Lecture Notes in

Classical Mechanics

Artemio González López

Madrid, September 23, 2023

Dedicated to the loving memory of my parents, Artemio and Matilde,
to whom I owe everything I am

Artemio González López

Departamento de Física Teórica
Facultad de Ciencias Físicas
Plaza de las Ciencias, 1
28040 Madrid

Contents

0 Preliminaries

A working knowledge of elementary linear algebra, multivariable calculus, and ordinary differential equations is essential to follow a course in classical mechanics, even at a basic level. In this chapter we shall briefly review a few fundamental results and identities, usually taught in freshman courses in Linear Algebra and Calculus, that we shall often use throughout the course.

0.1 Vectors in $\mathbb{R}^3$

We shall mostly work with vectors in $\mathbb{R}^3$ (the most notable exception is the last chapter on relativistic mechanics, which shall require the use of four-vectors). Throughout these notes vectors will be typeset in *roman boldface,* like $\mathbf{a}$, $\mathbf{b}$, etc., while when writing at the blackboard we shall use instead the notation $\vec{a}$, $\vec{b}$, etc. Unless otherwise stated, the components of a vector $\mathbf{a}$ shall be denoted by a_i ($i = 1, 2, 3$). We shall normally use the notation $|\mathbf{a}|$ for the length (or magnitude) $\sqrt{a_1^2 + a_2^2 + a_3^2}$ of the vector $\mathbf{a}$, and $\mathbf{a} \cdot \mathbf{b}$ or simply $\mathbf{ab}$ to denote the scalar product $a_1b_1 + a_2b_2 + a_3b_3$ of the vectors $\mathbf{a}$ and $\mathbf{b}$.

Vector product of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$:

$$\mathbf{a} \times \mathbf{b} := \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}, \tag{0.1}$$

where $\mathbf{i} := \mathbf{e}_1 = (1, 0, 0)$, $\mathbf{j} := \mathbf{e}_2 = (0, 1, 0)$, $\mathbf{k} := \mathbf{e}_3 = (0, 0, 1)$ are the unit vectors of the standard orthonormal frame in $\mathbb{R}^3$. In particular, note that

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \qquad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \qquad \mathbf{k} \times \mathbf{i} = \mathbf{j}\,.$$

The vector product is obviously *antisymmetric* ($\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$), so that $\mathbf{a} \times \mathbf{a} = 0$, and *distributive* in both of its arguments:

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}, \qquad (\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$$

(note that the second distributive law actually follows from the first by antisymmetry).

Triple product of three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$:

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} =: \det(\mathbf{a}, \mathbf{b}, \mathbf{c}). \tag{0.2}$$

In particular, from the properties of determinants it follows that $\mathbf{a} \times \mathbf{b}$ is *perpendicular* to both $\mathbf{a}$ and $\mathbf{b}$. The triple product is *invariant under cyclic permutations* of its arguments (again by the elementary properties of determinants):

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b}\,.$$

Geometrically, $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ is equal to the volume of the parallelepiped spanned by the three vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ if the frame $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is positively oriented, and minus this volume otherwise.

Useful identities:

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c} \tag{0.3}$$
$$(\mathbf{a}\times\mathbf{b})^2 = \mathbf{a}^2\mathbf{b}^2 - (\mathbf{a}\cdot\mathbf{b})^2\,. \tag{0.4}$$

From the identity

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta_{\mathbf{ab}},$$

where $\theta_{\mathbf{ab}} \in [0,\pi]$ is the *angle* between the vectors **a** and **b**, and Eq. (0.4) it follows that

$$|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta_{\mathbf{ab}}\,,$$

0.2 Vector calculus

A *scalar function* (also called a *scalar field*) is a mapping $f : \mathbb{R}^3 \to \mathbb{R}$. Given a scalar function $f(x_1,x_2,x_3)$ of class C^1 (i.e., with continuous partial derivatives), we define its **gradient** ∇f by

$$\nabla f := \frac{\partial f}{\partial x_1}\mathbf{i} + \frac{\partial f}{\partial x_2}\mathbf{j} + \frac{\partial f}{\partial x_3}\mathbf{k}\,. \tag{0.5}$$

