis at a microscopic level. Indeed as of now the detailed distribution of the velocity was not required and only the average of its squared amplitude has been used to compute both *pressure* and *internal energy*. We will introduce one additional characteristic of this distribution related to the amount of uncertainty or randomness. It will give another view on one essential hypothesis used in these developments, the so-called *equipartition of energy* introduced for diatomic gases[4].

[3]Indeed the ideal gas thermometer had been commonly used in the past and is still used as a standard for a wide range of temperature.

[4]For mono-atomic gases the *isotropy hypothesis* actually implies *equipartition of energy* with respect to the three translational degrees of freedom.

Chapter 12

Thermodynamics I: Temperature and Heat

Thermodynamics consists of four laws numbered from 0 to 3 and, as opposed to Newton's Laws which were introduced as a single masterpiece, it took more than a century for the laws of thermodynamics to be properly stated. It begun at the beginning of the XIXth century with some very different and seemingly unrelated forms of the intriguing *second law* studied in the next chapter. Then, in the course of the XIXth century, the *first law* was proposed as extension of the already-known *conservation of energy*. Then the so-called *zeroth law* had been added to provide a well-posed framework defining *thermodynamic equilibrium*, *state variables* and in particular *temperature*. Finally the *Third law* was added during the XXth century as a last refinement of the *second law*. This *third law* will not be studied in this course.

12.1 Zeroth Law of thermodynamics

From the perspective of the *kinetic theory of ideal gases*, studied in Chapter 11, *Temperature* is related to the *mean kinetic energy* of particles in an ideal gas. However, it is limited to ideal gases, though *Temperature* is measured in many other cases, including real gases, liquids, solids or even with gases made of massless particles: *photons*. Fortunately, ideal-gas *thermometers* can be built and used to measure the temperature of any object - generically referred to as *thermodynamic system*. When the *ideal-gas thermometer* can give a reliable measurement after some time, the system is said to be in *thermal equilibrium* with the *ideal-gas thermometer* and the *temperature* of the system is defined as the *temperature* of the *ideal-gas thermometer*. When a reliable measurement cannot be performed, the system is considered not being at *thermal equilibrium* and its temperature cannot be defined. Laws of thermodynamics simply do not apply to it. Hence, the *Zeroth Law* of thermodynamics set the conditions under which other laws can be used and, in particular, *temperature* can be defined[1]:

Definition 14 ***Zeroth law of thermodynamics:*** *If two systems are in thermal equilibrium with a third system, they are in thermal equilibrium with each other.*

[1] It may look surprising that *temperature* may not be defined in some cases. However since we have seen that even for ideal gases *temperature* is a statistical quantity, to have any significance some stability in the averaging process is rerquired. For instance temperature cannot be defined when the density of particles is too low. Hence, it is a common mistake to claim that there is a breach in the Laws of thermodynamics when these law simply do not apply to the system under study. Actually, this mistake has turned crazy many generations of physicists.

12.1.1 Thermodynamic system

Thermodynamic systems play a central role in thermodynamics though its definition is quit loose. A *thermodynamic systems* is supposed to occupy a certain region of space, which can evolve over time, but which contains a "large" amount[2] of matter or more accurately a large number of *micro-states of energy*. Both *matter* and *energy* can be exchanged with the outside world throughout *boundaries* or *walls*.

Definition 15 (Closed systems) *are* thermodynamic system *that are not exchanging any matter/mass with the outside. Exchanges only consist of energy throughout* work *or* heat.

When a *thermodynamic system* is assumed be at *thermal equilibrium* its *state* is assumed to be fully characterised by *state-variables* and related *state functions* that are global quantities controlled from outside of the system. These global quantities are average values on a large number of microscopic quantities. Hence:

Property 5 (The *state* of the *thermodynamic system*) *is fully characterised by the values of the* state-variables *and does not depend on the path followed to reach this state.*

12.1.2 Thermodynamic states and state-variables

Pressure P, *Temperature* T, *Volume* V, quantity of substance n are such *state-variables* for gases though not the only ones. *Internal Energy*, *Entropy*, *Enthalpy* will also be defined as *state-variables*.

Intensive and extensive state-variables State variables are of two types: *intensive* or *extensive*. *Intensive* variables are not changed when the system is replicated or divided. On the contrary, *Extensive* variables are rescaled by a factor of 2 when the system is replicated. *Temperature* T and *Pressure* P are obviously *intensive* state-variables. *Volume* V, *quantity of substance* n, *Energy* are *extensive* state-variables.

State-variables are not independent. Indeed, for instance, the *law of ideal gases*:

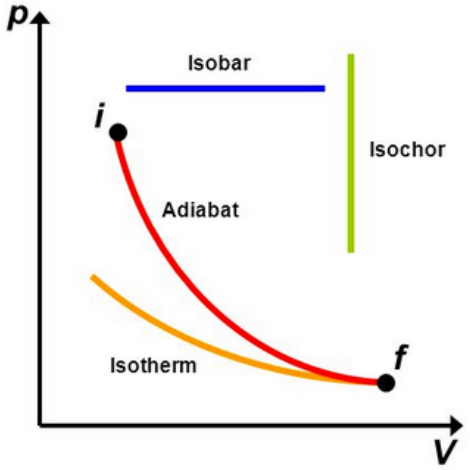
$$PV = nRT$$

allows to express any of these four *state-variables* as a function of the three others. Hence the state of an ideal gas - or of a real one - is fully characterised by only three *state-variables*, the remaining one being expressed as functions of the three others.

State-function: when one *state-variable* is expressed as a function of other *state-variables*, it is referred to as a *state-function* of these *state-variables*.

[2] "Large" can be made more explicit when considering that a thermodynamic system is, at a microscopic level, made of a large number of interacting particles or energy levels. This large number allows statistical averaging to occur. Let us recall that one mol of gas contains $N_A \approx 610^{23}$ molecules.

Diagrams are two-dimension graphs used to represent the evolution of the state of a system during a given process. Classical processes are defined in Section 12.3. For gases the usual diagrams are: (P,V)-diagrams, (P,T)-diagrams or (V,T)-diagrams. On the (P,V)-diagrams, a constant-temperature process known as *isothermal* is an hyperbola ($PV = \text{cst}$) for an ideal-gas. A process with constant-pressure known as an *isobaric process* is an horizontal line and a process with constant volume *isochoric* is a vertical line.



Cycles: are closed paths in the *state space* defined by any chosen state-variables, the *state* of the system is kept unchanged after a given cycle.

Statistical perspective As already shown for *pressure* and *temperature* or *internal energy* of an ideal-gas, *State-variables* are statistical quantities - average values - over a large number of microscopic states. Chapter 13 will come back on *thermal equilibrium* viewed as *equipartition of energy* at microscopic level.

12.2 First law of thermodynamics: energy conservation

Nowadays conservation of energy is something well accepted even beyond the scientific community. It has even been related to the invariance of laws of physics with respect to time[3]. Despite of this well accepted law, it is often found in the news or on the internet that "energy is created". Even scientists or engineers would talk about "renewable energies" as opposed to other "sources of energy" that could not "be renewed". Actually the problem scientists were facing in the XVIIIth and XIXth century is deeply related to this common-sense idea that energy can be created and destroyed, when at the same time it has to remain constant. The reason for it is that common-sense only accounts for the "useful" part of energy, sometimes called *exergy*, that basically consists in *mechanical work*, *electrical energy*, *chemical energy* or *macroscopic kinetic energy*. The remaining part of *energy* was mainly ignored by XVIIIth-century scientists as useless or was given another name: *heat*. In addition, for a long time *heat* had been considered as a particular fluid and not as a form of energy in itself. However, when it appeared that some *work* can be done from *heat* with the rise of the steam engin at the end of the XVIIIthe century, the picture changed and the first law of thermodynamics could be stated.

Definition 16 (First law of thermodynamics) *states that* energy *is conserved and can be exchanged through different means:*

- Work*: either electrical, mechanical or chemical;*
- Matter*:* internal energy *of matter exchanged between the system and the outside world,*

[3]Emmy Noether's theorem in 1915

- heat.

Actually this *First Law* could also be viewed as a mere definition of *heat*.

