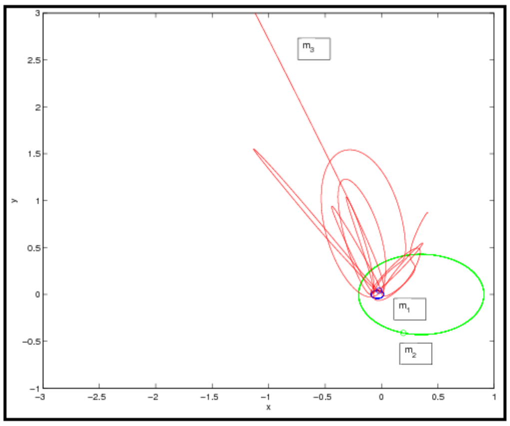
University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

School of Mathematical Sciences
Scoil na nEolaíochtaí Matamaitice
Mechanics and Special Relativity (ACM 10030)



Dr. Lennon Ó Náraigh

Lecture notes in Mechanics and Special Relativity, January 2011

Mechanics and Special Relativity (ACM10030)

- Subject: Applied and Computational Maths
- School: Mathematical Sciences
- Module coordinator: Prof. Adrian Ottewill
- Credits: 5
- Level: 1
- Semester: Second

This course develops the theory of planetary and satellite motion. It discusses the work of Kepler and Newton that described the elliptic orbits of planets around the earth and which can be applied to the elliptic motion of satellites around the earth. We examine the dynamics of spacecraft. Einstein's Special Theory of Relativity is then introduced. His two basic postulates of relativity are discussed and we show how space and time appear to two observers moving relative to each other. We derive, and discuss the meaning of, Einstein's famous formula $E = mc^2$.

What will I learn?

On completion of this module students should (be able to)

1. Explain the concepts of planetary and satellite orbits, Kepler's laws and how to boost an earth satellite from one orbit to another;
2. Solve orbit problems in mathematical term;
3. Describe Einstein's postulates and derive the results of special relativity on simultaneity, length contraction, time dilation and relative velocity;
4. Describe Einstein's postulates and derive the results of special relativity on simultaneity, length contraction, time dilation and relative velocity;
5. Solve problems in special relativity in mathematical terms.

Editions

First edition: January 2010

This edition: January 2011

Contents

Chapter 1

Introduction

1.1 Outline

Here is an executive summary of the aims of this course. If you cannot remember the more detailed outline that follows, at least keep the following in mind:

This course contains two topics:

1. In the first part, we show that Newton's law, Force = mass $\times$ acceleration implies Kepler's Laws, or rather, everything we know about the motion of planets around the sun.

2. In the second part, we show that Newton's laws need to be modified at speeds close to the speed of light. These modifcations enable us to solve lots of problems involving particle accelerators, cosmic rays, and radioactive decay.

Or, in more detail,

1. *Advanced Newtonian mechanics:* We will formulate Newton's equation in polar co-ordinates and discuss central forces. Using these two ideas, we will write down, and solve, the equations of motion for two bodies interacting through gravity. This enables us to prove Kepler's empirical laws of planetary motion from first principles. Using the same ideas, we will examine the dynamics of spacecraft, showing how they can be boosted from one orbit to another. The twin notions of *inertial frames* and *Galilean invariance* are at the heart of this study.

2. The principle of Galilean invariance says that the laws of physics are the same in all inertial frames. Using this idea, together with *Einstein's second postulate of relativity*, we will write down the basis of Einstein's special theory of relativity. These postulates allow us to derive the

Lorentz transformations[1]. As a consequence, we will discuss length contraction, time dilation, and relative velocity. We will also discuss the equation $E = mc^2$.

Some definitions:

- We call the problem of determining the trajectories of two bodies interacting via Newtonian gravity the *orbit problem*.
- We use the notation $\dot{x}$ and dx/dt interchangeably, to signify the derivative of the quantity x with respect to time. For the second derivative, we employ two dots.

1.2 Learning and Assessment

Learning:

- Thirty six classes, three per week.
- In some classes, we will solve problems together or look at supplementary topics.
- One of the main goals of a mathematics degree is *to solve problems autonomously*. This will be accomplished through homework exercises, independent study, and through independently practising the problems in the recommended textbooks.

Assessment:

- One end-of-semester exam, 60%.
- Two in-class exams, for a total of 20%.
- Four homework assignments, for a total 20%.

Policy on late submission of homework:

All uncertified late submissions will attract a penalty of 50%, that is, if the late assignment receives a grade X, only the grade $X/2$ will be entered.

Textbooks:

- Lecture notes will be put on the web. These are self-contained.
- However, here are some books for extra reading:

[1]These are referred to as the *Fitzgerald–Lorentz* transformations in another Dublin university, in honour of George Francis Fitzgerald.

– *An Introduction To Mechanics*, D. Kleppner, R. J. Kolenkow (one copy in library).

– For the second part of the course, look at the later chapters in *Sears and Zemansky's University Physics With Modern Physics*, H. D. Young, R. A. Freedman, T. R. Sandin, A. L. Ford (Many copies in library).

– You might also be interested in A. Einstein's book for 'the man (and woman) on the street', *Relativity: The Special and the General Theory*. This is out of copyright and I have placed it on Blackboard for downloading.

1.3 A modern perspective on classical mechanics

Before beginning the lecture course, let us discuss a problem in Newtonian mechanics that is the focus of continuing research. The problem also involves orbits.

The three-body problem: We shall show in this lecture course that an analytical solution in terms of integrals exists for a system of two particles interacting via gravity. No such solution exists for three particles, and the motion can become very complex and is not yet understood fully. Even the motion of such a system in the plane $\mathbb{R}^2$ is complicated, but can be tackled numerically. This is what we examine here. Newton's law of gravity states that the gravitational force on particle i due to particle j $\boldsymbol{F}_{ij}$ is given by

$$\boldsymbol{F}_{ij} = -Gm_i m_j \left(\frac{\boldsymbol{x}_i - \boldsymbol{x}_j}{|\boldsymbol{x}_i - \boldsymbol{x}_j|^3} \right) \tag{1.1}$$

In the schematic diagram (Fig. 1.1), particle (1) experiences a force from particle (0) and particle (2), to give a net force

$$\boldsymbol{F}_{12} + \boldsymbol{F}_{10} = -Gm_1 m_2 \left(\frac{\boldsymbol{x}_1 - \boldsymbol{x}_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|^3} \right) - Gm_1 m_0 \left(\frac{\boldsymbol{x}_1 - \boldsymbol{x}_0}{|\boldsymbol{x}_1 - \boldsymbol{x}_0|^3} \right). \tag{1.2}$$

Newton's law of motion, force=mass×acceleration gives $m_1\ddot{\boldsymbol{x}}_1 = \boldsymbol{F}_{12} + \boldsymbol{F}_{10}$, or,

$$m_1\ddot{\boldsymbol{x}}_1 = -Gm_1 m_2 \left(\frac{\boldsymbol{x}_1 - \boldsymbol{x}_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|^3} \right) - Gm_1 m_0 \left(\frac{\boldsymbol{x}_1 - \boldsymbol{x}_0}{|\boldsymbol{x}_1 - \boldsymbol{x}_0|^3} \right). \tag{1.3a}$$

The equations for the other two bodies can be written down immediately by permutations $0 \to 1$, $1 \to 2$, $2 \to 0$:

$$m_2\ddot{\boldsymbol{x}}_2 = -Gm_2 m_0 \left(\frac{\boldsymbol{x}_2 - \boldsymbol{x}_0}{|\boldsymbol{x}_2 - \boldsymbol{x}_0|^3} \right) - Gm_2 m_1 \left(\frac{\boldsymbol{x}_2 - \boldsymbol{x}_1}{|\boldsymbol{x}_2 - \boldsymbol{x}_1|^3} \right). \tag{1.3b}$$

$$m_0\ddot{\boldsymbol{x}}_0 = -Gm_0 m_1 \left(\frac{\boldsymbol{x}_0 - \boldsymbol{x}_1}{|\boldsymbol{x}_0 - \boldsymbol{x}_1|^3} \right) - Gm_0 m_2 \left(\frac{\boldsymbol{x}_0 - \boldsymbol{x}_2}{|\boldsymbol{x}_0 - \boldsymbol{x}_2|^3} \right). \tag{1.3c}$$

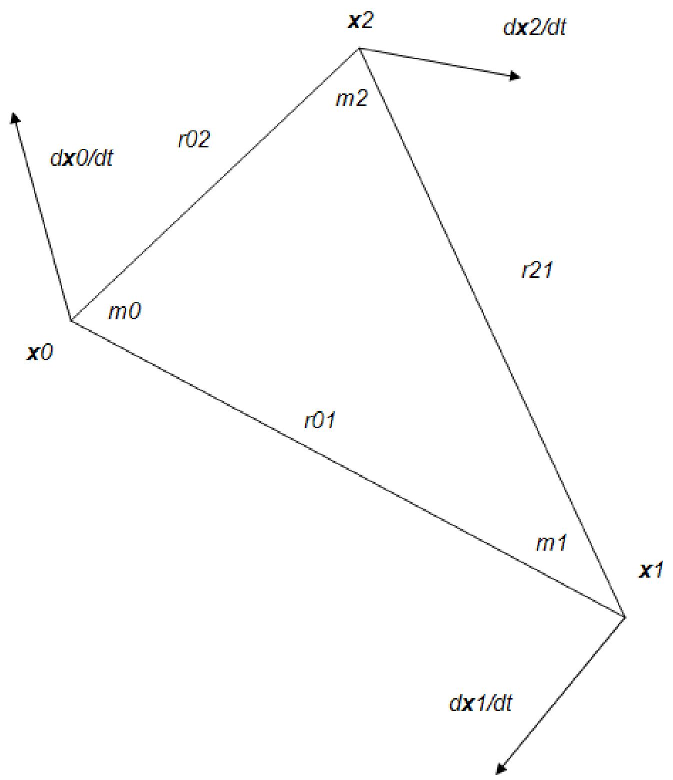


Figure 1.1: Schematic diagram of three-body problem. Position vectors and their time derivatives lie in the plane.

Radius vectors:

$$r_{ij} = r_{ji} = \sqrt{[(\boldsymbol{x}_i - \boldsymbol{x}_j) \cdot (\boldsymbol{x}_i - \boldsymbol{x}_j)]}.$$

Equations (1.3) can be solved numerically as an ODE problem by defining the vector $\boldsymbol{Y}$:

$$\boldsymbol{Y} = [x_0, y_0, \dot{x}_0, \dot{y}_0, x_1, y_1, \dot{x}_1, \dot{y}_1, x_2, y_2, \dot{x}_2, \dot{y}_2]^T, \tag{1.4}$$

such that

$$\frac{d\boldsymbol{Y}}{dt} = \begin{pmatrix} \dot{x}_0 \\ \dot{y}_0 \\ -Gm_1\left(x_0 - x_1\right)/r_{01}^3 - Gm_2\left(x_0 - x_2\right)/r_{02}^3 \\ -Gm_1\left(y_0 - y_1\right)/r_{01}^3 - Gm_2\left(y_0 - y_2\right)/r_{02}^3 \\ \dot{x}_1 \\ \dot{y}_1 \\ -Gm_2\left(x_1 - x_2\right)/r_{12}^3 - Gm_0\left(x_1 - x_0\right)/r_{10}^3 \\ -Gm_2\left(y_1 - y_2\right)/r_{12}^3 - Gm_0\left(y_1 - y_0\right)/r_{10}^3 \\ \dot{x}_2 \\ \dot{y}_2 \\ -Gm_0\left(x_2 - x_0\right)/r_{20}^3 - Gm_1\left(x_2 - x_1\right)/r_{21}^3 \\ -Gm_0\left(y_2 - y_0\right)/r_{20}^3 - Gm_1\left(y_2 - y_1\right)/r_{21}^3 \end{pmatrix}.$$

We perform a numerical simulation with the following initial conditions ($G = 1$):

$$m_0 = 1000,\ m_1 = 100,\ m_2 = 0.001, \tag{1.5a}$$

$$\boldsymbol{x}_0\left(t=0\right) = \left(0, 0\right),\ \boldsymbol{x}_1\left(t=0\right) = \left(1, 0\right),\ \boldsymbol{x}_2\left(t=0\right) = \tfrac{1}{2}\left(1, \sqrt{3}\right), \tag{1.5b}$$

$$\dot{\boldsymbol{x}}_0\left(t=0\right) = \left(0, -10\right),\ \dot{\boldsymbol{x}}_1\left(t=0\right) = \left(0, 10\right),\ \dot{\boldsymbol{x}}_2\left(t=0\right) = \left(1, -4\right). \tag{1.5c}$$

(See Fig. 1.2.)

- The particle orbits are shown in Fig. 1.3.
- An animation of the orbits is on the web.
- A *phase portrait* ($\dot{x}$ versus x) for m_1 is shown in Fig. 1.4.
- The conserved energy

  $$E = \tfrac{1}{2}\sum_{i=1}^{3} m_i \dot{\boldsymbol{x}}_i^2 - \frac{Gm_1 m_2}{r_{12}} - \frac{Gm_1 m_0}{r_{10}} - \frac{Gm_2 m_1}{r_{21}}$$

  is shown in Fig. 1.5.

The features of these graphs are discussed below:

- Trajectory: two heavy masses rotate around each other in closed or almost-closed orbits.
- 'Almost closed': phase portraits $\dot{x}$ versus x do not form closed loops the start and end-points of a quasi-period almost match up.

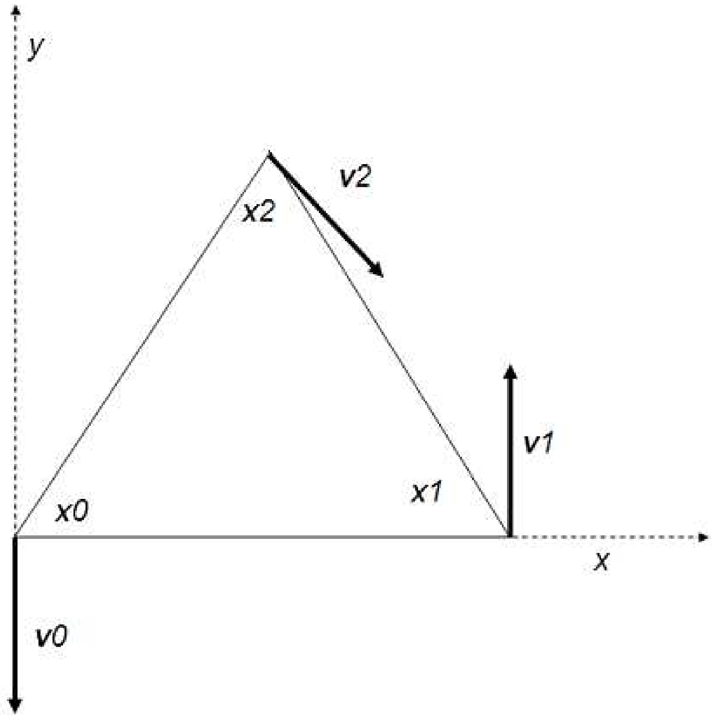


Figure 1.2: Initial conditions for the three-body problem.

- Third point-like mass executes complicated motion ('chaos') for a while but is ejected from the system eventually.
- Plots of *conserved quantities* such as the energy tell us if our numerical method is performing well. For example, our energy plot shows us that the energy is conserved to a high degree of accuracy.
- For a long-time ('secular') simulation, small errors such in the conserved energy can build up, leading to non-conserved energies and hence, unreliable results.
- A challenge in the numerical simulation of many-body problems is the derivation of integrators that manifestly conserve energy and other quantites. These are the so-called *symplectic* integrators.

Although this problem is *conservative*, there are some scenarios in which energy might be dissipated. This can dramatically change the character of the *equilibria* of the system. An equilibrium is a configuration $\{\boldsymbol{x}_i\}$ such that the forces and velocities vanish. For example, three identical masses located at the vertices of an equillateral triangle are an equilibrium of the three-body problem. Suppose therefore that the planets experience some kind of drag, which could be due to their interaction with interplanetary dust, or strong tidal friction:

$$\boldsymbol{F}_{D,i} \propto -\boldsymbol{v}_i,$$

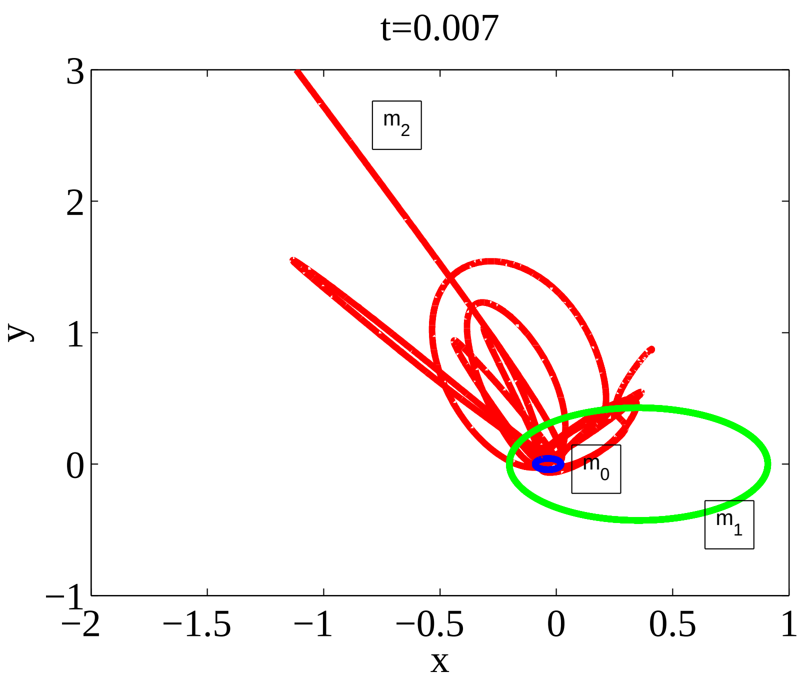
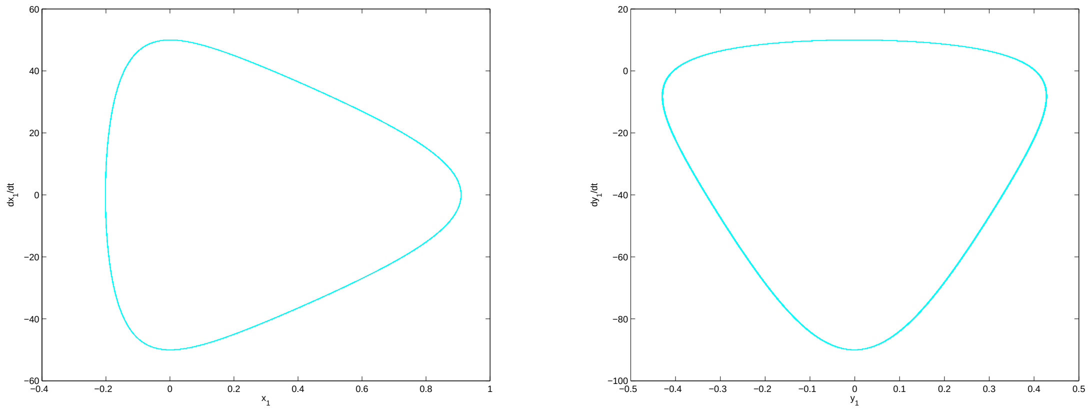


Figure 1.3: Orbits of three-body system

where i labels the objects. Then, given the nonlinearities in the problem, there is the possibility that the equilibria of the system become *strange attractors*. An example of a strange attractor, albeit in an entirely different setting, is the so-called *Lorenz attractor* in atmospheric sciences (Fig 1.6).



(a) (b)

Figure 1.4: Phase portrait of mass m_1

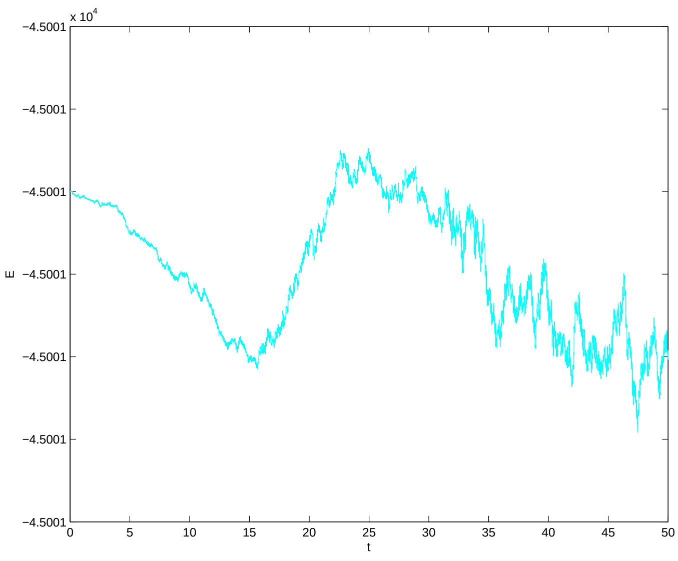


Figure 1.5: Energy conservation test

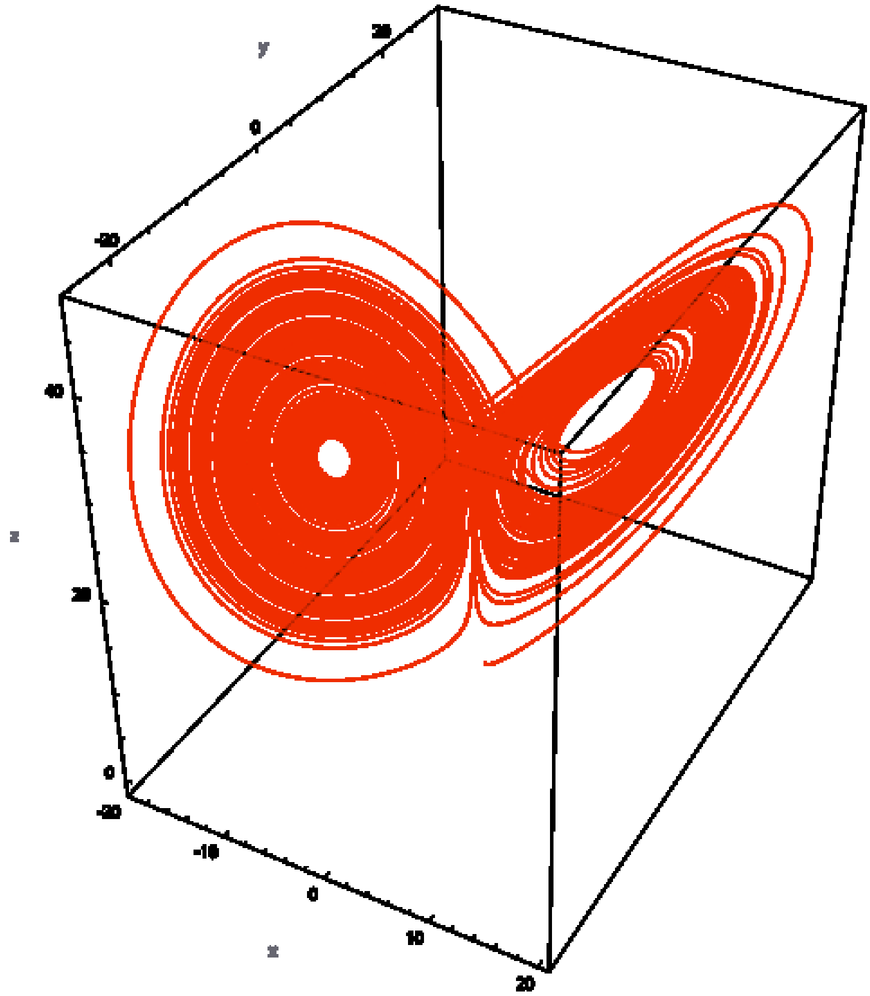


Figure 1.6: Strange attractor of the Lorenz system (three degrees of freedom)

Chapter 2

Vector operations

2.1 Summary

This is the first formal chapter of the module. Here, we introduce some background material that is vital for developing the mathematics of celestial mechanics. Central to this chapter are the notions of the *dot product* and the *cross product*.

2.2 The dot product

Consider two vectors with Cartesian components

$$\boldsymbol{a} = (a_1, a_2, a_3)\,, \qquad \boldsymbol{b} = (b_1, b_2, b_3)\,.$$

The *dot product* $\boldsymbol{a} \cdot \boldsymbol{b}$ is a real-number combination of these two vectors:

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 = \sum_{i=1}^{3} a_i b_i = \boldsymbol{a}^T \boldsymbol{b}.$$

The *length* of a vector is determined by the dot product of a vector with itself:

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

The dot product is *rotationally invariant* in the following sense: Let $\boldsymbol{R}$ be an *orthogonal matrix*:

$$\boldsymbol{R}^T \boldsymbol{R} = \mathbb{I}, \qquad \det(\boldsymbol{R}) = +1.$$

Then we define rotated vectors

$$\boldsymbol{a}' = \boldsymbol{R}\boldsymbol{a}, \qquad \boldsymbol{b}' = \boldsymbol{R}\boldsymbol{b},$$

The dot product is rotationally invariant because

$$\boldsymbol{a}' \cdot \boldsymbol{b}' = \boldsymbol{a} \cdot \boldsymbol{b}.$$

Proof:

$$\begin{aligned} \boldsymbol{a}' \cdot \boldsymbol{b}' &= (R\boldsymbol{a})^T (\boldsymbol{R}\boldsymbol{b}), \\ &= \boldsymbol{a}^T \boldsymbol{R}^T \boldsymbol{R}\boldsymbol{b}, \\ &= \boldsymbol{a}^T \mathbb{I} \boldsymbol{b}, \\ &= \boldsymbol{a}^T \boldsymbol{b}, \\ &= \boldsymbol{a} \cdot \boldsymbol{b}. \end{aligned}$$

A *scalar* is a real number that does not change under rotations. Therefore, it is appropriate to call the dot product the *scalar product* because it takes two vectors and forms a scalar. Because we have defined length in terms of the scalar product, it follows that length does not change under rotations, which is thankfully consistent with reality.

As a consequence of rotational invariance, we can prove the following claim.

Theorem: The dot product satisfies the following relationship:

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta,$$

where θ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$, measured in the sense of turning from $\boldsymbol{a}$ to $\boldsymbol{b}$ and chosen such that $0 \le \theta \le \pi$.

Proof: Because the dot product is rotationally invariant, we apply a rotation matrix $\boldsymbol{R}$ to our system of Cartesian axes such that the matrices $\boldsymbol{a}$ and $\boldsymbol{b}$ now lie in the x-y plane;

$$\boldsymbol{a}' = (a_1, a_2, 0), \qquad \boldsymbol{b}' = (b_1, b_2, 0).$$

The proof in this reduced frame is left as an exercise.

Note: The dot product, generalized to any dimension gives rise to a definition of *orthogonality*: vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are *orthogonal* if

$$\boldsymbol{a} \cdot \boldsymbol{b} = 0.$$

2.3 The vector product

Given vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we have seen how to form a scalar. We can also form a third vector from these two vectors, using the cross or vector product:

$$\begin{aligned} \boldsymbol{a} \times \boldsymbol{b} &= \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}, \qquad (2.1) \\ &= \hat{\boldsymbol{x}}\left(a_2 b_3 - a_3 b_2\right) - \hat{\boldsymbol{y}}\left(a_1 b_3 - a_3 b_1\right) + \hat{\boldsymbol{z}}\left(a_1 b_2 - a_2 b_1\right), \\ &= \hat{\boldsymbol{x}}\left(a_2 b_3 - a_3 b_2\right) + \hat{\boldsymbol{y}}\left(a_3 b_1 - a_1 b_3\right) + \hat{\boldsymbol{z}}\left(a_1 b_2 - a_2 b_1\right), . \end{aligned}$$

Properties of the vector or cross product:

1. Skew-symmetry: $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$,
2. Linearity: $(\lambda \boldsymbol{a}) \times \boldsymbol{b} = \boldsymbol{a} \times (\lambda \boldsymbol{b}) = \lambda\,(\boldsymbol{a} \times \boldsymbol{b})$, for $\lambda \in \mathbb{R}$.
3. Distributive: $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}$.

These results readily follow from the determinant definition. Result (1) is particularly weird. Note:

$$\begin{aligned} \boldsymbol{a} \times \boldsymbol{a} &= -\boldsymbol{a} \times \boldsymbol{a}, \qquad \text{Result (1)}, \\ 2\boldsymbol{a} \times \boldsymbol{a} &= 0, \\ \boldsymbol{a} \times \boldsymbol{a} &= 0. \end{aligned}$$

2.3.1 Numerical examples

1. Let

$$\boldsymbol{a} = \hat{\boldsymbol{x}} + 3\hat{\boldsymbol{y}} + \hat{\boldsymbol{z}}, \qquad \boldsymbol{b} = 2\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}} + 2\hat{\boldsymbol{z}}.$$

   Then

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ 1 & 3 & 1 \\ 2 & -1 & 2 \end{vmatrix} = 7\hat{\boldsymbol{x}} - 7\hat{\boldsymbol{z}}.$$

2. The so-called *orthonormal triad*

$$\begin{aligned} \hat{\boldsymbol{x}} &= (1,0,0), \\ \hat{\boldsymbol{y}} &= (0,1,0), \\ \hat{\boldsymbol{z}} &= (0,0,1) \end{aligned}$$

satisfies the relations

$$\begin{aligned}\hat{\boldsymbol{x}} \times \hat{\boldsymbol{y}} &= \hat{\boldsymbol{z}},\\ \hat{\boldsymbol{y}} \times \hat{\boldsymbol{z}} &= \hat{\boldsymbol{x}},\\ \hat{\boldsymbol{z}} \times \hat{\boldsymbol{x}} &= \hat{\boldsymbol{y}}.\end{aligned} \tag{2.2}$$

2.3.2 Geometrical treatment of cross product

So far, our treatment of the cross product has been in terms of a particular choice of Cartesian axes. However, the definition of the cross product is in fact independent of any choice of such axes. To demonstrate this, we re-construct the cross product.

Step 1: Finding the length of $\boldsymbol{a} \times \boldsymbol{b}$ Note that

$$\begin{aligned}|\boldsymbol{a} \times \boldsymbol{b}|^2 + (\boldsymbol{a} \cdot \boldsymbol{b})^2 &= (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2\\ &+ (a_1b_1 + a_2b_2 + a_3b_3)^2,\\ &= \left(a_1^2 + a_2^2 + a_3^2\right)\left(b_1^2 + b_2^2 + b_3^2\right),\\ &= |\boldsymbol{a}|^2|\boldsymbol{b}|^2.\end{aligned}$$

Hence,

$$\begin{aligned}|\boldsymbol{a} \times \boldsymbol{b}|^2 &= |\boldsymbol{a}|^2|\boldsymbol{b}|^2 - (\boldsymbol{a} \cdot \boldsymbol{b})^2,\\ &= |\boldsymbol{a}|^2|\boldsymbol{b}|^2\left(1 - \cos^2\theta\right),\\ &= |\boldsymbol{a}|^2|\boldsymbol{b}|^2 \sin^2\theta\end{aligned}$$

and

$$|\boldsymbol{a} \times \boldsymbol{b}| = |\boldsymbol{a}||\boldsymbol{b}|\sin\theta,$$

where $0 \le \theta \le \pi$, so $|\boldsymbol{a} \times \boldsymbol{b}| \ge 0$.

Step 2: Finding the direction of $\boldsymbol{a} \times \boldsymbol{b}$ Note that

$$\begin{aligned}\boldsymbol{a} \cdot (\boldsymbol{a} \times \boldsymbol{b}) &= a_1(a_2b_3 - a_3b_2) + a_2(a_3b_1 - a_1b_3) + a_3(a_1b_2 - a_2b_1),\\ &= 0.\end{aligned}$$

Similarly, $\boldsymbol{b} \cdot (\boldsymbol{a} \times \boldsymbol{b}) = 0$. Hence, $\boldsymbol{a} \times \boldsymbol{b}$ is a vector perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$. It remains to find the sense of $\boldsymbol{a} \times \boldsymbol{b}$. Indeed, this is arbitrary and must be fixed. We fix it such that we have a right-handed system, and such that the following rule-of-thumb is satisifed (Fig. 2.1).

Choosing a right-hand rule means that relations (2.2) are satisfied ($\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ form a 'right-handed'

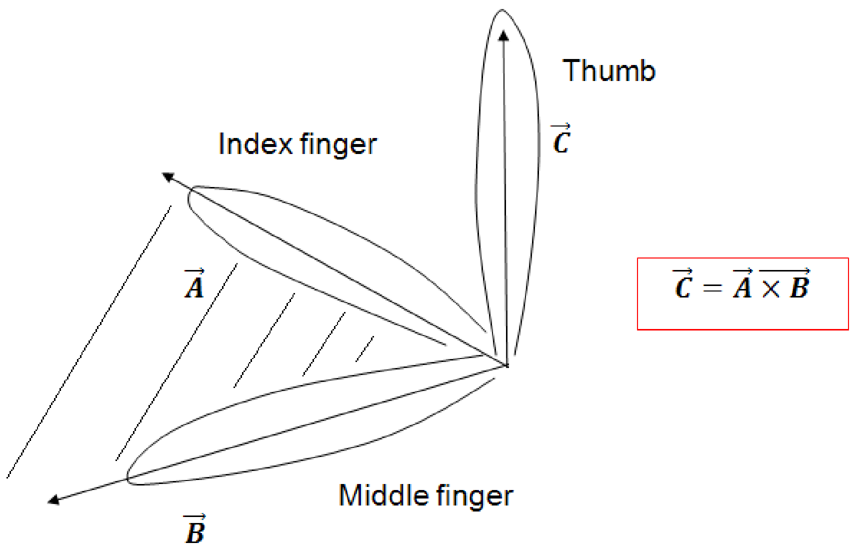


Figure 2.1: The right-hand rule.

system). This also corresponds to putting a plus sign in front of the determinant in the original definition of the cross product.

In summary, $\boldsymbol{a} \times \boldsymbol{b}$ is a vector of magnitude $|\boldsymbol{a}||\boldsymbol{b}| \sin\theta$, that is normal to both $\boldsymbol{a}$ and $\boldsymbol{b}$, and whose sense is determined by the right-hand rule.

The cross product as an area: Consider a parallelogram, whose two adjacent sides are made up of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ (Fig. 2.2). The area of the parallelogram is

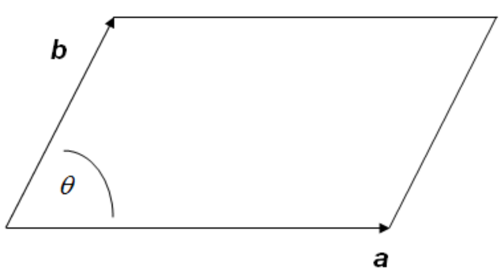


Figure 2.2: The cross product as an area

$$\begin{aligned} A &= (\text{base length}) \times (\text{perpendicular height})\,, \\ &= (\text{base length})\, |\boldsymbol{b}| \sin\theta, \\ &= |\boldsymbol{a}||\boldsymbol{b}| \sin\theta, \\ &= |\boldsymbol{a} \times \boldsymbol{b}|. \end{aligned}$$

The scalar triple product and volume: We can form a scalar from the three vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ by combining the operations just defined:

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})\,. \tag{2.3}$$

This is the so-called 'scalar triple product'.

Theorem: The vector triple product $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$ is identically equal to

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Proof: By brute force,

$$\begin{aligned} & (a_1\hat{\boldsymbol{x}} + a_2\hat{\boldsymbol{y}} + a_3\hat{\boldsymbol{z}}) \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \\ = \;& (a_1\hat{\boldsymbol{x}} + a_2\hat{\boldsymbol{y}} + a_3\hat{\boldsymbol{z}}) \cdot \left[(b_2c_3 - c_2b_3)\,\hat{\boldsymbol{x}} + (b_3c_1 - b_1c_3)\,\hat{\boldsymbol{y}} + (b_1c_2 - b_2c_1)\,\hat{\boldsymbol{z}}\right] \\ = \;& a_1\,(b_2c_3 - c_2b_3) + a_2\,(b_3c_1 - b_1c_3) + a_3\,(b_1c_2 - b_2c_1)\,, \end{aligned}$$

which is the determinant of the theorem.

Now consider a parallelepiped spanned by the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ (Fig. 2.3)

$$\begin{aligned} \text{Volume of parallelepiped} &= (\text{Perpendicular height}) \times (\text{Base area}) \\ &= (|\boldsymbol{a}| \cos\varphi) \times (|\boldsymbol{b}||\boldsymbol{c}| \sin\theta)\,, \\ &= (|\boldsymbol{a}| \cos\varphi)\,(|\boldsymbol{b} \times \boldsymbol{c}|)\,, \\ &= \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})\,. \end{aligned}$$

Corollary: Three nonzero vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are coplanar if and only if $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = 0$.

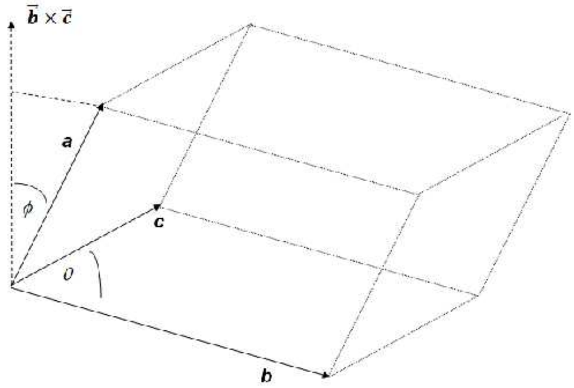


Figure 2.3: The scalar triple product as a volume

Proof: The volume of the parallelepiped spanned by the three vectors is zero iff the perpendicular height is zero, iff the three vectors are coplanar.

2.4 The vector triple product

Given three vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$, we can form yet another vector,

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})\,. \tag{2.4}$$

The brackets are important because the cross product is not associative, e.g.

$$\hat{\boldsymbol{x}} \times (\hat{\boldsymbol{x}} \times \hat{\boldsymbol{y}}) = \hat{\boldsymbol{x}} \times \hat{\boldsymbol{z}} = -\hat{\boldsymbol{y}},$$

but

$$(\hat{\boldsymbol{x}} \times \hat{\boldsymbol{x}}) \times \hat{\boldsymbol{y}} = 0 \times \hat{\boldsymbol{y}} = 0.$$

Theorem: The vector triple product satisfies

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}\,(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{c}\,(\boldsymbol{a} \cdot \boldsymbol{b})\,, \tag{2.5}$$

a result that can be recalled by the mnemonic 'BAC minus CAB'.

Proof: Without loss of generality, we prove the result in a frame wherein the x- and y-axes of our frame lie in the plane generated by $\boldsymbol{b}$ and $\boldsymbol{c}$. In fact, we may take

$$\begin{aligned} \boldsymbol{c} &= \hat{\boldsymbol{x}} c_1, \\ \boldsymbol{b} &= b_1 \hat{\boldsymbol{x}} + b_2 \hat{\boldsymbol{y}}, \end{aligned}$$

and

$$\boldsymbol{a} = a_1 \hat{\boldsymbol{x}} + a_2 \hat{\boldsymbol{y}} + a_3 \hat{\boldsymbol{z}}.$$

The result then follows by a brute-force calculation of the LHS and the RHS of Eq. (2.5).