Note that ∇f is a *vector field*, i.e., a mapping from $\mathbb{R}^3$ to $\mathbb{R}^3$. In other words, $(\nabla f)(x_1,x_2,x_3)$ is a *vector* at each point (x_1,x_2,x_3). Given a vector field $\mathbf{A} = (A_1,A_2,A_3) : \mathbb{R}^3 \to \mathbb{R}^3$ of class C^1, we define its **divergence** $\nabla\cdot\mathbf{A}$ as the *scalar function*

$$\nabla\cdot\mathbf{A} := \frac{\partial A_1}{\partial x_1} + \frac{\partial A_2}{\partial x_2} + \frac{\partial A_3}{\partial x_3}\,. \tag{0.6}$$

The gradient and the divergence can be easily generalized to scalar functions and vector fields in $\mathbb{R}^n$ with $n > 3$. For instance, the gradient of a scalar function $f(x_1,\dots,x_n)$ (that shall be needed in Chapter 3) is defined by

$$\nabla f(x_1,\dots,x_n) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\mathbf{e}_i,$$

where $\mathbf{e}_i = (0,\dots,0,\underset{\substack{\downarrow\\ i}}{1},0,\dots,0)$ is the i-th vector of the canonical basis of $\mathbb{R}^n$. There is, however, a differential operator which can only be defined in three dimensions, namely the **curl** $\nabla\times\mathbf{A}$ of a vector field (of class C^1) $\mathbf{A}(x_1,x_2,x_3)$ in $\mathbb{R}^3$:

$$\nabla\times\mathbf{A} := \left(\frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3}\right)\mathbf{i} + \left(\frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}\right)\mathbf{j} + \left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}\right)\mathbf{k} \equiv \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ A_1 & A_2 & A_3 \end{vmatrix}\,. \tag{0.7}$$

Another important differential operator is the **Laplacian** of a scalar function $f : \mathbb{R}^3 \to \mathbb{R}$ of class C^2, defined by

$$\nabla^2 f := \nabla\cdot(\nabla f) = \sum_{i=1}^{3}\frac{\partial^2 f}{\partial x_i^2}\,. \tag{0.8}$$

Note that $\nabla^2 f$ is again a *scalar function*. The Laplacian operator ∇^2 is often denoted as Δ.

Properties

- The vector $\nabla f(x_1,\dots,x_n)$ is *orthogonal* at each point to the level (hyper)surfaces $f(x_1,\dots,x_n) = c$ of the scalar function f (where c is a real constant). In other words, ∇f is orthogonal to the *tangent vectors* of curves lying on level surfaces of f.

- If $f(x_1, x_2, x_3)$ is *any* scalar function of class C^2 we have

$$\nabla \times (\nabla f) = 0. \tag{0.9}$$

Conversely, if $\mathbf{A}(x_1, x_2, x_3)$ is a vector field of class C^1 on *all* of $\mathbb{R}^3$ such that

$$\nabla \times \mathbf{A} = 0,$$

then $\mathbf{A} = \nabla f$ for some scalar function f of class C^2.

- Similarly, if $\mathbf{A}(x_1, x_2, x_3)$ is *any* vector field of class C^2 then

$$\nabla \cdot (\nabla \times \mathbf{A}) = 0. \tag{0.10}$$

Conversely, if $\mathbf{B}(x_1, x_2, x_3)$ is a vector field of class C^1 on all of $\mathbb{R}^3$ such that

$$\nabla \cdot \mathbf{B} = 0$$

then $\mathbf{B} = \nabla \times \mathbf{A}$ for some vector field $\mathbf{A}(x_1, x_2, x_3)$ of class C^2.

Some useful identities:

$$\nabla \cdot (f\mathbf{A}) = (\nabla f) \cdot \mathbf{A} + f \nabla \cdot \mathbf{A}, \tag{0.11}$$
$$\nabla \times (f\mathbf{A}) = (\nabla f) \times \mathbf{A} + f \nabla \times \mathbf{A}, \tag{0.12}$$
$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}. \tag{0.13}$$

In the first two identities f and $\mathbf{A}$ are respectively a scalar function and a vector field of class C^1 in $\mathbb{R}^3$, while in the third one $\mathbf{A}$ is a vector field of class C^2 and $\nabla^2 \mathbf{A}$ is the vector field

$$\nabla^2 \mathbf{A} := (\nabla^2 A_1)\mathbf{i} + (\nabla^2 A_2)\mathbf{j} + (\nabla^2 A_3)\mathbf{k}.$$

As we shall see in the next chapter, the identity (0.13) plays an important role in the formulation of electromagnetic theory.