Definition 17 (Heat) *Considering a* closed thermodynamic system *with* internal energy *evolving from U_o to $U_o + \Delta U$ with work W performed on the system by the outside world, the* heat Q *given to the system by the outside word in the process is:*

$$Q = \Delta U - W \tag{12.1}$$

Since *internal energy* is a *state-function*, the change ΔU does not depend on the path but only on the initial and final states. In particular any small change in the *internal energy* dU corresponds to a change in other state variables. Hence dU is called an exact differential satisfying for all states 1 and 2:

$$\int_{\text{state 1}}^{\text{state 2}} dU = U_2 - U_1$$

As opposed to *internal energy*, *work* performed on the system and *heat* transferred to it are controlled by outside parameters such as external pressure or external temperature. As a consequence they are *path dependent* and are called *path functions*. Any small increment of work δW and increment of *heat* δQ are referred to as *inexact differentials* and thus have to be integrated along a specific path[4]. Next sections will give some practical examples of this difficult and generic concepts of *work* and *heat*.

12.2.1 Work

As studied in Chapter 4, the increment of *Work* δW done by a *force* $\overrightarrow{F}_{\text{ext}}$ on a given point B is the scalar product of this force with the displacement experienced by this point $\overrightarrow{dl}$:

$$\delta W = \overrightarrow{F} \cdot \overrightarrow{dl}$$

When force and displacement vary over the path followed by point B the *network* is the sum over this given path:

$$W = \int_{\text{path}} \overrightarrow{F} \cdot \overrightarrow{dr}$$

Considering now a standard piston of area A and the force $\overrightarrow{F}$ induced by a constant external pressure P_{ext} on the piston:

$$\overrightarrow{F} = \int_A -P_{\text{ext}} \overrightarrow{n}_{\text{ext}} dA = -A \; P_{\text{ext}} \overrightarrow{n}_{\text{ext}}$$

where $\overrightarrow{n}_{\text{ext}}$ is the outer normal vector of the piston. Assuming that the piston moves by $-\delta x$ along the same normal vector $\overrightarrow{n}_{\text{ext}}$, then the work performed onto the piston is:

$$\delta W = P_{\text{ext}} \cdot A \cdot \delta x$$

Assuming now that the system is a gas-enclosure with fixed walls but the piston (see Figure 12.1). The change in volume of this enclosure is then:

$$dV = -A \cdot \delta x$$

[4]To better grasp this difficult mathematical concept of *exact* and *inexact* differentials, readers could reconsider the definition of conservative and non conservatives forces in mechanics. Conservative forces are exact differentials of a potential energy whereas non-conservative forces are inexact differentials.

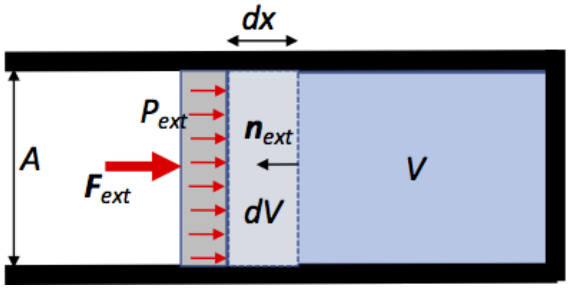


Figure 12.1: Work done on a piston.

finally leading to:

$$\delta W = -P_{\text{ext}} dV \tag{12.2}$$

where it is assumed that the work transferred to the piston is transferred to the gas in the enclosure. It is worth noticing that this formula can be generalised to any system with deformable boundaries but subjected to a constant pressure field P_{ext}, each part of the boundary playing the role of a small piston.

There are several strategies to compute the work W on a given path that are detailed below.

Constant external pressure is simplest case. Indeed then:

$$W = \int_{V_1}^{V_2} -P_{\text{ext}} dV = -P_{\text{ext}} \int_{V_1}^{V_2} dV = -P_{\text{ext}} \Delta V$$

Adiabatic processes also referred to as *isentropic* are defined as processes without any *heat exchange* with the outside. Hence the increment of work δW is fully balanced by the change in *internal energy* dU

$$dU = \overbrace{\delta Q}^{=0} + \delta W$$

and in this very particular case *work* becomes a full differential leading to:

$$W = U_2 - U_1$$

The particular case of ideal gases then gives:

$$W = C_v (T_2 - T_1)$$

where $C_v = \frac{3}{2} nR$ for monoatomic ideal-gases. Using the *ideal-gas* law the work can be expressed in terms of initial and final volume and pressure:

$$W = \frac{C_v}{nR} (P_2 V_2 - P_1 V_1)$$

Note that W can only be effectively computed when the initial and final states are known. Note also that *adiabatic process* is a special case of a *reversible process* for which the work can also be computed.

Reversible process is such that the external pressure P_{ext} slowly varies and is balanced by the internal pressure P. Thanks to some internal state laws for the system, the internal pressure can be expressed as a function of V leading to:

$$W = -\int_{V_1}^{V_2} P(V)dV$$

Two particular cases for ideal-gases are worth mentioning : *isothermal* processes and once again *adiabatic* ones.

Isothermal processes are defined as processes at constant temperature T. Hence, for ideal gases, Boyle-Mariotte Law $PV = P_1V_1$ directly gives the expected function $P(V)$ leading to:

$$W = -P_1V_1 \int_{V_1}^{V_2} \frac{dV}{V} = -P_1V_1 \ln\left(\frac{V_2}{V_1}\right) = -nRT \ln\left(\frac{V_2}{V_1}\right)$$

Adiabatic processes for ideal gases are characterised by specific paths in the $(P-V)$ diagram. Indeed, the special form of the *internal energy* leads to:

$$PdV = -\delta W = -dU = -C_v dT$$

Differentiating now the ideal-gas law for a closed system ($dn = 0$) leads to:

$$PdV + VdP = nRdT$$

Eliminating dT between these two equations gives:

$$(C_v + nR)PdV + C_vVdP = 0$$

or

$$\gamma\frac{dV}{V} + \frac{dP}{P} = 0 \qquad \text{with} \qquad \gamma = \frac{C_p}{C_v} \qquad \text{and} \qquad C_p = C_v + nR$$

and finally once integrated:

$$PV^\gamma = \text{cst} = P_1V_1^\gamma.$$

Hence P is defined as a function of V leading to:

$$W = -P_1V_1^\gamma \int_{V_1}^{V_2} \frac{dV}{V^\gamma} = -P_1V_1^\gamma \frac{1}{1-\gamma}\left(V_2^{1-\gamma} - V_1^{1-\gamma}\right) = \frac{1}{\gamma - 1}\left(P_2V_2 - P_1V_1\right)$$

which we have already obtained. However now the path to the final state is known.

12.2.2 Heat

For a long time, *Heat* had been a mysterious physical quantity even though common sense gives a good idea of it. Anyone can feel the *heat* from a fire, the Sun or any object being at a higher temperature than our own one. Hence *heat* was figured out as a special fluid flowing from hot systems to cooler ones. The most common analogy was an hydraulic system with an upper reservoir of "heat" at high temperature T_1 and a lower reservoir at lower temperature $T_2 < T_1$.

The picture offered by the first law of thermodynamics is somehow simpler: *heat is the part of energy to be given to or removed from the system to reach energy balance once work has been accounted for.* Though simpler, it remains a little more abstract concepts as it is defined as what

is missing in the balance of energy. In addition it does not give any clue as to how *Heat can never "flow" from a cold body to a warmer one, without some other change, connected therewith, occurring at the same time.* Actually this last statement is the way Clausius, in 1850, stated the Second Law further studied in next chapter.

One first practical question that can be put about heat is how much heat can be released or gained by the thermodynamic system when its temperature changes. This internal property of a thermodynamic system is called the *Heat capacity.*

Heat Capacity

As already mentioned, *Heat* exchanged by a thermodynamic system is path dependent as it is for *Work*. However, when the volume of a gas does not change, no work is done on the system ($\delta W = 0$) and thanks to the first law:

$$\delta Q + \overbrace{\delta W}^{=0} = dU$$

Hence:

$$C_v = \lim_{\Delta T \to 0} \left(\frac{\Delta Q}{\Delta T} \right)_{V,n} = \left(\frac{\partial U}{\partial T} \right)_{V,n} \tag{12.3}$$

Since the volume does not change, C_v is referred to as the *heat capacity at constant volume.* In general C_v is a function of Temperature and other state variables. Hence at different temperatures the heat released for a certain drop of temperature is not constant. However for an ideal gas it is independent of T:

$$C_v = \frac{n_{\text{dof}}}{2} nR \tag{12.4}$$

with $n_{\text{dof}} = 3$ for monoatomic and $n_{\text{dof}} = 5$ for diatomic ideal gases below some critical temperature. Hence for ideal gases, the released heat is proportional to the drop of temperature. Actually this should not be seen as a surprising result, since it is simply a consequence of the way temperature is defined using an ideal-gas thermometer.