Chapter 3

Inertial frames of reference

3.1 Summary

In this chapter, we follow the vector formalism introduced in Ch. 2 (Vector operations) and carry out the following programme of work:

- Define inertial frames;
- Introduce Galilean invariance;
- Discuss Galilean transformations;
- Recall Newton's laws.

3.2 Inertial frames of reference

A measurement of position, velocity or acceleration must be made with respect to a *frame of reference*. Loosely speaking, a frame of reference consists of an observer S, equipped with a metre stick and a stopwatch. *Examples:*

- Consider a person, S, standing on the platform of a train station. He measures displacements from where he is standing, and he uses the origin of a Cartesian set of axes to measure the (x, y, z) position of any object. These axes do not move relative to the platform. Any object at rest in the station has zero velocity relative to his frame of reference.
- A second person, S', is sitting on a train moving at velocity $100\,\mathrm{km/hr}$ through the station. He has his own set of axes (x', y', z') that are fixed to the train. Any object at rest in the train has zero velocity relative to the frame of reference of person S'.

- Therefore, a third person walking through the train at $5\,\mathrm{km/hr}$ in the direction of the train's motion has a velocity of $5\,\mathrm{km/hr}$ relative to person S' and a velocity of $105\,\mathrm{km/hr}$ relative to person S.

The last point is a formalized statement of something intuitive: that velocities add. However, it is WRONG. In the second part of the course, we shall find that velocities do not add in this way. This non-additivity happens in the so-called *relativistic limit*, where the velocity in question v approaches the speed of light c. When $v \ll c$, the additivity holds approximately. This is the subject of the first half of the course.

3.3 Galilean invariance

In the last section, the examples we described involved observers at rest or in uniform motion with respect to one another. This leads to a definition of *the class of inertial frames*:

> Describing frames of reference as at rest or in uniform motion relative to one another defines an equivalence relationship between frames, which in turn defines an equivalence class. A representative from this class is called an *intertial frame*.

No inertial frame can accelerate with respect to any other inertial frame: any frame that accelerates with respect to another frame is outside of this class. It is in inertial frames that we wish to formulate the laws of physics.

> The principle of Galilean invariance states that the form of all physical laws must be the same in all inertial frames of reference.

If observer S on the platform and observer S' on the train in constant motion conduct an experiment, they must obtain the same answer. Thus, if observer S' were blindfolded (but still able to do and to interpret experiments), he would have no way of knowing if his train were in motion relative to the platform in the station.

The results of an experiment performed in a frame of reference S, by Galilean invariance, must be transformable into the results of an identical experiment, performed in a frame S', moving at a velocity $\boldsymbol{V}$ relative to the first frame. This is done through a *Galilean transformation*. This is done as follows (Fig. 3.1):

- Position vector of some object relative to frame S: $\boldsymbol{x}$;

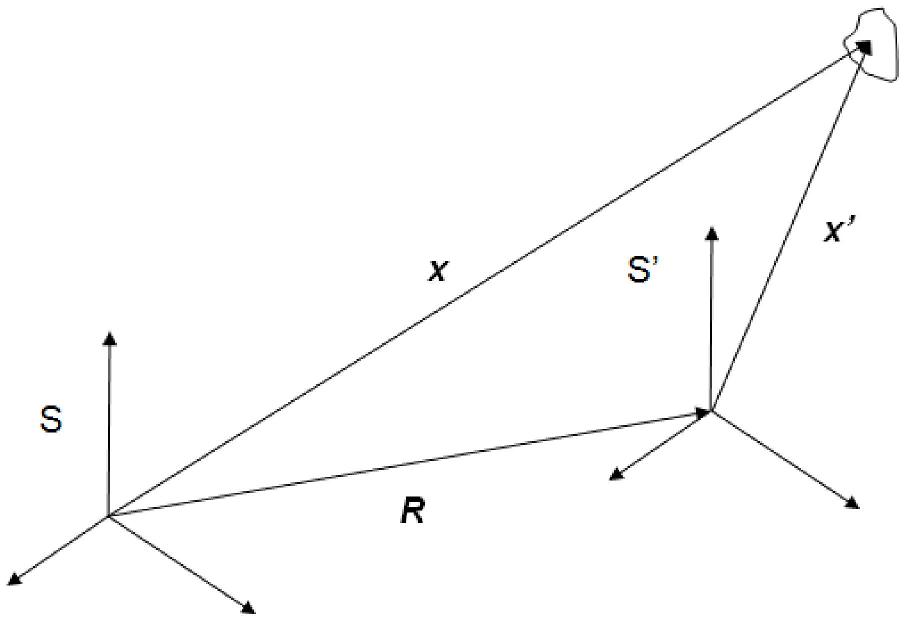


Figure 3.1: Relationship between position vectors in two distinct frames of reference.

- Position vector of the same object relative to frame S': $\boldsymbol{x}'$;
- Position vector of the origin of S' relative to S: $\boldsymbol{R}$.

Vector addition:

$$\boldsymbol{x} = \boldsymbol{R} + \boldsymbol{x}'. \tag{3.1}$$

Take the derivatives. Assume

$d/dt = d/dt'$; time is absolute.

Then,

$$\frac{d\boldsymbol{x}}{dt} = \frac{d\boldsymbol{R}}{dt} + \frac{d\boldsymbol{x}'}{dt}. \tag{3.2}$$

Introduce new notation:

$$\boldsymbol{v}_S = \boldsymbol{V}_{S'S} + \boldsymbol{v}_{S'} \tag{3.3}$$

In words,

Velocity of object relative to frame S =

Velocity of frame S' relative to frame S + Velocity of object relative to frame S'. (3.4)

Let us introduce some final notation:

$$\frac{d\boldsymbol{R}}{dt} = \boldsymbol{V}_{S'S} := \boldsymbol{V}. \tag{3.5}$$

We now specialise to inertial frames and demand that S' be in uniform motion with respect to S. Thus,

$$\frac{d\boldsymbol{V}}{dt} = 0 \text{ if and only if } \boldsymbol{V} = \text{Const.} \tag{3.6}$$

This has two consequences:

1. The acceleration is the same in both frames:

$$\frac{d\boldsymbol{v}_S}{dt} = \frac{d\boldsymbol{v}_{S'}}{dt}. \tag{3.7}$$

2. The two frames are linked via a transformation that involves $\boldsymbol{V}$. For, let us solve $\boldsymbol{V} = d\boldsymbol{R}/dt$ (Eq. (3.5)). If both frames coincide at time $t = 0$, then $\boldsymbol{R} = \boldsymbol{V}t$. Using the vector addition law $\boldsymbol{x} = \boldsymbol{R} + \boldsymbol{x}'$ (Eq. (3.1)), obtain

$$\boldsymbol{x} = \boldsymbol{V}t + \boldsymbol{x}'. \tag{3.8}$$

Let us apply this formal derivation to the train-and-platform scenario. Then, $\boldsymbol{V} = \hat{\boldsymbol{x}}V$, and

$$\begin{aligned} x &= Vt + x', \\ y &= y', \\ z &= z', \\ t &= t'. \end{aligned} \tag{3.9}$$

The velocity transformation is, as before:

$$\begin{aligned} v_x &= V + v'_x, \\ v_y &= v'_y, \\ v_z &= v'_z. \end{aligned} \tag{3.10}$$

The object we choose to specify with these coordinate systems is the third passenger on board the train , walking at a velocity $5\,\mathrm{km/hr}$ in the direction of motion of the train, and relative to observer S' (thus, this observer corresponds to the 'blob' in Fig. 3.1). Thus,

- $v'_x = 5\,\mathrm{km/hr}$,
- $V = 100\,\mathrm{km/hr}$,
- Using the Galilean transformation, $v_x = 105\,\mathrm{km/hr}$.

3.4 Newton's laws of motion

Newton's laws of motion are valid only in an intertial frame of reference:

1. A body will continue in its state of rest or of uniform motion in a straight line unless an external force is applied to it.
2. The vector sum of forces acting on a body is proportional to its rate of change of momentum. For a body of constant mass this becomes
$$m\frac{d^2\boldsymbol{x}}{dt^2} = \boldsymbol{F}. \tag{3.11}$$
3. If a body A exerts a force on a body B then B exerts an equal and opposite force on A.

3.4.1 Laws (1) and (2) do not hold in all frames of reference.

We can observe an object to accelerate not because of any forces acting upon it but because of the acceleration of our own reference frame. Consider a car accelerating on a straight road relative to the earth. An observer in the car who is equipped with a pendulum will see that pendulum deflected in a direction opposite to the acceleration. This is not due to any force on the pendulum itself, but rather due to a force acting on the entire frame of reference. Some more examples of non-inertial frames:

- The earth is not an inertial frame. Distant stars rotate about the earth but not because of any forces acting upon them, merely because of the earth's rotation. An observer on the equator undergoes centripetal acceleration with $\omega = 2\pi\,\mathrm{rad/day}$ and $r = 6,400\,\mathrm{km}$, giving
$$a = \omega^2 r = 0.03\,\mathrm{m}/s^2 = 3.45\times 10^{-3} g.$$

Since this is small compared to g, in practice we take the earth to be an inertial reference frame.

A demonstration of the non-inertial nature of a reference frame attached to the earth's surface is the pendulum of Foucault: the direction along which a simple pendulum swings rotates with time because of Earth's daily rotation.

- Even if the earth did not turn on its own axis we would still not be in an inertial reference frame since the earth goes around the sun with $\omega = 2\pi\,\mathrm{rad/year}$ and $r = 1.5\times 10^8\,\mathrm{km}$, hence

$$a = \omega^2 r = 5.94\times 10^{-3}\,\mathrm{m/s^2} = 6.05\times 10^{-4} g$$

- The sun also rotates around the centre of the Milky Way galaxy which, in turn, accelerates through the Local Group of galaxies and so on.

So, we imagine a hypothetical distant observer in deep space on whom no forces act. Such an observer will only see an object accelerating if it feels a resultant force. Any reference frame moving with constant velocity relative to this observer is also in an inertial frame.

3.4.2 Galilean invariance again

We have seen that the acceleration $\ddot{\boldsymbol{x}}$ is left invariant by a change of inertial frame under the transformations (3.8). Since the laws of physics are the same in all inertial frames, the force must also be left invariant by this transformation. Thus, $\boldsymbol{F}$ is not arbitrary in Eq. (3.11). We can check that central forces

$$\boldsymbol{F}_{ij} = -\lambda\left(\frac{\boldsymbol{x}_i - \boldsymbol{x}_j}{|\boldsymbol{x}_i - \boldsymbol{x}_j|^p}\right), \tag{3.12}$$

are Galilean invariant, provided λ is a constant scalar. Here $\boldsymbol{F}_{ij}$ means the force on particle i due to particle j. Under a Galilean transformation,

$$\boldsymbol{x}'_i = \boldsymbol{x}_i - \boldsymbol{V}t,$$
$$\boldsymbol{x}'_j = \boldsymbol{x}_j - \boldsymbol{V}t,$$

hence,

$$\boldsymbol{x}'_i - \boldsymbol{x}'_j = \boldsymbol{x}_i - \boldsymbol{x}_j,$$

and $\boldsymbol{F}'_{ij} = \boldsymbol{F}_{ij}$.

3.4.3 Laws (1)–(3) imply the principle of conservation of momentum.

Consider a system of two particles specified by an inertial frame of reference. The total momentum of the system is

$$\boldsymbol{P} = m_1\dot{\boldsymbol{x}}_1 + m_2\dot{\boldsymbol{x}}_2.$$

Hence,

$$\frac{d\boldsymbol{P}}{dt} = m_1\frac{d\dot{\boldsymbol{x}}_1}{dt} + m_2\frac{d\dot{\boldsymbol{x}}_2}{dt}.$$

Using the Second Law,

$$\frac{d\boldsymbol{P}}{dt} = \boldsymbol{F}_{12} + \boldsymbol{F}_{21} + \boldsymbol{F}_{1,\mathrm{ext}} + \boldsymbol{F}_{2,\mathrm{ext}},$$

where ‘ext’ denotes external forces. If these are zero (a ‘closed system’), then

$$\frac{d\boldsymbol{P}}{dt} = \boldsymbol{F}_{12} + \boldsymbol{F}_{21},$$

and by the Third Law ,this vector sum is zero. This proof is readily generalized to collections of arbitrary numbers of particles, and hence, to extended bodies.

3.5 Final remarks

Einstein in his theory of special relativity also stated that all laws of physics have the same form in all inertial reference frames but that the Galilean transformations (which are based on the notion of absolute time) are wrong. They are approximately correct at speeds much slower than that of light, $c = 3 \times 10^8\,\mathrm{m/s}$ and break down completely as speeds approach c. We will examine this theory in detail during the second half of the course.

For the moment we restrict ourselves to Newtonian mechanics, where the Galilean transformation is correct.

Chapter 4

Worked examples: Vectors and Galilean velocity addition

In the following questions, you may use the equations of motion for trajectory motion in a uniform gravitational field g, in the absence of air friction:

$$x = x_0 + ut, \tag{4.1}$$

$$y = y_0 + vt - \tfrac{1}{2}gt^2. \tag{4.2}$$

where (x_0, y_0) is the initial location of the particle relative to a given inertial frame and (u, v) is the initial velocity.

1. Consider the matrix

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

and the vector $\boldsymbol{x} = (x, y)^T$.

(a) Compute

$$\boldsymbol{e}_1' = R\begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\boldsymbol{e}_2' = R\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

(b) Suppose that $0 \le \theta \le \pi/2$. Draw these new vectors relative to the old frame $\boldsymbol{e}_1 = (1, 0)^T$, $\boldsymbol{e}_2 = (0, 1)^T$. Demonstrate in words, and using your sketch, that R, when applied to the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, rotates those axes, and so generates a new frame of reference.

(c) Compute vectors $\boldsymbol{x}_1' = R\left(x_1, y_1\right)^T$ and $\boldsymbol{x}_2 = R\left(x_2, y_2\right)^T$. Show that $\boldsymbol{x}_1' \cdot \boldsymbol{x}_2' = \boldsymbol{x}_1 \cdot \boldsymbol{x}_2$.

(d) Given a vector $\boldsymbol{x} = \left(x_1, y_1\right)^T$, through what angle must the vector be rotated such that $\boldsymbol{x}' = (x_1', 0)$?.

(e) Show that $R^{-1} = R^T$, and that $\det R = +1$.

2. Given the dot product definition $(x_1, y_1) \cdot (x_2, y_2) = x_1 x_2 + y_1 y_2$, show that

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = |\boldsymbol{x}_1||\boldsymbol{x}_2| \cos\theta,$$

where θ is the angle between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, in the sense of going from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$, and such that $0 \leq \theta \leq \pi$. Hint: Use the Law of Cosines.

3. Two piers A and B are located on a river: B is at a distance H downstream of A (Fig. 4.1). Two friends must make round trips from pier A to pier B and return. One rows a boat at a constant speed v relative to the water; the other walks on the shore at the same constant speed v. The velocity of the river is $v_{\text{current}} < v$ in the direction from A to B. How much time does it take each person to do the round trip?

4. Consider a rainstorm wherein the rain is falling at right angles to the ground.

(a) Discuss what determines the best angle (relative to the ground) which to hold the umbrella? Can you find a formula for this angle?

(b) Suppose the person wants to travel a fixed distance at constant velocity. Assume, rather unrealistically, that a person can be modelled as a two-dimensional sheet. Is there an optimal speed of travel to minimize how wet he gets?

Hint: The concept of *flux* might be helpful here. The flux $\mathrm{d}\Phi$ of rain through a surface $\mathrm{d}S$ is the amount of rain that flows through the surface per unit time per unit area:

$$\mathrm{d}\Phi = \boldsymbol{v} \cdot \hat{\boldsymbol{n}} \, \mathrm{d}S,$$

where $\hat{\boldsymbol{n}}$ is normal to the surface element $\mathrm{d}S$, and where $\boldsymbol{v}$ is the velocity viewed in the appropriate frame of reference.

(c) How does the formula in (a) change if the rain is not falling at right angles to the ground?

5. A military helicopter on a training mission is flying horizontally at a speed V relative to the ground. It accidentally drops a bomb at an elevation H.

(a) How much time is required for the bomb to reach the earth?

(b) How far does it travel horizontally while falling?

(c) Find the horizontal and vertical components of its velocity just before it strikes the earth?

(d) If the velocity of the helicopter remains constant, where is the helicopter when the bomb hits the ground?

Answers involve V, H, and g only.

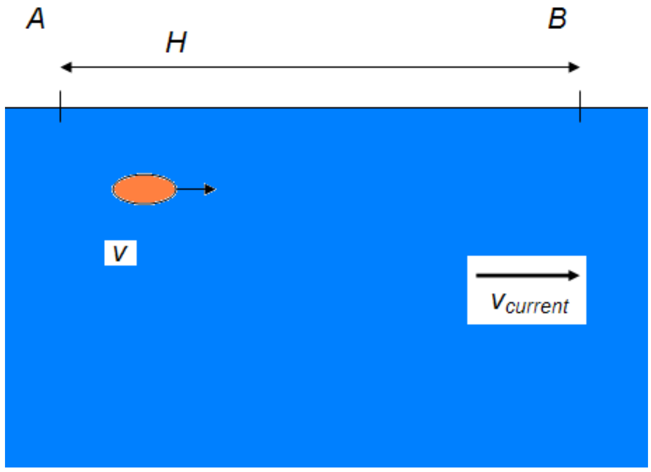


Figure 4.1: Problem 3: The boatrace

1. (a) Compute $\boldsymbol{e}_1' = R\,(1,0)^T$ and $\boldsymbol{e}_2' = R\,(0,1)^T$.

$$
\begin{aligned}
\boldsymbol{e}_1' &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \\
\boldsymbol{e}_2' &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}
\end{aligned}
$$

(b) Suppose that $0 \le \theta \le \pi/2$. Draw these new vectors relative to the old frame $\boldsymbol{e}_1 = (1,0)^T$, $\boldsymbol{e}_2 = (0,1)^T$. Demonstrate in words, and using your sketch, that R, when applied to the vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, rotates those axes, and so generates a new frame of reference.

Refer to the sketch (Fig. 4.2). Dot product: $\boldsymbol{e}_1' \cdot \boldsymbol{e}_2' = -\cos\theta\sin\theta + \sin\theta\cos\theta = 0$, hence $\boldsymbol{e}_1'$ and $\boldsymbol{e}_2'$ are orthogonal. They clearly have unit norm too, so they form a new basis of vectors. From the sketch, $\boldsymbol{e}_1'$ is the vector $\hat{\boldsymbol{x}}$ rotated through θ, and $\boldsymbol{e}_2'$ is the vector $\hat{\boldsymbol{y}}$ rotated through the same angle. Therefore, $(\boldsymbol{e}_1', \boldsymbol{e}_2')$ represent a rotated coordinate system.

(c) Compute vectors $\boldsymbol{x}_1' = R\,(x_1, y_1)^T$ and $\boldsymbol{x}_2 = R\,(x_2, y_2)^T$. Show that $\boldsymbol{x}_1' \cdot \boldsymbol{x}_2' = \boldsymbol{x}_1 \cdot \boldsymbol{x}_2$.

$$
\begin{aligned}
\boldsymbol{x}_1' = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} &= \begin{pmatrix} \cos\theta x_1 - \sin\theta y_1 \\ \sin\theta x_1 + \cos\theta y_1 \end{pmatrix} \\
\boldsymbol{x}_2' &= \begin{pmatrix} \cos\theta x_2 - \sin\theta y_2 \\ \sin\theta x_2 + \cos\theta y_2 \end{pmatrix}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\boldsymbol{x}_1' \cdot \boldsymbol{x}_2' = {} & (\cos\theta x_1 - \sin\theta y_1)(\cos\theta x_2 - \sin\theta y_2) \\
& + (\sin\theta x_1 + \cos\theta y_1)(\sin\theta x_2 + \cos\theta y_2).
\end{aligned}
$$

Expanding,

$$
\begin{aligned}
\boldsymbol{x}_1' \cdot \boldsymbol{x}_2' = {} & \cos^2\theta x_1 x_2 + \sin^2\theta y_1 y_2 - \cos\theta\sin\theta x_1 y_2 - \sin\theta\cos\theta x_2 y_1 \\
& + \sin^2\theta x_1 x_2 + \cos^2\theta y_1 y_2 + \sin\theta\cos\theta x_1 y_2 + \cos\theta\sin\theta y_1 x_2.
\end{aligned}
$$

Effecting the cancellations and using $\cos^2\theta + \sin^2\theta = 1$,

$$
\boldsymbol{x}_1' \cdot \boldsymbol{x}_2' = x_1 x_2 + y_1 y_2 = \boldsymbol{x}_1 \cdot \boldsymbol{x}_2.
$$

(d) Given a vector $\boldsymbol{x} = (x_1, y_1)^T$, through what angle must the vector be rotated such that $\boldsymbol{x}' = (x_1', 0)$?. From (c),

$$\boldsymbol{x}_1' = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \cos\theta x_1 - \sin\theta y_1 \\ \sin\theta x_1 + \cos\theta y_1 \end{pmatrix},$$

and we require that $\sin\theta x_1 + \cos\theta y_1 = 0$, hence $\tan\theta = -y_1/x_1$.

(e) Show that $R^{-1} = R^T$, and that $\det R = +1$.

$$\begin{aligned} R &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \\ R^{-1} &= \frac{1}{\cos^2\theta + \sin^2\theta} \begin{pmatrix} \cos\theta & +\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \\ &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = R^T. \end{aligned}$$

From the definition, it follows that $\det R = 1$. Further, since $R^{-1} = R^T$, it follows that $\mathbb{I} = RR^T$, hence

$$1 = \det\mathbb{I} = \det\left(RR^T\right) = \det R \det R^T = \det R \det R \implies \det R = \pm 1.$$

Orthogonal matrices whose determinants are positive are said to live in the group $SO(2)$, the group of rotations. Orthogonal matrices with negative determinants correspond to other length-preserving transformations that are not rotations, such as reflections.

2. Given the dot product definition $(x_1, y_1) \cdot (x_2, y_2) = x_1 x_2 + y_1 y_2$, show that

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = |\boldsymbol{x}_1||\boldsymbol{x}_2|\cos\theta,$$

where θ is the angle between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, in the sense of going from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$, and such that $0 \leq \theta \leq \pi$.

Since the dot product is rotation-invariant (Q1), we can rotate our coordinate system such that the vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ live in the plane described. Now refer to Fig. 4.3. Using the Law of Cosines,

$$L^2 = |\boldsymbol{x}_1|^2 + |\boldsymbol{x}_2|^2 - 2|\boldsymbol{x}_1||\boldsymbol{x}_2|\cos\theta.$$

But

$$L^2 = |\overrightarrow{P_2P_1}|^2 = |\boldsymbol{x}_1 - \boldsymbol{x}_2|^2 = (\boldsymbol{x}_1 - \boldsymbol{x}_2)^2 = |\boldsymbol{x}_1|^2 + |\boldsymbol{x}_2|^2 - 2\boldsymbol{x}_1 \cdot \boldsymbol{x}_2.$$

Equating both expressions for L^2 gives

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = |\boldsymbol{x}_1||\boldsymbol{x}_2| \cos\theta,$$

as required.

3. Consider the walker first. The time to go from A to B is $T_{AB} = H/v$. Thus, the round-trip time is $T_{\text{walker}} = 2H/v$. Next, consider the rower. On the first leg of the trip, the velocity of the rower relative to the bank is $v + v_{\text{current}}$, and thus $T_{AB} = H/\left(v + v_{\text{current}}\right)$. On the way back, the velocity of the rower relative to the bank is $v - v_{\text{current}}$, and hence, $T_{BA} = H/\left(v - v_{\text{current}}\right)$. Thus,
$$T_{\text{rower}} = T_{AB} + T_{BA} = \frac{H}{v + v_{\text{current}}} + \frac{H}{v - v_{\text{current}}} = \frac{2H}{v}\frac{v^2}{v^2 - v_{\text{current}}^2}.$$
Hence,
$$\frac{T_{\text{rower}}}{T_{\text{walker}}} = \frac{v^2}{v^2 - v_{\text{current}}^2} = \frac{1}{1 - \left(v_{\text{current}}/v\right)^2} \geq 1.$$
Hence, it takes the rower longer to complete the round trip.

4. (a) Refer to Fig. 4.4 (a). In the FOR of the person, the velocity of the person is zero by definition, and the velocity of the rain is $\boldsymbol{v}_R = -v\hat{\boldsymbol{x}} - w\hat{\boldsymbol{z}}$. In this frame, the angle between the rain and the horizontal is φ, and

$$|\hat{x} \cdot \boldsymbol{v}_R| = |\boldsymbol{v}_R| \cos\varphi = v.$$

Hence,
$$\cos\varphi = \frac{v}{|\boldsymbol{v}_R|} = \frac{v}{\sqrt{v^2 + w^2}}.$$

(b) In the FOR of the person, the velocity of the person is zero by definition, and the velocity of the rain is $\boldsymbol{v}_R = -v\hat{\boldsymbol{x}} - w\hat{\boldsymbol{z}}$. In this frame, the volume or rain received by the person per unit time is clearly (Fig. (b))

$$\mathrm{d}V = |\boldsymbol{v}_R| A \mathrm{d}t \cos\varphi,$$

where A is the surface area of the front of the person, and φ is the angle between $\boldsymbol{v}_R$ and the ground. But $|\boldsymbol{v}_R| \cos\varphi = v = |\boldsymbol{v}_R \cdot \hat{\boldsymbol{x}}|$, hence

$$\frac{dV}{dt} = vA.$$

Assuming a constant velocity v, $V = vAt$. Now consider a journey of length L from A to B at a constant speed v (w.r.t. the ground). The time of this journey is $T = L/v$.

Hence,

$$V = vA\,(L/v) = LA,$$

independent of v. Hence, all walking speeds are equally bad. Perhaps the real answer is therefore to move to Australia.

(c) Now, in the FOR of the ground, the rain has velocity $-w\,(\cos\alpha, \sin\alpha)$ and, using velocity addition,

$$\boldsymbol{v}_R = -v\hat{\boldsymbol{x}} - w\,(\cos\alpha, \sin\alpha)\,.$$

where the rain makes an angle $\pi/2 - \alpha$ with the ground, in the FOR of the ground itself. The analysis is the same as in the previous case, if we let $v \to v + w\cos\alpha$, and $w \to w\sin\alpha$, hence

$$\tan\varphi = \frac{w\sin\alpha}{v + w\cos\alpha}.$$

5. (a) How much time is required for the bomb to reach the earth?

   We do the calculation in the FOR of the earth. The initial speed of bomb is $\boldsymbol{v}_0 = (V, 0)$. Its initial location in our chosen FOR is $(x_0, y_0) = (0, H)$. Using the trajectory formulas,

$$x = Vt, \qquad y = H - \tfrac{1}{2}gt^2. \tag{4.3}$$

   The final time is when $y = 0$, i.e. $t_\mathrm{f} = \sqrt{2H/g}$.

   (b) How far does it travel horizontally while falling? Using $x = Vt_\mathrm{f}$, obtain $x = V\sqrt{2H/g} = \sqrt{2HV^2/g}$.

   (c) Find the horizontal and vertical components of its velocity just before it strikes the earth. Using Eq. (4.3),

$$\dot{x} = V, \qquad \dot{y} = -gt,$$

   obtain

$$\boldsymbol{v}\,(t_\mathrm{f}) = (V, -gt_\mathrm{f}) = \left(V, -g\sqrt{2H/g}\right) = \left(V, -\sqrt{2Hg}\right).$$

   (d) If the velocity of the helicopter remains constant, where is the helicopter when the bomb hits the ground?

   Using $dx_\mathrm{hel}/dt = V$, obtain $x_\mathrm{hel}\,(t_\mathrm{f}) = Vt_\mathrm{f}$, which is the same location as the bomb. Therefore, H must be large in order for the helicopter to avoid the impact of the blast.

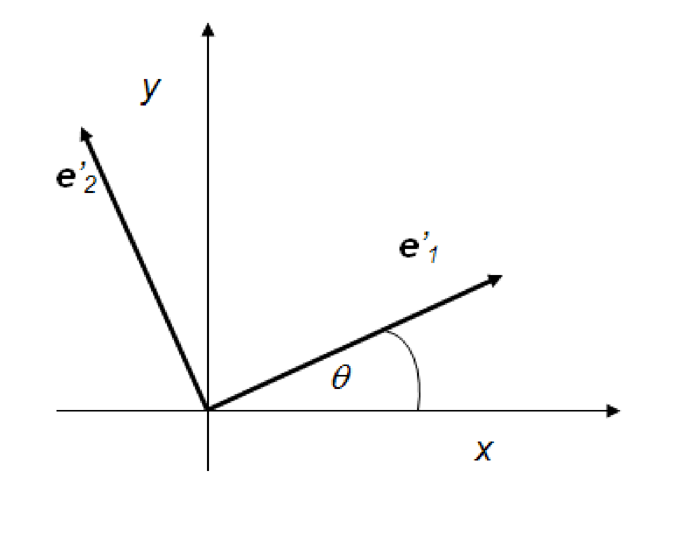


Figure 4.2: Problem 1: Rotation matrices

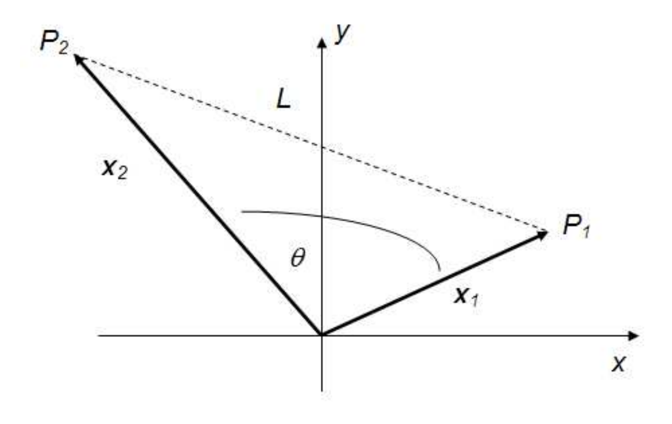


Figure 4.3: Problem 2: The dot product in $\mathbb{R}^2$

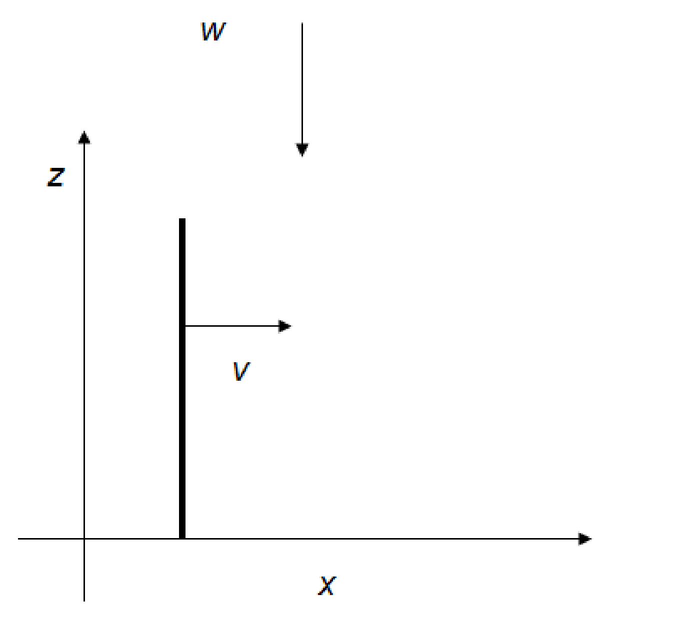
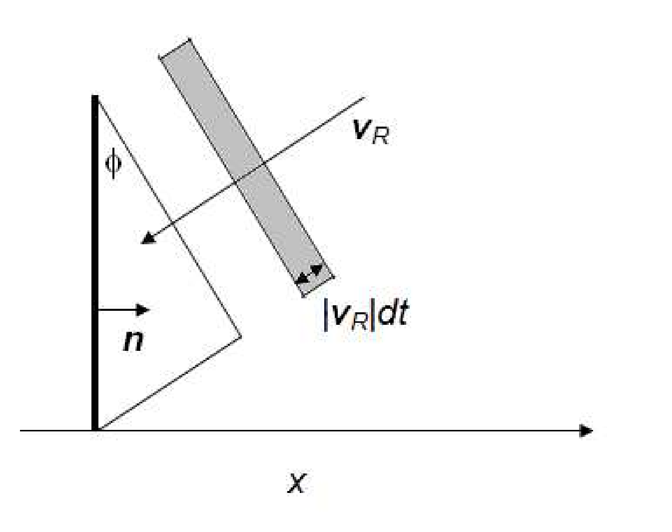


Figure 4.4: Problem 4: The rainstorm



Chapter 5

Conservative forces in one dimension

5.1 Overview

We look at some general principles of one-dimensional, single-particle motion. This will prepare us for problems in two dimensions. Indeed, the approach to the orbit problem involves a series of clever substitutions to reduce the problem to a one-dimensional one. We assume that we are in an inertial frame so that the equation of motion is

$$m\frac{d^2x}{dt^2} = F. \tag{5.1}$$

Against this backdrop, we consider the following concepts:

- Conservative forces;
- Energy conservation;
- Small oscillations;
- Stability;
- Solution by quadrature.

5.2 Conservative forces

In one dimension, a conservative force is a force that depends only on position.

In particular, this means that the force does not depend on time or velocity. Thus, the drag force

$$F_D = -k|v|^n v$$

that you learn about in ACM10020 is nonconservative. Now, associated with the conservative force $F(x)$ is the *potential* $\mathcal{U}$,

$$\mathcal{U}(x) = -\int_a^x F(s)\,\mathrm{d}s. \tag{5.2}$$

Note that the potential $\mathcal{U}$ is defined only up to an arbitrary constant, which depends on the lower limit of integration in Eq. (5.2). This is irrelevant, since the dynamics are governed by the derivative of $\mathcal{U}$, not $\mathcal{U}$ itself.

Two typical conservative forces:

1. The constant-field force $F = -mg$, where m is the particle mass and g is the field strength (positive or negative). Thus,

$$\mathcal{U} = \int^x mg\,\mathrm{d}s = mgx + \mathrm{Const.} \tag{5.3}$$

   Now x has the interpretation of the height above the reference level in the constant force field.

2. The spring force ('Hooke's Law') $F = -kx$, where k is the spring constant. Thus,

$$\mathcal{U} = \int^x ks\,\mathrm{d}s = \tfrac{1}{2}kx^2 + \mathrm{Const.} \tag{5.4}$$

Question: Are forces (1) and (2) invariant under a Galilean transformation? If so, why?

5.3 Energy

Consider a particle of mass m experiencing the conservative force F, with potential $\mathcal{U}(x) = -\int^x F(s)\,\mathrm{d}s$. The total energy E is the sum

$$E = \tfrac{1}{2}m\dot{x}^2 + \mathcal{U}(x) \tag{5.5}$$

The energy E of a conservative system $\mathcal{U}(x) = -\int^x F(s)\,\mathrm{d}s$ is a constant, $dE/dt = 0$.

Proof: First notice that the force F is the gradient of the potential:

$$F = -\frac{d\mathcal{U}}{dx}.$$

Now differentiate E with respect to time:

$$\begin{aligned}
\frac{dE}{dt} &= \frac{d}{dt}\left[\tfrac{1}{2}m\dot{x}^2 + \mathcal{U}(x)\right], \\
&= \tfrac{1}{2}m\left(2\dot{x}\frac{d\dot{x}}{dt}\right) + \frac{d\mathcal{U}}{dx}\frac{dx}{dt}, \\
&= \dot{x}\left(m\ddot{x} + \frac{d\mathcal{U}}{dx}\right), \\
&= \dot{x}\left(m\ddot{x} - F(x)\right), \\
&= 0.
\end{aligned}$$

Note that this result would break down if $\mathcal{U}$ were an explicit function of t, $\mathcal{U} = \mathcal{U}(x,t)$.

This result has many useful corollaries. An immediate one is the so-called *work-energy relation*. Recall the definition of work in one dimension. The work done in moving a particle at x through an infinitesimal distance dx is

$$dW = F(x)\,dx.$$

Thus ,the work done in bringing a particle from x_1 to x_2 is

$$W = \int_{x_1}^{x_2} F(x)\,\mathrm{d}x.$$

We have the work-energy relation:

> Let m be a particle experiencing a conservative force. The work done in bringing the particle from x_1 to x_2 equals the change in kinetic energy between these two states.

Proof: Since the energy is constant, this can be evaluated at x_1 or x_2, giving the same result:

$$\begin{aligned}
E(x_1) &= E(x_2), \\
\tfrac{1}{2}m\dot{x}^2\big|_{x_1} + \mathcal{U}(x_1) &= \tfrac{1}{2}m\dot{x}^2\big|_{x_2} + \mathcal{U}(x_2).
\end{aligned}$$

Re-arranging,

$$\begin{aligned}
\tfrac{1}{2}m\dot{x}^2\big|_{x_2} - \tfrac{1}{2}m\dot{x}^2\big|_{x_1} &= -\mathcal{U}(x_2) + \mathcal{U}(x_1), \\
&= -\left[\mathcal{U}(x_2) - \mathcal{U}(x_1)\right], \\
&= -\int_{x_1}^{x_2} \frac{d\mathcal{U}}{ds}\mathrm{d}s, \\
&= \int_{x_1}^{x_2} F(s)\,\mathrm{d}s, \\
&= W.
\end{aligned}$$

Conservation of energy is also useful in the construction of *energy diagrams*. Suppose a particle experiences a force whose potential is given by the figure (Fig. 5.1). The force is directed against

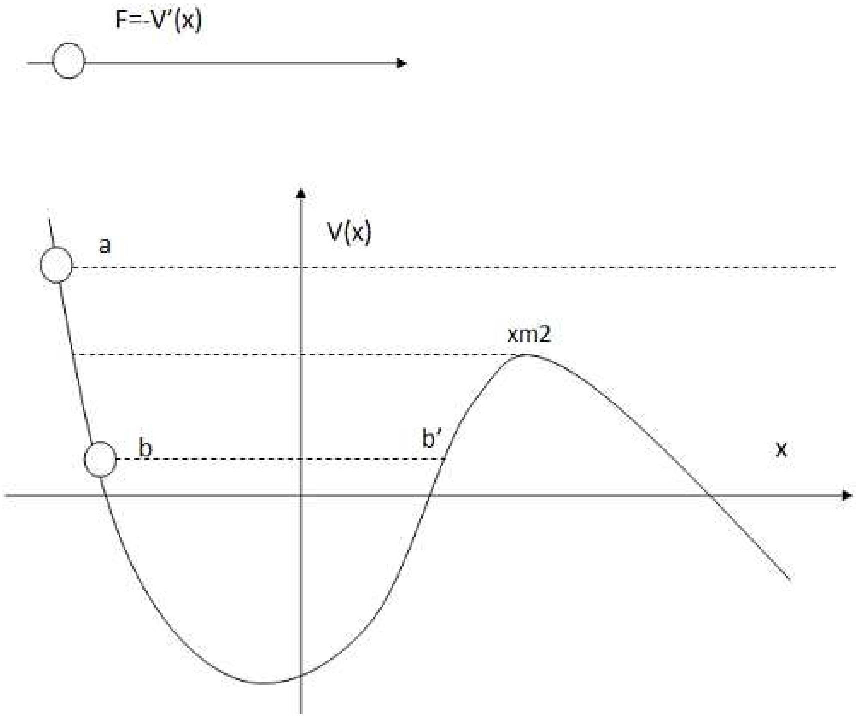


Figure 5.1: Relationship between force and potential.

the gradient of the potential. Thus, a particle will move from areas of high potential, to areas of lower potential.