Exercise. If $\mathbf{A}$ and $\mathbf{B}$ are C^1 vector fields in $\mathbb{R}^3$, express $\nabla \cdot (\mathbf{A} \times \mathbf{B})$ in terms of $\nabla \times \mathbf{A}$ and $\nabla \times \mathbf{B}$.

0.3 Chain rule

A working knowledge of the chain rule is essential to perform even the most basic calculations in classical mechanics (or, as a matter of fact, in any branch of physics). From a formal point of view, the chain rule simply asserts that the derivative $D(f \circ g)$ of the *composition* $f \circ g$ of two differentiable functions $g : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}^p$ is the composition $Df \circ Dg$ of their derivatives. This abstract formulation yields, however, simple and intuitive formulas for the *partial derivatives* of the composition $f \circ g$, which are the ones we shall use in practice throughout the course.

Consider, for instance, a scalar function $f(x_1, \dots, x_n)$ of the n variables $(x_1, \dots, x_n)$, and suppose that each of these variables x_i is in turn a function of m real variables $(y_1, \dots, y_m)$. Then $f(x_1, \dots, x_n)$ is *implicitly* a function of the y_j's through the x_i's. In other words, when we write (for simplicity's sake) $f(x_1, \dots, x_n)$ what we really mean is $f(x_1(y_1, \dots, y_m), \dots, x_n(y_1, \dots, y_m))$. The partial derivative of this function with respect to any of its *independent* variables y_k can be computed using the following variant of the chain rule:

$$\frac{\partial}{\partial y_k} f(x_1, \dots, x_n) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial y_k}. \tag{0.14}$$

This is a straightforward generalization of the well-known elementary formula for the $n = m = 1$ case, namely

$$\frac{\mathrm{d}}{\mathrm{d}y} f(x) = \frac{\mathrm{d}f}{\mathrm{d}x} \frac{\mathrm{d}x}{\mathrm{d}y}.$$

In classical mechanics (x_1, x_2, x_3) are usually the coordinates of the *position vector* $\mathbf{r}$ of a moving particle, so *they depend implicitly on time* t. In this case, therefore, *any function* $f(x_1, x_2, x_3)$ *depends implicitly on time through the coordinates* x_i. In other words, $f(x_1, x_2, x_3)$ should be normally understood as shorthand for $f(x_1(t), x_2(t), x_3(t))$. Applying Eq. (0.14) with $n = 3$, $m = 1$ and $y_k = y_1 \equiv t$ we obtain the important formula

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x_1, x_2, x_3) = \sum_{i=1}^{3} \frac{\partial f}{\partial x_i} \dot{x}_i, \tag{0.15}$$

where (as is standard in classical mechanics) we have used Newton's notation

$$\dot{x}_i \equiv \frac{\mathrm{d}x_i}{\mathrm{d}t}$$

for the time derivative. Since $\mathbf{r} = (x_1, x_2, x_3)$, we can write $f(x_1, x_2, x_3) = f(\mathbf{r})$ and $\dot{\mathbf{r}} = (\dot{x}_1, \dot{x}_2, \dot{x}_3)$. It is also customary to use the mnemonic notation $\frac{\partial f}{\partial \mathbf{r}}$ for the gradient of the scalar function $f(\mathbf{r})$, i.e.,

$$\frac{\partial f(\mathbf{r})}{\partial \mathbf{r}} \equiv \nabla f(x_1, x_2, x_3).$$

We can then rewrite the previous formula for $\frac{\mathrm{d}f}{\mathrm{d}t} \equiv \dot{f}$ in vector form as

$$\dot{f} = \frac{\partial f}{\partial \mathbf{r}} \cdot \dot{\mathbf{r}}. \tag{0.16}$$

Exercise. Compute the gradient of a scalar function $g(r)$ that depends on $\mathbf{r} = (x_1, x_2, x_3)$ only through $r := \sqrt{x_1^2 + x_2^2 + x_3^2}$.