Heat capacity at constant volume is an important parameter when the system is constrained to have a fixed volume. We may wonder whether the same amount of heat is released when keeping a constant pressure P_{ext}, the atmospheric pressure for instance, and letting the gas expanding freely. Of course, this will involve some work

$$\delta W = -P_{\text{ext}} dV$$

being negative in case of expansion, positive in case of contraction. Assuming a quasi-static process, the inside pressure is balanced by the outside pression. Taking now the derivative of the ideal-gas law for a closed system ($dn = 0$):

$$nRdT = d(PV) = V \overbrace{dP}^{=0} + \overbrace{P_{\text{ext}} dV}^{-\delta W}$$

Hence, first law of thermodynamic gives in this case:

$$\delta Q = dU - \delta W = d(U + PV) - V dP$$

Hence, as $dP = 0$:

$$C_p = \lim_{\Delta T \to 0} \left(\frac{\Delta Q}{\Delta T} \right)_{P,n} = \left(\frac{\partial (U + PV)}{\partial T} \right)_{P,n} \tag{12.5}$$

where C_p is the *heat capacity at constant pressure*. It is larger than C_v since when T increases so does the volume V. Work done by the outside work is negative. Hence some work has been done by the system that has to be compensated by some extra *heat* coming from outside to reach the same final internal energy. For an ideal gas:

$$C_p = C_v + nR$$

As opposed to C_v being the partial derivative of the U with respect to temperature at constant volume, C_p is not a partial derivative of the internal energy U at fixed pressure. However we can define a new state function $H(P,T)$ called the *Enthalpy*:

$$H = U + PV \tag{12.6}$$

such that:

$$C_p = \left(\frac{\partial H}{\partial T}\right)_{P,n} \tag{12.7}$$

Indeed, n being fixed:

$$dH = \delta Q + \overbrace{\delta W + PdV}^{0} + VdP$$

Actually H is the *Internal Energy* of the gas plus the energy that is required to empty a volume V under a constant pressure P. Similarly to the case of a particle under a gravity field for which the potential energy of gravity is added to the energy of the system, in H we add the potential energy of the surrounding environment at pressure P to the internal energy of the system.

12.3 Special processes

12.3.1 Isochoric processes ($V = \text{cst}$)

The isochoric pressure coefficient

$$\gamma_v = \frac{1}{P}\left(\frac{\partial P}{\partial T}\right)_V$$

For ideal gases:

$$VdP + PdV = nRdT \qquad \Rightarrow \qquad \gamma_v = \frac{1}{T}$$

12.3.2 Isobaric processes ($P = \text{cst}$)

The isobaric expansion coefficient

$$\gamma_p = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P$$

For ideal gases:

$$\gamma_p = \frac{1}{T}$$

12.3.3 Isothermal processes ($T = \text{cst}$)

The isothermal compressibility

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$$

For ideal gases:

$$\kappa_T = \frac{1}{P}$$

12.3.4 Adiabatic processes ($\delta Q = \text{cst}$)

The adiabatic compressibility

$$\kappa_S = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_S$$

For ideal gases:

$$\kappa_T = \gamma \kappa_S$$

Chapter 13

The second law of thermodynamics

Let us start this chapter with a set of seemingly unrelated statements:

- *Heat never flows from a cold bath to an hot one without any external work*
- *There is no engine more efficient than the Carnot Engine*
- *T is an integrating factor of the inexact differential form δQ*
- *Randomness always increases as a result of a non-reversible process*
- *Energy spreads to reach equipartition*
- *As time goes the amount of information to fully characterize any isolated system increases*
- *The amount of information to describe a system is maximum when states are equiprobable and is equal to the logarithm of the number of states.*
- *Entropy of the universe increases*

Now the question is "which one of these statements is the second law of thermodynamics?"

The answer is : "all but the last one".

This puzzling answer stresses on both the ubiquity and the complexity of the second law of thermodynamics studied in this chapter. As already mentioned, the story began with the efficiency of engins that was first studied by Saadi Carnot.

13.1 Efficiency of the Carnot cycle

The Carnot cycle of an ideal-gas heat engin is made of two reversible isothermal processes at hight and low temperatures T_h and $T_l < T_h$ closed by two reversible adiabatic processes. During the first isothermal process I, the work done by the outside is balanced by the transfer of heat as the internal energy remains constant. At high temperature T_h the gas expands from V_1 to V_2 with:

$$-W_{1\to 2} = nRT_1 \ln\left(\frac{V_2}{V_1}\right) = Q_{1\to 2}$$

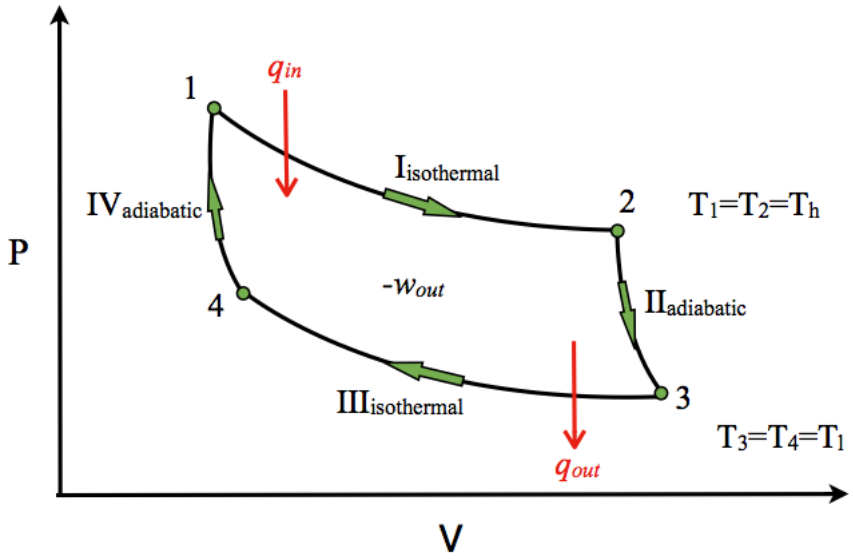


Figure 13.1: The Carnot Cycle

During the adiabatic cooling from $T_2 = T_1 = T_h$ to $T_3 = T_l$ no heat is exchanged and the work is:

$$W = C_v(T_h - T_l)$$

this work will be exactly balanced when heating from $T_4 = T_l$ to $T_1 = T_h$.

During isothermal compression at lower temperature T_l one obtains:

$$-W_{3\to4} = nRT_2 \ln\left(\frac{V_4}{V_3}\right) = Q_{3\to4}$$

The total work in the cycle is:

$$W_{\text{cycle}} = W_{1\to2} + W_{3\to4} = -(Q_{1\to2} + Q_{3\to4})$$

The efficiency η is defined as the ratio between the work provided to the outside world in the cycle $-W_{\text{cycle}}$ and the heat brought to the system by the hot bath $Q_{1\to2}$:

$$\eta = 1 + \frac{Q_{3\to4}}{Q_{1\to2}} = 1 + \frac{T_3}{T_1}\frac{\ln\left(\frac{V_4}{V_3}\right)}{\ln\left(\frac{V_2}{V_1}\right)}$$

But since along adiabatic transforms we have:

$$(V_2/V_3) = (T_3/T_1)^{1/(\gamma-1)}$$

and:

$$(V_4/V_1) = (T_1/T_3)^{1/(\gamma-1)}$$

Hence:

$$\frac{\ln\left(\frac{V_4}{V_3}\right)}{\ln\left(\frac{V_2}{V_1}\right)} = \frac{\ln\left(\frac{V_1}{V_3}\right) + (1/(\gamma-1))\ln(T_3/T_1)}{\ln\left(\frac{V_3}{V_1}\right) - (1/(\gamma-1))\ln(T_3/T_1)} = -1$$

and finally:

$$\eta = 1 + \frac{Q_{3\to4}}{Q_{1\to2}} = 1 - \frac{T_h}{T_l}$$

As a conclusion the efficiency of the Carnot cycle is solely given by the ratio between the temperature of the two heat baths. 100% efficiency can never be reached as $T_l > 0$. This was a major practical results from Carnot as it shows the role of the cold source on the efficiency of a particular heat engin, the Carnot's heat engin operating with an ideal gas. This was observed in the past by James Watt who noticed the importance of adding a condenser in the steam engin. What Carnot cleverly showed is that this property is ubiquitous for all engins performing a so-called *Carnot cycle* and not restricted to the Ideal-Gas Engin.