- Suppose a particle starts from rest at point b. It will move towards the minimum of the potential function and then towards the point b'. At the point b' the particle's kinetic energy vanishes, and the particle cannot move beyond this point. The particle therefore falls back into the well, and continues indefinitely with motion bounded between the states b and b'.
- If, instead, the particle starts from rest at point a, the particle will still possess finite kinetic

energy at points b' and x_{m2}. The particle will therefore escape from the neighbourhood of the potential minimum and execute unbounded motion.

5.4 Simple-harmonic motion by energy methods

Recall the equation of motion for a particle experiencing Hooke's force:

$$m\ddot{x} = -kx.$$

In your ODEs class, you will probably learn that this equation has the solution

$$x = A\cos(\omega t) + B\sin(\omega t), \qquad \omega = \sqrt{k/m},$$

where A and B are constants of integration. The way this solution is usually proposed is by guesswork, which is a little unsatisfactory. It can, however, be derived using energy methods.

The energy of simple-harmonic motion is

$$E = \frac{1}{2}m\dot{x}^2 + \tfrac{1}{2}kx^2,$$

where E is a numerical constant that depends on the initial conditions. We can invert this equation to give

$$\begin{aligned}\left(\frac{dx}{dt}\right)^2 &= \frac{2E}{m} - \frac{k}{m}x^2,\\ \frac{dx}{dt} &= \sqrt{\frac{2E}{m} - \frac{k}{m}x^2}.\end{aligned}$$

Hence,

$$dt = \frac{dx}{\sqrt{\frac{2E}{m} - \frac{k}{m}x^2}}.$$

Re-arrange and integrate:

$$\begin{aligned}dt &= \frac{1}{\sqrt{\frac{2E}{m}}}\frac{\mathrm{d}x}{\sqrt{1-\frac{k}{2E}x^2}},\\ dt &= \frac{1}{\sqrt{\frac{2E}{m}}}\frac{\mathrm{d}\left(\sqrt{k/(2E)}x\right)}{\sqrt{1-\left(\sqrt{k/(2E)}x\right)^2}} \times \sqrt{\frac{2E}{k}}, \qquad s = \sqrt{k/(2E)}x,\\ dt &= \sqrt{\frac{m}{k}}\frac{\mathrm{d}s}{\sqrt{1-s^2}}, \qquad \omega = \sqrt{k/m}.\end{aligned}$$

$$t = \omega^{-1} \int_{s_0}^{s} \frac{\mathrm{d}s}{\sqrt{1-s^2}}, \qquad \omega = \sqrt{k/m}.$$

This is a standard integral:

$$\int^{x} \frac{\mathrm{d}s}{\sqrt{1-s^2}} = \sin^{-1} s$$

Hence,

$$\omega t = \sin^{-1} s - \underbrace{\sin^{-1} s_0}_{=\varphi}$$

Or,

$$s = \sin\left(\omega t + \varphi\right).$$

But $s = \sqrt{k/\left(2E\right)}x$, hence

$$x = \sqrt{\frac{2E}{k}} \sin\left(\omega t + \varphi\right). \tag{5.6}$$

The constants φ and $\sqrt{2E/k}$ can be related to A and B by applying the trigonometric addition formula [exercise].

5.5 Small oscillations

Equilibrium corresponds to those points where the force is zero:

$$F\left(x_{\mathrm{m}}\right) = 0, \qquad \mathcal{U}'\left(x_{\mathrm{m}}\right) = 0$$

(See Fig. 5.2). Hence, equilibrium points correspond to local minima or maxima of the potential function. If a particle is at equilibrium and at rest then it will remain at the equilibrium point. What happens is initially close the the equilibrium point x_{m}?

- If a particle initially in the vicinity of the equilibrium stays in the vicinity of the equilbrium, said equilibrium is *stable*.

- If a particle initially in the vicinity of the equilibrium moves away from the equilibrium and into another part of the domaini, said equilibrium is *unstable*.

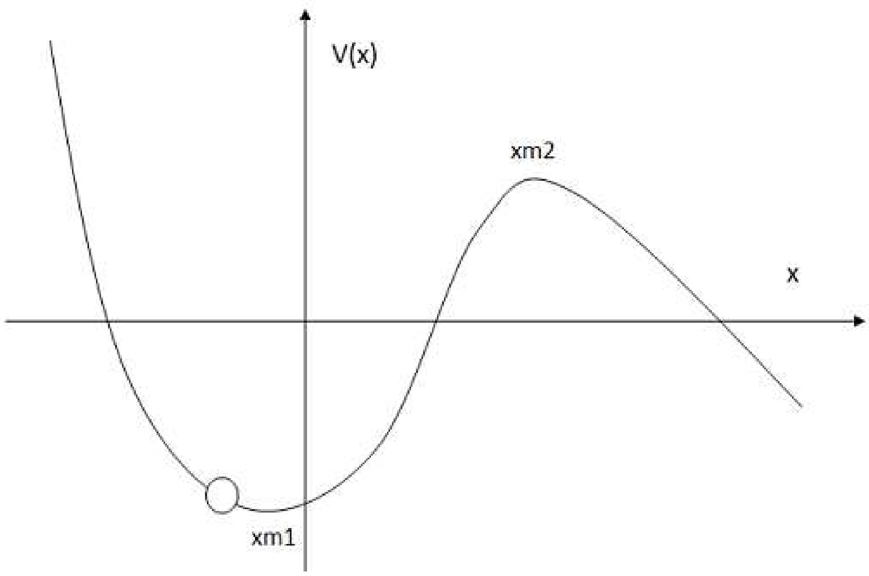


Figure 5.2: Maxima and minima of the potential function.

Mathematically, we can classify the equilibria as follows: Let us take a particle that is at a position x which is very close to an equilibrium point x_{m}. By Taylor's expansion,

$$\begin{aligned} F(x) &\approx F(x_{\mathrm{m}}) + F'(x_{\mathrm{m}})(x - x_{\mathrm{m}}) + \text{H.O.T.}, \\ &= -\mathcal{U}'(x_{\mathrm{m}}) - \mathcal{U}''(x_{\mathrm{m}})(x - x_{\mathrm{m}}), \\ &= -\mathcal{U}''(x_{\mathrm{m}})(x - x_{\mathrm{m}}). \end{aligned}$$

The equation of motion for the small disturbance $y = x - x_{\mathrm{m}}$ away from the equilibrium point is given by

$$\frac{d^2y}{dt^2} = -\frac{\mathcal{U}''(x_{\mathrm{m}})}{m} y. \tag{5.7}$$

Three possibilities for the motion:

1. $\mathcal{U}''(x_{\mathrm{m}}) > 0$ (local minimum). Identify a frequency

$$\omega = \sqrt{\mathcal{U}''(x_{\mathrm{m}})/m}. \tag{5.8}$$

Then, using the theory of simple harmonic motion,

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

where A and B are fixed by the initial conditions. The particle oscillates around the local minimum. The point $x_{\mathrm{m}} = x_{\mathrm{m1}}$ is a *stable* eq^{m}.

2. $\mathcal{U}''(x_{\mathrm{m}}) < 0$ (local maximum). Identify a rate $\sigma = \sqrt{-\mathcal{U}''(x_{\mathrm{m}})/m}$. Doing a similar integral to the one for SHM,

$$y = Ae^{\sigma t} + Be^{-\sigma t},$$

where A and B are fixed by the initial conditions. The particle moves exponentially fast away from the local maximum. The point $x_{\mathrm{m}} = x_{\mathrm{m2}}$ is an *unstable* eq^{m}.

3. If $\mathcal{U}''(x_{\mathrm{m}}) = 0$, one has to look at $\mathcal{U}'''(x_{\mathrm{m}})$, or possibly higher-order derivatives.

5.6 Solution by quadrature

So far we have learned some very important concepts in dynamics:

1. Energy conservation;
2. Simple-harmonic motion;
3. Linear stability of equilibria.

One subject we touched on was the *solution of equations of motion by quadrature*, wherein we solved for simple-harmonic motion via the energy method. In this final section, we do this for arbitrary potentials.

When the energy is conserved, the motion can be re-expressed in the following form:

$$\frac{dx}{dt} = \Phi(x), \tag{5.9}$$

where we take care only to work in an interval I of the potential landscape where the kinetic energy is positive or zero. The points where $\Phi = 0$ are called the *turning points*, and there can be no motion beyond these points. Re-writing Eq. (5.9), we obtain

$$dt = \frac{\mathrm{d}x}{\Phi(x)}.$$

If the particle begins with a position x_0 at $t = 0$ then

$$t = \int_{x_0}^{x} \frac{\mathrm{d}s}{\Phi(s)}. \tag{5.10}$$

If we can perform the integration (???) this gives us $t(x)$ which, if we can invert (???), gives us $x(t)$. We will solve the orbit problem by reducing it to a one-dimensional system, and then by solving the reduced system by quadrature.

> A system of N particles that can be solved by a reduction to a finite number of integrals of type (5.10) is called *integrable*.

Chapter 6

Worked example: Conservative forces in one dimension

A particle with mass m moves in one dimension. The potential-energy function is $\mathcal{U}(x) = \alpha x^{-2} - \beta x^{-1}$, where α and β are positive constants. The particle is released from rest at $x_0 = \alpha/\beta$.

1. Show that $\mathcal{U}(x)$ can be written as

$$\mathcal{U}(x) = \frac{\alpha}{x_0^2}\left[\left(\frac{x_0}{x}\right)^2 - \frac{x_0}{x}\right].$$

   Graph $\mathcal{U}(x)$; calculate $\mathcal{U}(x_0)$ and thereby locate the point x_0 on the graph. If a potential-well minimum exists, calculate the period of small oscillations about that minimum.

2. Calculate $v(x)$, the speed of the particle as a function of position. Graph the result and give a qualitative description of the motion.

3. For what value of x is the speed of the particle maximal? What is the value of that minimum speed?

4. If, instead, the particle is released at $x_1 = 3\alpha/\beta$, compute $v(x)$ and give a qualitative description of the motion. Locate the point x_1 on the graph of $\mathcal{U}$.

5. For each release point (x_0 and x_1), what are the maximum and minimum values of x reached during the motion?

1. We have

$$\mathcal{U}(x) = \frac{\alpha}{x_0^2}\left[\left(\frac{x_0}{x}\right)^2 - \frac{x_0}{x}\right].$$

Get rid of β: $\beta = \alpha/x_0$. Hence,

$$\begin{aligned}
\mathcal{U} &= \frac{\alpha}{x^2} - \frac{\alpha}{x_0}\frac{1}{x}, \\
&= \frac{\alpha}{x_0^2}\frac{1}{x/x_0}^2 - \frac{\alpha}{x_0^2}\frac{1}{x/x_0}, \\
&= \frac{\alpha}{x_0^2}\left[\left(\frac{x_0}{x}\right)^2 - \frac{x_0}{x}\right].
\end{aligned}$$

We now analyse $\mathcal{U}(x)$.

$$\begin{aligned}
\mathcal{U} &\sim +\infty, \qquad \text{as } x \to 0^+, \\
\mathcal{U} &\sim -0, \qquad \text{as } x \to \infty.
\end{aligned}$$

This suggests that $\mathcal{U}$ has a zero. Setting $\mathcal{U} = 0$ gives $(x_0/x)^2 - (x_0/x) = 0$, hence $x = x_0$. We should also try to find the extreme points of $\mathcal{U}$:

$$\begin{aligned}
\frac{d\mathcal{U}}{dx} &= \frac{\alpha}{x_0^2}\left[-\frac{2x_0^2}{x^3} + \frac{x_0}{x^2}\right], \\
\frac{d^2\mathcal{U}}{dx^2} &= \frac{\alpha}{x_0^2}\left[\frac{6x_0^2}{x^4} - \frac{2x_0}{x^3}\right].
\end{aligned}$$

The extreme point is at $d\mathcal{U}/dx = 0$, i.e. $x = 2x_0$. This is a minimum because

$$\begin{aligned}
\left.\frac{d^2\mathcal{U}}{dx^2}\right|_{x=2x_0} = \frac{2\alpha}{x_0^4}\left[3\left(\frac{x_0}{x}\right)^4 - \left(\frac{x_0}{x}\right)^3\right]_{x=2x_0} &= \frac{2\alpha}{x_0^4}\left(\tfrac{3}{2^4} - \tfrac{1}{2^3}\right) \\
&= \frac{2\alpha}{x_0^4}\left(\tfrac{3}{16} - \tfrac{1}{8}\right) = +\frac{\alpha}{8x_0^4} > 0.
\end{aligned}$$

Putting all these facts together, we draw a curve like Fig. 6.1.

Going back to Ch. 5, the frequency of small oscillations around the well minimum is

$$\omega = \sqrt{\mathcal{U}''(x_{\text{eq}})/m} = \sqrt{\frac{\alpha}{8x_0^4 m}}$$

But the period is $T = 2\pi/\omega$, hence

$$T = 2\pi\sqrt{\frac{8mx_0^4}{\alpha}}.$$

2. Calculate $v(x)$, the speed of the particle as a function of position. Graph the result and give a qualitative description of the motion.

We consider the energy associated with the system. Newton's equation is $d^2x/dt = -\mathcal{U}'(x)$, so the energy is

$$E = \tfrac{1}{2} m \left(\frac{dx}{dt}\right)^2 + \mathcal{U}(x) = \tfrac{1}{2} m v^2 + \mathcal{U}(x).$$

The particle starts from rest at $x = x_0$, so $E = 0 + \mathcal{U}(1) = 0$. Therefore, the speed $v(x)$ is

$$v(x) = \sqrt{-2\mathcal{U}(x)/m} = \sqrt{\frac{2\alpha}{m x_0^2}\left[-\left(\frac{x_0}{x}\right)^2 + \frac{x_0}{x}\right]},$$

which is valid for $x_0 \le x < \infty$.

Therefore, in qualitative terms, the particle starts from rest at $x = x_0$ and moves down the gradient of potential towards the minimum at $x = 2x_0$. It overshoots this point because $v(2x_0) = \sqrt{-2\mathcal{U}(2x_0)} = \left(1/\sqrt{2}\right)\sqrt{\alpha/(m x_0^2)}$. It travels towards the second turning point at $x = \infty$, which it never reaches. Thus, the motion is unbounded, but just barely so.

3. For what value of x is the speed of the particle maximal? What is the value of that minimum speed? Since $v = \sqrt{-2\mathcal{U}(x)/m}$, the speed is maximal when the potential is minimal, i.e. $x = 2x_0$, and $v(2x_0) = \left(1/\sqrt{2}\right)\sqrt{\alpha/(m x_0^2)}$. To prove this formally, compute dv/dx and its derivative at $x = 2x_0$, using $d\mathcal{U}/dx = 0$ there:

$$\begin{aligned}
\sqrt{m}\frac{dv}{dx}\bigg|_{x=2x_0} &= -\left(\frac{1}{\sqrt{-2\mathcal{U}(x)}}\frac{d\mathcal{U}}{dx}\right)_{x=2x_0} = 0. \\
\sqrt{m}\frac{d^2v}{dx^2}\bigg|_{x=2x_0} &= \left[-\frac{1}{\sqrt{-2\mathcal{U}(x)}}\frac{d^2\mathcal{U}}{dx^2} - \tfrac{1}{2\sqrt{2}}\frac{1}{(-\mathcal{U}(x))^{3/2}}\left(\frac{d\mathcal{U}}{dx}\right)^2\right]_{x=2x_0}, \\
&= \left(-\frac{1}{\sqrt{-2\mathcal{U}(x)}}\frac{d^2\mathcal{U}}{dx^2}\right)_{x=2x_0} < 0.
\end{aligned}$$

4. If, instead, the particle is released at $x_1 = 3\alpha/\beta$, compute $v(x)$ and give a qualitative description of the motion. Locate the point x_1 on the graph of $\mathcal{U}$.

Note: $\alpha/\beta = x_0$, hence the particle starts at $x = 3x_0$, from rest. The energy is $E = \mathcal{U}(3x_0) = -(2/9)(\alpha/x_0^2)$, and

$$v(x) = \sqrt{-\frac{2}{m}\left[\mathcal{U}(x) + \frac{2\alpha}{9x_0^2}\right]}.$$

The turing points are the the zeros of this function, and lie at

$$\mathcal{U}(x) + \frac{2\alpha}{9x_0^2} = 0,$$

or

$$\left(\frac{x_0}{x}\right)^2 - \frac{x_0}{x} = -\tfrac{2}{9}.$$

Letting $s := x/x_0$, this is

$$\begin{aligned}
\frac{1}{s^2} - \frac{1}{s} &= -\tfrac{2}{9}, \\
1 - s &= -\tfrac{2}{9}s^2, \\
2s^2 - 9s + 9 &= 0, \\
s &= \frac{9 \pm \sqrt{81 - 4 \times 2 \times 9}}{2 \times 2}, \\
&= \frac{9 \pm \sqrt{81 - 72}}{4} = 3 \text{ or } \tfrac{3}{2}.
\end{aligned}$$

Qualitatively, the motion is periodic and oscillates between turning points, $\frac{3}{2}x_0 \le x \le 3x_0$.

5. For each release point (x_0 and x_1), what are the maximum and minimum values of x reached during the motion?

   We have done this already:

   - Case 1: $x_0 \le x < \infty$.
   - Case 2: $\frac{3}{2}x_0 \le x \le 3x_0$.

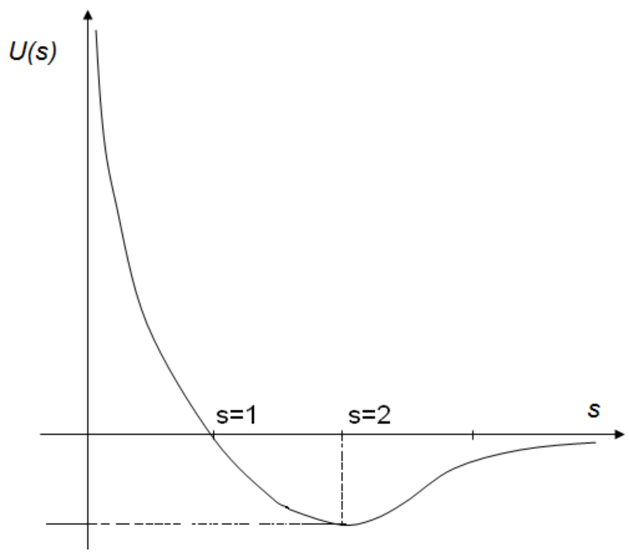


Figure 6.1: Sketch of potetnial function in worked example

Chapter 7

Motion in a plane

7.1 Overview

In Ch. 5 (Conservative forces in one dimension), we examined one-dimensional single-particle motion. Several important concepts carry over into planar motion:

1. Conservative forces and energy conservation;

2. Using the energy method to reduce the solution to an integral.

To carry out a similar reduction in two dimensions, appropriate coordinates are needed. For the orbit problem, these coordinates are the *circular polar coordinates*. In this chapter, we will

- Define circular polar coordinates;

- Define directional unit vectors in this system;

- Derive the velocity and acceleration along these directions.

The last point is nontrivial because these directions change with time. We will also solve some problems where life is made easier by the use of polar coordinates.

7.2 Circular polar coordinates

In two dimensions, and in an inertial frame, two Cartesian components x and y are necessary and sufficient to specify the position of a particle:

- Position: $\boldsymbol{x} = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}}$.

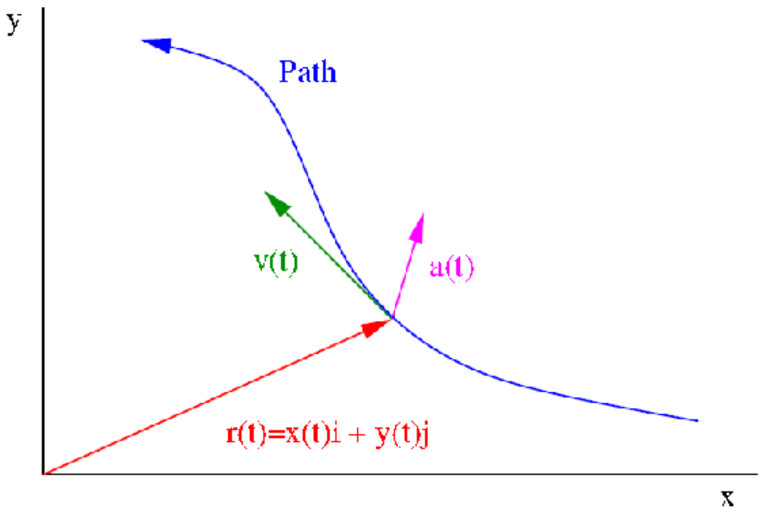


Figure 7.1: Trajectories in the plane.

- Velocity: $\boldsymbol{v} = d\boldsymbol{x}/dt = (dx/dt)\,\hat{\boldsymbol{x}} + (dy/dt)\,\hat{\boldsymbol{y}}$.
- Acceleration: $\boldsymbol{a} = d^2\boldsymbol{x}/dt^2 = (d^2x/dt^2)\,\hat{\boldsymbol{x}} + (d^2y/dt^2)\,\hat{\boldsymbol{y}}$.

(See Fig. 7.1.) We could use these coordinates to solve the equations of motion in a plane. They are not always appropriate however (e.g. circular motion). In some situations, it is more appropriate to use two further quantities:

- The distance a particle is away from the origin of the inertial frame, r;
- The angle the displacement vector $\boldsymbol{x}$ makes with the positive sense of the x-axis, θ.

These are the *circular polar coordinates*. From Fig. 7.2,

$$
\begin{aligned}
x &= r\cos\theta, \\
y &= r\sin\theta.
\end{aligned}
\tag{7.1}
$$

[Derive the inverse transformation.] We also introduce a direction vector $\hat{\boldsymbol{r}}$ in the direction of increasing r, and a direction vector $\hat{\boldsymbol{\theta}}$ in the direction of increasing θ. From the figure,

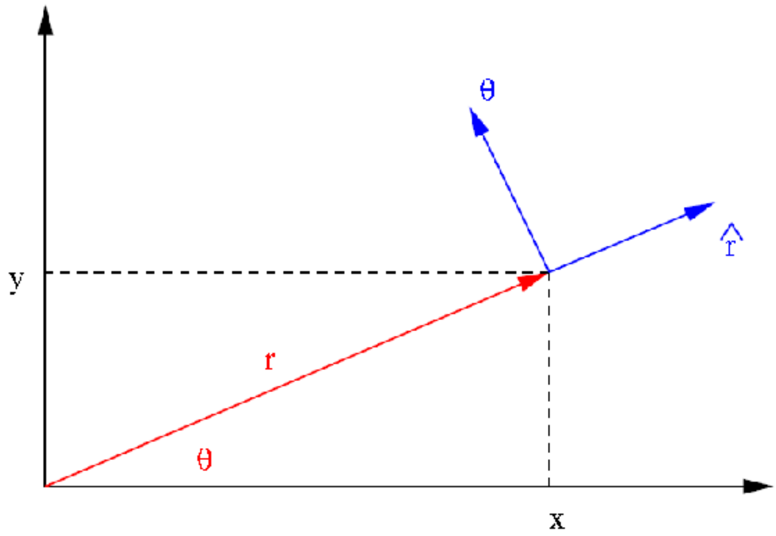


Figure 7.2: Polar coordinates

$$\begin{aligned} \hat{\boldsymbol{r}} &= \cos\theta \hat{\boldsymbol{x}} + \sin\theta \hat{\boldsymbol{y}}, \\ \hat{\boldsymbol{\theta}} &= -\sin\theta \hat{\boldsymbol{x}} + \cos\theta \hat{\boldsymbol{y}}, \end{aligned} \tag{7.2}$$

and

$$\boldsymbol{x} = r\hat{\boldsymbol{r}}.$$

Note: $\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{\theta}} = 0$ (orthogonality). Note further the matrix relations

$$\begin{pmatrix} \hat{\boldsymbol{r}} \\ \hat{\boldsymbol{\theta}} \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}}_{=R} \begin{pmatrix} \hat{\boldsymbol{x}} \\ \hat{\boldsymbol{y}} \end{pmatrix},$$

with inverse

$$\begin{pmatrix} \hat{\boldsymbol{x}} \\ \hat{\boldsymbol{y}} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{r}} \\ \hat{\boldsymbol{\theta}} \end{pmatrix}. \tag{7.3}$$

Exercise: Show that $\det(R) = 1$, and that $R^{-1} = R^T$. Compute R^2 and reduce it to a simple form using trigonometric identities.

Derive the velocity and acceleration in the new directions: We compute the velocity and acceleration in the radial and tangential directions. This is a little complicated, because these directions change with time, along with r and θ. We will need the relations

$$\frac{\partial \hat{\boldsymbol{r}}}{\partial \theta} = \frac{\partial}{\partial \theta}\left(\cos\theta \hat{\boldsymbol{x}} + \sin\theta \hat{\boldsymbol{y}}\right) = -\sin\theta \hat{\boldsymbol{x}} + \cos\theta \hat{\boldsymbol{y}} = \hat{\boldsymbol{\theta}}, \tag{7.4}$$

$$\frac{\partial \hat{\boldsymbol{\theta}}}{\partial \theta} = \frac{\partial}{\partial \theta}\left(-\sin\theta \hat{\boldsymbol{x}} + \cos\theta \hat{\boldsymbol{y}}\right) = -\cos\theta \hat{\boldsymbol{x}} - \sin\theta \hat{\boldsymbol{y}} = -\hat{\boldsymbol{r}}. \tag{7.5}$$

Derive the velocity:

$$\begin{aligned} \boldsymbol{v} &= \frac{d\boldsymbol{x}}{dt} = \frac{d}{dt}\left(r\hat{\boldsymbol{r}}\right), \\ &= \dot{r}\hat{\boldsymbol{r}} + r\frac{d\hat{\boldsymbol{r}}}{dt}. \end{aligned}$$

Using Eq. (7.4) and the chain rule,

$$\begin{aligned} \frac{d\hat{\boldsymbol{r}}}{dt} &= \dot{\theta}\frac{\partial \hat{\boldsymbol{r}}}{\partial \theta}, \\ &= \dot{\theta}\left(-\sin\theta \hat{\boldsymbol{x}} + \cos\theta \hat{\boldsymbol{y}}\right), \\ &= \dot{\theta}\hat{\boldsymbol{\theta}}. \end{aligned}$$

Hence,

$$\boldsymbol{v} = \dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}}.$$

Identify

- $v_r = \dot{r}$, the velocity in the radial (r-) direction;
- $v_\theta = r\dot{\theta}$, the velocity in the tangential (θ-) direction.

Derive the acceleration:

$$\begin{aligned} \boldsymbol{a} = \frac{d\boldsymbol{v}}{dt} &= \frac{d}{dt}\left(\dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}}\right), \\ &= \ddot{r}\hat{\boldsymbol{r}} + \dot{r}\dot{\theta}\frac{\partial \hat{\boldsymbol{r}}}{\partial \theta} + \left(r\ddot{\theta} + \dot{r}\dot{\theta}\right)\hat{\boldsymbol{\theta}} + r\dot{\theta}^2\frac{\partial \hat{\boldsymbol{\theta}}}{\partial \theta}, \\ &= \ddot{r}\hat{\boldsymbol{r}} + \dot{r}\dot{\theta}\hat{\boldsymbol{\theta}} + \left(r\ddot{\theta} + \dot{r}\dot{\theta}\right)\hat{\boldsymbol{\theta}} - r\dot{\theta}^2\hat{\boldsymbol{r}}. \end{aligned}$$

Hence, identify

- $a_r = \ddot{r} - r\dot{\theta}^2$, the acceleration in the radial (r-) direction;
- $a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta}$, the acceleration in the tangential (θ-) direction.

New names for the terms:

1. The *linear acceleration*, $\ddot{r}\hat{\boldsymbol{r}}$, that a particle has when it accelerates radially.
2. The *angular acceleration*, $r\ddot{\theta}\hat{\boldsymbol{\theta}}$, that occurs as a result of a change to the rate at which the particle is rotating.
3. The *centripetal acceleration*, $-r\dot{\theta}^2\hat{\boldsymbol{r}}$, which appears in the context of circular motion.
4. The *Coriolis acceleration*, $2\dot{r}\dot{\theta}\hat{\boldsymbol{\theta}}$, which a particle has when both r and θ change with time – even at uniform rates.

7.3 Dynamical situations where polar coordinates are appropriate

Circular motion in the absence of external forces: Consider ordinary circular motion, wherein a particle of mass m, held in tension by a string, undergoes uniform circular motion. We write down the equations of motion in polar coordinates and compute the tension N.

In the absence of constraints (a 'free particle'), the equations of motion are

$$m\left(\ddot{r} - r\dot{\theta}^2\right) = 0, \tag{7.6}$$

$$m\left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right) = 0. \tag{7.7}$$

In this system however, r is constant, so a constraint force must enter into the equations. The constraint acts against any change in the radial acceleration $\ddot{r}$, and so we add to Eq. (7.6) a force $-N$, and enforce $\dot{r} = 0$:

$$m\left(-r\dot{\theta}^2\right) = -N.$$

Writing $\dot{\theta} = \omega = \mathrm{Const.}$, this becomes

$$N = mr\omega^2,$$

while Eq. (7.7) reduces to the trivial expression $0 = 0$. Thus, in circular motion, the 'centripetal acceleration' and the tension are in balance.

A particle on a rotating groove: Consider a particle on a rotating groove, as shown in Fig. 7.3. The rate of rotation is constant and equal to ω. Calculate its motion as a function of time, and evaluate the constraining forces.

In the absence of any constraining forces (a 'free particle'), the equations of motion are unchanged from Eqs. (7.6) and (7.7). Now, however, there is a constraint in the $\hat{\boldsymbol{\theta}}$ direction, since the acceleration in this direction is forced to zero by the constant rotation. Thus, the constrained equations

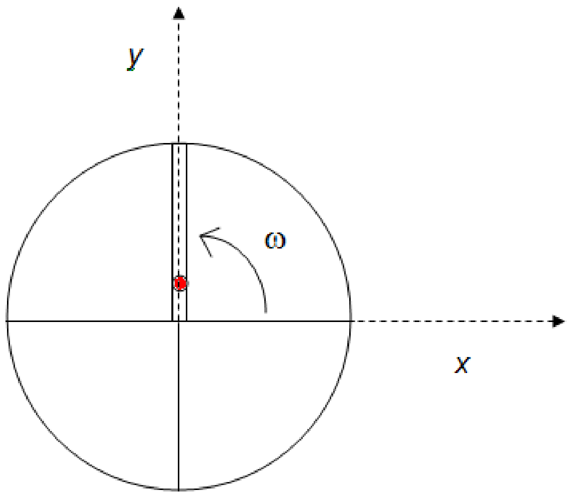


Figure 7.3: A particle in a rotating groove

of motion are

$$m\left(\ddot{r} - r\omega^2\right) = 0, \tag{7.8}$$
$$m\left(r\ddot{\theta} + 2\dot{r}\omega\right) = N_\theta. \tag{7.9}$$

We therefore solve Eq. (7.8)

$$\ddot{r} = r\omega^2.$$

It is readily shown (using energy methods) that the solution to this equation is

$$r = Ae^{\omega t} + Be^{-\omega t},$$

where A and B are constants of integration, for which we now solve. Suppose that the wheel starts from rest, so that the initial radial velocity is zero. Thus, $r(0) = r_0$, and $\dot{r}(0) = 0$. We have a matrix system:

$$\begin{pmatrix} 1 & 1 \\ \omega & -\omega \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} r_0 \\ 0 \end{pmatrix}.$$

Inverting gives $A = B = r_0/2$. Hence,

$$r = \tfrac{1}{2}r_0\left(e^{\omega t} + e^{-\omega t}\right) = r_0\cosh(\omega t).$$

Note that the derivative is thus $\dot{r} = r_0\omega\sinh(\omega t)$. Substitution of this relation into Eq. (7.9) gives the normal force:

$$N_\theta = 2m\omega^2 r_0\sinh(\omega t).$$

Circular motion in a gravitational field: Consider the problem of a particle executing circular motion in a gravitational field (See Fig. 7.4). The particle is given an initial 'kick' of velocity v_0 and starts from the bottom of the circle or hoop. What is the minimum value of v_0 necessary for the particle

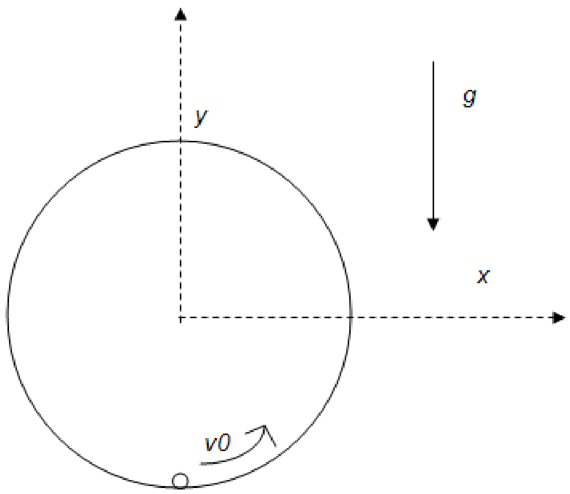


Figure 7.4: Circular motion in a gravitational field

not to fall off the circle? Note: You may have seen this problem before. Here, we will use the polar-coordinate formulation and energy methods to solve the problem.

Step 1: Forces As in problem (1), the particle is constrained to reside on the hoop, such that $\dot{r} = 0$, and such that an additional constraint force N_r is introduced into the equations of motion. Recall the free-space equations of motion:

$$\text{Radial equation :} \qquad m\left(\ddot{r} - r\dot{\theta}^2\right) = 0,$$
$$\text{Tangential equation :} \qquad m\left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right) = 0.$$

Introduce gravity and N_r. The gravitational force is

$$-g\hat{\boldsymbol{y}}.$$

Using the matrix inversion (7.3), this is

$$-g\left(\sin\theta\hat{\boldsymbol{r}} + \cos\theta\hat{\boldsymbol{\theta}}\right).$$

For an unconstrained particle in a gravitational field, the equations are thus

$$
\begin{aligned}
m\left(\ddot{r}-r\dot{\theta}^2\right) &= -mg\sin\theta,\\
m\left(r\ddot{\theta}+2\dot{r}\dot{\theta}\right) &= -mg\cos\theta.
\end{aligned}
$$

Impose the constraint: $\dot{r}=0$, constraint force in radial direction:

$$
\begin{aligned}
m\left(-r\dot{\theta}^2\right) &= -mg\sin\theta - N_r,\\
m\left(r\ddot{\theta}\right) &= -mg\cos\theta.
\end{aligned}
\tag{7.10}
$$

Hence,

$$
N_r = mr\dot{\theta}^2 - mg\sin\theta. \tag{7.11}
$$

The particle falls of the hoop when the constraint N_r vanishes. To compute N_r as a function of known quantities, we must use *energy methods*.

Step 2: Energy methods Take the equation of motion in the tangential direction (Eq. (7.10)) and multiply it by $\dot{\theta}$:

$$
m\dot{\theta}\left(r\ddot{\theta}\right) = -\dot{\theta}mg\cos\theta.
$$

Integrate:

$$
\frac{d}{dt}\left(\tfrac{1}{2}mr\dot{\theta}^2\right) = -\frac{d}{dt}mg\sin\theta.
$$

Hence,

$$
E = \tfrac{1}{2}mr^2\dot{\theta}^2 + mgr\sin\theta = \mathrm{Const.}
$$

(We have multipled the conserved quantity by the constant r to obtain an energy). Since E is constant, it must equal its initial value ($\theta=-\pi/2$):

$$
\tfrac{1}{2}mv_0^2 = \tfrac{1}{2}mr^2\dot{\theta}^2 + mgr\left(1+\sin\theta\right).
$$

Hence,

$$
mr\dot{\theta}^2 = \frac{mv_0^2}{r} - 2mg\left(1+\sin\theta\right). \tag{7.12}
$$

Combining Eqs. (7.11) and (7.12),

$$N_r = \frac{mv_0^2}{r} - mg\left(2 + 3\sin\theta\right).$$

The constraint N_r is minimal at $\theta = \pi/2$, $N_{r,\min} = (mv_0^2/r) - 5mg$. The condition for the particle to fall off the hoop is the vanishing of the constraint. Thus, if $N_{r,\min} \geq 0$, then the particle stays on the hoop:

$$\begin{aligned} \left(mv_0^2/r\right) &\geq 5mg, \\ |v_0| &\geq \sqrt{5gr}. \end{aligned}$$

Chapter 8

The theory of partial derivatives

8.1 Overview

This chapter contains two topics that are probably new to you:

- Partial derivatives;
- The gradient operator;

8.2 Partial derivatives

A function $f(x,y)$ if two variables is a map from a subset of $\mathbb{R}^2$ to $\mathbb{R}$:

$$\begin{array}{rcl} f : \left(A \subset \mathbb{R}^2\right) & \to & \mathbb{R} \\ (x,y) & \to & f(x,y). \end{array}$$

Examples:

- The elevation above sea level at any point in Ireland is a function of latitude and longitude;
- The pressure of an ideal gas is a function of temperature and density (Boyle's Law);
- The GDP of an economy is a function of the quantity of money in circulation, and its velocity (Quantity theory of money).

Henceforth, we shall consider functions of two coordinates (x, y). When we are given such a function, it is natural to ask how the function varies as x changes, and as y changes. Equivalently, we want to know how the function changes as we move in the 'x-direction', and in the 'y'-direction'. Thus, we make small variations in the x-coordinate, keeping y fixed:

$$f\left(x+\delta x, y\right).$$

Then, we form the quotient

$$\frac{f\left(x+\delta x, y\right)-f\left(x, y\right)}{\delta x}.$$

Taking $\delta x \rightarrow 0$, we obtain the *partial derivative of f w.r.t. x (keeping y fixed)*:

$$\frac{\partial f}{\partial x}\left(x, y\right)=\lim_{\delta x \rightarrow 0} \frac{f\left(x+\delta x, y\right)-f\left(x, y\right)}{\delta x}.$$

Similarly, we have a partial derivative with w.r.t. y keeping x fixed: First, we form the quotient

$$\frac{f\left(x, y+\delta y\right)-f\left(x, y\right)}{\delta y},$$

then we take the limit as $\delta y \rightarrow 0$:

$$\frac{\partial f}{\partial y}\left(x, y\right)=\lim_{\delta y \rightarrow 0} \frac{f\left(x, y+\delta y\right)-f\left(x, y\right)}{\delta y}.$$

Thus, to form a partial derivative in the x-direction, you treat y as a constant and do ordinary differentiation on the x-variable.