Solution. Applying the chain rule we obtain

$$\frac{\partial g(r)}{\partial x_i} = g'(r) \frac{\partial r}{\partial x_i} = g'(r) \frac{x_i}{r}.$$

Hence

$$\nabla[g(r)] \equiv \frac{\partial g(r)}{\partial \mathbf{r}} = \sum_{i=1}^{3} \frac{\partial g(r)}{\partial x_i} \mathbf{e}_i = \frac{g'(r)}{r} \sum_{i=1}^{3} x_i \mathbf{e}_i = g'(r) \frac{\mathbf{r}}{r} \equiv g'(r) \mathbf{e}_r.$$

0.4 Total and partial time derivatives

Understanding the difference between total and partial time derivatives of a scalar function $f(t, x_1, x_2, x_3) \equiv f(t, \mathbf{r})$ is again crucial to follow even the simplest arguments in classical mechanics. To begin with, as usual the notation $\frac{\partial f}{\partial t}$ stands for the *partial derivative* of f with respect to t considering f as a function of the *independent* variables (t, x_1, x_2, x_3). For instance, if $f(t, \mathbf{r}) = c^2 t^2 + \mathbf{r}^2$ (where c is a constant with the dimension of velocity) then

$$\frac{\partial f}{\partial t} = 2ct.$$

On the other hand, if —as is customary in classical mechanics— we consider each coordinate x_i as an implicit function of t (i.e., we write x_i as shorthand for $x_i(t)$), then $f(t, \mathbf{r})$ *becomes a function of t only,* through the explicit dependence of f on t *plus* the implicit dependence on t of $\mathbf{r} = (x_1, x_2, x_3)$. The **total time derivative** of f is the derivative with respect to t of the latter function, i.e., of $f(t, \mathbf{r})$ *considered as* $f(t, \mathbf{r}(t))$. Applying again the chain rule (with $n = 4$, $m = 1$) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t, \mathbf{r}) \equiv \dot{f} = \frac{\partial f}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}t} + \sum_{i=1}^{3} \frac{\partial f}{\partial x_i}\dot{x}_i = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{r}} \cdot \dot{\mathbf{r}} . \tag{0.17}$$

This is of course a generalization of Eq. (0.16), to which it reduces when f does not depend *explicitly* on time. For instance, if $f(t, \mathbf{r}) = c^2 t^2 + \mathbf{r}^2$ then

$$\frac{\mathrm{d}f}{\mathrm{d}t} \equiv \dot{f} = 2ct + \sum_{i=1}^{3} 2x_i \dot{x}_i = 2ct + 2\mathbf{r} \cdot \dot{\mathbf{r}} \neq \frac{\partial f}{\partial t} .$$

Note that, in general, $\dot{f}$ is a scalar function of the variables $(t, \mathbf{r}, \dot{\mathbf{r}})$, i.e., it depends not only on time and the coordinates of the particle, but also on its *velocity*.

1 Review of Newtonian mechanics

1.1 Kinematics

Position vector of a particle moving in ordinary space ($\mathbb{R}^3$):

$$\mathbf{r} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \equiv \sum_{i=1}^{3} x_i\mathbf{e}_i \,,$$

where $\mathbf{e}_i$ is the i-th coordinate unit vector in a *Cartesian* orthogonal coordinate system:

$$\mathbf{e}_1 = (1,0,0) \equiv \mathbf{i}\,, \quad \mathbf{e}_2 = (0,1,0) \equiv \mathbf{j}\,, \quad \mathbf{e}_3 = (0,0,1) \equiv \mathbf{k}\,.$$

Velocity and acceleration:

$$\mathbf{v} := \dot{\mathbf{r}}\,, \qquad \mathbf{a} := \dot{\mathbf{v}} = \ddot{\mathbf{r}}\,,$$

where the dot denotes derivative with respect to *time* t. In Cartesian coordinates,

$$\mathbf{v} = \sum_{i=1}^{3} v_i\mathbf{e}_i \,, \qquad \mathbf{a} = \sum_{i=1}^{3} a_i\mathbf{e}_i \,,$$

where (since the coordinate vectors $\mathbf{e}_i$ are *constant*)