His reasoning is indeed very elegant and is briefly reproduced hereunder. Since the Carnot cycle is a reversible cycle, it can be performed in reverse. That is to say using external work to transfer heat from the cold bath to the hot bath: a refrigerator or an air-conditioning system. Hence were there be an engin more efficient than the Carnot engin it would be possible to couple these two engins as a single thermodynamic system using all the work of the first engin to transfer heat from the cold source to the hot source using a reverse Carnot cycle. No work will be produced by such an coupled engin and the net result will only be heat transfer with the two heat bathes. Then there are only two options left:

- no engin is more efficient than the Carnot engin as it would result in a net positive heat exchange by the coupled hybrid system from the cold bath to the hot one,
- If there exists an engin more efficient than the Carnot engin operating between two heat bathes then it is possible to build a system extracting heat from a cold source and transferring it to an hotter one without any external work and matter exchange.

The latter statement is contradicted by all experimental results so far. Hence the first statement is the correct one. It shows the equivalence between the maximum efficiency of the Carnot Engin and *Clausius* statement of the second law saying that *no heat can flow from a cold source to an hotter one without external energy supply.*

Hence Carnot practical statement appears to be much more profound than originally thought of. It is at the heart of the second law of thermodynamics introducing a new concept called Entropy.

13.2 Entropy

Looking back to Carnot's result on efficiency it is very easy to derive a very simple consequence that is:

$$\frac{Q_{3\to4}}{T_3} + \frac{Q_{1\to2}}{T_1} = 0$$

Noticing in addition that the heat exchange vanishes on the two adiabatic branches of the Carnot cycle, it can be concluded that:

$$\int_{\text{Cycle}} \frac{\delta Q}{T} = 0$$

Moreover, since any reversible cycle can be approximate by a series of consecutive arbitrary small isothermal and adiabatic processes, this result can be generalised to any cycle. Moreover it can be generalised to any reversible engin or process whatever the substance involved in it. Indeed any engin can be coupled to the reversed reversible-Carnot Engin performing the same cycle. As a consequence:

$$dS = \frac{\delta Q}{T} \tag{13.1}$$

is an exact differential associated with a new state variable S that can be defined for all thermodynamic system.

This state-variable was called *Entropy* by *Clausius* in the mid of the XIXth century and many discussions have been occurring since then about this quantity, its meaning and even the name that has been chosen. Clausius chose it to express with ancient Greek terms the concept of *inside transformation* or a mysterious *power of transformation*. Actually it has also been argued that the main interest in choosing this name is that it more or less means *nothing*. It is only in the course of the XXth century that further developments have given some insights on this rather abstract concept of *Entropy*.

One of the many reasons why *entropy* remains so mysterious for students and scientists alike is that as opposed to other state variables such as *Temperature*, *volume*, *pressure*, *entropy* cannot be straight-forwardly measured with some apparatus named an *entropy-meter*.

Whatsoever, in this attempt to decipher this intricate concept it is worth noticing first that *Entropy* is an *extensive* state variable. When putting together two identical thermodynamic systems, the *entropy* of the overall system doubles as compared to the one of each individual system since the heat doubles and the temperature remains the same. From the mathematical point of view, expressing *entropy* as a function of other extensive state variables under the form $S(U, V, n)$, this function has to satisfy the homogeneity condition:

$$S(\lambda U, \lambda V, \lambda n) = \lambda S(U, V, n) \qquad \lambda > 0 \tag{13.2}$$

13.3 Irreversible processes

As of now only *reversible* or *quasi-static processes* have been considered. That is to say processes that are slow enough to allow a *thermal equilibrium* at each step of the process. This implies in particular homogeneous temperature and pressure states in the system. Such an equilibrium is not satisfied for spontaneous expansion of a gas in vacuum as the flow of particles during this quick expansion is going to be very complex and not uniform. Hence during this process none of the state variables, including *entropy* can be defined. When equilibrium is finally reached after some time all the state variables can be defined again. Amongst these, *entropy* is the only one on which a general statement can be made. In the absence of any work, heat or matter exchange with the outside, the change in *entropy* has to be positive. In other words:

Definition 18 (Second principle of thermodynamics) *In the absence of any work, heat and matter exchange no thermal equilibrium can be reached by a thermodynamic system with a decrease of entropy:*

$$\Delta S \geq 0$$

The underlying question is how this increase in entropy can be quantified. There are actually two options. The first one consists in having access to the state function itself $S(U, V, n)$, assuming that the other state variables in the initial and final states are known. Hence:

$$\Delta S = S(U_2, V_2, n_2) - S(U_1, V_1, n_1)$$

However, such a state function is only known in a very limited number of specific cases, including ideal gases as studied in Section 13.4. The other option left consists in finding for the system a reversible process starting from the same state 1 and ending at the same state 2. This reversible process involves heat or work exchange with the outside and the final entropy change is then given by:

$$\Delta S = \int_{\text{Path}} \frac{\delta Q}{T}$$

13.3.1 Spontaneous expansion of an ideal gas

Let us consider an ideal gas confined inside a container of volume V_1 with temperature T_1 and a fixed number of moles n. Pressure P_1 is given by the law of ideal gases. Moreover the internal energy is given by *heat capacity* $U_1 = C_v T_1$.

This container is equipped with a removable wall expending its entire volume to $V_2 > V_1$. The other part of the container $V_2 - V_1$ is assumed to be empty. No heat, work or matter are exchanged with the outside world ($n_2 = n_1 = n$). Hence the internal energy is preserved during the expansion process ($U_2 = U_1 = U$) and so is temperature ($T_2 = T_1 = T$). Hence the final pressure is:

$$P_2 = P_1 \frac{V_1}{V_2}$$

One possible reversible process to reach this state is thus an isothermal process for which:

$$\Delta Q = -\Delta W = nRT \ln\left(\frac{V_2}{V_1}\right)$$

So finally the change in *entropy* can be computed based on the reversible heat exchange:

$$\Delta S = \int_{\text{Path}} \frac{\delta Q}{T} = \frac{\Delta Q}{T} = nR \ln\left(\frac{V_2}{V_1}\right)$$

One important feature here is that the change in entropy is on the first hand proportional to the number of particle in the system and on the second hand proportional to the logarithm of the volume occupied by these particles. Note that nothing can yet be said on the dependency of *entropy* with respect to *temperature* or *Internal energy* as both are fixed in the considered experiment.

13.3.2 Mixing of two gases at different temperature

Let us now observe the change in entropy induced by mixing two gases each one contained in two enclosures of volume V at two different temperatures T_1 and T_2 with the same number of particles N in each compartment. The total internal energy:

$$U = U_1 + U_2 = C_V(T_1 + T_2)$$

is preserved giving the final temperature:

$$T_f = \frac{U}{2C_V} = \frac{T_1 + T_2}{2}$$

This state could be reached in each of the two containers through a reversible isochoric process from T_1 to T_f and T_2 to T_f respectively. Hence, no work is exchanged in this process and for the first enclosure:

$$\delta Q_1 = dU_1 = C_v dT$$

as a consequence:

$$\Delta S_1 = \int_{T_1}^{T_f} C_v \frac{dT}{T} = C_v \ln\left(\frac{T_f}{T_1}\right)$$

and finally:

$$\Delta S = C_v \ln\left(\frac{T_f}{T_1}\right) + C_v \ln\left(\frac{T_f}{T_2}\right) = 2C_v \ln\left(\frac{T_1 + T_2}{2\sqrt{T_1 T_2}}\right)$$

or considering the *internal energy* as a primary state-variable together with the number of mols n:

$$\Delta S_1 = nR \ln \left(\left(\frac{U}{U_1} \right)^{\frac{3}{2}} \right)$$

for a monoatomic gas in the first enclosure. Hence, the change in entropy is also proportional to the logarithm of temperature or internal energy. However the exponent $3/2$ on the *internal energy* is different from the one unveiled for the dependency on volume.

Section 13.4 gives a detailed derivation of the *entropy* of ideal gases generalising the features introduced in studying spontaneous expansion or mixing on ideal gases, based on a macroscopic approach. Appendix **??** gives some insight from a statistical perspective at microscopic level based on information theory. In particular it will explain the logarithmic dependency with respect to volume and energy in terms of *measure of information* required to locate one particle inside equally probable micro-states.