Examples

The function $f\left(x, y\right)=x^2+y^2$. Let us hold y fixed and differentiate w.r.t. x:

$$\frac{\partial f}{\partial x}=\frac{\partial}{\partial x}\left(x^2+y^2\right)=\frac{\partial}{\partial x}\left(x^2+\text{Const.}\right)=\frac{\partial}{\partial x}\left(x^2\right)=2x.$$

Now hold x fixed and differentiate w.r.t. y:

$$\frac{\partial f}{\partial y}=\frac{\partial}{\partial y}\left(x^2+y^2\right)=\frac{\partial}{\partial y}\left(\text{Const.}+y^2\right)=\frac{\partial}{\partial y}\left(y^2\right)=2y.$$

The function $f(x,y) = xy$. Let us hold y fixed and differentiate w.r.t. x:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} xy = \frac{\partial}{\partial x} x \,\mathrm{Const.} = \mathrm{Const.} = y.$$

Now hold x fixed and differentiate w.r.t. y:

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y} xy = \frac{\partial}{\partial y} \mathrm{Const.} y = \mathrm{Const.} = x.$$

The function $f(x,y) = x/y$. Let us hold y fixed and differentiate w.r.t. x:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\frac{x}{y} = \frac{\partial}{\partial x}\frac{x}{\mathrm{Const.}} = \frac{1}{\mathrm{Const.}} = \frac{1}{y}$$

Now hold x fixed and differentiate w.r.t. y:

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\frac{x}{y} = \frac{\partial}{\partial y}\frac{\mathrm{Const.}}{y} = -\frac{\mathrm{Const.}}{y^2} = -\frac{x}{y^2}$$

The function $f(x,y) = 1/\sqrt{x^2+y^2}$. Let us hold y fixed and differentiate w.r.t. x:

$$\begin{aligned}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\frac{1}{\sqrt{x^2+y^2}} = \frac{\partial}{\partial x}\frac{1}{\sqrt{x^2+\mathrm{Const.}}} &= \frac{\partial}{\partial x}\left(x^2+\mathrm{Const.}\right)^{-1/2} \\ &= -\tfrac{1}{2}\left(x^2+\mathrm{Const.}\right)^{-3/2}(2x) = -\frac{x}{\left(x^2+\mathrm{Const.}\right)^{3/2}} = -\frac{x}{\left(x^2+y^2\right)^{3/2}}\end{aligned}$$

Now hold x fixed and differentiate w.r.t. y:

$$\begin{aligned}\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\frac{1}{\sqrt{x^2+y^2}} = \frac{\partial}{\partial x}\frac{1}{\sqrt{\mathrm{Const.}+y^2}} &= \frac{\partial}{\partial x}\left(\mathrm{Const.}+y^2\right)^{-1/2} \\ &= -\tfrac{1}{2}\left(\mathrm{Const.}+y^2\right)^{-3/2}(2y) = -\frac{y}{\left(\mathrm{Const.}+y^2\right)^{3/2}} = -\frac{y}{\left(x^2+y^2\right)^{3/2}}\end{aligned}$$

Pedantic notation

1. When the function f is in fact a function of a single variable only ($f = f(x)$, say) there is no difference between $\partial/\partial x$ and d/dx. In that case, $\partial f/\partial x = df/dx = f'(x)$. So far, this is

the context in which the partial derivative sign has been used in class, especially to highlight the use of the chain rule, e.g.

$$\frac{d}{dt}\mathcal{U}\left(x\right)=\frac{dx}{dt}\frac{\partial\mathcal{U}}{\partial x}=\dot{x}\frac{\partial\mathcal{U}}{\partial x}.$$

But this is nothing other than $(d/dt)\mathcal{U}=\dot{x}\mathcal{U}'\left(x\right)$.

2. To save chalk, we will sometimes write $\partial f/\partial x$ as $\partial_x f$ or even f_x. A similar notation holds for partial derivatives w.r.t. y.

8.3 The gradient operator in two dimensions

Let f be a function of two variables,

$$\begin{aligned} f:\left(A\subset\mathbb{R}^2\right) &\rightarrow \mathbb{R} \\ (x,y) &\rightarrow f\left(x,y\right). \end{aligned}$$

Then the gradient operator acting on f is a vector with with the following form:

$$\mathrm{grad}\, f := \hat{\boldsymbol{x}}\frac{\partial f}{\partial x}+\hat{\boldsymbol{y}}\frac{\partial f}{\partial y}.$$

In class, we will write this vector as ∇f, and call it 'grad f' or 'nabla f'.

Examples

The function $f\left(x,y\right)=x^2+y^2$. We know that $\partial_x f=2x$ and $\partial_y=2y$. Hence,

$$\nabla f=\hat{\boldsymbol{x}}\partial_x f+\hat{\boldsymbol{y}}\partial_y f=2\hat{\boldsymbol{x}}x+2\hat{\boldsymbol{y}}y=2\left(x,y\right)=2\boldsymbol{x},$$

where $\boldsymbol{x}$ is a position vector.

The function $f\left(x,y\right)=1/\sqrt{x^2+y^2}$. We know that

$$\begin{aligned} \partial_x f &= -\frac{x}{\left(x^2+y^2\right)^{3/2}}, \\ \partial_y f &= -\frac{y}{\left(x^2+y^2\right)^{3/2}}. \end{aligned}$$

Hence,

$$\nabla f = \hat{\boldsymbol{x}}\left(-\frac{x}{\left(x^2+y^2\right)^{3/2}}\right) + \hat{\boldsymbol{y}}\left(-\frac{y}{\left(x^2+y^2\right)^{3/2}}\right) = -\frac{(x,y)}{\left(x^2+y^2\right)^{3/2}} = -\frac{\boldsymbol{x}}{|\boldsymbol{x}|^3}.$$

Definition: In two or more dimensions, a force $\boldsymbol{F}(\boldsymbol{x})$ is called conservative if it can be written in terms of the gradient of a potential:

$$\boldsymbol{F} = -\nabla\mathcal{U}(\boldsymbol{x}).$$

Examples

The two-dimensional harmonic oscillatior $\boldsymbol{F} = -k\boldsymbol{x} = -k(x,y)$ is conservative because

$$\mathcal{U} = \tfrac{1}{2}k\left(x^2+y^2\right),$$

and

$$-\nabla\mathcal{U} = -(kx, ky) = -k\boldsymbol{x} = \boldsymbol{F}.$$

The two-dimensional gravitational force

$$\boldsymbol{F} = -\frac{A}{|\boldsymbol{x}|^3}, \qquad \boldsymbol{x} = (x,y)$$

with $|\boldsymbol{F}| = A/|\boldsymbol{x}|^2$ is conservative because

$$\mathcal{U} = -\frac{A}{\sqrt{x^2+y^2}},$$

and

$$\begin{aligned}-\nabla\mathcal{U} = -\left(-A\left(-\tfrac{1}{2}\frac{2x}{\left(x^2+y^2\right)^{3/2}}\right), -A\left(-\tfrac{1}{2}\frac{2y}{\left(x^2+y^2\right)^{3/2}}\right)\right),\\ -A\left(\frac{x}{\left(x^2+y^2\right)^{3/2}}, \frac{y}{\left(x^2+y^2\right)^{3/2}}\right) = -A\frac{\boldsymbol{x}}{|\boldsymbol{x}|^3} = \boldsymbol{F}.\end{aligned}$$

Chapter 9

Angular momentum and central forces

9.1 Overview

Previously, in Ch. 7 (Motion in a plane),

- We studied motion in the plane (two dimensions).
- We were particularly concerned with problems with circular symmetry. Thus, we wrote down the equations of motion in polar coordinates.
- We solved several problems involving motion constrained to a circle.

In this chapter we introduce two further concepts which will enable us to solve far more general problems of motion in a plane: *angular momentum* and *central forces*.

9.2 Angular momentum

Consider a particle of mass m, position vector $\boldsymbol{x}$, and velocity $\boldsymbol{v} = \dot{\boldsymbol{x}}$ relative to an inertial frame. The angular momentum $\boldsymbol{J}$ relative to that frame is

$$\boldsymbol{J} = \boldsymbol{x} \times (m\boldsymbol{v}) \,. \tag{9.1}$$

Let us differentiate this expression and apply Newton's equation:

$$\begin{aligned}\frac{d\boldsymbol{J}}{dt} &= \frac{d\boldsymbol{x}}{dt} \times (m\boldsymbol{v}) + \boldsymbol{x} \times \left(m\frac{d\boldsymbol{v}}{dt}\right) \\ &= \boldsymbol{v} \times (m\boldsymbol{v}) + \boldsymbol{x} \times (m\boldsymbol{a}), \\ &= 0 + \boldsymbol{x} \times \boldsymbol{F},\end{aligned}$$

where $\boldsymbol{F}$ is the vector sum of forces acting on the particle. We identify the torque:

$$\boldsymbol{\tau} = \text{Torque} := \boldsymbol{x} \times \boldsymbol{F},$$

the equation of angular momentum:

$$\frac{d\boldsymbol{J}}{dt} = \boldsymbol{\tau}, \tag{9.2}$$

and the principle of conservation of angular momentum:

> If the sum of the torques acting on a particle is zero, then the angular momentum is constant.

Some points about torque:

- Torque depends on the origin we choose, but force does not;
- $\boldsymbol{\tau} = \boldsymbol{x} \times \boldsymbol{F}$: the torque and the force are mutually perpendicular;
- For a system of more than one particle, there can be torque without force, and force without torque, but in general both are present (See Fig. 9.1).

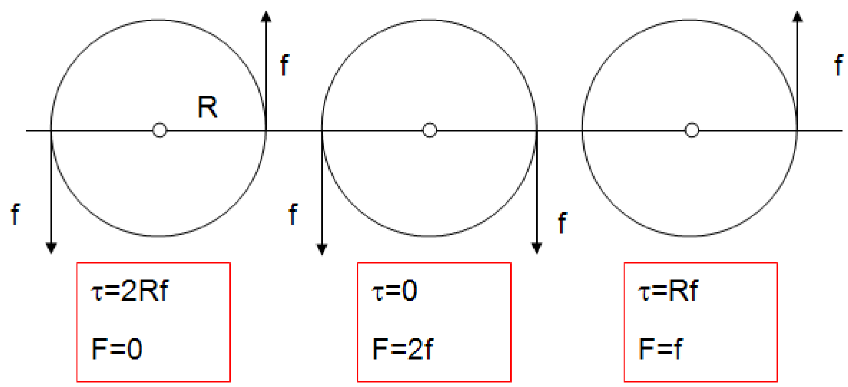


Figure 9.1: Relationship between torque and force. The torques are pointing out of the page.

Example: Computing the angular momentum of the conical pendulum Consider the conical pendulum shown in Fig. 9.2 (a). Calculate the angular momentum about the points $P1$ and $P2$.

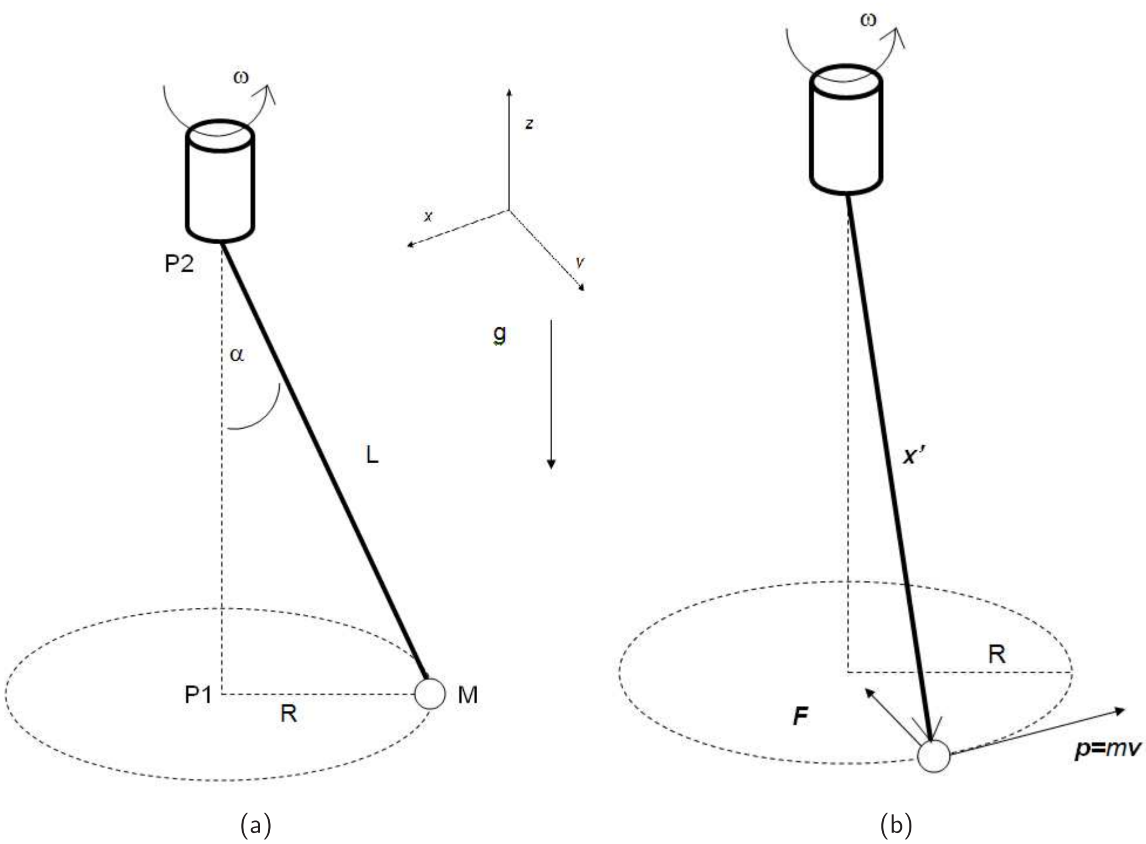


Figure 9.2: The conical pendulum.

1. The point $P1$: We use polar coordinates in the x-y plane. The particle executes circular motion in this plane, with velocity $\boldsymbol{v} = \omega R\hat{\boldsymbol{\theta}}$ in the tangential direction. The radius vector is $R\hat{\boldsymbol{r}}$. Thus,
$$\boldsymbol{J}_{P1} = \omega MR^2\hat{\boldsymbol{r}} \times \hat{\boldsymbol{\theta}} = \omega MR^2\hat{\boldsymbol{z}}.$$

2. The point $P2$: Refer to Fig. 9.2. Call the vector from $P1$ to M to $\boldsymbol{x}'$. Clearly,
$$\boldsymbol{x}' = \sin\alpha L\hat{\boldsymbol{r}} - \cos\alpha L\hat{\boldsymbol{z}},$$
and the velocity $\boldsymbol{v}$ is in the $\hat{\boldsymbol{\theta}}$ direction, where $\left(\hat{\boldsymbol{r}}, \hat{\boldsymbol{\theta}}\right)$ are polar coordinates in the x–y plane.

Thus,

$$\begin{aligned} \boldsymbol{J}_{P2} &= M\boldsymbol{x}' \times \boldsymbol{v}, \\ &= M\left(\sin\alpha L\hat{\boldsymbol{r}} - \cos\alpha L\hat{\boldsymbol{z}}\right) \times \left(v_\theta \hat{\boldsymbol{\theta}}\right), \\ &= ML\sin\alpha v_\theta \left(\hat{\boldsymbol{r}} \times \hat{\boldsymbol{\theta}}\right) - ML\cos\alpha v_\theta \left(\hat{\boldsymbol{z}} \times \hat{\boldsymbol{\theta}}\right), \\ &= ML\sin\alpha v_\theta \hat{\boldsymbol{z}} + ML\cos\alpha v_\theta \hat{\boldsymbol{r}}. \end{aligned}$$

Taking the absolute magnitude and using $\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{r}} = 0$, this gives $|\boldsymbol{J}_{P2}| = MLv_\theta = ML\omega R$.

Computing the torque of the conical pendulum In the z-direction, a component of the tension is in balance with gravity:

$$T\cos\alpha = Mg,$$

hence, there is no net force in the z-direction. In the x-y plane, the force is given by a component of the tension, and is radially inwards:

$$\boldsymbol{F} = -T\sin\alpha\hat{\boldsymbol{r}},$$

which balances with the centripetal acceleration.

1. The point $P1$: The total force is $\boldsymbol{F} = -T\sin\alpha\hat{\boldsymbol{r}}$. The torque is therefore

$$\boldsymbol{\tau}_{P1} = R\hat{\boldsymbol{r}} \times (-T\sin\alpha\hat{\boldsymbol{r}}) = 0.$$

   Hence, $d\boldsymbol{J}_{P1}/dt = 0$, consistent with $\boldsymbol{J}_{P1} = \omega MR^2\hat{\boldsymbol{z}} = \mathrm{Const.}$

2. The point $P2$:

$$\begin{aligned} \boldsymbol{\tau}_{P2} &= \left(\sin\alpha L\hat{\boldsymbol{r}} - \cos\alpha L\hat{\boldsymbol{z}}\right) \times \left(-T\sin\alpha\hat{\boldsymbol{r}}\right), \\ &= +TL\cos\alpha\sin\alpha\left(\hat{\boldsymbol{z}} \times \hat{\boldsymbol{r}}\right), \\ &= MgL\sin\alpha\hat{\boldsymbol{\theta}}. \end{aligned}$$

9.3 Central forces in three dimensions

Consider a particle experiencing a force $\boldsymbol{F}$ in a coordinate system (x, y, z). The force $\boldsymbol{F}$ is said to be central with respect to the coordinate system if

$$\boldsymbol{F} = F(r)\,\hat{\boldsymbol{r}}, \tag{9.3}$$

where r is the radial distance and $\hat{\boldsymbol{r}}$ is the radial direction. We call the coordinate origin the *force centre*.

In three dimensions, the radial vector is $\hat{\boldsymbol{r}} = \boldsymbol{x}/|\boldsymbol{x}|$, where $|\boldsymbol{x}| = \sqrt{x^2+y^2+z^2} = r$.

In three dimensions, a force $\boldsymbol{F}$ is called *conservative* if it can be written as the gradient of a function $\mathcal{U}(x,y,z)$:

$$\boldsymbol{F} = -\nabla\mathcal{U}(x,y,z). \tag{9.4}$$

Here ∇ is the *gradient* operator,

$$\nabla = \hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z}.$$

Theorem: All central forces are conservative.

Proof: We write down what we think to be the potential:

$$\mathcal{U} = -\int^{|\boldsymbol{x}|} F(s)\,\mathrm{d}s.$$

Now

$$\begin{aligned}
\nabla\mathcal{U} &= \frac{\partial\mathcal{U}}{\partial x}\hat{\boldsymbol{x}} + \frac{\partial\mathcal{U}}{\partial x}\hat{\boldsymbol{y}} + \frac{\partial\mathcal{U}}{\partial x}\hat{\boldsymbol{z}}, \\
&= \frac{\partial\mathcal{U}}{\partial|\boldsymbol{x}|}\frac{\partial|\boldsymbol{x}|}{\partial x}\hat{\boldsymbol{x}} + \cdots, \\
&= -\hat{\boldsymbol{x}}F(|\boldsymbol{x}|)\frac{\partial}{\partial x}\left(x^2+y^2+z^2\right)^{1/2} + \cdots, \\
&= -\hat{\boldsymbol{x}}F(|\boldsymbol{x}|)\frac{x}{\sqrt{x^2+y^2+z^2}} + \cdots, \\
&= -F(|\boldsymbol{x}|)\frac{\boldsymbol{x}}{|\boldsymbol{x}|}.
\end{aligned}$$

Hence, $\boldsymbol{F} = -\nabla\mathcal{U}$, as required.

Theorem: For a particle experiencing a conservative force in three dimensions, the energy

$$E = \tfrac{1}{2}m\dot{\boldsymbol{x}}^2 + \mathcal{U}(|\boldsymbol{x}|)$$

is conserved, provided the motion satisfies Newton's equation, $m\ddot{\boldsymbol{x}} = -\nabla\mathcal{U}$.

Proof: Left as an exercise.

9.3.1 Central forces for a system of two particles

Consider a system of two particles of masses m_1 and m_2, interacting via a force $\boldsymbol{F}$, in any spatial dimension. For such a system, the force is central if the force on particle 1 due to particle 2 is

$$\boldsymbol{F}_{12} = F\left(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\right)\frac{\boldsymbol{x}_1 - \boldsymbol{x}_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|}.$$

By Newton's Third Law,

$$\boldsymbol{F}_{21} = -\boldsymbol{F}_{12} = F\left(|\boldsymbol{x}_1 - \boldsymbol{x}_2|\right)\frac{\boldsymbol{x}_2 - \boldsymbol{x}_1}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|}.$$

The most important example is gravity (See Fig. 9.3):

$$\boldsymbol{F}_{12} = -\frac{Gm_1m_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|^2}\left(\frac{\boldsymbol{x}_1 - \boldsymbol{x}_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|}\right), \tag{9.5}$$

where G is the gravitational constant. When one of the particles is much more massive than the

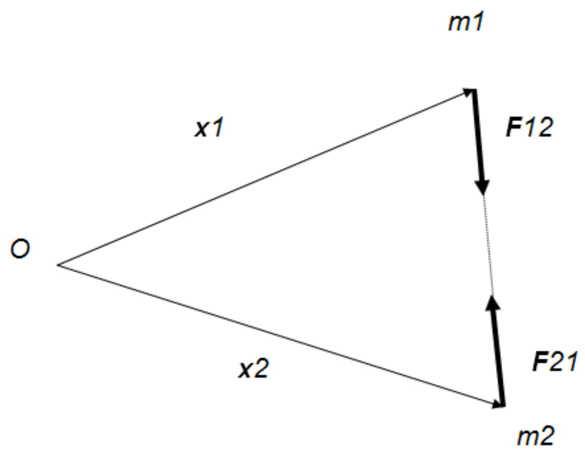


Figure 9.3: Force diagram for a two-particle gravitational interaction.

other ($m_2 \gg m_1$, say), then m_2 can be regarded as having infinite inertia, and to be unmoved by the interaction force. The location of m_2 can then be treated as a force centre, and we may regard m_1 as experiencing a central force field relative to this centre. An example of such a scenario is satellite motion around the earth. Later on we shall formulate this approximation in a rigorous way.

Coulomb's law for interactions between charged particles is also a central force

$$\boldsymbol{F}_{12} = \frac{e_1e_2}{4\pi\varepsilon_0|\boldsymbol{x}_1 - \boldsymbol{x}_2|^2}\left(\frac{\boldsymbol{x}_1 - \boldsymbol{x}_2}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|}\right). \tag{9.6}$$

Unlike Newton's law, wherein m_1 and m_2 must be positive, the charges e_1 and e_2 can have either sign. If both charges have the same sign, then the force (9.6) is *repulsive* (Fig. 9.4).

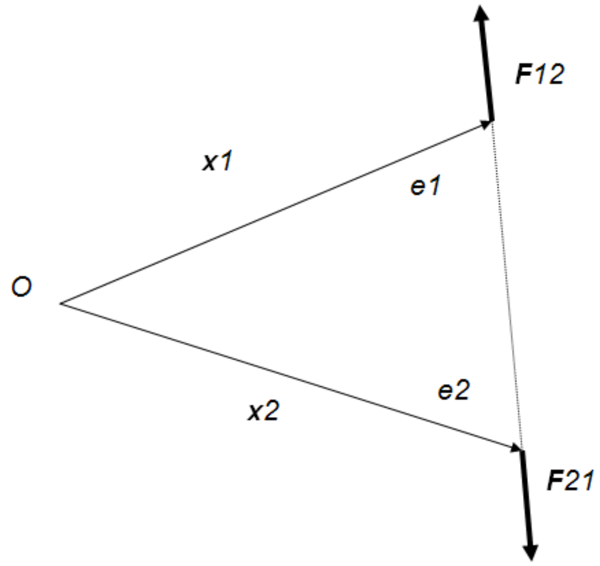


Figure 9.4: Force diagram for a repulsive Coulomb interaction.

> The law of gravitation as written down here (Eq. (9.5)) is for point particles. However, it also holds for spherical bodies at finite separations. This follows from the point-particle law by integration. Specifically, consider a particle P experiencing a gravitational force from a collection of point particles. The net force on P is obtained by summing over these particles. If these particles are arranged in a spherical distribution, and if the number of these particles tends to infinity while their separation tends to zero, then we the net force on P is an integral, whose form is identical to the original law for point particles. Similarly, point P can be replaced by a spherical distribution and another integral performed, and the final result is the original force law, with the point particle masses replaced by the masses of the extended bodies.

See Note[1].

[1] This proof is not on the syllabus but can be found in University Physics, Ch. 12 (10^{th} edition), and in other places, like Wikipedia: http://en.wikipedia.org/wiki/Shell_theorem

Chapter 10

Central forces reduce to one-dimensional motion

10.1 Overview

In Ch. 9 (Angular momentum and central forces),

- We defined angular momentum: $\boldsymbol{J} = \boldsymbol{x} \times (m\boldsymbol{v})$. This depends on the coordinate origin.
- We formulated the principle of conservation of angular momentum: in the absence of net torque, the angular momentum is conserved.
- We defined central forces and showed that they were conservative.

In this chapter we demonstrate that central forces possess zero torque, when all measurements are made relative to the force centre. This enables us to reduce the problem of particle motion under central forces to a one-dimensional one.

10.2 The master theorem

Theorem: Consider a particle experiencing a central force. When measured with respect to the force centre, the angular momentum is conserved.

Proof: Consider the angular momentum $\boldsymbol{J} = \boldsymbol{x} \times (m\boldsymbol{v})$, as measured from the force centre. Recall that we showed that the rate of change of angular momentum is the torque:

$$\frac{d\boldsymbol{J}}{dt} = \boldsymbol{\tau} = \boldsymbol{x} \times \boldsymbol{F} = r\hat{\boldsymbol{r}} \times \boldsymbol{F}.$$

But $\boldsymbol{F} = F(r)\,\hat{\boldsymbol{r}}$, hence

$$\boldsymbol{x} \times \boldsymbol{F} = r\hat{\boldsymbol{r}} \times \boldsymbol{F} = rF(r)\,\hat{\boldsymbol{r}} \times \hat{\boldsymbol{r}} = 0,$$

and the *vector* angular momentum is conserved. Since the direction of angular momentum is conserved, it follows that $\boldsymbol{x}$ and $\boldsymbol{v}$ lie in the same plane for all times, and that motion under a central force is in fact two dimensional. Thus, we consider central forces in two dimensions only.

10.3 Central forces in Newton's equations

We have seen that a central force can written as $\boldsymbol{F} = F(r)\,\hat{\boldsymbol{r}}$. There is therefore no force in the tangential direction. Newton's laws for such a system are therefore the following:

$$m\left(\ddot{r} - r\dot{\theta}^2\right) = F(r), \tag{10.1}$$

$$m\left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right) = 0. \tag{10.2}$$

- The force in Eq. (10.1) can be re-written as

$$F = -\frac{\partial \mathcal{U}}{\partial r}, \qquad \mathcal{U}(r) = -\int^{r} F(s)\,\mathrm{d}s.$$

- The second equation (Eq. (10.2)) can be re-written as

$$\frac{m}{r}\frac{d}{dt}\left(r^2\dot{\theta}\right) = 0,$$

  and immediate consequence of which is that the the quantity $h := r^2\dot{\theta}$ is conserved.

The quantity h is related to the magnitude of the angular momentum:

$$\begin{aligned}
\boldsymbol{J} &= r\hat{\boldsymbol{r}} \times (m\boldsymbol{v}), \\
&= mr\hat{\boldsymbol{r}} \times \left(\hat{\boldsymbol{r}}\,v_r + \hat{\boldsymbol{\theta}}\,v_\theta\right), \\
&= mr^2\dot{\theta}\,\hat{\boldsymbol{r}} \times \hat{\boldsymbol{\theta}}, \\
&= mr^2\dot{\theta}\,\hat{\boldsymbol{z}}, \\
&= mh\hat{\boldsymbol{z}}.
\end{aligned}$$

hence $h = J/m$.

Problem: Re-write Eqs. (10.1)–(10.2) as

$$m\left(\ddot{r}-r\dot{\theta}^2\right)=-\frac{\partial\mathcal{U}}{\partial r},$$
$$m\left(r\ddot{\theta}+2\dot{r}\dot{\theta}\right)=0. \tag{10.3}$$

Identify the kinetic energy

$$K=\tfrac{1}{2}m\boldsymbol{v}^2=\tfrac{1}{2}m\left(\dot{r}^2+r^2\dot{\theta}^2\right).$$

Prove that Newton's equations (10.3) satisfy the *Euler–Lagrange equations*

$$\frac{d}{dt}\frac{\partial K}{\partial\dot{r}}-\frac{\partial K}{\partial r}=-\frac{\partial\mathcal{U}}{\partial r},$$
$$\frac{d}{dt}\frac{\partial K}{\partial\dot{\theta}}-\frac{\partial K}{\partial\theta}=-\frac{\partial\mathcal{U}}{\partial\theta}.$$

Next year, you will have to study these equations in more generality. For now, take them as a handy mnemonic for remembering the components of acceleration in the polar-coordinate system.

10.3.1 Reduction of the central force equations to a single equation

Take the equation (10.1)

$$m\left(\ddot{r}-r\dot{\theta}^2\right)=-\frac{\partial\mathcal{U}}{\partial r}. \tag{10.4}$$

Since $h=r^2\dot{\theta}$ is conserved, the 'weird' quantity $mr\dot{\theta}^2$ can be eliminated:

$$mr\dot{\theta}^2=mr\left(\frac{h}{r^2}\right)^2=\frac{mh^2}{r^3}.$$

Hence, Eq. (10.4) becomes

$$m\ddot{r}-\frac{h^2}{r^3}=-\frac{\partial\mathcal{U}}{\partial r}.$$

By introducing the *effective potential*

$$\mathcal{U}_{\text{eff}}:=\tfrac{1}{2}\frac{mh^2}{r^2}+\mathcal{U}, \tag{10.5}$$

the equation of motion becomes quasi one-dimensional:

$$m\ddot{r}=-\frac{\partial\mathcal{U}_{\text{eff}}}{\partial r}.$$

The central-force problem in three dimensions has been reduced to a problem of one-dimensional motion.

Example: In Chapter 8, we showed that conservative forces in many dimensions conserve the energy

$$E = \tfrac{1}{2}m\boldsymbol{v}^2 + \mathcal{U}(r), \qquad r := |\boldsymbol{r}|.$$

Let us study the energy for the gravitational interaction between a particle of mass m and a much more massive particle of mass M, $M \gg m$. We may approximate the particle m as experiencing a central force with a force centre at the location of M,

$$\boldsymbol{F} = -\frac{GMm}{r^2}\hat{\boldsymbol{r}} := F(r)\,\hat{\boldsymbol{r}}.$$

Recall, from Chapter 8, the potential associated with such a central force is

$$\mathcal{U}(\boldsymbol{r}) = -\int^r F(s)\,\mathrm{d}s = GMm\int^r s^{-2}\mathrm{d}s = -\frac{GMm}{r}.$$

The energy is therefore

$$E = \tfrac{1}{2}m\boldsymbol{v}^2 - \frac{GMm}{r}.$$

We can re-write this in several ways, depending on the problem we wish to solve. It is therefore helpful to be familiar with each form:

$$\begin{aligned} E &= \tfrac{1}{2}m\boldsymbol{v}^2 - \frac{GMm}{r}, \\ &= \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{GMm}{r}, \\ &= \tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{2}m\frac{h^2}{r^2} - \frac{GMm}{r}, \\ &= \tfrac{1}{2}m\dot{r}^2 + \mathcal{U}_{\text{eff}}. \end{aligned}$$

Note that this last expression is precisely the expression one would obtain by multiplying the quasi one-dimensional equation $m\ddot{r} = -\mathcal{U}'_{\text{eff}}(r)$ by $\dot{r}$ and integrating w.r.t. time.

More on the effective potential: Consider the effective potential for the basic potential $\mathcal{U} = -\lambda r^{-1}$:

$$\mathcal{U}_{\text{eff}} = \tfrac{1}{2}\frac{mh^2}{r^2} - \frac{\lambda}{r}.$$

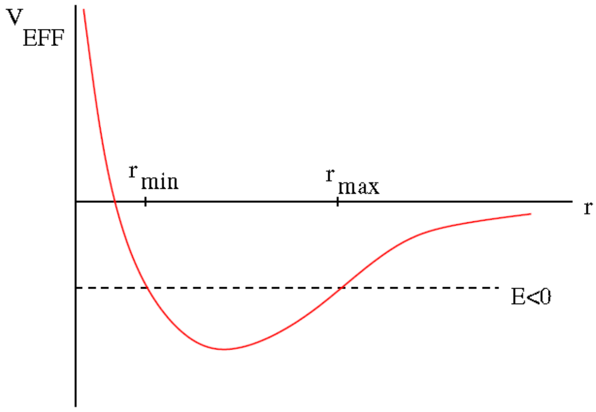


Figure 10.1: Effective potential of an attractive force (sign-negative potential).

- The initial conditions determine the magnitude of h and hence the well depth. They also control the energy E.
- If $E < 0$ the particle is constrained to lie between $r_{\min}$ and $r_{\max}$.
- The particle cannot therefore spiral in towards $r = 0$. The repulsive part of the effective potential prevents this. This is intimately related to the conservation of angular momentum.

Note further, that if the particle resides at the well minimum r_0, then $\mathcal{U}'_{\text{eff}}(r_0) = 0$, $\ddot{r} = 0$, and the particle radius remains constant. This corresponds to a stable circular orbit. There is precisely one such orbit for each h-value.

Escape velocity of a particle in the earth's gravitational field: Suppose that a rocket of mass m is launched from the surface of the earth and that its velocity relative to the centre of the earth is $\boldsymbol{v}_0$. Find the condition that the rocket escapes the earth's pull (ignoring the effect of the atmosphere and the rotation of the earth).

The energy of the rocket-earth system is

$$E = \tfrac{1}{2}m\boldsymbol{v}_0^2 - \frac{GMm}{R}$$

where R is the initial radial location of the rocket (equal to the earth's radius). From considering the effective-potential diagram (Fig. 10.1), the condition for the rocket to escape the earth's pull is that $E = 0$, since then the motion is just barely unbounded. Hence,

$$\tfrac{1}{2}m\boldsymbol{v}_0^2 = \frac{GMm}{R},$$

and

$$|\boldsymbol{v}_0| = \sqrt{\frac{2GM}{R}}, \tag{10.6}$$

which is independent of m.

Now let us include the effects of rotation. Then, $\boldsymbol{v}_0$ contains both a radial part $\dot{r}$, and a tangential part $r_p\dot{\theta}$. The radius r_p is not the distance R to the earth's centre, but rather the distance to the earth's axis of rotation. At the equator however, $r_p = R$, and $|\boldsymbol{v}_0|$ is maximized:

$$\boldsymbol{v}_0^2 = \underbrace{\dot{r}^2}_{\text{Boost given to rocekt from engines}} + \underbrace{R^2\dot{\theta}^2}_{\text{Boost given to rocket from earth's rotation, at equator}}$$

Thus, the best launch sites are near the equator.

Note finally that Eq. (10.6) is independent of mass. We may re-arrange it to give R as a function of $v := \boldsymbol{v}_0$:

$$R = \frac{2GM}{v^2},$$

and apply it to massless particles. In particular, consider the photon, which travels at the speed of light. Then, the radius is

$$R_c = \frac{2GM}{c^2}.$$

For the earth, $R_c \approx 2\,\mathrm{cm}$. Thus, if all the earth's mass were concentrated at a point, no photon within a two $2\,\mathrm{cm}$ radius of this point could escape from that point's gravitational field, and we would have a small black hole.

Chapter 11

Interlude: Physical units

The module has reached its half-way point. It is a good place to stop for a break! In this chapter and in the next one, we therefore return to some basic concepts. First, in this chapter, we review the notion of physical units. Then, in Ch. 12 we revisit the concept of energy and *energy landscapes*. We will then be in a position to tackle the derivation of Kepler's Laws.

In science, there are only seven basic physical units, and all other physical units can be constructed from these seven by multiplication. For mechanics, we are only interested in the first three (or possibly four).

Quantity	SI unit	Symbol
Mass	Kilogram	kg
Length	Metre	m
Time	Second	s
Charge	Coulomb	C
Temperature	Kelvin	K
Amount of substance	Mol	mol
Luminous intensity	Candela	cd

The kilogram is defined with respect to a 'standard' kilogram held in a vault in Paris. The other units are defined relative to physical processes. For example,

> One second is the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom,

and

> One metre is the distance travelled by light in free space in $1/299,792,458$ of a second.

Examples

Energy: The units of energy are not fundamental. For example, the kinetic energy of a body is $mv^2/2$, hence the units of energy are

$$[E] = \frac{\text{Mass} \times \text{Length}^2}{\text{Time}^2},$$

where the square brackets mean 'dimensions of'. Thus, the SI units of energy are $\mathrm{kg\,m^2\,s^{-2}}$, also called the *Joule*.

Force: The units of force are not fundamental. Since $F = ma$,

$$[F] = \frac{\text{Mass} \times \text{Length}}{\text{Time}^2},$$

and the SI units of force are $\mathrm{kg\,m\,s^{-2}}$. This combination of basic units is referred to as the *Newton*.

Hooke's constant: The constant in Hooke's law is not fundamental. Since $F = -kx$, and $[F] = \text{Mass} \times \text{Length}/\text{Time}^2$, we have

$$\frac{\text{Mass} \times \text{Length}}{\text{Time}^2} = [k]\text{Length},$$

hence

$$[k] = \frac{\text{Mass}}{\text{Time}^2}.$$

Note also,

$$[(k/m)^{1/2}] = \frac{1}{\text{Time}},$$

and thus, $\sqrt{k/m}$ is a frequency.

Newton's constant, G The gravitational constant G is a derived quantity, because $F = Gm_1m_2/r^2$, hence

$$\frac{\text{Mass} \times \text{Length}}{\text{Time}^2} = [G]\frac{\text{Mass}^2}{\text{Length}^2},$$

and

$$[G] = \frac{\text{Length}^3}{\text{Time}^2 \times \text{Mass}};$$

the SI units are thus $\mathrm{m^3\,s^{-2}\,kg^{-1}}$.

Dimensional analysis

Sometimes it is possible to solve a problem in mechanics without solving any differential equations. For example, to estimate the energy of a system from the system parameters, we combine those parameters in such a way as to give an energy. This estimate usually gives the correct answer, up to a nondimensional prefactor.

The ground-state energy of hydrogen: In a hydrogen atom, an electron and a proton interact via the Coulomb force

$$F = -\frac{Ke^2}{r^2},$$

where the minus sign indicates attraction, and e is the unit charge. Hence,

$$[Ke^2] = [F] \times \text{Length}^2 = \frac{\text{Mass} \times \text{Length}^3}{\text{Time}^2}$$

In every quantum mechanics problem, Planck's constant h appears, and

$$[h] = \text{Momentum} \times \text{Length} = \frac{\text{Mass} \times \text{Length}^2}{\text{Time}}$$

We expect the mass of the electron to be important too. Thus, we write

$$E = m_e^a h^b \left(Ke^2\right)^c,$$

and solve for a, b, and c. The dimensions must match in this equation:

$$\frac{ML^2}{T^2} = M^a \left(\frac{ML^2}{T}\right)^b \left(\frac{ML^3}{T^2}\right)^c.$$

Or,

$$ML^2T^{-2} = M^{a+b+c}L^{2b+3c}T^{-b-2c}.$$

This gives a system of equations to solve:

$$\begin{aligned} 1 &= a+b+c, \\ 2 &= 2b+3c, \\ -2 &= -b-2c. \end{aligned}$$

Solving the last two equations together gives $c = 2$, $b = -2$. Backsubstitution into the first equation gives $a = 1$. Hence, the binding energy of the hydrogen atom is

$$E_{\text{da}} = \frac{m_e \left(Ke^2\right)^2}{h^2}$$

The true energy from solving the Schrödinger equation is

$$E = -2\pi^2 \frac{m_e \left(Ke^2\right)^2}{h^2},$$

which is not far off our guesstimate!