$$v_i = \dot{x}_i \,, \qquad a_i = \ddot{x}_i \,.$$

Notation: $\quad r = |\mathbf{r}|\,, \quad v = |\mathbf{v}|\,.$

Exercise. Show that if $\mathbf{v} \neq 0$ then

$$\dot{v} = \frac{\mathbf{v}\cdot\mathbf{a}}{v}\,.$$

In particular, if v is constant the velocity and acceleration vectors are *orthogonal*. Note also that the latter formula can be written as

$$\dot{v} = \mathbf{a}\cdot\mathbf{t}\,,$$

where $\mathbf{t} := \mathbf{v}/v$ is the *unit tangent vector* along the trajectory. Thus the *tangential acceleration* $a_t := \mathbf{a}\cdot\mathbf{t}$ is equal to the time derivative of v.

Exercise. Show that $\mathbf{r}\dot{\mathbf{r}} = r\dot{r}$.

1.2 Curvilinear coordinate systems

Consider the system of curvilinear coordinates $\mathbf{q} = (q_1, q_2, q_3)$ defined by a bijective transformation $\mathbf{r} = \mathbf{r}(\mathbf{q})$ with a *non-vanishing Jacobian*

$$\det\left(\frac{\partial\mathbf{r}}{\partial\mathbf{q}}\right) \equiv \det\left(\frac{\partial x_i}{\partial q_j}\right)_{1\leqslant i,j\leqslant 3}\,.$$

For example, in the case of *spherical coordinates* we have $\mathbf{q} = (r, \theta, \varphi)$, with

$$r \geqslant 0\,, \quad 0 \leqslant \theta \leqslant \pi\,, \quad 0 \leqslant \varphi < 2\pi,$$

and

$$\boxed{\mathbf{r} = r(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)} \tag{1.1}$$

(cf. Fig. 1.1). The **unit coordinate vectors** $\{\mathbf{e}_{q_i}\}_{i=1}^3$ are the unit vectors tangent to the *coordinate curves*, in which one of the curvilinear coordinates q_i varies while the rest are held constant. We thus have

$$\boxed{\mathbf{e}_{q_i} = \frac{1}{h_i}\frac{\partial \mathbf{r}}{\partial q_i}\,, \qquad \text{with} \quad h_i := \left|\frac{\partial \mathbf{r}}{\partial q_i}\right|.} \tag{1.2}$$

Note that $h_i(\mathbf{q}) > 0$ for all i, since the vector $\dfrac{\partial \mathbf{r}}{\partial q_i}$ is the i-th column of the Jacobian matrix $\dfrac{\partial \mathbf{r}}{\partial \mathbf{q}}$, whose determinant —the *Jacobian* of the mapping $\mathbf{r}(\mathbf{q})$— is nonvanishing by hypothesis. It is also important to realize that in general the unit coordinate vectors $\mathbf{e}_{q_i}$ are *not* constant, but depend on the curvilinear coordinates $\mathbf{q}$ of the point at which they are evaluated. In other words, each $\mathbf{e}_{q_i}$ is a *vector field* in $\mathbb{R}^3$.

We shall normally deal with **orthogonal** coordinate systems, whose unit coordinate vectors are mutually orthogonal and thus make up an *orthonormal frame* at each point of $\mathbb{R}^3$. Since $\mathbf{e}_{q_i}$ is a unit vector proportional to $\dfrac{\partial \mathbf{r}}{\partial q_i}$, the necessary and sufficient condition for a coordinate system to be orthogonal is that

$$\boxed{\frac{\partial \mathbf{r}}{\partial q_i}\cdot\frac{\partial \mathbf{r}}{\partial q_j} = 0, \qquad \forall\, i \neq j\,.}$$

We shall also assume that the coordinate system $\mathbf{q}$ is *positively oriented*, by which we mean that

$$\boxed{\mathbf{e}_{q_1} \times \mathbf{e}_{q_2} = \mathbf{e}_{q_3}\,,}$$

or equivalently

$$\det(\mathbf{e}_{q_1}, \mathbf{e}_{q_2}, \mathbf{e}_{q_3}) = 1$$

(cf. second exercise on p. 12). Since

$$\det\left(\frac{\partial \mathbf{r}}{\partial \mathbf{q}}\right) = \det\left(\frac{\partial \mathbf{r}}{\partial q_1}, \frac{\partial \mathbf{r}}{\partial q_2}, \frac{\partial \mathbf{r}}{\partial q_3}\right) = \det\left(h_1\mathbf{e}_{q_1}, h_2\mathbf{e}_{q_2}, h_3\mathbf{e}_{q_3}\right) = h_1h_2h_3\det\left(\mathbf{e}_{q_1}, \mathbf{e}_{q_2}, \mathbf{e}_{q_3}\right)$$