13.4 Entropy of an ideal monoatomic gas

Entropy being a *state variable*, it can be expressed as a function of other state variables. Focussing first on closed systems - that is to say a fixed number of particles N- the other extensive variables are *Energy* and *Volume*. Dividing the *first law*:

$$dU = \delta Q - PdV$$

by the non vanishing temperature T gives:

$$dS = \frac{dU}{T} + \frac{PdV}{T}$$

Using the expression of the internal energy and law of ideal gas to substitute for $T = \frac{2U}{n_{\text{dof}} N k_B}$ and P leads to:

$$dS = k_B N \left(\frac{3}{2} \frac{dU}{U} + \frac{dV}{V} \right)$$

Giving access to definite expression of the partial derivatives of S with respect to U and V:

$$dS = \left(\frac{\partial S}{\partial U} \right)_{V,n} dU + \left(\frac{\partial S}{\partial V} \right)_{U,n} dV$$

It finally gives the value of the *entropy* up to a multiplicative term in the logarithm that does not depend on U and V but possibly depends on the number of particles N:

$$S(U, V, N) = k_B N \ln \left(\frac{V U^{3/2}}{f(N)} \right)$$

Since the entropy is an extensive variable, last condition in Equation (13.2) leads to:

$$f = C N^{5/2}$$

where C is a dimensional constant with dimension $[M]^{3/2}[L]^6[T]^{-3}$. However the mass m of the particles is the only dimensional parameter left in the system itself. Fortunately quantum physics and in particular the Heisenberg uncertainty principle[1] gives the last parameter to close

[1] $$\Delta x \Delta p \geq \hbar$$

the analysis, namely the Planck's constant $\hbar$ with the physical dimension of an action $[M][L]^2[T]^{-1}$. Hence:

$$C = \alpha \left(\frac{\hbar^2}{m}\right)^{3/2}$$

The final result is known as the *Sackur-Tetrode* equation published in 1912 based on Boltzmann's statistics:

$$S(U,V,N) = k_B N \left(\ln\left(\frac{V}{N}\left(\frac{U}{N}\right)^{3/2}\left(\frac{4\pi m}{3\hbar^2}\right)^{3/2}\right) + \frac{5}{2}\right) \tag{13.3}$$

where $\frac{V}{N}$ is the average volume occupied by one particle and $\frac{U}{N}$ is its average kinematic energy. Hence the *Entropy* of an ideal gas is the sum of the *entropy* of the N particles. The fact that this *entropy* of one single particle scales as the logarithm of the average volume plus the average energy will be discussed in the Appendix **??** in terms of Shannon's Measure of Information, together with the value of the constant term α.

13.5 Properties of entropy and other state functions

Taking the maximum of Shannon's entropy as the second law of thermodynamics, whatever substance or system considered, temperature can now be defined as the inverse of the rate of change of entropy with respect to internal energy for fixed volume and number of particles:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{V,N} > 0 \tag{13.4}$$

Actually factor K_B in the definition of entropy is set to end up with the absolute Kelvin temperature. For a monoatomic ideal gas temperature is found to be a linear function of the internal energy, rescaled by the number of particles:

$$T = \left(\frac{\partial U}{\partial S}\right)_{V,N} = \left(\frac{\partial S}{\partial U}\right)_{V,n}^{-1} = \left(\frac{3}{2}\frac{NK}{U}\right)^{-1} = \frac{2}{3}\frac{U}{NK}$$

Another interesting property derived from the fact that at thermodynamic equilibrium Shannon's entropy is maximum with respect to extensive state variables (U,V,N) is that S is a concave function of the state variables. In particular:

$$0 \geq \left(\frac{\partial^2 S}{\partial U^2}\right)_{V,n} = \left(\frac{\partial (1/T)}{\partial U}\right)_{V,n} = -\frac{1}{T^2}\left(\frac{\partial T}{\partial U}\right)_{V,n} = -\frac{1}{C_v T^2}$$

showing the the heat capacity at constant volume is strictly positive:

$$C_v \geq 0$$

Similar properties can be obtained taking the partial derivatives with respect to the other state variables. However, the corresponding expression takes an unusual form as temperature is not taken as a primary state variable.

Helmholtz free energy is defined as:

$$\Phi(T,V,N) = U - TS$$

and is a convex function of the state variables (T, V, N). Hence one possible statement of the second law of thermodynamics is that the *Helmholtz Free energy* of the system is minimum at thermodynamic equilibrium.

Moreover since:

$$d\Phi = \left(\frac{\partial \Phi}{\partial T}\right)_{V,n} dT + \left(\frac{\partial \Phi}{\partial V}\right)_{T,n} dV + \left(\frac{\partial \Phi}{\partial n}\right)_{T,V} dn = \overbrace{\underbrace{-PdV + \mu dn}_{\delta W} + \underbrace{TdS}_{\delta Q}}^{dU} \overbrace{-TdS - SdT}^{-d(TS)}$$

Hence:

$$S = -\left(\frac{\partial \Phi}{\partial T}\right)_{V,n}, \qquad P = -\left(\frac{\partial \Phi}{\partial V}\right)_{T,n}, \qquad \mu = \left(\frac{\partial \Phi}{\partial n}\right)_{T,V}$$

Where μ is the chemical potential. As a consequence. One can link the change in pressure with temperature to the change of entropy with volume:

$$\left(\frac{\partial P}{\partial T}\right)_{V,n} = -\left(\frac{\partial^2 \Phi}{\partial T \partial V}\right)_{n} = \left(\frac{\partial S}{\partial V}\right)_{T,n} \overset{\text{ideal gas}}{=} \frac{NK}{V}$$

Taking the derivatives with respect to other extensive state variables V and n gives access to the two adjoint intensive variables :

$$0 < \left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{P}{T}$$

and:

$$0 < \left(\frac{\partial S}{\partial N}\right)_{U,V} = \frac{\mu}{T}$$

where μ is the chemical potential.

$$0 \geq \left(\frac{\partial^2 S}{\partial V^2}\right)_{U,N} = \left(\frac{\partial (P/T)}{\partial V}\right)_{U,N} = \frac{1}{T}\left(\frac{\partial P}{\partial V}\right)_{U,N} - \frac{P}{T^2}\left(\frac{\partial T}{\partial V}\right)_{U,N}$$

and:

$$0 < \left(\frac{\partial S}{\partial N}\right)_{U,V} = \frac{\mu}{T}$$

where μ is the chemical potential.

Part III

Dynamics of continuous systems

Chapter 14

Waves

In physics, waves can be of many types: quantum waves, electro-magnetic waves/light, acoustic waves, seismic waves, flexure waves, water waves. Though in these cases different physical quantities vary as a function of space and time: the wave function, the electric field, the pressure field, the displacement of the ground, of a string or of the water free-surface, the way they evolve shares common features that can be studied in particular cases and then generalised. In Mechanics, when restricting the analysis to one dimensional propagation, the vibrating string and the acoustic pipe are the two examples we are going to study in this chapter to explore some of these features.

14.1 The vibrating string

Let us consider the displacements of an horizontal string being tight at both ends with an uniform tension T_o. Let x be the longitudinal coordinate along the string, and y its vertical displacement. It is assumed that the horizontal displacement is negligible as compared to the vertical one. Hence the displacement along the string and over time will be fully described by the function of two variables $y(x,t)$. We also assume that the vertical displacement at rest due to gravity is negligible[1].

At any point along the string - hence for any x - let $\overrightarrow{T}(x)$ be the force vector applied by the right part of the string on its left part. Let μ be the mass per unit length - L being the length, M the total mass then $\mu = M/L$.

The 2nd Newton's Law to a small part of the string between x and $x + \Delta x$ is first considered. The force applied on the right side is:

$$\overrightarrow{T}(x + \Delta x) = \overrightarrow{T} + \overrightarrow{\Delta T}$$

Due to the action-reaction Law - 3rd Newton's Law - the force applied on the left side is $-\overrightarrow{T}$. Other external forces are neglected including gravity and friction with the surrounding air.