The gravitational self-energy of the earth: We know that there is a gravitational potential energy between two particles:

$$U = -Gm_1 m_2 r^{-1},$$

where r is the particles' relative separation and m_1 and m_2 denote their masses. However, an extended body (such as the earth) can be treated as a collection of infinitesimal particles. What if we integrate over this collection of particles? We will get an energy that is called the *self-energy* of the extended body. This is the binding energy associated with all the attractive forces interacting within the system.

If we were to estimate the self-energy of the earth, there are only three parameters to play with: G, M_e, the earth mass, and R_e, the earth radius. Based on dimensional analysis, we would guess that

$$U_{\text{self}} = -\frac{GM_e^2}{R_e}.$$

The true value, obtained from integration, is

$$U_{\text{self}} = -\tfrac{3}{5}\frac{GM_e^2}{R_e}.$$

The energy required for a solar system to form from interstellar dust: Imagine a collection of interstellar dust in deep space. Suppose that it has uniform density ρ and that this collection of dust is of extent ℓ. Then its mass is approximately

$$M \approx \rho \left(\tfrac{4}{3}\pi\ell^3\right),$$

and its gravitational self-energy is approximately

$$E_{\text{self}} \approx \frac{GM^2}{\ell}.$$

The particles will have some random motion associated with the finite temperature of the dust. This gives rise to an energy

$$E_{\text{thermal}} \approx N k_B T,$$

where T is the temperature, k_B is Boltzmann's constant, and N is the number of dust particles. Now

$$\rho = \frac{M}{V} = \frac{N m_0}{V} \implies N = \frac{\rho}{m_0}\left(\tfrac{4}{3}\pi\ell^3\right),$$

where m_0 is the mass of each dust particle. Thus we have two energies:

$$\begin{aligned} E_{\text{self}} &= \rho^2 \left(4\pi/3\right)^3 \ell^5, \\ E_{\text{thermal}} &= \frac{\rho\left(\tfrac{4}{3}\pi\ell^3\right) k_B T}{m_0}. \end{aligned}$$

Gravitational collapse occurs if the binding energy can overcome the random thermal motion:

$$E_{\text{self}} > E_{\text{thermal}},$$

and collapse just barely occrus if

$$E_{\text{self}} = E_{\text{thermal}},$$

or

$$\rho\left(\tfrac{4}{3}\pi\ell^3\right)\left(\rho G \tfrac{4}{3}\pi\ell^2\right) = \frac{\rho\left(\tfrac{4}{3}\pi\ell^3\right) k_B T}{m_0}.$$

Effecting the cancellations implied gives

$$\ell = \sqrt{\frac{3 k_B T}{4\pi m_0 \rho G}}.$$

Thus, if the extent of the dust cloud exceeds this critical length (called the Jeans length), gravitational collapse will occur, possibly giving rise to a primordial solar system.

Chapter 12

Interlude: Energy revisited

In this chapter we continue our 'break' before tackling the derivation of Kepler's Laws. Here, we review the concept of energy and *energy landscapes*.

12.1 Some old definitions

- The *position vector* of a particle is a triple $\boldsymbol{x} := (x, y, z) \in \mathbb{R}^3$ that specifies the location of a particle relative to the origin $O = (0, 0, 0)$. The set of Cartesian axes generated in this description can be called a *frame of reference*, S.

- The velocity of a particle relative to the frame of reference S is simply $\boldsymbol{v} := d\boldsymbol{x}/dt$.

- The acceleration of a particle relative to the frame of reference S is $\boldsymbol{a} = d\boldsymbol{v}/dt = d^2\boldsymbol{x}/dt^2$.

- A *force* is anything that causes acceleration.

- Two frames of reference S and S' are *inertial* if they move at constant velocity w.r.t. one another. That is, the position vector $\boldsymbol{R}$ going from origin O to origin O' is such that $d^2\boldsymbol{R}/dt^2 = 0$.

- Relative to an inertial frame, a particle continues to move in a straight line, or stays at rest, unless acted on by an external force.

- When acted on by external forces, a particle satisfies Newton's second law,

$$m\frac{d^2\boldsymbol{x}}{dt^2} = \boldsymbol{F}.$$

- When a particle experiences an external force, *work* is done. The amount of work done by the external force $\boldsymbol{F}$ on the particle as the particle goes from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ is equal to

$$W = \int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{x}.$$

  This is a *path integral*, which you haven't done yet, but in one dimension, it reduces to

$$W = \int_{x_1}^{x_2} F \mathrm{d}x.$$

- *Energy* is the ability to do work. It comes in two kinds: *kinetic* energy, and *potential* energy. Kinetic energy is due to the motion of the particle:

$$K = \tfrac{1}{2} m \boldsymbol{v}^2 = \tfrac{1}{2} m \dot{\boldsymbol{x}}^2.$$

  Potential energy is energy that the particle possesses due to its position.

- Two simple forces we have looked at so far: The *spring force*, $F = -kx$, and the uniform force field, $F = -mg$, where m is the particle's mass and g is a constant field strength.

12.2 Work and energy

Let us take the example of the spring force:

$$F = -kx.$$

That is, a particle at a distance x from the origin experiences a force that tends to pull the particle back towards the origin ('restoring force'). Let us calculate the work done in going from location x_1 to x_2:

$$W = \int_{x_1}^{x_2} F \mathrm{d}x = \int_{x_1}^{x_2} -kx \mathrm{d}x = -\tfrac{1}{2}\left(x_2^2 - x_1^2\right).$$

Let us define a new function

$$\mathcal{U}(x) = -\int^{x} F(s)\,\mathrm{d}s.$$

This is the *potential energy*.

$$W(x_1 \to x_2) = \mathcal{U}(x_1) - \mathcal{U}(x_2) = -\Delta\mathcal{U}.$$

Thus, the work done by a particle in going from x_1 to x_2 is minus the change in potential energy.

One final relation: The work done in going from x_1 to x_2 is the change in kinetic energy:

$$\begin{aligned}
W(x_1 \to x_2) &= \int_{x_1}^{x_2} F \mathrm{d}x, \\
&= \int_{x_1}^{x_2} m\ddot{x} \mathrm{d}x, \\
&= \int_{x_1}^{x_2} m\frac{dv}{dt} dx, \\
&= \int_{x_1}^{x_2} m\frac{dv}{dt}\frac{dx}{dt} dt, \\
&= \int_{x_1}^{x_2} mv\frac{dv}{dt} dt, \\
&= \int_{x_1}^{x_2} \frac{d}{dt}\left(\tfrac{1}{2}mv^2\right), \\
&= \tfrac{1}{2}mv^2|_{x_2} - \tfrac{1}{2}mv^2|_{x_1}.
\end{aligned}$$

Thus, the work done in going from x_1 to x_2 is minus the change in potential energy, OR, it is the change in kinetic energy. So,

$$\text{Change in kinetic energy} = -\text{Change in potential energy},$$

or

$$\left[\tfrac{1}{2}mv^2 + \mathcal{U}\right]_{x_2} = \left[\tfrac{1}{2}mv^2 + \mathcal{U}\right]_{x_1}.$$

Hence, the total energy

$$E = \tfrac{1}{2}mv^2 + \mathcal{U}(x)$$

is constant ('conserved').

12.3 Force and energy

Force is an intuitive concept. Energy is not. It is intuitive to derive a potetnial energy from a force field. However, in applications, we are often given an energy and need to derive the force. Energy is the more natual, albeit abstract, quantity with which to formulate the dynamics. Therefore, let us demonstrate how to derive a force from a potential.

We know that the work done in moving through an increment $\Delta x = x_2 - x_1$ is minus the change in potetnial energy:

$$W = -\Delta\mathcal{U}.$$

But

$$W = F\Delta x.$$

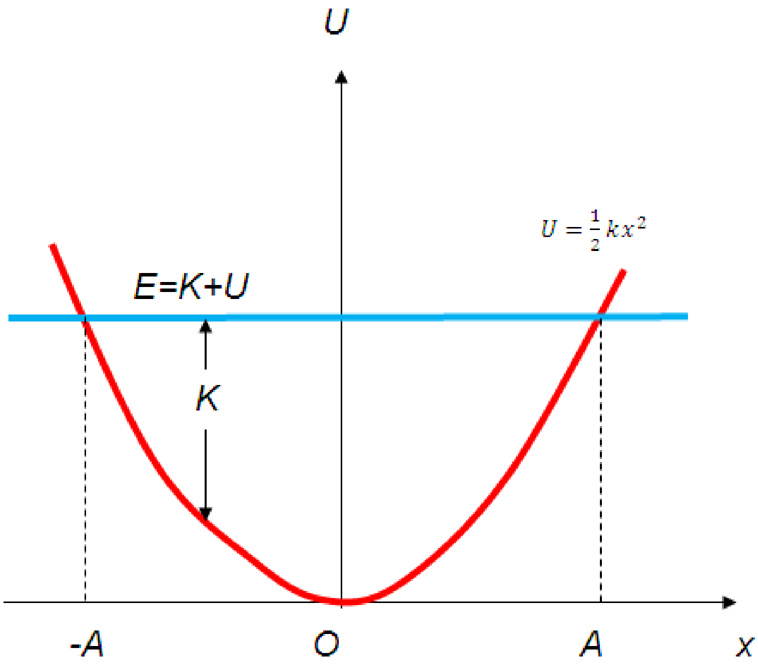


Figure 12.1: Potential-energy plot for the harmonic oscillator.

Hence,

$$F\Delta x = -\Delta\mathcal{U}.$$

Dividing by Δx and letting $\Delta x \to 0$ gives

$$F = -\frac{\Delta\mathcal{U}}{\Delta x} \to -\frac{\mathrm{d}\mathcal{U}}{\mathrm{d}x}.$$

Hence, *force is minus the gradient of potential*. Check this for the spring force, for which $\mathcal{U} = kx^2/2$:

$$F = -\frac{\mathrm{d}}{\mathrm{d}x}\left(\tfrac{1}{2}kx^2\right) = -kx.$$

Let's plot the potential (Fig. 12.1).

- Energy constant and equal to E.
- Kinetic energy greatest at $x = 0$ (where PE is least).
- PE is greatest at $x = \pm A$ (where KE is zero).
- PE cannot be greater than E because KE is positive. Hence, particle cannot pass beyond $x = \pm A$. These are the *turning points* of the motion.
- Result is back-and-forth motion between A and $-A$.

A more general potential energy function is plotted in Fig. 12.2. This is also called the 'energy landscape'.

- Points x_1, x_2, x_3, and x_4 are *equilibria* because $F = 0$ there (hence $\mathrm{d}\mathcal{U}/\mathrm{d}x$ is either a maximum or a minimum.
- The force is directed against the slope of $\mathcal{U}$. Thus, if $\mathcal{U}$ is an increasing function somewhere, then the force at that point is in the negative x-direction.
- Thus, the force near the points x_1 and x_3 is directed towards those points. A particle starting near these points will get 'pushed' towards those points. Hence, *stable equilibria*.
- The force near points x_2 and x_4 is directed away from those points. *Unstable equilibria*.

12.4 Solving equations of motion by quadrature

Consider a particle experiencing the force $F = kx$, a repulsive spring force.

1. Write down the equation of motion and the energy.

   Newton's equation: $m\ddot{x} = F = kx$. Identify $\sigma = \sqrt{k/m}$, hence

   $$\ddot{x} - \sigma^2 x = 0.$$

2. Reduce the motion to an integral using the energy.

   Multiply the equation of motion (EOM) by $\dot{x}$ and integrate w.r.t. time. The result is

   $$\tfrac{1}{2}\dot{x}^2 - \tfrac{1}{2}\sigma^2 x^2 = \varepsilon = E/m,$$

   a constant with units of $[\mathrm{Energy}][\mathrm{mass}]^{-1}$. Solve for dx/dt:

   $$\left(\frac{dx}{dt}\right)^2 = \sigma^2 x^2 + 2\varepsilon \geq 0,$$

   hence ε is required to be nonnegative. Inverting for dt/dx,

   $$\begin{aligned} \frac{dx}{dt} &= \sqrt{2\varepsilon}\sqrt{1 + \frac{\sigma^2}{2\varepsilon}x^2}, \\ \frac{dt}{dx} &= \frac{1}{\sqrt{2\varepsilon}}\frac{1}{\sqrt{1 + \frac{\sigma^2}{2\varepsilon}x^2}}, \\ t &= \frac{1}{\sqrt{2\varepsilon}}\int_{x_0}^{x}\frac{\mathrm{d}x'}{\sqrt{1 + \frac{\sigma^2}{2\varepsilon}x^2}}. \end{aligned}$$

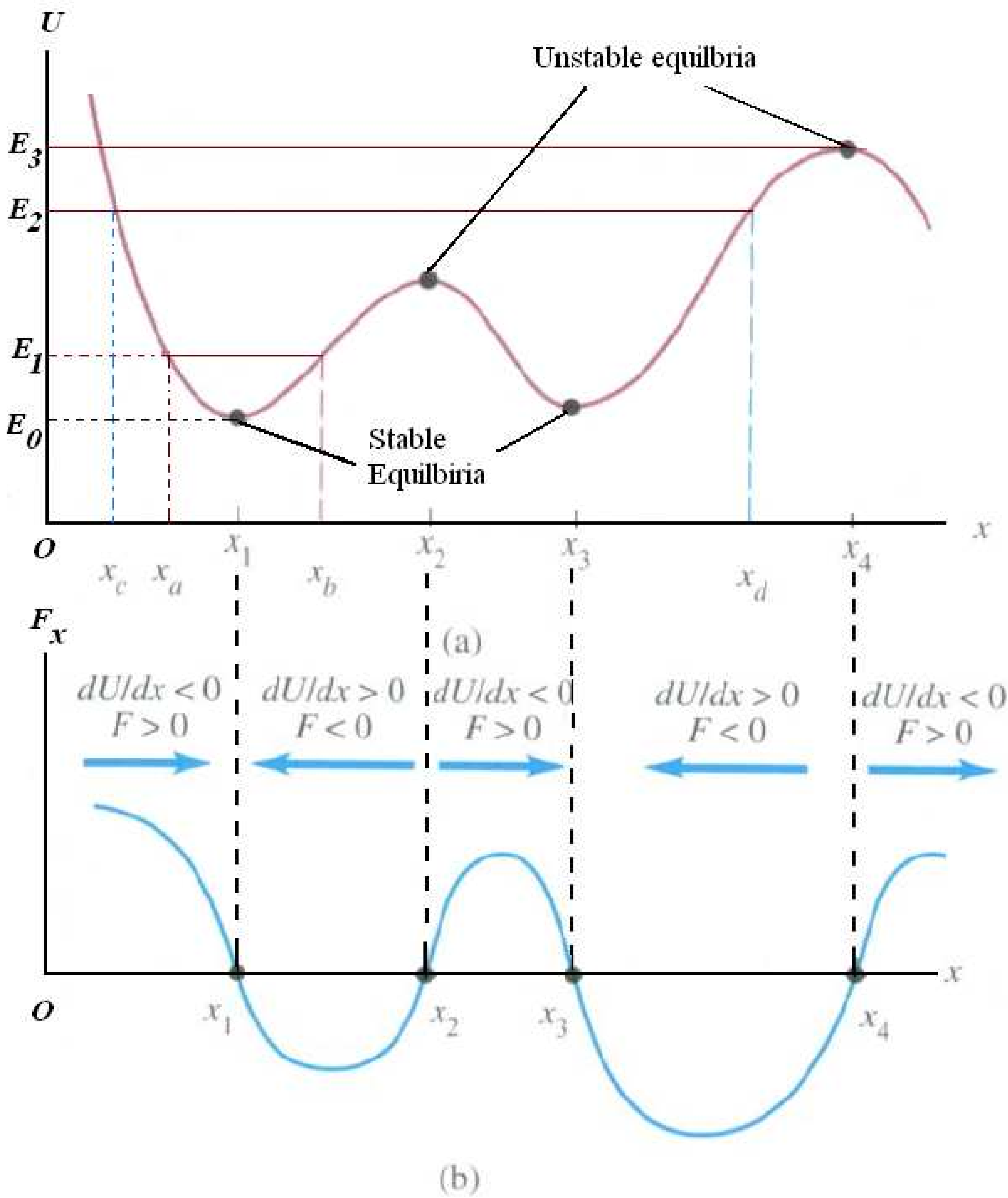


Figure 12.2: Energy landscape of a more general potential function.

3. Solve this integral using any method you know.

   First, we transform to dimensionless variables: $s = \sigma x/\sqrt{2\varepsilon}$. The integral is thus

$$t = \frac{1}{\sigma} \int_{\sigma x_0/\sqrt{2\varepsilon}}^{\sigma x/\sqrt{2\varepsilon}} \frac{\mathrm{d}s}{\sqrt{1+s^2}}.$$

   Looking this integral up in the tables,

$$t = \frac{1}{\sigma} \left[\sinh^{-1} \frac{\sigma x}{\sqrt{2\varepsilon}} - \sinh^{-1} \frac{\sigma x_0}{\sqrt{2\varepsilon}} \right].$$

   Define a constant of integration $\tilde{A}$,

$$\tilde{A} = \sinh^{-1} \frac{\sigma x_0}{\sqrt{2\varepsilon}}.$$

   Hence,

$$\sigma t = \sinh^{-1} \frac{\sigma x}{\sqrt{2\varepsilon}} - \tilde{A} \iff x = \frac{\sqrt{2\varepsilon}}{\sigma} \sinh\left(\sigma t + \tilde{A}\right).$$

   Defining a further constant of integration

$$\tilde{B} = \frac{\sqrt{2\varepsilon}}{\sigma},$$

   the solution is

$$x = \tilde{B} \sinh\left(\sigma t + \tilde{B}\right).$$

   Using $\sinh s = \left(e^s - e^{-s}\right)/2$, this is

$$x = \tfrac{1}{2}\tilde{B}e^{\tilde{A}}e^{\sigma t} - \tfrac{1}{2}\tilde{B}e^{-\tilde{A}}e^{-\sigma t}.$$

   Defining further constants of integration $A = \tilde{B}e^{\tilde{A}}/2$ and $B = -\tilde{B}e^{-\tilde{A}}/2$, this is

$$x = Ae^{\sigma t} + Be^{-\sigma t},$$

   and A and B can be fixed by the initial conditions on x and $\dot{x}$.

Chapter 13

Kepler's Laws

13.1 Overview

In Ch. 10 (Central forces),

- We showed how the motion of a particle experiencing a central force can be reduced to a problem in one-dimensional mechanics, using the effective potential.
- We tackled some problems involving the gravitational force between two point masses in the limit where one particle was more massive than the other. Thus, we treated the lighter particle as one experiencing a central force from a fixed force centre.

In this chapter we shall show how such a reduction is possible for all pairs of particles, regardless of their mass. With this final piece of the jigsaw in place, we shall show how Newton's laws provide the theoretical explanation for the empirical laws of planetary motion observed by the astronomer Kepler.

13.2 Kepler's Laws

The following historical notes are taken from the website

`http://www.hps.cam.ac.uk/starry/`

Nicholas Copernicus (1473-1543), a Polish monk, formulated a *heliocentric view* of the solar system. In this theory, the planets move around the sun in circular orbits. The reasons for this theory were both practical, since it gave a solution to the problem of the 'retrograde motion' of the planets, and aesthetic, since a theory based on circles seemed geometrically 'pure'. Two months after publishing his book, Copernicus died.

Later, Tycho Brahe, (1546-1601)[1], a Danish astronomer, built up an enormous and very precise database over his lifetime concerning the motion of the planets then known[2]. He initially worked at Hven in Denmark but after disagreements with the king, moved to Prague. Brahe's assistant, Johannes Kepler (1571-1630), analyzed these data and formulated three empirical laws of planetary motion [3]:

1. Each planet moves in an ellipse with the sun at one focus

2. The radius vector form the sun to a planet sweeps out equal areas in equal times.

3. The period of revolution, T, is related to the semi-major axis of the ellipse, a, by $T^2 \propto a^3$.

13.3 Centre-of-mass coordinates

We begin by explaining how *all* two-body central-force problems can be reduced to a quasi one-dimensional, quasi one-particle problem. Consider two particles acting under gravity and refer to Fig. 13.1. The force on particle 1 due to particle 2 is

$$\boldsymbol{F}_{12} = -\frac{Gm_1m_2}{|\boldsymbol{x}_1-\boldsymbol{x}_2|^2}\left(\frac{\boldsymbol{x}_1-\boldsymbol{x}_2}{|\boldsymbol{x}_1-\boldsymbol{x}_2|}\right). \tag{13.1a}$$

By Newton's Third Law ,the force on particle 2 due to particle 1 is

$$\boldsymbol{F}_{21} = -\boldsymbol{F}_{12} = -\frac{Gm_1m_2}{|\boldsymbol{x}_1-\boldsymbol{x}_2|^2}\left(\frac{\boldsymbol{x}_2-\boldsymbol{x}_1}{|\boldsymbol{x}_1-\boldsymbol{x}_2|}\right). \tag{13.1b}$$

Let us introduce a new coordinate

$$\boldsymbol{r} = \boldsymbol{x}_1 - \boldsymbol{x}_2,$$

which points from mass 2 to mass 1. We also introduce a radial vector

$$\hat{\boldsymbol{r}} = \boldsymbol{r}/|\boldsymbol{r}| = \frac{\boldsymbol{x}_1-\boldsymbol{x}_2}{|\boldsymbol{x}_1-\boldsymbol{x}_2|}.$$

[1]Brahe lost his nose in a duel in Rostock in 1566. He wore a replacement nose made out of brass. He is one of several scientists to have fought an unsuccessful duel (*cf.* Evariste Galois).

[2]Mercury–Saturn

[3]The precise circumstances under which Kepler gained full access to Brahe's database are not clear.

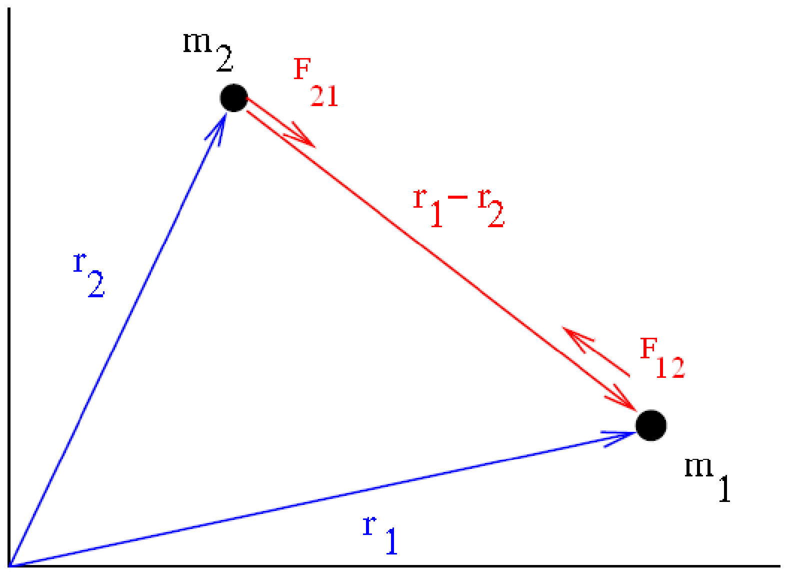


Figure 13.1: Force diagram for a two-particle gravitational interaction.

Then, the force laws (13.1) can be rewritten as

$$\begin{aligned} \boldsymbol{F}_{12} &= -\frac{Gm_1m_2}{|\boldsymbol{r}|^3}\boldsymbol{r}, \\ \boldsymbol{F}_{21} &= +\frac{Gm_1m_2}{|\boldsymbol{r}|^3}\boldsymbol{r}, \end{aligned} \tag{13.2}$$

and the equations of motion are

$$\begin{aligned} m_1\ddot{\boldsymbol{x}}_1 &= -\frac{Gm_1m_2}{|\boldsymbol{r}|^3}\boldsymbol{r}, \\ m_2\ddot{\boldsymbol{x}}_2 &= +\frac{Gm_1m_2}{|\boldsymbol{r}|^3}\boldsymbol{r}. \end{aligned} \tag{13.3}$$

Let's add these equations (Eqs. (13.3)) together:

$$m_1\ddot{\boldsymbol{x}}_1 + m_2\ddot{\boldsymbol{x}}_2 = 0.$$

This shows that the *centre-of-mass vector*

$$\boldsymbol{R} := \frac{m_1\boldsymbol{x}_1 + m_2\boldsymbol{x}_2}{m_1 + m_2}$$

does not accelerate. That is, *a frame of reference in which the centre of mass and the origin coincide is an inertial frame*. Next, let us subtract Eqs. (13.3) from each other:

$$\begin{aligned}\ddot{\boldsymbol{x}}_1 &= -\frac{Gm_2}{|\boldsymbol{r}|^3}\boldsymbol{r},\\ \ddot{\boldsymbol{x}}_2 &= +\frac{Gm_1}{|\boldsymbol{r}|^3}\boldsymbol{r}.\end{aligned}$$

Let's SUBTRACT these equations one from the other:

$$\ddot{\boldsymbol{r}} = \ddot{\boldsymbol{x}}_1 - \ddot{\boldsymbol{x}}_2 = -\frac{G\left(m_1+m_2\right)}{|\boldsymbol{r}|^3}\boldsymbol{r}. \tag{13.4}$$

Define the *reduced mass*:

$$\mu := \frac{m_1 m_2}{m_1+m_2}. \tag{13.5}$$

Multiply Eq. (13.4) by the reduced mass:

$$\mu\ddot{\boldsymbol{r}} = -\frac{Gm_1m_2}{|\boldsymbol{r}|^3}\boldsymbol{r}. \tag{13.6}$$

Therefore, if we can solve the problems $\ddot{\boldsymbol{R}} = 0$ and Eq. (13.6), we can reconstruct $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ via the transformations

$$\begin{aligned}\begin{pmatrix}\boldsymbol{r}\\ \boldsymbol{R}\end{pmatrix} &= \begin{pmatrix}1 & -1\\ \frac{m_1}{m_1+m_2} & \frac{m_2}{m_1+m_2}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{pmatrix},\\ \begin{pmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{pmatrix} &= \begin{pmatrix}\frac{m_2}{m_1+m_2} & 1\\ -\frac{m_1}{m_1+m_2} & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{r}\\ \boldsymbol{R}\end{pmatrix}\end{aligned}$$

Moreover, the angular momentum

$$\boldsymbol{J} = \mu\boldsymbol{r}\times\boldsymbol{v}, \qquad \boldsymbol{v} = \dot{\boldsymbol{r}},$$

as measured relative to the location of m_2, is conserved:

$$\begin{aligned}\frac{d\boldsymbol{J}}{dt} &= \mu\dot{\boldsymbol{r}}\times\boldsymbol{v} + \mu\boldsymbol{r}\times\frac{d\boldsymbol{v}}{dt},\\ &= \boldsymbol{r}\times\left(\mu\ddot{\boldsymbol{r}}\right),\\ &= \boldsymbol{r}\times\left(\frac{Gm_1m_2}{|\boldsymbol{r}|^3}\boldsymbol{r}\right),\\ &= 0.\end{aligned}$$

Similarly, the energy

$$E = \tfrac{1}{2}\mu \boldsymbol{v}^2 - \frac{Gm_1m_2}{r}, \qquad \mathcal{U} = -\frac{Gm_1m_2}{r},$$

is conserved:

$$\begin{aligned}
\frac{dE}{dt} &= \mu\boldsymbol{v}\cdot\dot{\boldsymbol{v}} + \boldsymbol{v}\cdot\nabla\mathcal{U}\left(|\boldsymbol{r}|\right), \\
&= \boldsymbol{v}\cdot\left(\mu\ddot{\boldsymbol{r}} + \mathcal{U}'\left(|\boldsymbol{r}|\right)\nabla|\boldsymbol{r}|\right), \\
&= \boldsymbol{v}\cdot\left(\mu\ddot{\boldsymbol{r}} + \mathcal{U}'\left(|\boldsymbol{r}|\right)\frac{\boldsymbol{r}}{|\boldsymbol{r}|}\right), \\
&= \boldsymbol{v}\cdot\left(\mu\ddot{\boldsymbol{r}} - \boldsymbol{F}\right), \\
&= 0.
\end{aligned}$$

Thus, the motion can be reduced to one-dimensional particle motion in the $\boldsymbol{r}$–$\boldsymbol{v}$ plane.

For the earth-sun system, the centre-of-mass vector $\boldsymbol{R}$ and the position vector of the sun $\boldsymbol{x}_2 := \boldsymbol{x}_\odot$ almost coincide:

$$\begin{aligned}
m_2 = M_\odot &= 2\times 10^{30}\,\mathrm{kg}, \\
m_1 = m_\mathrm{e} &= 6\times 10^{24}\,\mathrm{kg}, \\
\frac{m_\mathrm{e}}{m_\odot} &= 3\times 10^{-6}.
\end{aligned}$$

$$\begin{aligned}
\frac{|\boldsymbol{R}|}{|\boldsymbol{x}_1|} = \frac{|\boldsymbol{R}|}{|\boldsymbol{x}_\mathrm{e}|} = \frac{|m_\mathrm{e}\boldsymbol{x}_\mathrm{e} + M_\odot\boldsymbol{x}_\odot|}{|\boldsymbol{x}_\mathrm{e}|\left(m_\mathrm{e}+M_\odot\right)} &\le \frac{m_\odot}{M_\odot + m_\mathrm{e}}\frac{|\boldsymbol{x}_\odot|}{|\boldsymbol{x}_\mathrm{e}|} + \frac{m_\mathrm{e}}{m_\mathrm{e}+M_\odot} \\
&= \frac{1}{1+(m_\mathrm{e}/M_\odot)}\frac{|\boldsymbol{x}_\odot|}{|\boldsymbol{x}_\mathrm{e}|} + \frac{(m_\mathrm{e}/M_\odot)}{1+(m_\mathrm{e}/M_\odot)} \\
&= \frac{1}{1+3\times 10^{-6}}\frac{|\boldsymbol{x}_\odot|}{|\boldsymbol{x}_\mathrm{e}|} + \frac{3\times 10^{-6}}{1+3\times 10^{-6}} \approx \frac{|\boldsymbol{x}_\odot|}{|\boldsymbol{x}_\mathrm{e}|}.
\end{aligned}$$

Therefore, to simplify the solution of the equations $\ddot{\boldsymbol{R}} = 0$, $\mu\ddot{r} = -\left(GM_\odot m_\mathrm{e}/|\boldsymbol{r}|^3\right)\boldsymbol{r}$, we solve the equations in the *centre-of-mass coordinates*

$$\boldsymbol{R} = 0,$$

wherein the centre-of-mass location vector coincides with the origin. Hence, the origin or our coordinate system and the centre of the sun almost coincide:

$$\boldsymbol{x}_\odot \approx \boldsymbol{R}.$$

The same procedure works for the other planets too.

13.4 Conclusions

- Written down Kepler's Laws;
- Introduced centre of mass coordinates and the relative vector $\boldsymbol{r} = \boldsymbol{x}_1 - \boldsymbol{x}_2$ for two-body problems.
- Shown how the two-body gravitational interaction reduces to a quasi one-particle problem with a reduced mass.
- Shown how the the reduced mass gives rise to a conserved energy and angular momentum.
- Next, we will show how to solve this equation fully.

Chapter 14

Kepler's First Law

14.1 Overview

In Ch. 13 (Kepler's Laws),

- We wrote down Kepler's Laws;
- We introduced centre of mass coordinates and the relative vector for two-body problems (see Fig. 14.1).
- We showed how the two-body gravitational interaction reduces to a quasi one-particle problem with a reduced mass.
- We showed how the the reduced mass gives rise to a conserved energy and angular momentum.
- Now, we will show how to solve this equation fully.

14.2 A more detailed summary of the story so far

Recall the following quantities introduced in the last lecture:

- Relative vector $\boldsymbol{r} = \boldsymbol{x}_1 - \boldsymbol{x}_2$.
- COM vector $\boldsymbol{R} = (m_1\boldsymbol{x}_1 + m_2\boldsymbol{x}_2) / (m_1 + m_2)$.
- Gravitational force on particle 1 due to particle 2: $\boldsymbol{F}_{12} = -\left(Gm_1m_2/|\boldsymbol{x}_1 - \boldsymbol{x}_2|^3\right)(\boldsymbol{x}_1 - \boldsymbol{x}_2)$.
- Newton's equation: $m_1\ddot{\boldsymbol{x}}_1 = \boldsymbol{F}_{12}$, $m_2\ddot{\boldsymbol{x}}_2 = -\boldsymbol{F}_{12}$.

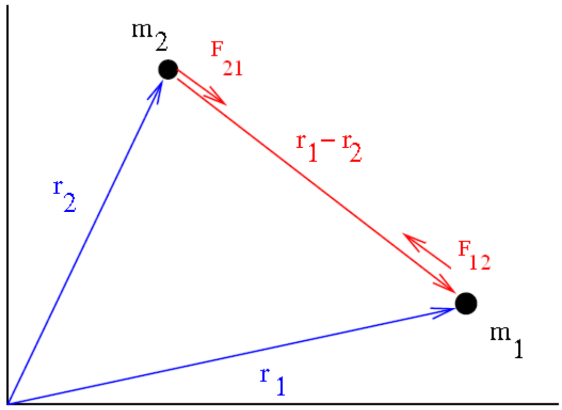


Figure 14.1: Force diagram for a two-particle gravitational interaction.

- Reduced mass: $\mu = m_1 m_2 / (m_1 + m_2)$.

- Reduced equation:

$$\mu \ddot{\boldsymbol{r}} = -\frac{G m_1 m_2}{|\boldsymbol{r}|^3} \boldsymbol{r}. \tag{14.1}$$

Writing Eq. (14.1) in polar coordinates, obtain

$$\mu \left(\ddot{r} - r\dot{\theta}^2 \right) = -\frac{G m_1 m_2}{r^2}, \tag{14.2}$$

$$\mu \left(r\ddot{\theta} + 2\dot{r}\dot{\theta} \right) = 0. \tag{14.3}$$

Writing Eq. (14.3) as

$$\frac{\mu}{r} \frac{d}{dt} \left(r^2 \dot{\theta} \right) = 0,$$

which shows that the angular momentum $J = \mu r^2 \dot{\theta}$ is conserved. Thus, $h = r^2 \dot{\theta}$ is also conserved.

Next, substitute $\dot{\theta} = h/r^2$ into Eq. (14.2):

$$\mu \ddot{r} - \frac{\mu h^2}{r^3} = -\frac{G m_1 m_2}{r^2}.$$

Multiply the equation by $\dot{r}$ and recast i.t.o. perfect derivatives:

$$\frac{d}{dt} \left(\tfrac{1}{2} \mu \dot{r}^2 \right) + \frac{d}{dt} \left(\frac{1}{2} \frac{\mu h^2}{r^2} \right) = \frac{d}{dt} \left(\frac{G m_1 m_2}{r} \right).$$

Identify the energy:

$$E = \tfrac{1}{2}\mu\dot{r}^2 + \mathcal{U}_{\text{eff}} = \text{Const.}, \qquad \mathcal{U}_{\text{eff}} = \tfrac{1}{2}\frac{\mu h^2}{r^2} - \frac{Gm_1m_2}{r}. \tag{14.4}$$

Of course, since $h = r^2\dot{\theta}$, the energy can be re-expressed as

$$\begin{aligned} E &= \tfrac{1}{2}\mu\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{Gm_1m_2}{r}, \\ &= \tfrac{1}{2}\mu\boldsymbol{v}^2 + \mathcal{U}, \qquad \boldsymbol{v} = \dot{\boldsymbol{r}} = \frac{d}{dt}\left(\boldsymbol{x}_1 - \boldsymbol{x}_2\right) \end{aligned}$$

14.3 Solution of the orbit problem

The formal solution of the equation of motion is given as follows: Take the energy equation (14.4):

$$E = \tfrac{1}{2}\mu\dot{r}^2 + \mathcal{U}_{\text{eff}}, \qquad \mathcal{U}_{\text{eff}} = \tfrac{1}{2}\frac{\mu h^2}{r^2} - \frac{Gm_1m_2}{r}.$$

Re-express dr/dt as

$$\begin{aligned} \frac{dr}{dt} &= \sqrt{\frac{2}{\mu}}\sqrt{E - \mathcal{U}_{\text{eff}}}, \\ \frac{dt}{dr} &= \sqrt{\frac{\mu}{2}}\frac{1}{\sqrt{E - \frac{1}{2}\frac{\mu h^2}{r^2} + \frac{Gm_1m_2}{r}}}, \\ t &= \sqrt{\frac{\mu}{2}}\int_0^r \frac{\mathrm{d}s}{\sqrt{E - \frac{1}{2}\frac{\mu h^2}{s^2} + \frac{Gm_1m_2}{s}}}. \end{aligned}$$

The solution can now be found using a table of integrals.

However, to prove Kepler's First Law, we need to find r as a function of θ. We use the chain rule:

$$\begin{aligned} \frac{dr}{d\theta} &= \frac{dr}{dt}\frac{dt}{d\theta} \\ &= \frac{dr}{dt}\frac{1}{\dot{\theta}} \\ &= \frac{\mu r^2}{J}\frac{dr}{dt}, \qquad \text{since } J = \mu r^2\dot{\theta} = \text{Const.} \end{aligned}$$

Using Eq. (14.5), this is

$$\begin{aligned} \frac{dr}{d\theta} &= \frac{\mu r^2}{J}\sqrt{\frac{2}{\mu}}\sqrt{E-\mathcal{U}_{\text{eff}}\left(r\right)}, \\ \frac{d\theta}{dr} &= \frac{J}{\sqrt{2\mu}}\frac{1}{r^2}\frac{1}{\sqrt{E-\mathcal{U}_{\text{eff}}\left(r\right)}}, \end{aligned}$$

$$\theta-\theta_0=\frac{J}{\sqrt{2\mu}}\int_{r_0}^{r}\frac{\mathrm{d}s}{s^2\sqrt{E-\mathcal{U}_{\text{eff}}\left(s\right)}}, \tag{14.5}$$

where we have chosen a coordinate system such that $\theta\left(r_0\right)=\theta_0$. Now, the solution to the orbit problem is reduced to calculating an integral.