with $h_i > 0$ for all i, the necessary and sufficient condition for the orthogonal coordinate system $\mathbf{q}$ to be positively oriented is that the Jacobian of the transformation $\mathbf{r}(\mathbf{q})$ be *positive*:

$$\boxed{\det\left(\frac{\partial \mathbf{r}}{\partial \mathbf{q}}\right) > 0\,.}$$

- The *line element* $\mathrm{d}s^2 := \mathrm{d}\mathbf{r}^2$ can be easily expressed in any orthogonal curvilinear system $\mathbf{q}$ using the chain rule:

$$\mathrm{d}s^2 = \mathrm{d}\mathbf{r}^2 = \left(\sum_{i=1}^3 \frac{\partial \mathbf{r}}{\partial q_i}\,\mathrm{d}q_i\right)^2 = \left(\sum_{i=1}^3 h_i\mathbf{e}_{q_i}\,\mathrm{d}q_i\right)^2 = \boxed{\sum_{i=1}^3 h_i^2\,\mathrm{d}q_i^2\,,}$$

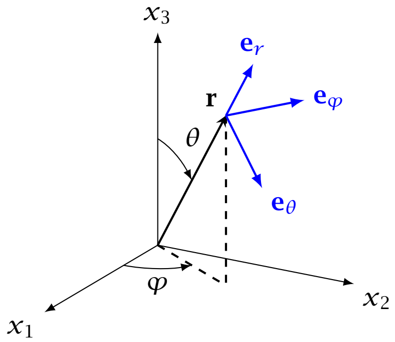


Figure 1.1. Spherical coordinate system.

where in the last equality we have used the fact that $\mathbf{e}_{q_i} \cdot \mathbf{e}_{q_j} = \delta_{ij}$. In particular, the line element along the i-th coordinate curve is given by

$$ds = h_i \, \mathrm{d}q_i, \qquad i = 1, 2, 3,$$

a formula that is often used to compute h_i by geometric arguments.

- If $\mathbf{A}$ is a vector field in $\mathbb{R}^3$ we define its components in an orthogonal coordinate system $\mathbf{q}$ by

$$\boxed{A_{q_i} := \mathbf{A} \cdot \mathbf{e}_{q_i},}$$

so that we can write

$$\boxed{\mathbf{A} = \sum_{i=1}^{3} A_{q_i} \mathbf{e}_{q_i} .}$$

The components of the *velocity vector* $\mathbf{v}$ in such a system can also be readily computed applying the chain rule. Indeed,

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \sum_{i=1}^{3} \frac{\partial \mathbf{r}}{\partial q_i} \dot{q}_i = \boxed{\sum_{i=1}^{3} h_i \dot{q}_i \mathbf{e}_{q_i} ,}$$

and hence

$$\boxed{v_{q_i} = h_i \dot{q}_i .} \tag{1.3}$$

1.2.1 Spherical coordinates

In this case

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial r} &= (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) , \\
\frac{\partial \mathbf{r}}{\partial \theta} &= r(\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta) , \\
\frac{\partial \mathbf{r}}{\partial \varphi} &= r\sin\theta(-\sin\varphi, \cos\varphi, 0),
\end{aligned}$$

and hence

$$\boxed{h_r = \left|\frac{\partial \mathbf{r}}{\partial r}\right| = 1, \qquad h_\theta = \left|\frac{\partial \mathbf{r}}{\partial \theta}\right| = r, \qquad h_\varphi = \left|\frac{\partial \mathbf{r}}{\partial \varphi}\right| = r\sin\theta ,}$$

so that

$$\begin{aligned}\mathbf{e}_r &= \frac{\partial\mathbf{r}}{\partial r} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \equiv \frac{\mathbf{r}}{r}\,,\\ \mathbf{e}_\theta &= \frac{1}{r}\frac{\partial\mathbf{r}}{\partial\theta} = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta)\,,\\ \mathbf{e}_\varphi &= \frac{1}{r\sin\theta}\frac{\partial\mathbf{r}}{\partial\varphi} = (-\sin\varphi, \cos\varphi, 0)\end{aligned} \tag{1.4}$$