Assuming that Δx is small enough, the acceleration is uniform over this small element and is given by:

$$\overrightarrow{\gamma} = \frac{\partial^2 y}{\partial t^2}\overrightarrow{e_y}$$

The mass being $\mu \Delta x$, the balance of momentum reads:

$$\overrightarrow{T} + \overrightarrow{\Delta T} - \overrightarrow{T} = \mu \Delta x \overrightarrow{\gamma}$$

[1] As an exercise you could show that its shape is an hyperbolic-cosine.

or once simplified and divided by Δx:

$$\frac{\overrightarrow{\Delta T}}{\Delta x} = \mu \frac{\partial^2 y}{\partial t^2} \overrightarrow{e_y}$$

For vanishing Δx

$$\frac{\overrightarrow{\Delta T}}{\Delta x} \to \frac{\partial \overrightarrow{T}}{\partial x}$$

yielding:

$$\frac{\partial \overrightarrow{T}}{\partial x} = \mu \frac{\partial^2 y}{\partial t^2} \overrightarrow{e_y} \tag{14.1}$$

Denoting by T_x and T_y the two components of force $\overrightarrow{T}$, the horizontal component of the previous vector equation gives:

$$\frac{\partial T_x}{\partial x} = 0$$

Showing that the horizontal component is constant and equal to the original traction T_o. The vertical component reads:

$$\frac{\partial T_y}{\partial x} = \mu \frac{\partial^2 y}{\partial t^2}$$

We now need to relate the force at x, and more specifically its unknown vertical component T_y with the vertical displacement y to close the latter equation. Actually we known that, as a consequence of the balance of angular momentum[2], the tension force $\overrightarrow{T}$ is along the tangent of the string:

$$\overrightarrow{T} \times \overrightarrow{t} = \overrightarrow{0}, \qquad \text{with} \qquad \overrightarrow{t} = \begin{pmatrix} 1 \\ \frac{\partial y}{\partial x} \end{pmatrix}$$

leading to:

$$T_x \frac{\partial y}{\partial x} - T_y = 0 \qquad \Rightarrow \qquad T_y = T_o \frac{\partial y}{\partial x}$$

Since T_o is constant, using this expression in the vertical component of the balance of momentum yields:

$$T_o \frac{\partial^2 y}{\partial^2 x} = \mu \frac{\partial^2 y}{\partial t^2} \tag{14.2}$$

Since T_o and μ and non negative quantities, defining:

$$c = \sqrt{\frac{T_o}{\mu}}$$

One can easily verify that c has the dimension of a velocity:

$$[c] = (\frac{[M][L][T]^{-2}}{[M][L]^{-1}} = [L][T]^{-1}$$

[2]Indeed, the balance of angular momentum with respect to the right end of the string part reads:

$$(\Delta x \overrightarrow{t}) \times \overrightarrow{T} = \mu \frac{(\Delta x)^3}{3} \frac{\partial^2 \theta}{\partial t^2} \overrightarrow{e_z}$$

with θ the rotation angle of the string part with respect to $\overrightarrow{e_z}$. Hence dividing by Δx and taking the limit $\Delta x \to 0$ yield:

$$\overrightarrow{t} \times \overrightarrow{T} = \overrightarrow{0}$$

and obtain the so-called one-dimensional wave equation:

$$c^2 \frac{\partial^2 y}{\partial^2 x} = \frac{\partial^2 y}{\partial t^2} \tag{14.3}$$

Note that to be solved properly, this second order partial differential equation has to be supplemented with initial and boundary conditions. For instance, the string is released from a given shape $y_o(x)$ at $t = 0$ and the two ends are fixed ($y(0,t) = 0 = y(L,t)$). This will be studied in subsequent sections.

14.2 Traveling waves

The wave equation is named as such because any function:

$$y(x,t) = f(x - ct)$$

is solution of this equation, for any twice differentiable function of a single variable f. Indeed:

$$\frac{\partial y}{\partial x} = f'(x - ct)\frac{\partial (x - ct)}{\partial x} = f'(x - ct)$$

Similarly:

$$\frac{\partial y}{\partial t} = f'(x - ct)\frac{\partial (x - ct)}{\partial t} = f'(x - ct)(-c)$$

Hence:

$$c^2 \frac{\partial^2 y}{\partial^2 x} = c^2 f''(x - ct) = (-c)^2 f''(x - ct) = \frac{\partial^2 y}{\partial t^2}$$

Let us also notice that this property holds also for:

$$y(x,t) = g(x + ct)$$

provided that g is twice differentiable. Let us now show that these particular solution behave like "waves", that is to say, that something remains constant once translated over space and time at a given velocity. Let us remark that at $t = 0$:

$$y(x, t = 0) = f(x)$$

Let us now observe what happen at $x' = x + ct$ at time t. One can easily show that:

$$y(x', t) = f(x' - ct) = f((x + ct) - ct) = f(x) = y(x, t = 0)$$

Hence, the vertical displacement at x and $t = 0$ "propagates" to x' at time t. And since the distance between x and x' is $d = ct$, the corresponding velocity is $d/t = c$. Note that the propagation is towards the positive x-axis. Similarly it can easily be noticed that $y(x,t) = g(x+ct)$ also propagates but in the opposite direction, from positive to negative x.

14.2.1 How to generate a travelling wave?

The question is now to known how to practically create such travelling waves. Let us start considering an infinitely long string with x ranging from $-\infty$ to $+\infty$. Let us try to hold the string at $t = 0$ so that it form a nice bell-shape function at $x = 0$. Hence the initial conditions are:

$$y(x, 0) = y_o(x) \qquad \text{and} \qquad \frac{\partial y}{\partial t}(x, 0) = 0$$

since we assume the string at rest at time $t = 0$.

Since we do not know in which direction waves are going to propagate - and they are likely to equally propagate in both directions as the orientation of the x axis is purely arbitrary here - let us assume that y is the sum of two waves propagating in opposite directions:

$$y(x,t) = f(x - ct) + g(x + ct)$$

The velocity is then given by:

$$\frac{\partial y}{\partial t}(x,t) = c(-f'(x - ct) + g'(x + ct))$$

and condition $\frac{\partial y}{\partial t}(x,0) = 0$ yields, for all x on the real line:

$$f'(x) = g'(x), \qquad \Rightarrow \qquad g(x) = f(x) + \text{cst}$$

This constant term cst is expected to vanish as a constant vertical displacement of an infinite string is not physically sound. Hence, putting this condition $f = g$ in the initial condition simply gives:

$$y(x,0) = f(x) + g(x) = 2f(x) = y_o(x)$$

and finally:

$$y(x,t) = \frac{1}{2}(y_o(x - ct) + y_o(x + ct))$$

Thus, the initial shape $y_o(x)$ is equally distributed onto a right-propagating wave and a left-propagating wave, propagating to $+\infty$ and $-\infty$ respectively at wave-velocity c.

It is worth noticing that though this solution has been found for positive times ($t \geq 0$) it remains valid for negative times too. Now the physical meaning changes as we have two waves having the same shape and amplitude, one coming from $-\infty$ the other one coming from $+\infty$ that meet and superpose at $t = 0$ around $x = 0$.

Interestingly too, one could consider this solution for only positive x - the right part of the infinite string. Once again the physical interpretation is different as we now have one incident wave propagating from $+\infty$ towards $x = 0$. Then it is reflected at $x = 0$ with a reflection coefficient equal to 1 by some boundary condition at $x = 0$. Actually, this physical interpretation is correct only if we assume that y_o is an even function. In this case ($y_o(-x) = y_o(x)$) it is easy to verify that:

$$\frac{\partial y}{\partial x}(0,t) = c(f'(-ct) + g'(+ct)) = \frac{c}{2}(y_o'(-ct) + y_o'(ct)) = 0$$

Hence, since $T_y = T_o \frac{\partial y}{\partial x}$, the vertical force vanishes at $x = 0$, corresponding to a free-end boundary condition, in this particular case. Actually we cannot, from the physical point of view rely on the assumption that the initial condition y_o is an even function of space, as it is very difficult to achieve experimentally. However, these last years researchers have been able to experimentally implement this boundary condition that is called a time-reversal mirror. This is presently a very active field of research raising fundamental questions about reversibility of time and giving rise to many very useful applications, from health monitoring and non invasive treatment to underwater communication, seismology and volcanic eruption and earthquake prediction.