14.4 Evaluation of the orbital integral

Consider the orbital integral (14.5). We re-express is slightly:

$$\begin{aligned} \theta-\theta_0 &= \frac{J}{\sqrt{2\mu}}\int_{r_0}^{r}\frac{\mathrm{d}s}{s^2\sqrt{E-\frac{1}{2}\frac{J^2}{\mu s^2}+\frac{Gm_1m_2}{s}}}, \\ &= J\int_{r_0}^{r}\frac{\mathrm{d}s}{s\sqrt{2\mu Es^2-J^2+2\mu Gm_1m_2s}}. \end{aligned}$$

This is in fact a standard integral that can be looked up in a table:

$$\mathcal{I}:=\int\frac{\mathrm{d}s}{s\sqrt{\gamma s^2+\beta s-\alpha}}. \tag{14.6}$$

Here

$$\begin{aligned} \alpha &= J^2, \\ \beta &= 2\mu Gm_1m_2:=2\mu B, \\ \gamma &= 2\mu E; \end{aligned}$$

α and β are both positive, while γ can be positive or negative, depending on where in the effective-potential well the particle sits. Abramowitz and Stegun (AS)[1] gives

$$\begin{aligned} \mathcal{I} &= \int \frac{\mathrm{d}s}{s\sqrt{\gamma s^2 + \beta s - \alpha}} = -\int \frac{\mathrm{d}t}{\sqrt{\gamma + \beta t - \alpha t^2}}, \qquad t = 1/s, \qquad \text{Eq. 3.3.38, AS,} \\ &= \frac{1}{\sqrt{\alpha}} \sin^{-1}\left(\frac{-2\alpha t + \beta}{\sqrt{\beta^2 + 4\alpha\gamma}}\right), \qquad \text{Eq. 3.3.36, AS.} \end{aligned}$$

Hence,

$$\begin{aligned} \theta - \theta_0 &= \sin^{-1}\left(\frac{-2J^2\left(1/r\right) + 2\mu B}{\sqrt{4\mu^2 B^2 + 8J^2\mu E}}\right), \\ &= \sin^{-1}\left(\frac{1}{r}\frac{\mu B r - J^2}{\sqrt{\mu^2 B^2 + 2\mu E J^2}}\right), \end{aligned}$$

or

$$\mu B r - J^2 = r\sqrt{\mu^2 B^2 + 2\mu E J^2}\,\sin\left(\theta - \theta_0\right).$$

Solving for r,

$$r = \frac{J^2/\mu B}{1 - \sqrt{1 + (2EJ^2/\mu B^2)}\sin\left(\theta - \theta_0\right)}. \tag{14.7}$$

The usual convention is to take $\theta_0 = -\pi/2$ and to introduce the following parameters:

$$\begin{aligned} r_0 &:= \frac{J^2}{\mu B}, \\ \epsilon &:= \sqrt{1 + \frac{2EJ^2}{\mu B^2}}. \end{aligned} \tag{14.8}$$

Physically, r_0 is the radius of the circular orbit corresponding to the minimum of the effective potential well $\mathcal{U}_{\text{eff}}\left(r\right)$, and to the given values of J, μ, and B. The dimensionless parameter ϵ is called the *eccentricity*, and determines the shape of the orbit. To see how, let us re-write Eq. (14.7) with the replacements (14.8):

$$r = \frac{r_0}{1 - \epsilon\cos\theta}.$$

[1] This is a famous table of integrals and special functions and is now out of copyright: http://www.math.ucla.edu/∼cbm/aands/

In Cartesian coordinates $x = r\cos\theta$, $y = r\sin\theta$,

$$\begin{aligned} r\left(1-\epsilon\cos\theta\right) &= r_0, \\ r-\epsilon r\cos\theta &= r_0, \\ \sqrt{x^2+y^2}=\epsilon x &= r_0, \end{aligned}$$

$$x^2\left(1-\epsilon^2\right)-2r_0\epsilon x+y^2=r_0^2. \tag{14.9}$$

Equation (14.9) is the equation of a *conic section*, which can be of several forms, depending on ϵ:

1. $\epsilon = 0$ means that Eq. (14.9) becomes $x^2+y^2=r_0^2=J^2/\left(\mu B\right)$, which is the equation of a circle.

2. $0 \le \epsilon < 1$. We re-write the conic-section equation by completing the square in x:

$$\left(1-\epsilon^2\right)\left(x+\frac{\epsilon r_0}{\left(1-\epsilon^2\right)}\right)^2+y^2=r_0^2+\frac{\epsilon^2 r_0^2}{1-\epsilon^2}=\frac{r_0^2}{1-\epsilon^2},$$

   which is of the form

$$\frac{\left(x-x_0\right)^2}{a^2}+\frac{y^2}{b^2}=1,$$

   which is the equation of an ellipse (Fig. 14.2) with *semimajor axis* a.

Note: In Fig. 14.2, we have defined

$$a=\frac{r_0}{1-\epsilon^2}, \qquad b=\frac{r_0}{\sqrt{1-\epsilon^2}}=a\sqrt{1-\epsilon^2}. \tag{14.10}$$

3. $\epsilon = 1$, which gives

$$-2r_0x+y^2=r_0^2, \implies x=\tfrac{1}{2}\left(\frac{y^2}{r_0}-r_0\right).$$

   This is the equation of a *parabola* (Fig. 14.3, which is not a closed curve. This cannot describe a particle in orbit, because it is not closed; it describes a particle which comes in from infinitely far away, is deflected by the sun but never comes back.

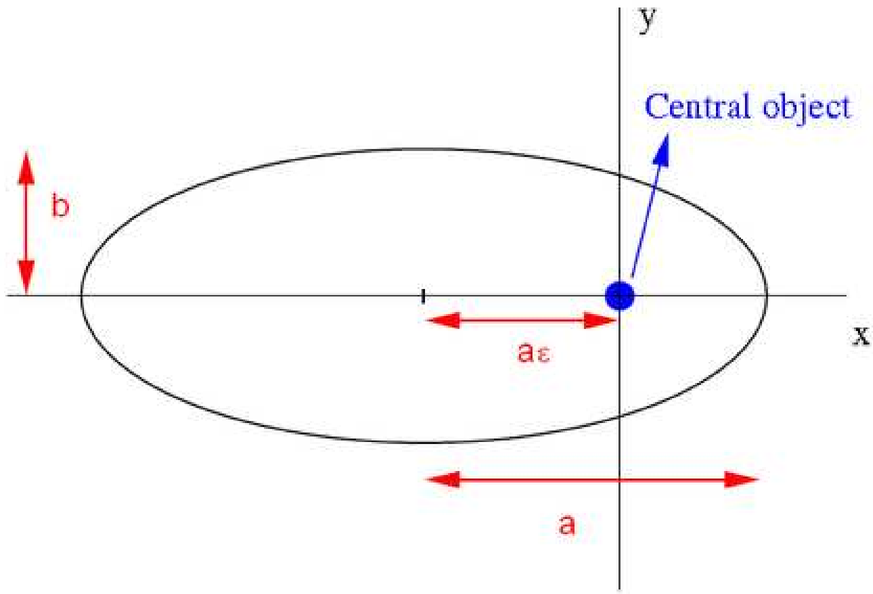


Figure 14.2: Ellipse: a is the semimajor axis, b is the semiminor axis, and ϵ is the eccentricity. The central object at $(x, y) = 0$ is at the *focus*.

4. $\epsilon > 1$ In this case, after some manipulation similar to the elliptic case [Exercise], we get the equation of a hyperbola,
$$\frac{(x - x_0)^2}{a^2} - \frac{y^2}{b^2} = 1,$$
and this is not a closed curve - it is describing a particle that is not bound to the central object but merely deflected by it.

14.4.1 More on the ellipse

The way in which we have defined the ellipse and introduced the semimajor axis a seems a little contrived. The more natural way of defining an ellipse is through the the specification of two *foci* O and O'. Then the ellipse is the set of all points P such that

$$|OP| + |O'P| = \text{Const.} = 2a. \tag{14.11}$$

Suppose that the focus O is at point $(0, 0)$ of a Cartesian system, and that the point O' is at point $(-2c, 0)$, where $0 < c < a$ (Fig. 14.4). Now,

$$\overrightarrow{O'P} = (2c + x, y)\,, \qquad \overrightarrow{OP} = (x, y)\,.$$

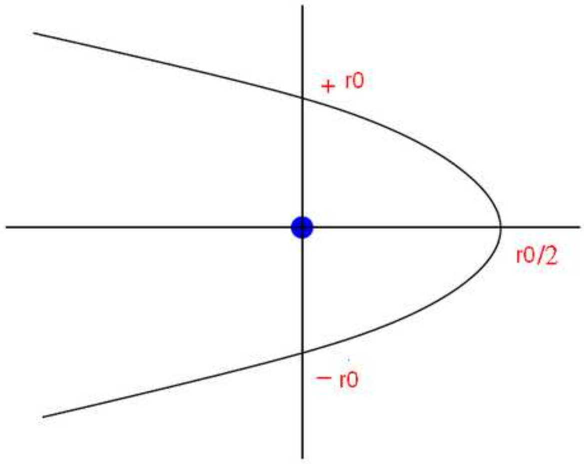


Figure 14.3: Parabola

Hence

$$\sqrt{(2c+x)^2+y^2}+\sqrt{x^2+y^2}=2a.$$

Bring the second term to the right-hand side and square both sides:

$$\begin{aligned}\sqrt{(2c+x)^2+y^2} &= 2a-\sqrt{x^2+y^2},\\ (2c+x)^2+y^2 &= 4a^2+x^2+y^2-4a\sqrt{x^2+y^2}.\end{aligned}$$

Solve for the square root:

$$\sqrt{x^2+y^2}=\frac{1}{4a}\left\{4a^2+x^2+y^2-\left[(2c+x)^2+y^2\right]\right\}$$

Simplify this result:

$$\sqrt{x^2+y^2}=\frac{a^2-c^2-cx}{a}.$$

Eliminate the final square root:

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

or,

$$x^2\left(1-\frac{c^2}{a^2}\right)+2c\left(1-\frac{c^2}{a^2}\right)x+y^2=a^2\left(1-\frac{c^2}{a^2}\right)^2. \tag{14.12}$$

Since $c<a$ by definition, define $\epsilon := c/a < 1$. Hence, Eq. (14.12) becomes

$$x^2\left(1-\epsilon^2\right)+2\epsilon a\left(1-\epsilon^2\right)x+y^2=a^2\left(1-\epsilon^2\right)^2. \tag{14.13}$$

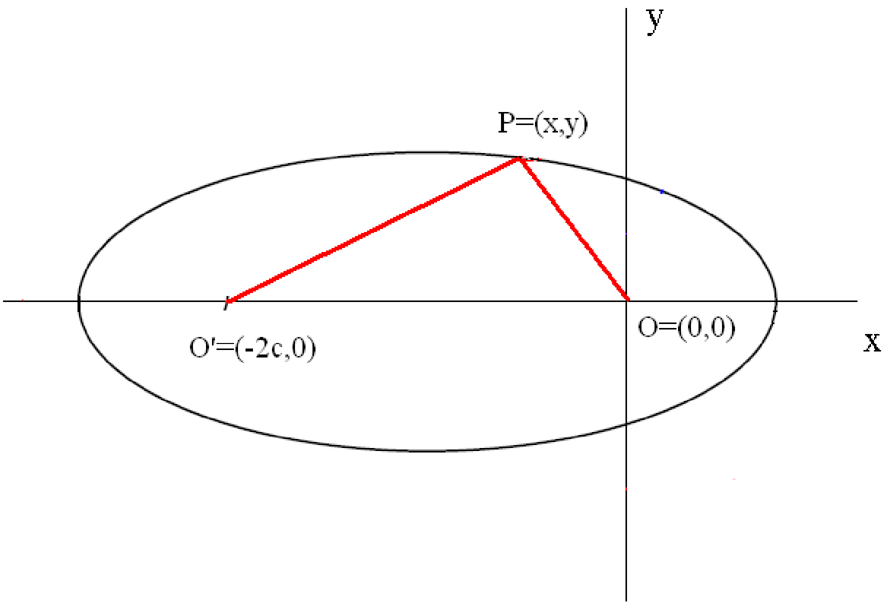


Figure 14.4: Ellipse: The sum $|OP|+|O'P|$ is constant and equal to $2a$. The constant c is necessarily less than a because when $P=(x>0,0)$, $2c+2|OP|=2c+2x=2a$, hence $c<a$.

Completing the square in x gives

$$(x+\epsilon a)^2+\frac{y^2}{1-\epsilon^2}=a^2,$$

which is the equation of an ellipse. Shifting the x-coordinate leftwards by

$$x'=x+\epsilon a=x+c,$$

and defining $b^2:=a^2-c^2=a^2\left(1-\epsilon^2\right)$ gives rise to the canonical form for the equation of the ellipse:

$$\frac{x'^2}{a^2}+\frac{y^2}{b^2}=1.$$

14.4.2 The Energy and ϵ

Recall formula (14.8):

$$\epsilon=\sqrt{1+\frac{2EJ^2}{\mu B^2}}, \qquad B=Gm_1m_2.$$

Square this formula and bring the one over to the LHS. Hence, invert for E:

$$\begin{aligned}
\epsilon^2 - 1 &= \frac{2EJ^2}{\mu B^2}, \\
&= \frac{2E}{\mu B^2} \times (\mu B r_0)\,, \ \ldots \text{ from Eq. (14.8)} \\
&= \frac{2Er_0}{B}, \\
\implies E &= \frac{B}{2r_0}\left(\epsilon^2 - 1\right),
\end{aligned}$$

$$E = \frac{Gm_1 m_2}{2r_0}\left(\epsilon^2 - 1\right).$$

Therefore, a particle's energy determines whether $\epsilon < 1$ or $\epsilon \geq 1$, as a result whether the particle goes into orbit

- $\epsilon \geq 1 \implies E \geq 0$: these particles have enough kinetic energy to equal, or exceed, potential energy. They are not bound to the central object and do not go into orbit but their paths are deflected into a parabolic or hyperbolic path.
- $\epsilon < 1 \implies E < 0$: these particles have a negative total energy, and their energy is dominated by the potential term. They are bound to the central object and follow an elliptic or circular path.

In the case of small eccentricies $\epsilon < 1$, we can substitute the semi-major axis:

$$E = -\frac{Gm_1 m_2}{2a}.$$

This is a very useful formula because we can use it to relate a planet or satellite's speed to its position by

$$E = -\frac{GMm}{2a} = \tfrac{1}{2}m\boldsymbol{v}^2 - \frac{GMm}{r}$$

(we have set $\mu = m$ here for an earth-satellite calculation). This gives (Eq. (14.10))

$$v^2 = GM\left(\frac{2}{r} - \frac{1}{a}\right).$$

14.4.3 Some numerical examples

- For the earth, $\epsilon \approx= 1/60$, $a \approx 1.5 \times 10^8\,\mathrm{km}$, and the semiminor axis is $b = a\sqrt{1-\epsilon^2} = 0.99986a$, so the earth orbit is almost circular.
- For Halley's comet, $\epsilon = 0.9674$.

14.5 Conclusions

We have now proved Kepler's first law, namely that the trajectories of the planets are ellipses, with the sun at one focus. It remains to prove the second and the third laws. These are relatively easy compared with the first law.

Chapter 15

Kepler's Second and Third Laws

15.1 Overview

In Ch. 14 (Kepler's First Law),

- We proved Kepler's first law from first principles, namely that the orbits of planets are ellipses with the sun at one focus.
- Now we prove the second and third laws.
- First, we need to determine the area of an ellipse.

15.2 Area integrals

In this section we derive some expressions regarding area that are needed to prove Kepler's second and third laws.

15.2.1 The area of an ellipse

Recall the definition of an ellipse: Given the two *foci* O and O', the ellipse is the set of all points P such that

$$|OP| + |O'P| = \text{Const.} = 2a. \tag{15.1}$$

Suppose that the focus O is at point $(0,0)$ of a Cartesian system, and that the point O' is at point $(-2c,0)$, where $0 < c < a$ (Fig. 15.1). Now,

$$\overrightarrow{O'P} = (2c + x, y)\,, \qquad \overrightarrow{OP} = (x, y)\,.$$

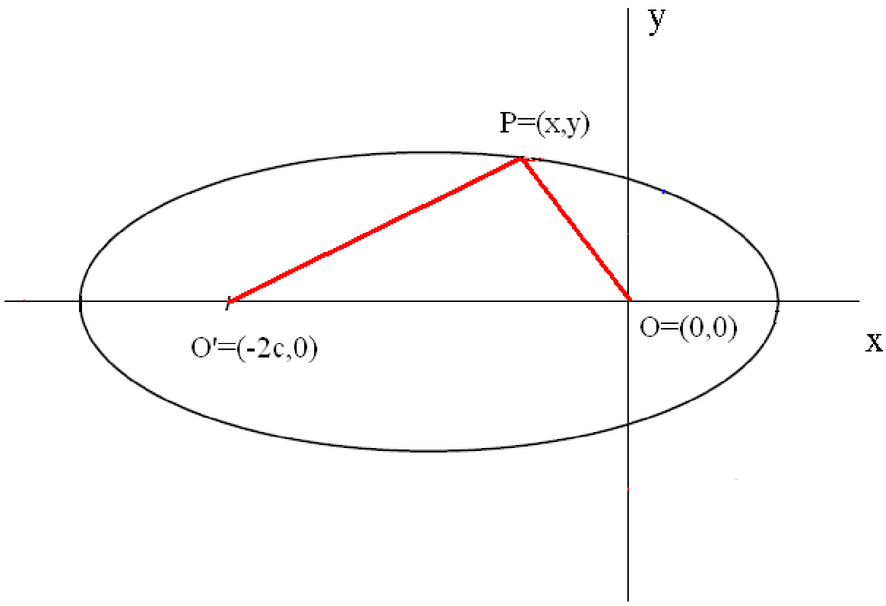


Figure 15.1: Ellipse: The sum $|OP|+|O'P|$ is constant and equal to $2a$. The constant c is necessarily less than a because when $P=(x>0,0)$, $2c+2|OP|=2c+2x=2a$, hence $c<a$.

Hence
$$\sqrt{(2c+x)^2+y^2}+\sqrt{x^2+y^2}=2a.$$
Bring the second term to the right-hand side and square both sides:
$$\begin{aligned}\sqrt{(2c+x)^2+y^2} &= 2a-\sqrt{x^2+y^2},\\ (2c+x)^2+y^2 &= 4a^2+x^2+y^2-4a\sqrt{x^2+y^2}.\end{aligned}$$
Solve for the square root:
$$\sqrt{x^2+y^2}=\frac{1}{4a}\left\{4a^2+x^2+y^2-\left[(2c+x)^2+y^2\right]\right\}$$
Simplify this result:
$$\sqrt{x^2+y^2}=\frac{a^2-c^2-cx}{a}.$$
Eliminate the final square root:
$$x^2+y^2=\frac{(a^2-c^2)^2-2c\,(a^2-c^2)\,x+c^2x^2}{a^2},$$
or,
$$x^2\left(1-\frac{c^2}{a^2}\right)+2c\left(1-\frac{c^2}{a^2}\right)x+y^2=a^2\left(1-\frac{c^2}{a^2}\right)^2. \tag{15.2}$$

Since $c < a$ by definition, define $\epsilon := c/a < 1$. Hence, Eq. (15.2) becomes

$$x^2\left(1-\epsilon^2\right) + 2\epsilon a\left(1-\epsilon^2\right)x + y^2 = a^2\left(1-\epsilon^2\right)^2. \tag{15.3}$$

Completing the square in x gives

$$(x+\epsilon a)^2 + \frac{y^2}{1-\epsilon^2} = a^2,$$

which is the equation of an ellipse. Shifting the x-coordinate leftwards by

$$x' = x + \epsilon a = x + c,$$

and defining $b^2 := a^2 - c^2 = a^2\left(1-\epsilon^2\right)$ gives rise to the canonical form for the equation of the ellipse:

$$\frac{x'^2}{a^2} + \frac{y^2}{b^2} = 1, \tag{15.4}$$

being the equation of an ellipse with foci at $x = \pm c = \epsilon a$, with semimajor axis a and semiminor axis b, $b^2 = a^2 - c^2$.

Theorem: The area enclosed by the curve (15.4) is πab.

Proof: Divide the ellipse into infinitesimal strips, like the one shown in the figure. The area of the strip is

$$dA = 2y\,dx = 2b\sqrt{1-\frac{x^2}{b^2}},$$

where $-a \le x \le a$. Integrate:

$$\begin{aligned} A &= \int dA = \int_{-a}^{a} 2b\sqrt{1-\frac{x^2}{a^2}}\mathrm{d}x, \\ &= 2ab\int_{-1}^{1}\sqrt{1-s^2}\,\mathrm{d}s, \\ &= \pi ab. \end{aligned}$$

15.2.2 The area of a sector

A sector of a curve is the area swept out by the radius vector in moving through an angle $\mathrm{d}\theta$. It is shown in Fig. 15.3. We compute it as follows.

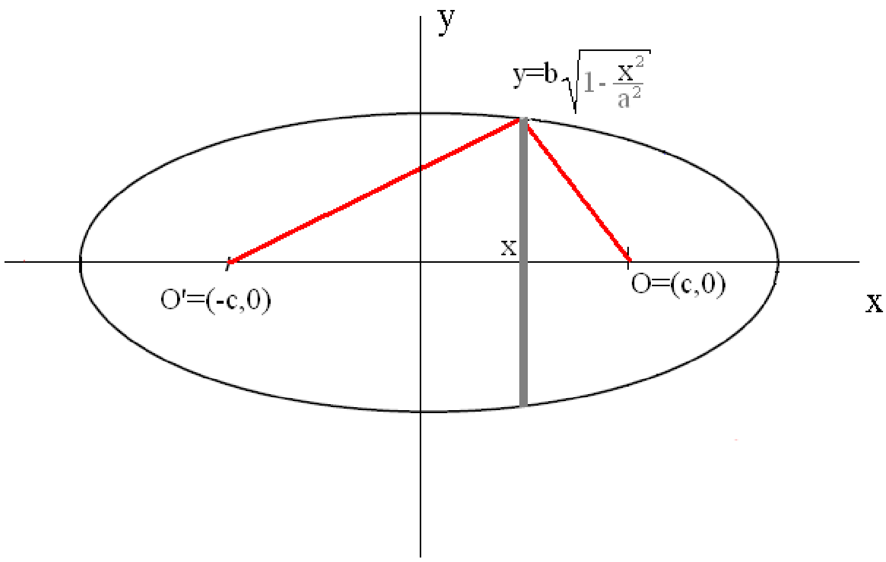


Figure 15.2: Computing the area enclosed by an ellipse.

The area element is

$$\mathrm{d}A = \mathrm{d}x\mathrm{d}y.$$

Using the Jacobian transformation (or geometric reasoning),

$$\mathrm{d}A = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{vmatrix} \mathrm{d}r\mathrm{d}\theta.$$

But $x = r\cos\theta$, $y = r\sin\theta$

$$\begin{aligned} \frac{\partial x}{\partial r} &= \cos\theta \quad , & \frac{\partial y}{\partial r} &= \sin\theta, \\ \frac{\partial x}{\partial \theta} &= -r\sin\theta \quad , & \frac{\partial y}{\partial r} &= r\cos\theta, \end{aligned}$$

hence

$$\mathrm{d}A = r\,\mathrm{d}r\mathrm{d}\theta.$$

A *sector* is the area between 0 and R, and between θ and $\theta + \mathrm{d}\theta$:

$$dA_{\text{sec}} = \int_0^R r\,\mathrm{d}r\mathrm{d}\theta = \tfrac{1}{2}R^2\mathrm{d}\theta.$$

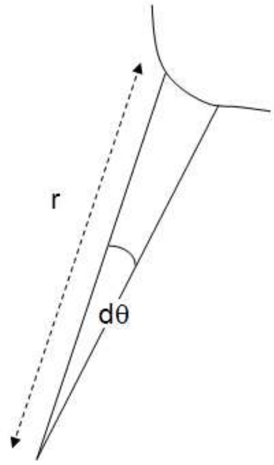


Figure 15.3: The sector is the area swept out by the radius vector in moving through an angle $d\theta$.

15.3 Kepler's second and third laws

2. The radius vector form the sun to a planet sweeps out equal areas in equal times.

3. The period of revolution, T, is related to the semi-major axis of the ellipse, a, by $T^2 \propto a^3$.

These laws follow directly from the conservation of angular momentum:

$$J = \mu r^2 \dot{\theta} = \text{Const.},$$

valid for any central force. Indeed, Law 2 is not unique to planets, it is true for any central-force interaction.

Law 2: The area swept out by the radius vector in a time dt is

$$dA = \tfrac{1}{2} r^2 d\theta = \tfrac{1}{2} r^2 \frac{d\theta}{dt} dt = \tfrac{1}{2} r^2 \left(\frac{J}{\mu r} \right) dt,$$

hence

$$\frac{dA}{dt} = \frac{J}{2\mu} = \text{Const.}$$

Law 3: The area swept out in a finite time interval t is therefore

$$A = \frac{J}{2\mu} t.$$

In a period of revolution, the area swept out by the radius vector must equal the area of the curve traced out by the radius vector – the area of an ellipse:

$$\begin{aligned} \pi ab &= \frac{J}{2\mu}T, & (15.5) \\ \pi a^2\sqrt{1-\epsilon^2} &= \frac{J}{2\mu}T. & (15.6) \end{aligned}$$

But the eccentricity ϵ is (Chapter 14, Kepler's First Law)

$$\epsilon = \sqrt{1 + \frac{2EJ^2}{\mu G^2 m_1^2 m_2^2}},$$

while the energy is expressible i.t.o. the semimajor axis (Chapter 14):

$$E = -\frac{Gm_1m_2}{2a}.$$

Hence,

$$\begin{aligned} 1-\epsilon^2 &= -\frac{2EJ^2}{\mu G^2 m_1^2 m_2^2}, \\ &= \frac{Gm_1m_2}{2a}\frac{2J^2}{\mu G^2 m_1^2 m_2^2}, \\ &= \frac{J^2}{\mu G m_1 m_2}a^{-1}, \\ \implies \sqrt{1-\epsilon^2} &= \frac{J}{\sqrt{\mu G m_1 m_2}}a^{-1/2}. \end{aligned}$$

Put this back into Eq. (15.6):

$$\begin{aligned} \pi a^2\sqrt{1-\epsilon^2} &= \frac{J}{2\mu}T, \\ \pi a^{3/2}\frac{J}{\sqrt{\mu G m_1 m_2}} &= \frac{J}{2\mu}T, \\ 2\pi a^{3/2}\frac{\sqrt{\mu}}{\sqrt{Gm_1m_2}} &= T. \end{aligned}$$

Now let $m_1 = M_\odot$ and $m_2 = m_{\rm p}$, a planetary mass. In this case, $\mu \approx m_{\rm p}$, the planet's mass, and

$$\sqrt{\mu/(m_1m_2)} \approx \sqrt{m_{\rm p}/(m_{\rm p}M_\odot)} = 1/\sqrt{M_\odot}.$$

Hence,

$$T = \frac{2\pi}{\sqrt{GM_\odot}} a^{3/2}, \tag{15.7}$$

and $T \propto a^{3/2}$, and the constant of proportionality is independent of the planet in question and only depends on the solar mass. Having now proved all of Kepler's laws, in the next Chapter, we turn to applications.

Chapter 16

Applications of Kepler's Laws in planetary orbits

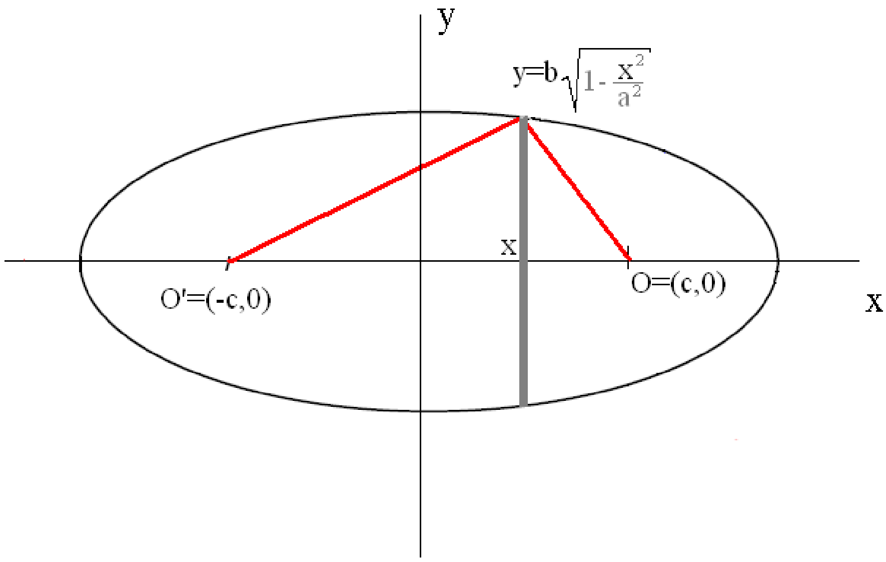


Figure 16.1: Computing the area enclosed by an ellipse.

16.1 An earth satellite

An earth satellite of mass m_0 is fired into orbit at horizontal speed $32,000\,\mathrm{km/hr}$ at an altitude of $640\,\mathrm{km}$ (Fig. 16.2). Discuss the motion. Assume that the earth radius is $R = 6,400\,\mathrm{km}$, and that $g = GM_\mathrm{e}R^{-2} = 9.8\,\mathrm{m/s^2}$.

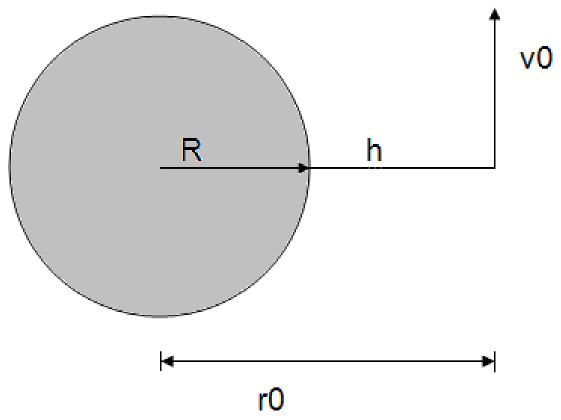


Figure 16.2: Satellite motion around the earth

Velocity:

$$\begin{aligned} v_0 &= 32,000\,\mathrm{km/hr} = \frac{3.2\times 10^7}{60\times 60}\mathrm{m/s}, \\ &= 8,900\,\mathrm{m/s}. \end{aligned}$$

Height above earth:

$$h = 640\mathrm{km} \; = 6.4\times 10^5\,\mathrm{m}.$$

Earth radius:

$$R = 6,400\mathrm{km} = 6.4\times 10^6\,\mathrm{m}.$$

Initial altitude of satellite orbit:

$$r_0 = R + h = 7.04\times 10^6\,\mathrm{m}.$$

Potential energy:

$$\begin{aligned} g &= GM_\mathrm{e}/R^2 \implies GM_\mathrm{e} = gR^2 = \left(9.864\,\mathrm{m/s^2}\right)\times\left(6.4\times 10^6\,\mathrm{m}\right)^2, \\ GM_\mathrm{e} &= 4.014\times 10^{14}\,\mathrm{m^3/s^2}, \\ \mathcal{U} &= -GM_\mathrm{e}m/r_0 = -m_0\left(GM\mathrm{e}/r_0\right) = -m_0\left(5.702\times 10^7\right). \end{aligned}$$

Angular momentum:

$$\begin{aligned} J/m_0 \approx J/\mu = r_0 v_0 &= 6.258\times 10^{10}\,\mathrm{m^2/\,s}, \\ &\approx 6.3\times 10^{10}\,\mathrm{m^2/\,s}. \end{aligned}$$

Total energy:

$$\begin{aligned} E &= \tfrac{1}{2}\mu v_0^2 - \frac{GM_\mathrm{e}m_0}{r_0}, \\ &\approx \tfrac{1}{2}m_0 v_0^2 - \frac{GM_\mathrm{e}m_0}{r_0}, \\ &= m_0\left[3.951\times 10^7 - 5.702\times 10^7\right], \\ &= -m_0\left(1.751\times 10^7\,\mathrm{m}^2/\,\mathrm{s}^2\right). \end{aligned}$$

The energy is negative, so the orbit is an ellipse.

The semimajor axis: Recall the formula

$$E = -\frac{GM_\mathrm{e}m_0}{2a} \implies a = -\frac{GM_\mathrm{e}m_0}{2E}$$

for the semimajor axis of an elliptical orbit. Hence,

$$\begin{aligned} a &= \frac{GM}{2\left(E/m_0\right)} = \frac{4.014\times 10^{14}\,\mathrm{m}^3/\mathrm{s}^2}{2\times 1.751\times 10^7\,\mathrm{m}^2/\,\mathrm{s}^2}, \\ &= 1.146\times 10^7\,\mathrm{m}, \\ &\approx 1.1\times 10^7\,\mathrm{m}. \end{aligned}$$

The eccentricity:

$$\begin{aligned} \epsilon &= \sqrt{1+\frac{2EJ^2}{\mu G^2 m_1^2 m_2^2}}, \\ &\approx \sqrt{1+\frac{2EJ^2}{\left(GM_\mathrm{e}\right)^2 m_0^3}}, \\ &\approx \sqrt{1+\frac{2\left(E/m_0\right)\left(J/m_0\right)^2}{\left(GM_\mathrm{e}\right)^2}}, \\ &= \sqrt{1-\frac{2\left(1.751\times 10^7\,\mathrm{m}^2/\,\mathrm{s}^2\right)\left(6.258\times 10^{10}\,\mathrm{m}^2/\,\mathrm{s}\right)^2}{\left(4.014\times 10^{14}\,\mathrm{m}^3/\mathrm{s}^2\right)^2}}, \\ &= 0.3857, \\ &\approx 0.39. \end{aligned}$$

For a generic ellipse $(x/a)^2 + (y/b)^2 = 1$ (Fig. 16.1), the maximum radius occurs for along the line $y = 0$. Hence, $x = \pm a$. Applying this result to the orbital problem in hand, the maximum distance from the earth is

$$r_{\mathrm{max}} = a + c = a + \epsilon a = a\left(1+\epsilon\right).$$

This is called the *aphelion* of the orbit – *the point furthest from the focus of the orbit* attained by the satellite. Similarly,

$$r_{\rm min} = a - c = a - \epsilon a = (1 - \epsilon)\,.$$

This is the *perihelion*[1]. [Exercise: Prove that $\dot{r} = 0$ at the perihelion and the aphelion. Hint: Work in polar coordinates.]

The maximum height of the satellite above the earth: This is

$$r_{\rm max} - R = 956\,{\rm km}.$$

The period of the orbit: From Kepler's Third Law,

$$\begin{aligned} T &= \frac{2\pi}{\sqrt{GM_{\rm e}}} a^{3/2}, \\ &= 1.223 \times 10^3\,{\rm s}, \\ &= 3.4\,{\rm hrs}. \end{aligned}$$

16.2 A satellite touching the orbits of Earth and Neptune

The orbits of Earth and Neptune are approximately circular and coplanar relative to a fixed force centre, the Sun. Neptune takes 164.8 years to complete an orbit around the Sun. Suppose that a satellite is given an orbit around the sun, extending from the orbit of Earth to that of Neptune, and just touching both. Find the orbital period of the satellite, assuming that the gravitational effects of the planets can be ignored. Refer to Fig. 16.3. Let a be the semimajor axis of the satellite's orbit and let ϵ be the eccentricity. Then, by construction,

$$\begin{aligned} r_{\rm min} &= a\,(1 - \epsilon) = a_E, \\ r_{\rm max} &= a\,(1 + \epsilon) = a_N. \end{aligned}$$

By adding these two equations to eliminate ϵ, obtain

$$a = \tfrac{1}{2}\,(a_E + a_N)\,.$$

But everything is now expressible in terms of T_E, the orbital period of the Earth:

$$\begin{aligned} T_E^2 &= \frac{4\pi^2}{GM_\odot} a_E^3, \\ T_N^2 &= \frac{4\pi^2}{GM_\odot} a_N^3 = (164.8)^2\, T_E^2 = (164.8)^2\, \frac{4\pi^2}{GM_\odot} a_E^3, \end{aligned}$$

[1]In Greek, 'apo' means 'from', 'peri' means 'around', and 'helios' means sun.

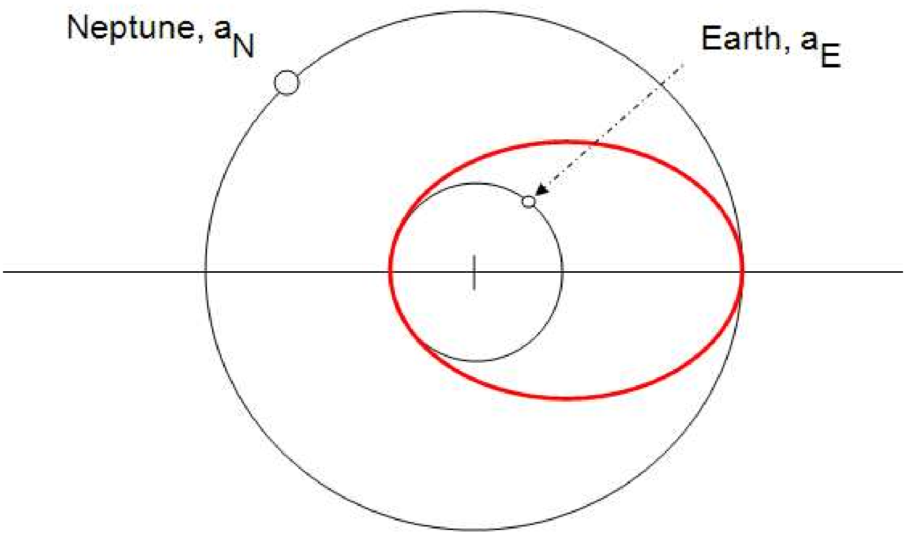


Figure 16.3: Satellite motion centred at the sun and touching the orbits of Earth and Neptune

or

$$a_N^3 = (164.8)^2\, a_E^3, \qquad a_N = (164.8)^{2/3}\, a_E.$$

Hence,

$$a = \tfrac{1}{2}\left[1 + (164.8)^{2/3}\right] a_E.$$

Finally,

$$\begin{aligned}
T^2 &= \frac{4\pi^2}{GM_\odot} a^3, \\
&= \frac{4\pi^2}{GM_\odot} a_E^3 \left[\tfrac{1}{2} + \tfrac{1}{2}(164.8)^{2/3}\right]^2, \\
&= T_E^2 \left[\tfrac{1}{2} + \tfrac{1}{2}(164.8)^{2/3}\right]^3, \\
T &= T_E \left[\tfrac{1}{2} + \tfrac{1}{2}(164.8)^{2/3}\right]^{3/2},
\end{aligned}$$

or $T = 61.205\, T_E = 61.20\,\mathrm{yrs}$.

16.3 Halley's comet

Halley's comet is in an elliptic orbit around the sun. The eccentricity of the orbit is 0.967 and the period is 75.3 years.

1. Using these data, determine the distance of Halley's comet from the sun at perihelion and

aphelion.