14.2.2 Reflection of waves

Having now a better understanding of wave propagation, reflection of such waves can be studied. To this aim let us consider a semi-infinite string corresponding to non-negative x and let us assume that the string is fixed at its left-end side, that is to say:

$$y(0,t) = 0, \qquad \forall t$$

Let us also assume that we known that an incident wave y_{inc} with pulse shape $y_o(x)$ is propagating from $+\infty$ towards this fixed point at $x = 0$[3]:

$$y_{\text{inc}}(x,t) = y_o(x+ct)$$

Obviously, this incident wave would only meet this boundary condition $y(0,t) = 0$ at all times when the incident wave itself vanishes. Hence let us add to the solution of the wave equation a right-propagating wave or reflected wave y_{refl} of unknown shape $f(x)$ so that:

$$y(x,t) = y_{\text{refl}}(x,t) + y_{\text{inc}}(x,t) = f(x-ct) + y_o(x+ct)$$

This first function on the right-hand side of this equation being unknown, when the second one is assumed to be known. The fixed boundary condition at $x = 0$ simply gives the shape of the reflected wave:

$$y(0,t) = 0 = f(-ct) + y_o(ct) \qquad \Rightarrow \qquad f(x) = -y_o(-x)$$

giving finally:

$$y(x,t) = -y_o(-x+ct) + y_o(x+ct)$$

Note that the first term on the right-hand side is indeed a right-propagating wave due to the $-x$ term. Thus the effect of the fixed end is twofold:

- the direction of propagation is reversed,
- the amplitude is reversed too.

Hence the reflection coefficient is -1. Note that at the fixed end the vertical force T_y doubles as compared to the one induced by the incident wave only. Indeed:

$$T_y(0,t) = T_o \frac{\partial y}{\partial x} = T_o(y_o'(-0+ct) + y_o'(0+ct)) = 2T_o y_o'(0+ct) = 2T_o y_{\text{inc}}(0,t)$$

Reflection on a free-end The same approach can be worked out to find the reflection on a free end. The end-condition then becomes:

$$T_y(0,t) = 0 \qquad \Rightarrow \qquad \frac{\partial y}{\partial x} = 0, \qquad \forall t$$

Leading to:

$$\frac{\partial y}{\partial x} = 0 = f'(-ct) + y_o'(ct) \qquad \Rightarrow \qquad f'(x) = -y_o'(-x)$$

and finally integrating:

$$f(x) = y_o(-x) + \text{cst}$$

Leading to a permanent displacement of the string, the constant term is dropped as non physical yielding:

$$y(x,t) = y_o(-x+ct) + y_o(x+ct)$$

Thus the effect of the free-end is again a change in the direction of the wave, and this will remain true for any standard end-condition. However, this time the amplitude is preserved and, as a consequence, the amplitude doubles at the free-end when the incident and the reflected waves coincide. The reflection coefficient is thus $+1$.

[3]Note that, though the wave is only defined on the real positive line, since is has to be defined for all times, y_o as to be defined on the entire real line.

14.3 Harmonic waves

When time dependency of the wave amplitude is harmonic in time at each and every point, the wave is called an harmonic wave. for a right-propagating wave it reads:

$$y(x,t) = A^+ \cos(\omega t + \phi(x)) = f(x - ct)$$

and as a consequence:

$$f(-ct) = A^+ \cos(\omega t + \phi_o^+) \qquad \Rightarrow \qquad f(x) = A^+ \cos(-kx + \phi_o^+), \qquad \phi(x) = -kx + \phi_o^+, \qquad k = \frac{\omega}{c}$$

where k is the *wave-number* having a physical dimension equal to $[\mathrm{rad}L^{-1}]$ and where A^+ is the amplitude.

Hence when the amplitude is harmonic with respect to time with a period $T = \frac{2\pi}{\omega}$, it is harmonic with respect to space with a wave-length:

$$\lambda = \frac{2\pi}{k} = \frac{2\pi c}{\omega} = cT$$

This means that the wavelength λ is the distance travelled by the wave during one period T. It is worth noticing that a harmonic wave has an unbounded support in both space and time.

Using now a left propagating wave yields:

$$g(ct) = A^- \cos(\omega t + \phi_o^-) \qquad \Rightarrow \qquad g(x) = A^- \cos(kx + \phi_o^-), \qquad \phi(x) = kx + \phi_o^-, \qquad k = \frac{\omega}{c}$$

These two solutions or any combination of these are actually the only solution for harmonic waves. Indeed, directly putting the general form:

$$y(x,t) = A \cos(\omega t + \phi(x))$$

in the wave equation yields for $A \neq 0$:

$$c^2 \phi''(x) \sin(\omega t + \phi(x)) = (c^2 \phi'(x)^2 - \omega^2) \cos(\omega t + \phi(x))$$

Since the sin and cos functions can never coincide on their entire range the leading coefficients on the right and left-hand sides of this equation have to vanish yielding:

$$\phi''(x) = 0 \qquad \Rightarrow \qquad \phi(x) = kx + \phi_o$$

and

$$c^2 k^2 - \omega^2 = 0 \qquad \Rightarrow \qquad k = \pm \frac{\omega}{c}$$

As a consequence the general solution reads:

$$y(x,t) = A^+ \cos(\omega t - kx + \phi_o^+) + A^- \cos(\omega t + kx + \phi_o^-)$$

$\pm$ superscript standing for the direction of propagation.

14.3.1 Standing waves

We have seen previously that for a fixed-end string the incident and reflected wave must have opposite amplitudes and have to be synchronised at $x = 0$.

Hence $\phi_o^+ = \phi_o^- = \phi_o$ and $A^+ = -A^- = -A$ yielding:

$$\begin{aligned} y(x,t) &= A\left(-\cos(\omega t - kx + \phi_o) + \cos(\omega t + kx + \phi_o)\right) && (14.4) \\ &= -2A\sin(\omega t + \phi_o)\sin(kx) && (14.5) \end{aligned}$$

Surprisingly, when these two propagating waves combine, propagation seemingly stops as the solution can be written as the product of one function of space and one function of time. Though y is varying both with respect to space and time, it happens to do so not truly independently but more correctly separately[4]. Hence this feature of waves combining in such a way that no propagation can be seen is called *standing waves*. As it will be observed in the next chapter this phenomenon is more in line with vibrations as observed in Chapter 9.

One remarkable feature of these standing waves is the existence of "nodal" and "bright" points. Indeed one can easily observe that all points periodically located at multiples of half the wavelength ($x = n\lambda/2$) have a vanishing amplitude at all time t. They are called nodal points. Conversely, all points located at $x = n\lambda/2+\lambda/4$ - namely all points located in the middle of the latter neighbouring nodal points -. experience the largest amplitudes of all points over time. They are called "bring" points. This feature showing fringes of low and large amplitudes is ubiquitous in physics anytime waves are involved and is known as *interferences*.

As a conclusion, let us observe that such standing waves also occur in the case of a free-end, showing the same bright and nodal points for the vertical force T_y, the bright and nodal points for the displacement being interchanged as compared to the fixed-end conditions.

[4]Indeed these two functions are not independents as they are both sin-functions and moreover ω and k are related throughout the wave velocity $k = \omega/c$.

Chapter 15

Resonant modes

Vibration of strings such as those studied in the previous chapter is the driving phenomenon for many applications and in particular for many music instruments, known as string instruments. Actually it even frames our perception of music and the definition of the tones and how it differs from frequency. Before studying this, we will first extend the analysis to wind instruments that basically share the same features once we replace the vibrating string by an acoustic pipe.

15.1 Acoustics in a pipe

Let us consider a straight pipe with a constant cross-section of area A, possibly circular but not necessarily, and of length L large with respect to the cross-section $L \gg \sqrt{A}$. This pipe is rigid, filled with air at room temperature with two openings at both ends. x is the longitudinal coordinate along the pipe ranging from 0 to L.

The air in the pipe is assumed to be characterised by three related physical quantities:

- p the excess pressure from atmospheric pressure in the room,
- ρ the mass density of air
- v the axial velocity of air in the pipe.

Thank to the large aspect ratio of the pipe, these quantities are assumed to be constant over the cross-section and thus will vary only as a function of the longitudinal coordinate x and time t.

These three quantities are related by two conservation laws:

- the conservation of mass, also known as the continuity equation
- the conservation or balance of momentum, that is to say the 2nd Newton's law.