We have $r = r_0/(1-\varepsilon\cos\theta)$, hence $r_{\min} = r_0/(1+\epsilon)$, and $r_{\max} = r_0/(1-\epsilon)$. But from the properties of the ellipse, $r_{\min} = a - c = a - \epsilon a = a(1-\epsilon)$, and $r_{\max} = a + c = a(1+\varepsilon)$. We know ϵ, so it suffices to compute a. From Kepler's Third Law,

$$T = \frac{2\pi a^{3/2}}{\sqrt{GM_\odot}},$$

hence

$$a = \left(\frac{T\sqrt{GM_\odot}}{2\pi}\right)^{2/3} = \left(\frac{76\times 365.25\times 24\times 60\times 60\times\sqrt{6.6730\times 10^{-11}\times 1.9889\times 10^{30}}}{2\pi}\right)^{2/3}$$
$$= 2.667620671686515\,(e+12)\ \mathrm{m} \approx 2.67\times 10^{12}\,\mathrm{m}$$

Hence,

$$\begin{aligned} r_{\min} &= a(1-0.967) = 8.80\times 10^{10}\,\mathrm{m} = 0.588\,\mathrm{AU}, \\ r_{\max} &= a(1+0.967) = 5.25\times 10^{12}\,\mathrm{m} = 35.1\,\mathrm{AU}, \end{aligned}$$

Note:

$$1\,\mathrm{AU} = 149,597,870.700 \pm 0.003\mathrm{km}.$$

2. What is the speed of Halley's comet when it is closest to the sun? We know

$$E = \tfrac{1}{2}m\boldsymbol{v}^2 - \frac{GM_\odot m}{r} = -\frac{GM_\odot m}{2a},$$

hence

$$\tfrac{1}{2}m\boldsymbol{v}^2 = \frac{GM_\odot m}{r_{\min}} - \frac{GM_\odot m}{2a}.$$

The m's cancel, and we are left with

$$\begin{aligned} v_0^2 &= 2GM_\odot\left(\frac{1}{r_{\min}} - \frac{1}{2a}\right), \\ &= 2GM_\odot\left(\frac{1}{a(1-\varepsilon)} - \frac{1}{2a}\right), \\ &= \frac{GM_\odot}{a}\left(\frac{1+\varepsilon}{1-\varepsilon}\right). \end{aligned}$$

Plugging in the numbers,

$$\begin{aligned}
v_0 &= \sqrt{\frac{GM_\odot}{a}\left(\frac{1+\varepsilon}{1-\varepsilon}\right)}, \\
&= \sqrt{\frac{6.6730\times10^{-11}\times1.9889\times10^{30}}{2.66762\times10^{12}}\left(\frac{1+0.967}{1-0.967}\right)}, \\
&= 5.445656070140230\,(e+04)\ \mathrm{m/s} \approx 5.45\times10^{4}\,\mathrm{m/s}.
\end{aligned}$$

Chapter 17

Mathematical analysis of orbits

17.1 Plausible solar systems, part 1

For what values of p are circular orbits stable with the potential energy $\mathcal{U}(r) = -C/(pr^p)$, where $C > 0$?

From the expression

$$\begin{aligned} \mathcal{U}_{\text{eff}}(r) &= \tfrac{1}{2}\frac{J^2}{mr^2} - \frac{C}{r^p}, \\ &= \frac{A}{2r^2} - \frac{C}{pr^p}. \end{aligned}$$

we see that in order to have a well, we need $p < 2$. Then, the potential behaves as r^{-2} as $r \to 0$, and as $-(C/p)\, r^{-p}$ as $r \to \infty$.

More formally, the potential $\mathcal{U}_{\text{eff}}(r)$ needs to have a minimum:

$$\frac{d\mathcal{U}_{\text{eff}}}{dr} = -\frac{A}{r^3} + \frac{C}{r^{p+1}}.$$

Setting this to zero gives

$$r_0^{2-p} = A/C.$$

Clearly we must have $p \neq 2$, since then

$$r_0 = \left(\frac{A}{C}\right)^{\frac{1}{2-p}},$$

and

$$\begin{aligned} \left.\frac{d^2\mathcal{U}_{\text{eff}}}{dr^2}\right|_{r_0} &= \frac{3A}{r_0^4} - \frac{(p+1)\,C}{r_0^{p+2}}, \\ &= \frac{1}{r_0}\left[\frac{3A}{r_0^3} - \frac{pC}{r_0^{p+1}} - \frac{C}{r_0^{p+1}}\right], \\ &= \frac{1}{r_0}\left[\frac{2A}{r_0^3} - \frac{pC}{r_0^{p+1}} + \frac{A}{r_0^3} - \frac{C}{r_0^{p+1}}\right], \\ &= \frac{1}{r_0}\left[\frac{2A}{r_0^3} - \frac{pC}{r_0^{p+1}} + \mathcal{U}'_{\text{eff}}\,(r_0)\right], \\ &= \frac{1}{r_0}\left[\frac{2C}{r_0^{p+1}} - \frac{pC}{r_0^{p+1}}\right], \\ &= \frac{1}{r_0}\frac{C}{r_0^{p+1}}\,(2-p)\,. \end{aligned}$$

Taking $p < 2$ therefore gives a well minimum, as opposed to a maximum, which would be unstable.

17.2 Plausible solar systems, part 2

Now, consider planar motion under the potential $\mathcal{U}\,(r) = Cr^4$ and angular momentum J. For what energy will the motion be circular, and what is the radius of the circle?

The effective potential is

$$\mathcal{U}_{\text{eff}}\,(r) = \tfrac{1}{2}\frac{J^2}{mr^2} + Cr^4,$$

Circular motion corresponds to a minimum in the effective potential,

$$\mathcal{U}'_{\text{eff}}\,(r_0) = 0 \iff \frac{J^2}{mr_0^3} = 4Cr_0^3,$$

hence

$$r_0 = \left(\frac{J^2}{4mC}\right)^{1/6}.$$

Since

$$E = \tfrac{1}{2}\dot{r}^2 + \mathcal{U}_{\text{eff}}\,(r)\,,$$

and since $\dot{r} = 0$ for circular motion, the energy of circular motion corresponds to the minimum in the effective potential:

$$E = \mathcal{U}_{\text{eff}}\,(r_0)\,,$$

$$
\begin{aligned}
E &= \tfrac{1}{2}\frac{J^2}{mr_0^2} + Cr_0^4, \\
&= \frac{1}{2r_0^2}\left(\frac{J^2}{m} + cr_0^6\right) \\
&= \tfrac{5}{8} r_0^{-2}\frac{J^2}{m}, \\
&= \tfrac{5}{8}\frac{J^2}{m}\left(\frac{4mC}{J^2}\right)^{1/3}.
\end{aligned}
$$

Chapter 18

Planning a trip to the moon

18.1 Overview

In this lecture, we come up with a simple (although not necessarily optimal) method to plan a trip to the moon. This consists of three steps:

1. We launch our rocket into a low-earth, circular orbit.
2. We boost our rocket by a velocity increment Δv_1 to give the rocket an elliptic trajectory around the earth.
3. When the rocket is near the moon, we give it another boost Δv_2, to put it back into a circular orbit around the earth, but a circular orbit that shadows the moon's orbit (a co-rotating orbit).

18.2 Hohmann transfer orbit

The schematic diagram of steps (1)–(3) is shown in Fig. 18.1. Note that particular care must be taken to launch the mission when the moon is at a certain inclination in the sky. This strategy is called a *Hohmann transfer orbit*, and is an approximate solution of the full, three-body problem. Another approximate strategy is called the *Method of patched conics*. During the Apollo 11 mission, the scientists at Parkes Observatory NSW radio telescope used patched conics to re-orient the telescope to receive transmissions, after a storm disrupted the operations.[1]

Computation of Δv at the first transfer point : We use the formula

[1]See: http://www.parkes.atnf.csiro.au/news_events/apollo11/

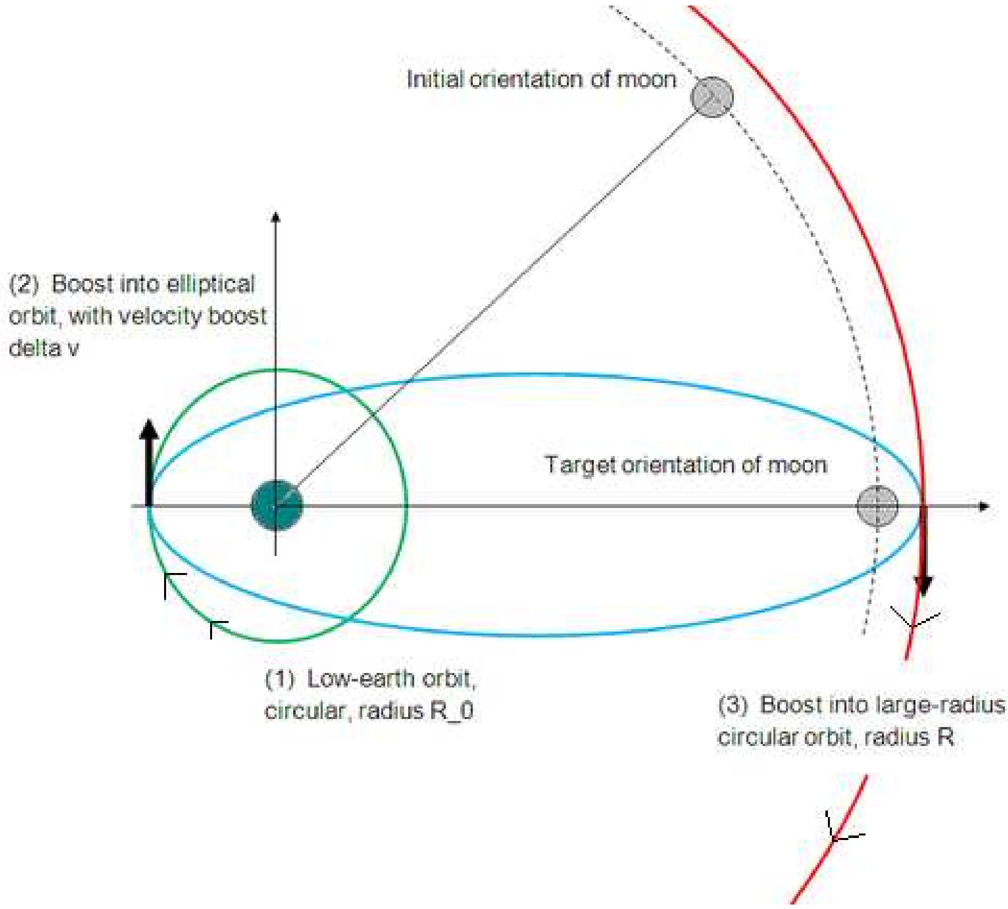


Figure 18.1: A Hohmann transfer orbit

$$E = \tfrac{1}{2}mv^2 - \frac{GM_\mathrm{e}m}{r} = -\frac{GM_\mathrm{e}m}{2a}$$

At the first transfer point, the radius vector is $r = R_0$, the orbit is circular, of radius R_0, and the energy is

$$E_{\mathrm{init},1} = \tfrac{1}{2}mv_{0,1}^2 - \frac{GM_\mathrm{e}m}{R_0} = -\frac{GM_\mathrm{e}m}{2R_0},$$

hence

$$v_{0,1}^2 = \frac{GM_\mathrm{e}}{R_0}.$$

After the *instantaneous* boost, the radius vector is still R_0, but the orbit is elliptical, and has the required semimajor axis $a = (R_0 + R)/2$. Hence,

$$E_{\mathrm{fin},1} = \tfrac{1}{2}mv_1^2 - \frac{GM_\mathrm{e}m}{R_0} = -\frac{GM_\mathrm{e}m}{2a}$$

Thus,

$$\begin{aligned} v_{0,1} &= \sqrt{\frac{GM_{\mathrm{e}}}{R_0}}, \\ v_1 &= \sqrt{\frac{2GM_{\mathrm{e}}}{R_0} - \frac{2GM_{\mathrm{e}}}{R_0 + R}}. \end{aligned}$$

The delta-vee required for the transfer is

$$\Delta v_1 = v_1 - v_{0,1} = \sqrt{\frac{2GM_{\mathrm{e}}}{R_0} - \frac{2GM_{\mathrm{e}}}{R_0 + R}} - \sqrt{\frac{GM_{\mathrm{e}}}{R_0}}.$$

Tidying this up gives

$$\Delta v_1 = \sqrt{\frac{GM_{\mathrm{e}}}{R_0}} \left[\sqrt{\frac{2R}{R + R_0}} - 1 \right],$$

and the direction of the delta-vee must be the same as the direction of $\boldsymbol{v}_0$.

Transiting between the circles After the instantaneous boost Δv_1 has been performed, we allow the rocket to coast for a time $T/2$, where

$$T = \frac{2\pi}{\sqrt{GM_{\mathrm{e}}}} \left[(R_0 + R) / 2 \right]^{3/2}.$$

Thus, the rocket completes one half-period of the elliptic orbit.

Computation of Δv at the second transfer point : The second transfer point is at the aphelion of the transfer ellipse, for which

$$E_{\mathrm{init},2} = \tfrac{1}{2} m v_{0,2}^2 - \frac{GM_{\mathrm{e}} m}{R} = -\frac{GM_{\mathrm{e}} m}{R_0 + R},$$

hence

$$v_{0,2}^2 = 2GM_{\mathrm{e}} \left(\frac{1}{R} - \frac{1}{R_0 + R} \right).$$

After the second boost, the velocity is v_2, and

$$E_{\mathrm{fin},2} = \tfrac{1}{2} m v_2^2 - \frac{GM_{\mathrm{e}} m}{R} = -\frac{GM_{\mathrm{e}} m}{2R},$$

hence

$$v_2^2 = \frac{GM_{\mathrm{e}}}{R}.$$

The delta-vee is then

$$\Delta v_2 = v_2 - v_{0,2} = \sqrt{\frac{GM_{\mathrm{e}}}{R}} - \sqrt{2GM_{\mathrm{e}} \left(\frac{1}{R} - \frac{1}{R_0 + R} \right)}.$$

Tidying this up gives

$$\Delta v_2 = \sqrt{\frac{GM_\mathrm{e}}{R}}\left[1 - \sqrt{\frac{2R_0}{R+R}}\right] > 0,$$

and Δv_2 is in the direction of $\boldsymbol{v}_{0,2}$.

When should the first boosters be fired? The first delta-vee is delivered when the moon is at an angle θ to the horizontal (the horizontal is given by the line between the earth and the target point). If the rocket is to hit the target point, coincide with the moon, and subsequently co-rotate with the moon, we must have

$$\theta = \omega\,(T/2)\,,$$

where $\omega = 2\pi\,\mathrm{rad}/[27.3\text{ days}]$ is the orbital period of the moon.

18.3 Numerical results

We have[2]

- Mean earth-moon distance $R_{M0} = (a_M + b_M)/2 = 3.8411 \times 10^8$ m.
- Lunar radius, $R_{M1} = 1.737 \times 10^6\,\mathrm{m}$.
- Earth mass $M_\mathrm{e} = 5.974 \times 10^{24}$ kg.
- $G = 6.673 \times 10^{-11}$ m^3 kg^{-1} s^{-2}
- Orbital period in LEO, $T_\mathrm{E} = 88.18$ minutes
- Minimum altitude above lunar surface, $h = 116.5\,\mathrm{km}$

Hence, compute Δv_1, Δv_2, the time in flight, and the launch angle of the moon.

1. Orbital period gives R_0:

$$88.18\text{ minutes} = 5.2908 \times 10^3\text{ sec} = \frac{2\pi R_0^{3/2}}{\sqrt{GM_\mathrm{e}}},$$

hence

$$R_0 = \left[\sqrt{GM_\mathrm{e}} \times \left(5.2908 \times 10^3\text{ sec}\right)/(2\pi)\right]^{2/3} = 6.562809005818279\,(e+06)\ \mathrm{m}.$$

[2]Orbital parameters of Apollo 11 mission,
http://history.nasa.gov/SP−4029/Apollo_00g_Table_of_Contents.htm

Using the altitude data,

$$R = h + R_{M0} + R_{M1} = 3.85963500\,(e+08)\ \mathrm{m}$$

2. Delta-vee at transfer point 1: We have

$$\Delta v_1 = \sqrt{\frac{GM_\mathrm{e}}{R_0}}\left[\sqrt{\frac{2R}{R+R_0}} - 1\right],$$

where

$$\sqrt{\frac{GM_\mathrm{e}}{R_0}} = 7.793782626291548\,(e+03)$$

and

$$\sqrt{\frac{2R}{R+R_0}} - 1 = 0.402341319502358,$$

hence

$$\Delta v_1 = 3.136 \times 10^3\ \mathrm{m/s},$$

to the numerical precision implied by the flight parameters.

3. Delta-vee at transfer point 2:

$$\Delta v_2 = \sqrt{\frac{GM_\mathrm{e}}{R}}\left[1 - \sqrt{\frac{2R_0}{R_0+R}}\right] > 0,$$

where

$$\sqrt{\frac{GM_\mathrm{e}}{R}} = 1.016295619237095\,(e+03)$$

and

$$1 - \sqrt{\frac{2R_0}{R+R_0}} = 0.817137145334582,$$

hence

$$\Delta v_2 = 8.304 \times 10^2\ \mathrm{m/s}.$$

to four significant figures.

4. Time of transfer:

$$T_T = \frac{\pi}{\sqrt{GM_{\mathrm{e}}}} \left[\left(R_0 + R \right) / 2 \right]^{3/2} = 4.326282387011411 \left(e + 05 \right) \text{ sec} = 5.007 \text{ days.}$$

5. Initial angle of moon relative to target axis:

$$\theta = \omega T_T = \frac{2\pi}{27.32} \times 5.007 = 1.151596388854147 \text{ Rad} = 65.98^o.$$

Important conclusion: In the orbit proposed, astronauts would spend five days on their way to the moon and (hopefully!), another five days on the way back. In contrast, in Apollo 11's orbit, the astronauts spent a total of 8d 3h in space.

Chapter 19

Introduction to Einstein's theory of Special Relativity

19.1 Introduction

- Relativity is not a new concept to us: we have formulated classical mechanics on the basis of Galilean relativity, which says that *the laws of physics are the same in all inertial frames*.
- One further postulate of Galilean relativity is that time is absolute: given a pair of inertial frames in uniform motion relative to one another, clocks in both frames run at the same rate.
- This assumption breaks down at velocities close to the speed of light, and this requires the introduction of a new postulate of relativity, namely that *the speed of light is the same in all inertial frames of reference*.
- These two postulates form the basis of Special Relativity (SR). It is called special because it is a special case wherein gravity is absent (cf. General Relativity).
- It is not a purely kinematic theory however, because the interaction of charged particles with electromagnetic fields can be described using SR.

19.2 Electromagnetic radiation

In 1861, James Clark Maxwell formulated the equations for electromagnetic radiation (light, radiowaves, microwaves &c). A solution to Maxwell's equations in free space is the wave equation:

$$\frac{1}{c^2}\frac{\partial\psi}{\partial t^2} = \nabla^2\psi,$$

where ψ is the $\boldsymbol{E}$ or $\boldsymbol{B}$ field, and c is the velocity of light, which is a universal constant. But this is inconsistent with the idea that velocities depend on the frame of reference with respect to which they are measured. Contemporary physicists explained this by postulating the existence of the *aether*, an elastic medium fixed in space. Vibrations in this medium would correspond to EM radiation propagating at a speed c w.r.t. the fixed aether. An experiment by Michelson and Morley in 1882 failed to detect any relative motion of the earth with respect to the aether. This made the entire concept of the aether problematic. Some solutions to the problem were proposed by Poincaré, Lorentz, and Fitzgerald[1], but the most complete and ultimately, the successful one, was given by Einstein in 1905.

Einstein proposed to dispense with the unobservable aether altogether, and introduce a second postulate of relativity:

1. The laws of physics have the same form in all inertial frames of reference.
2. The speed of light in a vacuum is a universal constant, independent of the velocity of the light source.

Note that the second postulate is INCONSISTENT with the Galilean law of velocity additions and so produces an entirely new physics. The first consequence of this new physics is that *time is not absolute*.

19.3 Time is not absolute

Consider the familiar scenario of a train in motion relative to the platform of a station. The train is equipped with a clock. During one 'tick' of the clock, a beam of light is emitted from a source located on the floor the carriage, travels upwards, hits a mirror on the carriage ceiling, and arrives back at the floor. Let the height of the carriage be d. The time for one tick, as measured in the carriage, is

$$t' = \frac{2d}{c}.$$

(See Fig. 19.1). Now in the FOR of the platform, the light beam travels a distance

$$s = 2\sqrt{d^2 + \tfrac{1}{4}v^2t^2},$$

where v is the velocity of the train w.r.t. the platform and t is the tick time measured by the observer on the platform. But $s = ct$, hence

$$t = \frac{s}{c} = \frac{2}{c}\sqrt{d^2 + \tfrac{1}{4}v^2t^2}.$$

[1]Bio: see http://www.ucc.ie/academic/undersci/pages/sci_georgefrancisfitzgerald.htm

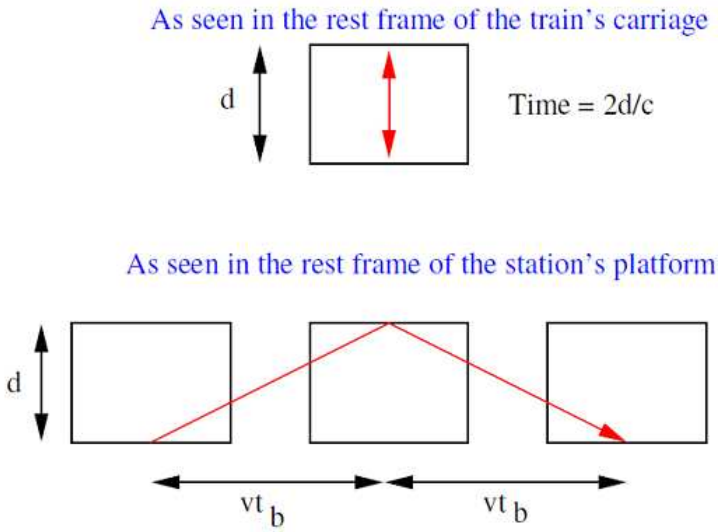


Figure 19.1: Schematic description of the light clock, $t_b = t/2$.

Re-arranging gives

$$t^2\left(1-\frac{v^2}{c^2}\right) = \frac{4d^2}{c^2} = t'^2,$$

hence

$$t = \frac{t'}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{19.1}$$

or

$$t' = t\sqrt{1-\frac{v^2}{c^2}}.$$

We shall see that $v < c$ always, so that

$$t' < t,$$

a result encapsulated in the phrase 'moving clocks run slow'.

Could we ever measure this for a real train ? If we take $d = 3\,\mathrm{m}$, then

$$t' = 6/c = 2\times 10^{-8}\,\mathrm{s}.$$

If we have a very fast train we could take $v = 200\,\mathrm{kmhr}^{-1}$, but this gives $v/c = 1.85\times 10^{-7}$, and, using the binomial theorem to estimate the denominator in Eq. (19.1),

$$t \approx 2\times 10^{-8}\left(1 + 1.7\times 10^{-14}\right).$$

Hence, the station master measures a duration equal to that measured by someone on the train plus 3.4×10^{-22} seconds – which is why we don't notice disagreements about time in our everyday experience.

Einstein's second postulate implies that we must revise the Galilean transformations and these are known as the Lorentz transformations.

19.4 Discussion question

Consider identical twin astronauts named Eartha and Astrid. Eartha remains on earth while her twin Astrid takes off on a high-speed trip through the galaxy. Because of time dilation, Eartha sees Astrid's heartbeat and all other life processes proceeding more slowly than her own. Thus, to Eartha, Astrid ages more slowly; when Astrid returns to earth she is younger (has aged less) than Eartha.

Now here is the paradox: All inertial frames are equivalent. Can't Astrid make exactly the same arguments to conclude that Eartha is in fact the younger? Then each twin measures the other to be younger when they're back together, and that is a paradox. What is the resolution to this apparent contradiction?

The answer is that the frames are inequivalent: Astra's frame is non-inertial becuase it undergoes accelearation at the point when Astra turns back towards earth. On the other hand, Earth's frame remains approximately inertial at all times. It turns out that Eartha is correct: Astra ages more slowly than Eartha, so that when she comes back, she is in fact younger than Eartha.

Chapter 20

The Lorentz Transformations

20.1 Overview

We want to derive the relation between an event occuring at position (x, y, z) and time t wrt observer S to the same event as seen by observer S' at (x', y', z') and time t'. Here frame S' moves at constant velocity V w.r.t. frame S. At speeds that are small compared to light we must recover the Galilean transformations but our new Lorentz transformations must be consistent with Einstein's second postulate which in turn implies that $t \neq t'$. Previously, we used the following diagram to derive the Galilean transformations: Thus, the relative motion is only in the x-direction,

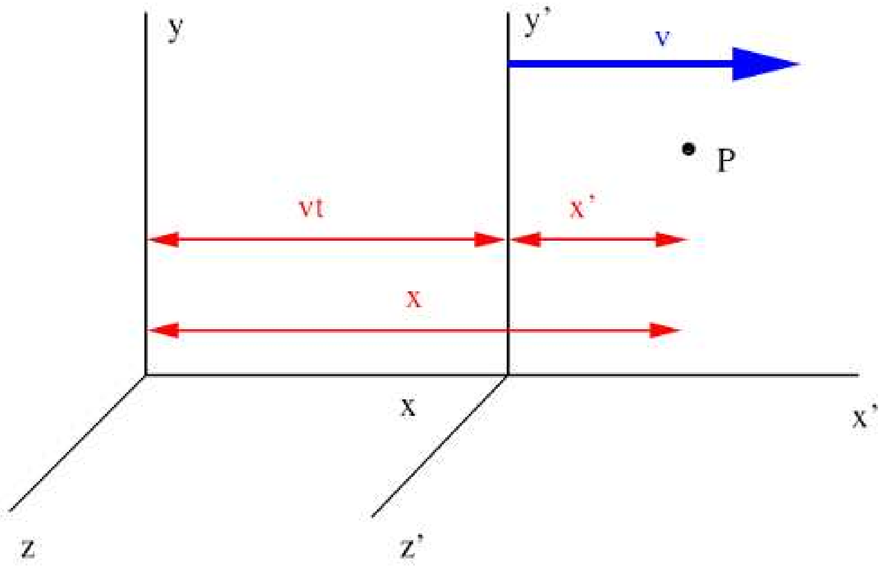


Figure 20.1: Relationship between two inertial frames of reference.

hence $y = y'$ and $z = z'$. As before, we arrange for the clocks to be synchronised so that when $x = x' = 0$, $t = t' = 0$. The transformations between frames S and S' must be linear, so we have the general form

$$\begin{aligned} x' &= Ax + Bt, \\ t' &= Cx + Dt. \end{aligned}$$

20.2 Fixing the coefficients in the linear transformation

We can determine the four constants, and how they depend on v , from the following four constraints.

1. Observer S sees the origin of S' moving along the x-axis with velocity V . We write this path for the origin of S' in each frame as

   $$x = Vt \text{ and } x' = 0.$$

   Therefore,

   $$x' = Ax + Bt \implies 0 = AVt + Bt \implies B = -AV,$$

   and now we have only three unknowns,

   $$\begin{aligned} x' &= A(x - Vt), \\ t' &= Cx + Dt. \end{aligned}$$

2. Observer S' sees the origin of S moving along the x' axis with velocity $-V$. The path for the origin of S is then

   $$x = 0 \text{ and } x' = -Vt',$$

   to give

   $$A(0 - Vt) = -V(0 + Dt) \implies D = A,$$

   which leaves two constants,

   $$\begin{aligned} x' &= A(x - Vt), \\ t' &= Cx + At. \end{aligned}$$

3. A pulse of light emitted from the origin in the x-direction at moves at speed c relative to both frames. Therefore, the path of this pulse of light in each frame is given by

   $$x = ct \text{ and } x' = ct'$$

to give

$$A\,(ct - Vt) = c\,(Cct + At) \implies C = -\frac{AV}{c^2},$$

and we are left with one remaining constant

$$\begin{aligned} x' &= A\,(x - Vt)\,, \\ t' &= A\left(t - \frac{Vx}{c^2}\right). \end{aligned}$$

4. A light pulse is sent along the y-axis at $t = 0$. In the S frame we have

$$x = 0, \text{ and } y = ct,$$

while in the S' frame the pulse also moves at speed c but now in the (x', y') plane:

$$x'^2 + y'^2 = c^2 t'^2,$$

which becomes

$$A^2\,(0 - Vt)^2 + (ct)^2 = c^2 A^2\,(0 + t)^2\,,$$

which tidies up to give

$$A = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$$

The quantity A is called the *Lorentz factor* and usually given the symbol γ:

$$\gamma := \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$$

We now have the equations which relate the coordinates and time for an event as measured by two inertial observers, the *Lorentz Transformations*:

$$\begin{aligned} x' &= \gamma\,(x - Vt)\,, \\ t' &= \gamma\left(t - \frac{Vx}{c^2}\right), \\ y' &= y, \\ z' &= z. \end{aligned}$$

Once we have the coordinates for an event in one frame we can work out the coordinates in the other inertial frame. Note that, as required by the principle of relativity, neither frame is more important than the other and the inverse transformations are obtained by replacing $-V$ by $+V$:

$$\begin{aligned} x &= \gamma\left(x' + Vt'\right), \\ t &= \gamma\left(t' + \frac{Vx'}{c^2}\right). \end{aligned}$$

20.3 Minkowski diagrams

Note that the Lorentz transformations for x' and t' are more alike that they look at first sight. If we were to change our units, from SI, to a system where $c = 1$ which is more 'natural', then we would have

$$\begin{aligned} x' &= \gamma\left(x - Vt\right), \\ t' &= \gamma\left(t - Vx\right), \end{aligned}$$

and we can represent both sets of axes on one diagram. Note that the t'-axis, for example, is defined by $x' = 0$, which becomes $t = x/V$. This is an example of a *Minkowski or Space-Time diagram*. Note that we will always work in SI so that $c = 3 \times 10^8$, and we can still draw these diagrams, but the t- and t'-axes would be very close to one another. In more advanced work, we typically work in units where $c = 1$. It becomes immediately apparent from this diagram that if two events are simultaneous in one frame then they will not be simultaneous in the other frame.

20.4 The Relativity of Simultaneity

As a thought experiment we imagine a person standing in the middle of a train carriage, of length $2L$ (in S') with a lamp that emits a pulse of light. This light arrives at each end of the carriage, A and B , after a time interval of L/c. The two events of light arriving at A and light arriving at B are therefore simultaneous relative to an observe on the train. However, an observer standing on the station's platform, S , observes the train to be moving at speed V but the forward and reverse pulse of light each move at speed c because of Einstein's second postulate. Therefore, observer S sees the backwards pulse hitting the back of the carriage before the forward pulse hits the front of the carriage. The events are not simultaneous in S. The Lorentz transformations make quantitative statements about this system. We take the origin of S' to be the centre of the carriage and the

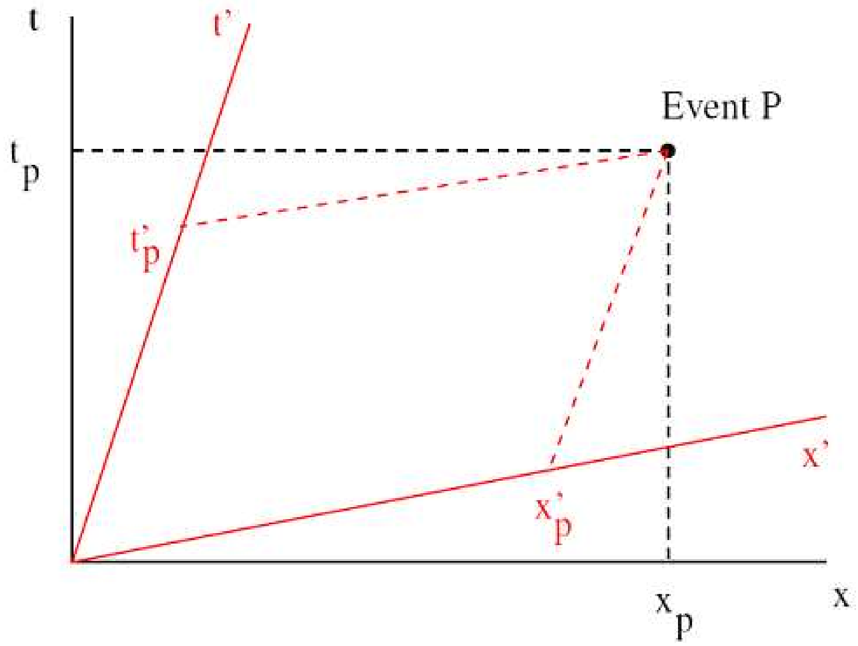


Figure 20.2: Minkowski diagram

flashes of light are emitted when $t' = 0$. Therefore, the flashes are emitted at $x' = 0$, $t' = 0$, which corresponds to $x = 0$, $t = 0$. We now look at each pulse arriving at the front and back.

Event 1: Pulse arrives at the back of the train. As seen by S', this event is

$$x'_1 = -L, \qquad t'_1 = \frac{L}{c} := T.$$

The time at which this event takes place as seen by S is

$$\begin{aligned} t_1 &= \gamma\left(t'_1 + \frac{Vx'_1}{c^2}\right), \\ &= \gamma\left(T - \frac{VL}{c^2}\right), \\ &= \sqrt{1-\frac{V^2}{c^2}}\left(T - \frac{V}{c}T\right), \\ &= T\sqrt{\frac{1-\frac{V}{c}}{1+\frac{V}{c}}}, \end{aligned}$$

which is less than T .

Event 2: Pulse arrives at the front of the train. As seen by S' this event is

$$x'_2 = L, \qquad t'_2 = \frac{L}{c} = T.$$

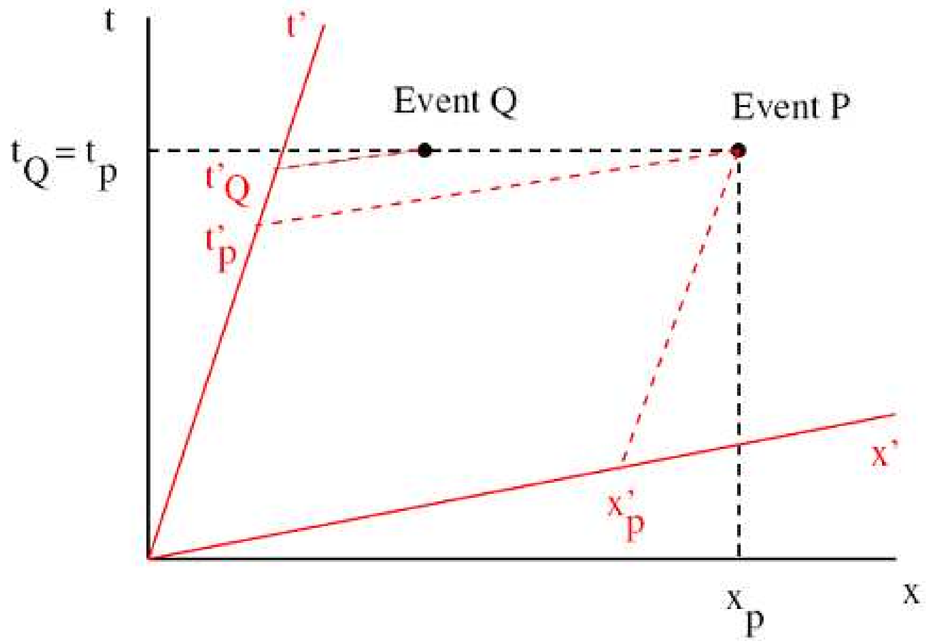


Figure 20.3: Minkowski diagram and the relativity of simulaneity: simultaneity is relative in special relativity because the transformed time axis is no longer parallel to the old time axis.

The time at which this event takes place as seen by S is

$$\begin{aligned} t_2 &= \gamma\left(t_2' + \frac{Vx_2'}{c^2}\right), \\ &= \gamma\left(T + \frac{VL}{c^2}\right), \\ &= \sqrt{1 - \frac{V^2}{c^2}}\left(T + \frac{V}{c}T\right), \\ &= T\sqrt{\frac{1+\frac{V}{c}}{1-\frac{V}{c}}}, \end{aligned}$$

which is greater than T.

Note: If we consider a very fast train then

$$V = 200\,\mathrm{km,hr}^{-1}, \qquad \frac{t_{1,2}}{T} \approx 1 \pm 1.85 \times 10^{-7},$$

which is a very small difference ! However, for $V = 0.5c$ (a very fast and very unrealistic rocket), then

$$t_1 = \frac{T}{\sqrt{3}}, \qquad t_2 = T\sqrt{3}.$$

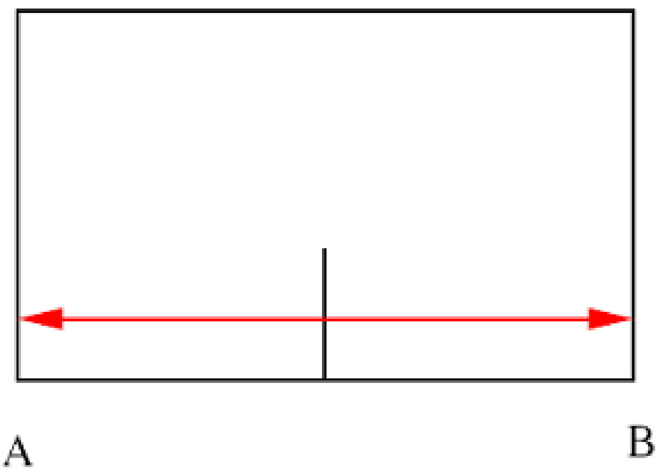


Figure 20.4: In frame S light hits A and B simultaneously.

20.5 A spacetime invariant

Although our two observers will disagree on the *when* and the *where* of events, there is a quantity that has the same value for each of them. Starting from the transformation law for an event,

$$\begin{aligned} x' &= \gamma\left(x - Vt\right), \\ t' &= \gamma\left(t - \frac{xV}{c^2}\right), \end{aligned}$$

we calculate $c^2t'^2 - x'^2$:

$$\begin{aligned} c^2t'^2 - x'^2 &= \gamma^2\left[c^2\left(t - \frac{xV}{c^2}\right)^2 - (x - Vt)^2\right], \\ &= \gamma^2\left[\left(ct - \frac{xV}{c}\right)^2 - (x - Vt)^2\right], \\ &= \gamma^2\left[c^2t^2 - 2\frac{xV}{c}ct + \left(\frac{V}{c}\right)^2 x^2 - x^2 + 2xVct - V^2t^2\right], \\ &= \gamma^2\left[\left(c^2 - V^2\right)t^2 - x^2 + \left(\frac{V}{c}\right)^2 x^2\right], \\ &= \gamma^2\left[\left(c^2 - V^2\right)t^2 - \left(1 - \frac{V^2}{c^2}\right)x^2\right], \\ &= c^2t^2 - x^2. \end{aligned}$$

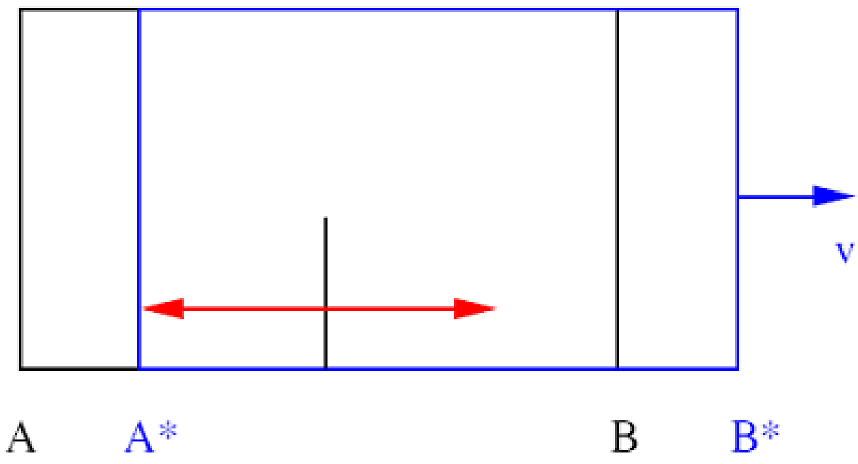


Figure 20.5: In frame S light hits back of train first.

Hence, we have a quantity that is the same for all inertial observers,

$$s^2 = c^2t^2 - x^2.$$

The quantity s is a Lorentz Invariant.

Chapter 21

Length contraction and time dilation

21.1 Overview

In Ch. 20 (The Lorentz transformations) we saw how events seen in one frame are transformed to a second frame. In the current chapter we look at the transformation of intervals in space and time. We also use the LTs to derive the transformation laws relating velocities in different frames.

21.2 Length contraction

Consider a stick at rest in frame S' lying along the x-axis. One end of the stick is at x'_A while the other is at x'_B . The length of the stick is given by

$$L_0 = x'_B - x'_A$$

and this is called the *proper length*; **i.e. the length of the stick in the frame where the stick is at rest.**

What is the length of the stick as measured by observer S ? The length of the stick is the distance between the end points measured at the same time !!. In other words the positions of the end points must be determined at the same time in S. If observer S determines the front of the stick to be at position x_B at time t and the back of the stick to be at x_A at the same time t then the length of the stick as measured by S is

$$L = x_B - x_A.$$

Moreveor, by the LTs we have

$$x'_B = \gamma\left(x_B - vt\right),$$
$$x'_A = \gamma\left(x_A - vt\right).$$

Subtracting we have

$$L = \frac{L_0}{\gamma} = L_0\sqrt{1 - \frac{v^2}{c^2}}.$$

The longest length a stick can have is when viewed in its own rest frame – in every other frame it is shorter. A stick has a shorter length the faster it travels.

Note that everything is relative – imagine observer S' measuring his carriage to be of length L_0; he is at rest relative to the carriage so this is the proper length. Observer S measures this carriage to be of length L_0/γ. Now place observer S in his own carriage which is at rest in the train station and also has proper length L_0 . Observer S' will now measure that carriage to be contracted to the length L_0/γ.

21.3 Time dilation

We have already seen two chapters ago that time intervals are not absolute. Let us make that more exact here. Consider two events occuring on the train *at the same position*:

- Event A at position x'_0 and time t'_A;
- Event B at position x'_0 and time t'_B.

The time interval between these events is known as the proper time since the events occur at the same position and is

$$T_0 = t'_B - t'_A.$$

In frame S we have a time interval of $T = t_B - t_A$ between these events, where

$$t_A = \gamma\left(t'_A + \frac{vx'_0}{c^2}\right),$$
$$t_B = \gamma\left(t'_B + \frac{vx'_0}{c^2}\right),$$

so that

$$T = \frac{T_0}{\sqrt{1 - \frac{v^2}{c^2}}},$$

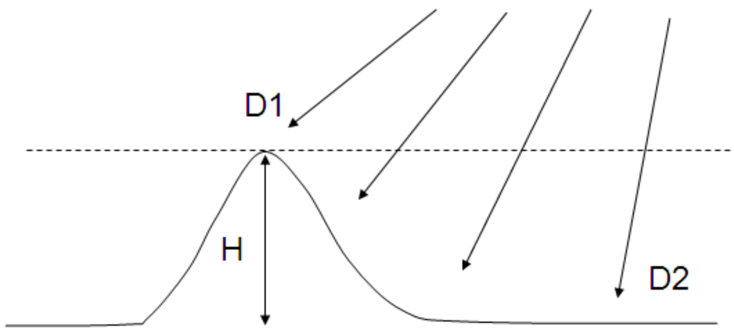


Figure 21.1: Schematic description of muon decay at two different locations.

with the result that $T > T_0$. The shortest interval of time between two events is when they are observed in a frame where they occur at the same position, their proper frame, in all other frames the time interval is longer.

Note that is not always possible to find a proper frame for two events. This occurs when one would have to travel faster than light for the two events to occur at the same position – such events are known as space-like.

Time dilation and meson decay

The muon is an unstable subatomic particle that decays into a more stable state. Muons constantly bombard us on earth through cosmic rays. The liftime of the particles depends on the frame in which the decay events are viewed, so providing an experimental verification of time dilation.

The muon decays into an electron and a neutrino-antineutrino pair. Symbolically,

$$\mu^{\pm} \rightarrow e^{\pm} + \nu + \overline{\nu}.$$

The decay of the muons follows the law of radioactive decay: if there are N_0 particles present at time $t = 0$ then at time t, there are $N(t)$ particles, where $N(t) = N_0 e^{-t/\tau_0}$, and τ_0 is the *lifetime* of the particles. The decay time can be measured in the laboratory, and it is found that $\tau_0 = 2.15 \times 10^{-6}\,\mathrm{s}$. Now, suppose that we have two muon detectors, one on top of a mountain, and one at sea level, as in Fig. 21.1. Let N_1 be the number of particles detected at the mountaintop per unit area of detector per unit time. Similarly, let N_2 be the number of particles detected at sea-level per unit area of detector per unit time. Now the muons are travelling at a velocity $v \lesssim c$. Because the particles decay in going through a distance H, and the time for this travel is $T \approx H/c$, we might expect that

$$\frac{N_2}{N_1} ?? = ?? e^{-T/\tau_0} \approx e^{-H/(c\tau_0)}.$$

However, because we make the observation in a frame that is different from the muon's rest frame, the observed lifetime is dilated, and equal to

$$\tau = \gamma\tau_0$$

where $\gamma = 1/\sqrt{1 - v^2/c^2}$. Thus, the correct value for the flux ratio is

$$\frac{N_2}{N_1} = e^{-T/\gamma\tau_0} \approx e^{-H/(\gamma c\tau_0)}.$$

Putting in a value $H = 2,000\,\mathrm{m}$ and $\gamma = 10$, we obtain

$$\frac{N_2}{N_1} = 0.7,$$

while $e^{-H/(c\tau_0)} = 0.045$. Therefore, time dilation means that fewer decays happen in the frame of the earth, and we see more particles at sea level than might otherwise be expected.

21.4 The Lorentz transformation law for velocities

Consider a particle moving with velocity u along the x-axis relative to observer S. Relative to observer S' this particle has a velocity u'. According to Galilean relativity we should have

$$u' = u - v$$

but, in the case of a photon with $u = c$ this would violate Einstein's second postulate. This law for transforming velocities must be incorrect.

The derivation: Recall that the definitions for velocity are

$$u = \frac{dx}{dt}, \qquad u' = \frac{dx'}{dt'}$$

and the Lorentz transformations allow us to relate space-time intervals from one frame to another. Firstly,

$$x' = \gamma\,(x - vt) \implies dx' = \gamma\,(dx - vdt)\,.$$

and for time,

$$t' = \gamma\left(t - \frac{vx}{c^2}\right), \qquad dt' = \gamma\left(dt - \frac{v}{c^2}dx\right).$$

Formally dividing these equations now gives

$$\frac{dx'}{dt'} = \frac{dx - vdt}{dt - \frac{v}{c^2}dx}.$$

Divide the RHS above and below by dt:

$$u' = \frac{dx'}{dt'} = \frac{u - v}{1 - \frac{uv}{c^2}}.$$

Now, in the case that our particle is a photon so that $u = c$ we get

$$u' = \frac{c - v}{1 - \frac{v}{c}} = c,$$

consistent with Einstein's second postulate. Moreover, as usual in relativity, at normal everyday terrestrial speeds where $u, v \ll c$ the above result reduces to the Galilean limit that $u' \approx u - v$.

Example: Consider an observer that sees rocket A moving along the positive x-axis with speed $0.9c$ and rocket B moving along the negative x-axis at speed $0.9c$. What is the speed of one rocket relative to the other?

According to Galileo it would be $1.8c$ but using the correct law for adding velocities we get

$$u' = \frac{0.9c + 0.9c}{1 + 0.81} = 0.994475c,$$

which is less than the speed of light.

The law for adding velocities always gives $u' < c$ when $u < c$, hence no material particle can move faster than the speed of light (exercise).

21.5 Velocity transformations in three spatial dimensions

What if our particle were not moving purely along the x-axis, i.e. the velocity in frame S had three components

$$\boldsymbol{u} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) = (u_x, u_y, u_z)\,,$$

with the corresponding velocity as measured in S'

$$\boldsymbol{u}' = \left(\frac{dx'}{dt'}, \frac{dy'}{dt'}, \frac{dz'}{dt'}\right) = \left(u'_x, u'_y, u'_z\right)\,,$$

We can now generalise the argument we presented above by noting that

$$\begin{aligned} dx' &= \gamma\,(dx - v dt)\,, \\ dt' &= \gamma\left(dt - \frac{v}{c^2}dx\right)\,, \\ dy' &= dy, \\ dz' &= dz. \end{aligned}$$

Straightforward division gives the transformation laws

$$
\begin{aligned}
u'_x &= \frac{u_x - v}{1 - \frac{u_x v}{c^2}}, \\
u'_x &= \frac{u_y/\gamma}{1 - \frac{u_x v}{c^2}}, \\
u'_x &= \frac{u_z/\gamma}{1 - \frac{u_x v}{c^2}},
\end{aligned}
$$

The laws for u'_y and u'_z are now substantially different from their simple Gallilean counterparts. It is possible to show (try it !) that if

$$u = \sqrt{u_x^2 + u_y^2 + u_z^2} = c,$$

then

$$u' = c$$

too, so that, as expected, a beam of light will move with speed c as measured in either frame but the direction of that beam will appear different to each observer.

Chapter 22

Relativistic momentum and energy

22.1 Overview

In this chapter we generalize the notions of momentum and kinetic energy in such a way that these concepts become relativistically correct. These quantities can then be used to solve a variety of *kinematic* problems, involving collisions between particles.

Recall the contrasting notions of kinematics and dynamics. Kinematics is the study of motion; dynamics is the study of changes to motion. In kinematics we simply want to describe motion (energy, velocity, momentum), without being interested in changes in this motion.

22.2 Relativistic momentum

It turns out that momentum is still conserved under special relativity. That is, when two bodies interact, the total momentum is conserved, provided that the net external force acting on the bodies in an inertial reference frame is zwero.

However, in SR, the momentum is NOT $\boldsymbol{p} = m\boldsymbol{v}$. For, it is possible to have a two-particle collision in which, relative to a frame S, the Newtonian total momentum is conserved, whereas in a different intertial frame S', it is not conserved. Thus, we need a new defition of relativistic momentum. It turns out that the following is the correct definition:

$$\boldsymbol{p} = \frac{m\boldsymbol{v}}{\sqrt{1-(v/c)^2}} := \gamma m\boldsymbol{v}.$$

Because of the ubiquity of the factor $\left(1-(v/c)^2\right)^{-1/2}$ in SR, we give it a special name, γ, and $\gamma \geq 1$.

Example: An electron (rest mass $9.11 \times 10^{-31}\,\mathrm{kg}$ is moving at velocity v in a linear accelerator. Find the momentum of the electron if (a) $v = 0.01c$; (b) $v = 0.99c$.

Now $p_a = \gamma_a m v_1 = \left(1 - 0.01^2\right)^{-1/2} \left(9.11 \times 10^{-31}\right) \left(0.01 \times 3.00 \times 10^8\right) = 2.7 \times 10^{-24}\,\mathrm{kgm/s}$.

Similarly, $p_b = \gamma_b m v_b = \left(1 - 0.99^2\right)^{-1/2} \left(9.11 \times 10^{-31}\right) \left(0.99 \times 3.00 \times 10^8\right) = 1.9 \times 10^{-21}\,\mathrm{kgm/s}$.

22.3 Relativistic kinetic energy

Recall the energy framework we built up in ordinary Newtonian mechanics. The kinetic energy was identified with a change in momentum through the work-energy relation:

$$K_2 - K_1 = \int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \frac{d\boldsymbol{p}}{dt} \cdot \mathrm{d}\boldsymbol{x}.$$

Why not do the same thing in the relativistic case?

$$\begin{aligned} K_2 - K_1 &= \int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \frac{d\boldsymbol{p}}{dt} \cdot \mathrm{d}\boldsymbol{x}. \\ &= \int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \frac{d}{dt} \left[\frac{m\boldsymbol{v}}{\sqrt{1 - (v/c)^2}} \right] \cdot \mathrm{d}\boldsymbol{x}, \\ &= \int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \frac{d}{dt} \left[\frac{m\boldsymbol{v}}{\sqrt{1 - (v/c)^2}} \right] \cdot \boldsymbol{v}\,\mathrm{d}t, \\ &= \int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \boldsymbol{v} \cdot \mathrm{d} \left[\frac{m\boldsymbol{v}}{\sqrt{1 - (v/c)^2}} \right]. \end{aligned}$$

But the integrand is $\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{p} = \mathrm{d}\left(\boldsymbol{v} \cdot \boldsymbol{p}\right) - \boldsymbol{p} \cdot \mathrm{d}\boldsymbol{v}$ (integration by parts). Hence,

$$\begin{aligned} K_2 - K_1 &= \left(\boldsymbol{v} \cdot \boldsymbol{p}\right)\Big|_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} - \int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \boldsymbol{p} \cdot \mathrm{d}\boldsymbol{v}, \\ &= \left.\frac{mv^2}{\sqrt{1 - v^2/c^2}}\right|_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} - \int_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} \frac{mv dv}{\sqrt{1 - v^2/c^2}}, \end{aligned}$$

where we have used $\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{v} = v\mathrm{d}v$. The integral can be looked up in a table, and we find

$$K_2 - K_1 = \left.\frac{mv^2}{\sqrt{1 - v^2/c^2}}\right|_{\boldsymbol{x}_1}^{\boldsymbol{x}_2} + mc^2 \sqrt{1 - v^2/c^2}\Big|_{\boldsymbol{x}_1}^{\boldsymbol{x}_2}$$

(note the sign change in the second term!). If we take the point $\boldsymbol{x}_2$ to be arbitary, and let the particle be at rest at $\boldsymbol{x}_1$, $\boldsymbol{v}(\boldsymbol{x}_1) = 0$, then

$$\begin{aligned} K &= \frac{mv^2}{\sqrt{1-v^2/c^2}} + mc^2\sqrt{1-v^2/c^2} - mc^2, \\ &= \frac{m\left[v^2 + c^2\left(1-v^2/c^2\right)\right]}{\sqrt{1-v^2/c^2}} - mc^2, \\ &= \frac{mc^2}{\sqrt{1-v^2/c^2}} - mc^2, \end{aligned}$$

hence

$$K = (\gamma - 1)\, mc^2. \tag{22.1}$$

This result bears little resemblance to its Newtonian analogue, $K = mv^2/2$. However, if we expand γ in small v/c, we obtain

$$\gamma = 1 + \tfrac{1}{2}\frac{v^2}{c^2} + \tfrac{3}{8}\frac{v^4}{c^4} + \cdots .$$

Keeping only terms to second order on v/c, we obtain

$$K = (\gamma - 1)\, mc^2 = \tfrac{1}{2}mv^2 + v^2 O\left(v^2/c^2\right).$$

Equation (22.1) for the kinetic energy – the energy due to the motion of the particle – contains an energy term $mc^2/\sqrt{1-v^2/c^2}$ that depends on the motion and a second energy term mc^2 that is independent of the motion. It seems that the kinetic energy of a particle is the difference between some **total energy** E and an energy mc^2 that the particle has when it is at rest. Thus, we re-write Eq. (22.1) as

$$E = K + mc^2 = \frac{mc^2}{1-v^2/c^2} = \gamma mc^2.$$

For a particle at rest ($K = 0$), we see that $E = mc^2$. The energy mc^2 associated with rest mass m rather tahn motion is called the **rest energy** of the particle. The total energy E is conserved in time-independent processes (conservation of mass-energy). Thus, in some processes, neither the sum of the rest masses of the particles nor the total energy (other than the rest energy) is separately cosnerved, but rather the sum implied in Eq. (22.3) is conserved.

The energy and the momentum can be related together as follows. We re-write the formula for E as

$$\left(\frac{E}{mc^2}\right)^2 = \frac{1}{1 - v^2/c^2},$$

and the formula for p as

$$\left(\frac{p}{mc}\right)^2 = \frac{v^2/c^2}{1 - v^2/c^2}.$$

Subtracting these equations gives

$$E^2 - (pc)^2 = \left(mc^2\right)^2.$$

This might be the most important equation of all in classical physics. Just as $x^2 - c^2t^2$ is an *invariant* under the Lorentz transformations, so too is $E^2 - p^2c^2$, since this difference is a constant in all frames.

Example: Two protons (each with mass $M = 1.67 \times 10^{-27}\,\mathrm{kg}$ are initally moving with equal speeds in opposite directions. They continue to exist after a head-on collision that also produces a neutral pion of mass $m = 2.40 \times 10^{-28}\,\mathrm{kg}$. If the protons and the pion are at rest after the collision, find the initial speed of the protons. Energy is conserved in this collision.

Solution: Try conservation of momenta: $P_{\mathrm{init}} = \gamma M\,(v - v) = 0$, and $P_{\mathrm{fin}} = 0$, since the particles emerge from the collision at rest. This gives no information. So we turn to conservation of energy. That is, the total energy $E = \sum_i \gamma_i m_i c^2$ is conserved. Initially, $E = 2\gamma M c^2$. After the collision, the energy is pure rest energy and is equal to $E = 2Mc^2 + mc^2$. Equating gives

$$\gamma = 1 + \frac{m}{2M},$$

where $\gamma = 1/\sqrt{1 - v^2/c^2}$ is the gamma-factor of the initial proton pair. Inverting for v/c gives

$$\frac{v}{c} = \left[1 - \left(1 + \frac{m}{2M}\right)^{-2}\right]^{1/2}.$$

Plugging in the numerical values gives

$$v = 0.360c.$$

22.4 Photons

There are two complementary ways of viewing light (electromagnetic radiation). In one picture, light is a wave, which satisfies a wave equation (Maxwell's equations). In the other picture, light behaves like mass-zero particles called *photons*. Both pictures are needed for a full description of light-related phenomena. Although these pictures are complementary, they are unfied on the level of quantum field theory (Quantum Electrodynamics). Whenever a particle picture of light is needed, SR provides the framework for this description. Here are some relations that relate the particle nature of light to its wave nature: In this table,

Mass of a photon	$m = 0$
SR formula for photon energy	$E = pc$
Quantum mechanics formula for photon energy	$E = h\nu$, ν =light frequency
Equating the energy formulas	$E = h\nu = pc \implies p = h\,(\nu/c) = h/\lambda.$

Table 22.1: Properties of the photon

$$h = 6.626068 \times 10^{-34}\,\mathrm{m^2 kg/s} = 6.626068 \times 10^{-34}\mathrm{J \cdot s}$$

is Planck's constant, a fundamental constant of quantum mechanics, [1] and $\lambda = \nu/c$ is the wavelength of the light.

Example: Compton scattering Consider very energtic photons of energy $E = h\nu$ incident on a target. Each photon strikes an electron and sets the electron into motion. During the collision, the photon's momentum changes. If $\lambda = c/\nu$ is the wavelength of a photon, compute the wavelength of the scattered photons.

Solution: Schematically, the process we are considering is this:

$$[\text{Photon}] + [\text{Electron at rest}] \to [\text{Scattered photon}] + [\text{Electron in motion}]\,.$$

We use the conservation of mass-energy, $E = \mathrm{Const.}$. The initial enegy is

$$pc + m_{\mathrm{e}}c^2 = h\nu + m_{\mathrm{e}}c^2.$$

The final energy is

$$p'c + \gamma mc^2 = h\nu' + \gamma m_{\mathrm{e}}c^2.$$

We are going to call the electron energy $\gamma m_{\mathrm{e}}c^2$ E. Thus,

$$h\,(\nu - \nu') + m_{\mathrm{e}}c^2 = E.$$

We can square this if we want to:

$$\left[h\,(\nu - \nu') + m_{\mathrm{e}}c^2\right]^2 = E^2.$$

[1] This leads to the joke: What's new? E/h. Don't all laugh at once.

Now let's look at momenta. The initial momentum is $(h\nu/c)\,\hat{\boldsymbol{x}}$. The final momentum is $\boldsymbol{p} + (h\nu'/c)\,\hat{\boldsymbol{n}}$, where $\hat{\boldsymbol{n}}$ is just some unit vector. By conservation of momentum,

$$(h\nu/c)\,\hat{\boldsymbol{x}} = \boldsymbol{p} + (h\nu'/c)\,\hat{\boldsymbol{n}}.$$

Re-arrange this:

$$h\nu\hat{\boldsymbol{x}} - h\nu'\hat{\boldsymbol{n}} = \boldsymbol{p}c.$$

Squaring gives

$$p^2c^2 = (h\nu)^2 + (h\nu')^2 - 2\,(h\nu)\,(h\nu')\cos\phi,$$

where $\cos\phi = \hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{n}}$ is an angle. Now $E^2 = (mc^2)^2 + (pc)^2$, and we have

$$\begin{aligned} E^2 &= \left[h\,(\nu - \nu') + m_{\mathrm{e}}c^2\right]^2, \\ (pc)^2 &= (h\nu)^2 + (h\nu')^2 - 2\,(h\nu)\,(h\nu')\cos\phi. \end{aligned}$$

Hence,

$$\left[h\,(\nu - \nu') + m_{\mathrm{e}}c^2\right]^2 = \underline{(mc^2)^2} + (h\nu)^2 + (h\nu')^2 - 2\,(h\nu)\,(h\nu')\cos\phi$$

Expand the LHS:

$$\begin{aligned} &(h\nu)^2 + (h\nu')^2 + (m_{\mathrm{e}}c^2)^2 + 2\,(h\nu)\,m_{\mathrm{e}}c^2 - 2\,(h\nu')\,m_{\mathrm{e}}c^2 - 2\,(h\nu)\,(h\nu') = \\ &\qquad (mc^2)^2 + (h\nu)^2 + (h\nu')^2 - 2\,(h\nu)\,(h\nu')\cos\phi \end{aligned}$$

Effecting cancellations and tidying up gives

$$\frac{c}{\nu'} - \frac{c}{\nu} = \frac{h}{mc}\,(1 - \cos\phi)\,.$$

But $c = \lambda\nu$ and $\lambda = c/\nu$ for waves, so

$$\lambda' - \lambda = \frac{h}{mc}\,(1 - \cos\phi)\,. \qquad (22.2)$$

Example: X-rays emanate from a source have a wavelength $\lambda = 0.124\,\mathrm{nm}$ and are Compton-scattered off a target. At what scattering angle ϕ is the wavelength of the scattered X-rays 1.0% longer than that of the incident X-rays?

Solution: In Eq. 22.2 we want $\Delta\lambda := \lambda' - \lambda$ to be 1.0% of $0.124\,\mathrm{nm}$. That is, $\Delta\lambda = 0.00124\,\mathrm{nm} =$

1.24×10^{-12}m. Using the value $h/mc = 2.426 \times 10^{-12}$ m, we find

$$\Delta\lambda = \frac{h}{mc}\left(1 - \cos\phi\right),$$

or

$$\cos\phi = 1 - \frac{\Delta\lambda}{(h/mc)} = 1 - \frac{1.24 \times 10^{-12}}{2.426 \times 10^{-12}} = 0.4889.$$

Hence, $\phi = 60.7^o$.

Chapter 23

Special topics in Special Relativity

In these questions, you may use the fact the following conversion factor relating the electron-volt to Joules: $1\,\mathrm{eV} = 1.60217646 \times 10^{-19}\,\mathrm{Joules}$, where the Joule is the SI unit of energy, $\mathrm{J} = \mathrm{kg}\,\mathrm{m}^2/\mathrm{s}^2$.

1. **Time dilation:** After being produced in a collision between elementary particles, a positive pion must travel down a $1.20\,\mathrm{km}$-long tube to reach a detector. A positive pion has an average lifetime (measured in its rest frame) of $\tau_0 = 2.60 \times 10^{-8}\,\mathrm{s}$; the pions we consider have this lifetime. (a) How fast must the pion travel if is not to decay before it reaches the end of the tube? (Since the speed v of the pion is close to c, write $v = (1-\Delta)\,c$ and give your answer in terms of Δ); (b) The pion has a rest energy of $139.6 \times 10^6\,\mathrm{eV} = 139.6\,\mathrm{MeV}$. What is the total energy of the pion at the speed calculated in part (a)?

   Solution: Time of travel: $T = H/v$, $v = (1-\Delta)\,c$. We require $T < \tau = \gamma\tau_0$. Hence,

   $$H = \tau_0 c \frac{1-\Delta}{\sqrt{1-(1-\Delta)^2}}.$$

   Using the binomial expansion $\sqrt{1-(1-\Delta)^2} \approx \sqrt{2\Delta}$. To lowest order then

   $$H = \frac{\tau_0 c}{\sqrt{2\Delta}},$$

   and

   $$\Delta = \tfrac{1}{2}\left(\frac{\tau_0 c}{H}\right)^2.$$

   Plug in the numbers: $\Delta = 2.11 \times 10^{-5}$.

   In the second part, $E = \gamma m c^2$, and $\gamma \approx 1/\sqrt{2\Delta}$. Hence,

   $$E \approx \frac{mc^2}{\sqrt{2\Delta}}.$$

Plugging in the numbers,

$$E = \frac{139.6 \times 10^6\,\mathrm{eV}}{0.0065} = 2.147692307692308e + 10\,\mathrm{eV} = 2.15 \times 10^4\,\mathrm{MeV}.$$

2. **Length contraction:** A rod of length ℓ_0 lies in the $x'y'$ plane of its rest ssytem and makes an angle θ_0 with the x' axis. What is the length and orientation of the rod in the lab system xy in which the rod moves to the right with velocity v?

   Call the ends of the rod A and B. In the rest system these points have coordinates

$$\begin{array}{lll} A: & x'_A = 0, & y'_A = 0 \\ B: & x'_B = \ell_0 \cos\theta_0, & y'_B = \ell_0 \sin\theta_0. \end{array}$$

   We Lorentz-transform the lengths using $x' = \gamma\,(x - vt)$, $y' = y$ and obtain

$$\begin{array}{lll} A: & x'_A = 0 = \gamma\,(x_A - vt)\,, & y'_A = 0 = y_A \\ B: & x'_B = \ell_0 \cos\theta_0 = \gamma\,(x_B - vt)\,, & y'_B = \ell_0 \sin\theta_0 = y_B. \end{array}$$

   Inverting for the unprimed quantities gives

$$\begin{aligned} x_B - x_A &= \frac{\ell_0 \cos\theta_0}{\gamma}, \\ y_B - y_A &= \ell \sin\theta_0. \end{aligned}$$

   The length is

$$\begin{aligned} \ell = \sqrt{(x_B - x_A)^2 + (y_A - y_B)^2} &= \ell_0 \sqrt{(1 - v^2/c^2)\cos^2\theta_0 + \sin^2\theta_0} \\ &= \ell_0 \sqrt{1 - \frac{v^2}{c^2}\cos^2\theta_0}. \end{aligned}$$

   The angle that the rod makes with the x-axis in the lab frame is

$$\theta = \arctan \frac{y_B - y_A}{x_B - x_A} = \arctan\left(\gamma \frac{\sin\theta_0}{\cos\theta_0}\right) = \arctan\left(\gamma \tan\theta_0\right).$$

3. **Kinematics: Pair production** Under certain conditions, a photon can spontaneously decay into an electron-positron pair. This cannot happen in empty space however. To see why, let us perform the energy-momentum budget for the reaction $\gamma \to e^+ + e^-$. Conservation of energy gives

$$h\nu = (\gamma_+ + \gamma_-)\,mc^2.$$

Conservation of momentum gives

$$h\nu/c = |\gamma_+ \boldsymbol{v}_+ + \gamma_- + \boldsymbol{v}_-| m$$

Equating the two expressions for $h\nu/c$ gives

$$(\gamma_+ + \gamma_-)\, c = |\gamma_+ \boldsymbol{v}_+ + \gamma_- + \boldsymbol{v}_-|.$$

But this violates the following string of inequalities:

$$\begin{aligned}
(\gamma_+ + \gamma_-)^2 c^2 &= |\gamma_+ \boldsymbol{v}_+ + \gamma_- + \boldsymbol{v}_-|^2, \\
&= \gamma_+ |\boldsymbol{v}_+|^2 + \gamma_- |\boldsymbol{v}_-|^2 + \gamma_+\gamma_- |\boldsymbol{v}_+||\boldsymbol{v}_-| \cos\phi, \\
&\leq \gamma_+ |\boldsymbol{v}_+|^2 + \gamma_- |\boldsymbol{v}_-|^2 + \gamma_+\gamma_- |\boldsymbol{v}_+||\boldsymbol{v}_-|, \\
&= (\gamma_+ |\boldsymbol{v}_+| + \gamma_- |\boldsymbol{v}_-|)^2, \\
&< (\gamma_+ + \gamma_-)^2 c^2.
\end{aligned}$$

Therefore, pair production is impossible in empty space. However, suppose that a photon spontaneously decays in the vicinity of a massive nucleus at rest (mass M). Schematically, the decay process is $\gamma + M \to e^+ + e^- + M$. Now suppose that the decaying photon gives the nucleus M a small kinetic energy, such that the post-collision nucleus travels in the same direction as the incident photon with a velocity $V \ll c$. Moreover, suppose that the pair produced is at rest in the lab frame. Then, the energy balance gives

$$h\nu = 2mc^2 + \tfrac{1}{2} M V^2, \tag{23.1}$$

while the momentum balance (for $V \ll c$) gives

$$\frac{h\nu}{c} = MV. \tag{23.2}$$

Substituting Eq. (23.2) into Eq. (23.1), we obtain

$$h\nu = 2m_0 c^2 + \tfrac{1}{2} \frac{(h\nu)^2}{Mc^2}.$$

However, the assumption that the nucleus behaves in a Newtonian way requires that the incident photon does not affect the nuculeus that much. Therefore, the energy of the photon must be much less than the nucleus rest energy, $h\nu \ll Mc^2$. Thus, to lowest order, the photon energy required for pair production in the vicinity of a very massive nucleus is

$$h\nu = 2m_0c^2.$$

Plugging in the electron mass, we require

$$h\nu_{\text{threshold}} = 1.02\,\text{MeV}.$$

For comparison, the rest energy of hydrogen is $940\,\text{MeV}$, and at threshold, the photon energy is therefore much less than the nucleus rest energy.

4. **Kinematics: Mass-energy equivalence** In 1932, Cokcroft and Walton successfully operated the first high-energy proton accelerator and succeeded in acausing a nuclear disintegration. Their experiment provded one of the earliest confirmations of the relativisitc mass-energy relation. They studied a reaction wherein protons hit a Lithium target at rest. This precipitated a nuclear reaction in which the proton-lithium pair were converted into two Helium atoms (α-particles). Schematically, they observed the reaction

$$^{1}\text{H} +^{7}\text{Li} \rightarrow^{4}\text{He} +^{4}\text{He}.$$

The mass-energy balance for this reaction is

$$K\left(^{1}\text{H}\right) + M\left(^{1}\text{H}\right)c^2 + M\left(^{7}\text{Li}\right)c^2 = 2K\left(^{4}\text{He}\right) + 2M\left(^{4}\text{He}\right).$$

Calling $K = 2K\left(^{4}\text{He}\right) - K\left(^{1}\text{H}\right)$ and $\Delta M = M\left(^{1}\text{H}\right) + M\left(^{7}\text{Li}\right) - 2M\left(^{4}\text{He}\right)$, we have

$$K = \Delta M c^2.$$

Cockroft and Walton [1] knew the mass of each of the elements from mass spectrometry. They had

$$\begin{aligned} M\left(^{1}\text{H}\right) &= 1.0072, \\ M\left(^{7}\text{Li}\right) &= 7.0104 \pm 0.0030, \\ M\left(^{4}\text{He}\right) &= 4.0011, \end{aligned} \qquad (23.3)$$

[1]E.T.S. Walton (1903 -1995) is known as Ireland's only nobel Lauteate in science. He studied at TCD and then worked at Cambridge, under Rutherford. He later taught back in TCD. Arguably his real *alma mater* is the Methodist College Belfast. Other famous Irish / British alumni of the Methodist College include W.M.F. Orr (1866-1934) and John Herivel (b. 1918). Orr was a professor at UCD and co-discovered the Orr–Sommerfield equation, an equation vital to the understanding of turbulence, while Herviel helped to break the Enigma Code in WWII.

where the units here are atomic mass units, and $1\,\mathrm{amu} = 931\,\mathrm{MeV}$. These numbers yield

$$\Delta M = (1.0072 + 7.0104) - 2\,(4.0011) = (0.0154 \pm 0.0030)\ \mathrm{amu},$$

and $\Delta Mc^2 = (14.3 \pm 2.7)\,\mathrm{MeV}$. From particle detectors, C&W measured $\Delta K = 17.2\,\mathrm{MeV}$, which implies a ΔMc^2 of $17.2\,\mathrm{MeV}$ also. The experimental value of ΔMc^2 is thus $17.2\,\mathrm{MeV}$, while the upper-limit of the theoretical value is $17.0\,\mathrm{MeV}$. With more modern techniques of mass spectrometry, the rest masses in Eq. (23.3) are better known, and this gap has been closed.

5. **Kinematics:** Two particles of rest mass m_0 approach each other with equal and opposite velocity, v, in the laboratory frame. What is the total energy of one particle as measured in the rest frame of the other?

   Vector addition: this is the analogue of the problem in Test 1. Let $V = \hat{\boldsymbol{x}}v$ be the velocity of frame A w.r.t. the lab. Let $v' = -\hat{\boldsymbol{x}}v$ be the velocity of particle B w.r.t. the lab frame. Then,

$$v' = \frac{(-v) - V}{1 - (-v)\,V/c^2} = \frac{-2v}{1 + v^2/c^2}.$$

   Use

$$E = \gamma\,(v')\,mc^2.$$

$$E = \frac{1}{\sqrt{1 - \frac{4v^2/c^2}{(1+v^2/c^2)^2}}}mc^2 = \left(\frac{1 + v^2/c^2}{1 - v^2/c^2}\right)mc^2.$$

   Alternative derivation: Let E_t and P_t denote the total energy in either frame. Now $E_t^2 - (P_t c)^2$ is invariant, so it is the frame in both the lab and the moving frames. In the lab frame, $E_t = 2\gamma_0 mc^2$, where $\gamma_0 = \gamma\,(v)$, and $P_t = 0$. In the moving frame, $E_t = mc^2 + \gamma_1 mc^2$ and $P_t = \gamma_1 mv'$, where $\gamma_1 = \gamma\,(v')$. Writing down the invariant, we have

$$\left(2\gamma_0 mc^2\right)^2 = \left(mc^2 + \gamma_1 mc^2\right)^2 - \left(\gamma_1 v'c\right)^2.$$

   Expanding the squares, using $\gamma_1\,(1 - v'^2/c^2) = 1$ and eliminating terms, we obtain

$$\gamma_1 = 2\gamma_0^2 - 1,$$

   hence

$$E = \gamma_1 mc^2 = \left(2\gamma_0^2 - 1\right)mc^2 = \left(\frac{1 + v^2/c^2}{1 - v^2/c^2}\right)mc^2,$$

   as before.

6. **Kinematics:** A particle of rest mass m and speed v_0 collides and sticks to a stationary particle

of mass M. What is the final speed of the composite particle?

Initial state:

$$\text{Energy:} \qquad E = \gamma_0 mc^2 + Mc^2,$$
$$\text{Momentum:} \qquad \gamma m v_0 \hat{x}$$

Final state:

$$\text{Energy:} \qquad E = \gamma_1 \mu c^2,$$
$$\text{Momentum:} \qquad \gamma \mu v_1 \hat{\boldsymbol{x}},$$

where $\mu \neq m + M$ is the final mass. NOTE THE NONEQUALITY! Equating energies gives $\gamma_0 mc^2 + Mc^2 = \gamma_1 \mu c^2$; equating momenta gives $\gamma_1 \mu v_1 = \gamma_0 m v_0$. The energy equation implies that $\gamma_1 \mu = \gamma_0 m + M$. Plug this into the momentum equation:

$$v_1 (\gamma_0 m + M) = \gamma_0 m v_0 \implies v_1 = \frac{\gamma_0 m v_0}{\gamma_0 m + M}.$$

Note that the final mass is

$$\mu = \frac{\gamma_0 m + M}{\gamma_1} = \frac{\gamma_0}{\gamma_1} m + \frac{1}{\gamma_1} M.$$

Let us compute the first-order correction to the mass assuming that $v/c \ll 1$. We have $\mu = \mu(x)$, where $x = v^2/c^2$. To first order,

$$\mu = \mu(0) + \frac{d\mu}{dx}(0)\, x.$$

Now $\mu(0) = m + M$.

$$\frac{d\mu}{dx} = \frac{m}{\gamma_1}\frac{d\gamma_0}{dx} - \frac{m\gamma_0{}^2}{\gamma_1}\frac{d\gamma_1}{dx} - \frac{M}{\gamma_1^2}\frac{d\gamma_1}{dx}.$$

Here $\gamma_0(x) = (1-x)^{-1/2}$, hence $\gamma_0'(0) = 1/2$, and $\gamma_0(0) = 1$. Moreover,

$$\gamma_1(x) = \frac{1}{\sqrt{1 - \frac{\gamma_0^2 m^2 x}{(\gamma_0 m + M)^2}}}.$$

Hence,

$$\frac{d\gamma_1}{dx} = \tfrac{1}{2}\frac{1}{\left[1-\frac{\gamma_0^2 m^2 x}{(\gamma_0 m+M)^2}\right]^{3/2}}\frac{d}{dx}\left(\frac{\gamma_0^2 m^2 x}{(\gamma_0 m+M)^2}\right)$$

$$= \tfrac{1}{2}\frac{1}{\left[1-\frac{\gamma_0^2 m^2 x}{(\gamma_0 m+M)^2}\right]^{3/2}}\left[\frac{\gamma_0^2 m^2}{(\gamma_0 m+M)^2} + x\frac{d}{dx}\left(\frac{\gamma_0^2 m^2}{(\gamma_0 m+M)^2}\right)\right].$$

Evaluating this at zero, it is

$$\gamma_1'(0) = \tfrac{1}{2}\frac{m^2}{(m+M)^2};$$

similarly, $\gamma_1(0) = 1$. Thus,

$$\mu'(0) = \tfrac{1}{2}m - \tfrac{1}{2}m\frac{m^2}{(m+M)^2} - \tfrac{1}{2}M\frac{m^2}{(m+M)^2}.$$

Thus,

$$\mu = (m+M) + \tfrac{1}{2}x\left[m - m\frac{m^2}{(m+M)^2} - M\frac{m^2}{(m+M)^2}\right] + O\left(x^2\right).$$

Tidying this up by introducing $\delta = m/M$ and restoring $x = v^2/c^2$, the correction to the mass is

$$\mu = m + M + \tfrac{1}{2}\frac{v^2}{c^2}\left[m - (m+M)\frac{\delta^2}{(1+\delta)^2}\right] + O\left(v^4/c^4\right).$$