Finally the air is assumed to have a constant compressibility:

$$c^{-2} = \frac{\partial \rho}{\partial p}$$

15.1.1 Conservation of mass

Let us consider the mass of the small volume of air $\Delta V = A\Delta x$ between the two cross sections at x and $x + \Delta x$. At time t this mass is:

$$M(t) = \rho(t) A \Delta x$$

At time $t + \Delta t$ it is given by:

$$M(t + \Delta t) = \rho(t + \Delta t)A\Delta x$$

but the conservation of mass states that any change in mass in this volume is balanced by inflows or outflows of air on both side of ΔV. Hence:

$$dM = M(t + \Delta t) - M(t) = \underbrace{\rho A v(x)\Delta t}_{\text{inflow left side}} - \underbrace{\rho A v(x + \Delta x)\Delta t}_{\text{outflow left side}}$$

Dividing this equation by $A\Delta x\Delta t$ together with the limit for vanishing Δx and Δt:

$$\frac{\partial \rho}{\partial t} = -\rho \frac{\partial v}{\partial x}$$

that is known as the one-dimensional continuity equation. Noticing now that[1]:

$$\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial p}\frac{\partial p}{\partial t}$$

the continuity equation becomes:

$$c^{-2}\frac{\partial p}{\partial t} = -\rho \frac{\partial v}{\partial x} \tag{15.1}$$

where ρ can now be considered constant and equal the the mass density of air in the room.

15.1.2 Balance of momentum

Let us now consider the balance of momentum of the same volume of air $\Delta V = A\Delta x$. p_o being the atmospheric pressure in the room, the force applied on the left-hand side is:

$$\overrightarrow{F}_{\text{left}} = (p_o + p(x))A\overrightarrow{e_x}$$

as p is the pressure excess only, whereas the force on the right-hand side is:

$$\overrightarrow{F}_{\text{right}} = -(p_o + p(x + \Delta x))A\overrightarrow{e_x}$$

with a minus sign since, on the upstream face, the pressure pushes the small volume backwards. The sum of the surface force applied by the tube on the air inside ΔV is denoted by $\overrightarrow{F}_{\text{pipe}}$. It is locally balanced by the air pressure on the tube and is thus normal to $\overrightarrow{e_x}$ as non friction or viscosity is considered:

$$\overrightarrow{F}_{\text{pipe}} \cdot \overrightarrow{e_x} = 0$$

Moreover, since the velocity field is assumed to be only in the longitudinal direction the rate of momentum is given by:

$$\frac{d\overrightarrow{P}}{dt} = \underbrace{\rho A\Delta x}_{M}\frac{\partial v}{\partial t}\overrightarrow{e_x}$$

Hence the balance of momentum gives:

$$p(x)A - p(x + \Delta x)A = \rho A\Delta x\frac{\partial v}{\partial t}, \qquad \text{and} \qquad \overrightarrow{F}_{\text{pipe}} = 0$$

Dividing the first equation by $\Delta V = A\Delta x$ and taking the limit for small Δx yields:

$$-\frac{\partial p}{\partial x} = \rho\frac{\partial v}{\partial t} \tag{15.2}$$

[1] It is shown in Chapter 12 that for rapid change with not heat exchange, density only depends on pressure and:

$$c^2 = \gamma RT/m$$

with R the Avogadro number, T the temperature, m the molar mass and γ the ratio of molar specific heat of an ideal-gas.

15.1.3 Acoustic wave equation

Taking the sum of the time derivative of the continuity equation (15.1) and the space derivative of the balance of momentum (15.2) finally gives:

$$c^{-2}\frac{\partial^2 p}{\partial t^2} - \frac{\partial^2 p}{\partial x^2} = 0$$

which happens to the the wave equation for the pressure excess p. As an analogy we could call the pipe as a string of air.

15.1.4 End conditions

Most of wing instruments have openings at both ends. One is close to the mouthpiece, the other one is the first open hole or key depending on the fingering of the player. Hence the fingering allows to change the length of pipe. Hence when there is an opening the pressure is equal to the atmospheric pressure and the pressure excess p vanishes. The velocity of air is unknown as the air is free to flow outside the pipe.

Few instruments such as clarinets and oboes do not have an opening close to the mouthpiece. Hence air is not allowed to flow and the velocity vanishes. Using the balance of momentum one can deduce that the derivative of the pressure excess with respect to space vanishes also at such ends. The case of brass instruments is a little more complex but they are considered as reed instruments, the reed being the lips of the player.

15.2 Vibrating modes

Let us consider a pipe of length L opened at both ends so that the pressure excess p vanishes. Note that this problem is similar to the one of a vibrating string of length L fixed at both ends. The equations to solve for p are as follows:

$$c^2\frac{\partial^2 p}{\partial x^2} = \frac{\partial^2 p}{\partial t^2} \tag{15.3}$$

$$p(x \in \{0, L\}, t) = 0 \tag{15.4}$$

As observed in last chapter, the existence of fixed ends will generate multiple reflections and is likely to give rise to standing waves. Hence we look for solutions with separable variables:

$$p(x,t) = \chi(x)\tau(t)$$

when substituted in the wave equation one obtains:

$$c^2\chi''\tau = \chi\ddot{\tau} \tag{15.5}$$

$$\chi(x \in \{0, L\}) = 0 \tag{15.6}$$

where $'$ stands for space derivatives and $\dot{}$ for time derivatives. Dividing the first equation by $\chi(x)\tau(t)$ yields:

$$c^2\frac{\chi''}{\chi}(x) = \frac{\ddot{\tau}}{\tau}(t) = -\omega^2$$

Since the first term is a function of space only and the second a function of time only, then there are both constant. A dimensional analysis tells us that this constant has the physical dimension of $[T^{-2}]$ and thus it is set equal to $-\omega^2$.

Solving for $\tau(t)$ leads to:

$$\tau(t) = A\cos(\omega t) + B\sin(\omega t)$$

Solving for χ leads to:

$$\chi(x) = C\cos(kx) + D\sin(kx)$$

with $k = w/c$. Thanks to (15.6), $C = 0$ and:

$$\sin(kL) = 0 \qquad \Rightarrow \qquad \omega_n = (n+1)\pi\frac{c}{L}$$

$\omega_o = \pi\frac{c}{L}$ is called the *fundamental circular frequency*, other ω_n are called *natural* or *eigen* frequencies. For strings or pipes they happened to be multiples of the fundamental one and are called *harmonics*. χ_n are the *natural* or *eigen* modes, the fundamental one consisting of half a wavelength fitting in the pipe, two wavelengths for the second one, three haves for the third and so on.

Hence the general solution for p is:

$$p(x,t) = \sum_{n=0}^{+\infty} \left(A_n\cos(\omega_n t) + B_n\sin(\omega_n t)\right)\chi_n(x) \tag{15.7}$$

$$\chi(x) = \sin\left((n+1)\pi\frac{x}{L}\right) \tag{15.8}$$

Coefficients A_n and B_n are defined by the initial conditions $p_o(x)$ and $q_o(x) = \frac{\partial p}{\partial t}(x,0)$. As a conclusion when playing any wind or string instrument, not only the fundamental frequency is played but also its harmonics. The superposition of these makes the "tone" having a different spectrum for different instruments depending on the amplitude/energy put on each of the modes.

As mentioned earlier, some of the wind instruments have no opening at the mouthpiece. Hence the problem reads:

$$c^2\frac{\partial^2 p}{\partial x^2} = \frac{\partial^2 p}{\partial t^2} \tag{15.9}$$

$$\frac{\partial p}{\partial x}(0,t) = 0 \tag{15.10}$$

$$p(L,t) = 0 \tag{15.11}$$

the solution process is the same except that it ends up with $D = 0$ and:

$$\cos(kL) = 0 \qquad \Rightarrow \qquad \omega_n = (n+\frac{1}{2})\pi\frac{c}{L}$$

As a consequence only the odd multiples of the fundamental frequency are natural frequencies. The spectrum will then only consist in these odd harmonics giving a different sound.

The modes are given by:

$$\chi(x) = \cos\left((n+1/2)\pi\frac{x}{L}\right)$$

with only a quarter a wavelength in the pipe for the fundamental mode. This allows to make the instrument twice shorter for the same tone.

For other instruments not being based on the resonance of a one-dimensional system such as bells, drums, the property of having higher natural frequencies being the harmonics - multiples - of the fundamental frequency is usually lost. Tuning those to be such is an art but this is usually achieved for only a few. Hence these instruments usually exhibit less "pure tones".

This is however only true for western music, since in many oriental music the fundamental frequency itself is slightly changed when playing a given tone, slightly changing the length, the fingering or the tension of the string. This had become quite common also with electric guitar.