(cf. Fig. 1.1). Spherical coordinates are *orthogonal*, as from the latter equations it is easily verified that

$$\mathbf{e}_r\cdot\mathbf{e}_\theta = \mathbf{e}_r\cdot\mathbf{e}_\varphi = \mathbf{e}_\theta\cdot\mathbf{e}_\varphi = 0\,.$$

Thus the vectors $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi\}$ form an orthonormal frame at each point. This frame is *positively oriented*, since

$$\mathbf{e}_r\times\mathbf{e}_\theta = \mathbf{e}_\varphi\,,$$

or, equivalently,

$$(\mathbf{e}_r\times\mathbf{e}_\theta)\cdot\mathbf{e}_\varphi = 1\,.$$

Velocity and acceleration in spherical coordinates.

The components of the velocity vector $\mathbf{v}$ in spherical coordinates can be easily computed using Eq. (1.3). It is also instructive to obtain them directly differentiating the relation $\mathbf{r} = r\mathbf{e}_r$, as we shall next explain. To this end, we first compute the time derivatives of the unit coordinate vectors. Note first of all that, since

$$\mathbf{e}_\alpha\cdot\mathbf{e}_\alpha = 1 \qquad (\alpha = r, \theta, \varphi)\,,$$

differentiating with respect to time we obtain

$$\dot{\mathbf{e}}_\alpha\cdot\mathbf{e}_\alpha = 0\,.$$

Since the vectors $\mathbf{e}_\alpha$ are mutually orthogonal, $\dot{\mathbf{e}}_r$ must be a linear combination of $\mathbf{e}_\theta$ and $\mathbf{e}_\varphi$, and similarly for the remaining coordinate vectors. More precisely, applying the *chain rule* we arrive at

$$\begin{aligned}\dot{\mathbf{e}}_r &= \frac{\partial\mathbf{e}_r}{\partial\theta}\dot\theta + \frac{\partial\mathbf{e}_r}{\partial\varphi}\dot\varphi = \dot\theta\mathbf{e}_\theta + \sin\theta\dot\varphi\mathbf{e}_\varphi\,,\\ \dot{\mathbf{e}}_\theta &= \frac{\partial\mathbf{e}_\theta}{\partial\theta}\dot\theta + \frac{\partial\mathbf{e}_\theta}{\partial\varphi}\dot\varphi = -\dot\theta\mathbf{e}_r + \cos\theta\dot\varphi\mathbf{e}_\varphi\,,\\ \dot{\mathbf{e}}_\varphi &= \frac{\partial\mathbf{e}_\varphi}{\partial\varphi}\dot\varphi = -(\cos\varphi, \sin\varphi, 0)\,\dot\varphi = -\dot\varphi(\sin\theta\mathbf{e}_r + \cos\theta\mathbf{e}_\theta)\,.\end{aligned} \tag{1.5}$$

From the above relations we easily deduce that

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(r\mathbf{e}_r) = \dot r\mathbf{e}_r + r\dot{\mathbf{e}}_r = \dot r\mathbf{e}_r + r\dot\theta\mathbf{e}_\theta + r\sin\theta\dot\varphi\mathbf{e}_\varphi\,, \tag{1.6}$$

and hence

$$v_r = \dot r\,, \qquad v_\theta = r\dot\theta\,, \qquad v_\varphi = r\sin\theta\,\dot\varphi\,. \tag{1.7}$$

Note that, since the coordinate vectors are *orthonormal*,

$$v^2 = v_r^2 + v_\theta^2 + v_\varphi^2 = \dot r^2 + r^2(\dot\theta^2 + \sin\theta^2\dot\varphi^2)\,. \tag{1.8}$$

Similarly, differentiating Eq. (1.6) with respect to t and using Eqs. (1.5) we obtain:

$$\begin{aligned}\mathbf{a} &= \ddot{r}\mathbf{e}_r + \dot{r}(\dot{\theta}\mathbf{e}_\theta + \sin\theta\dot{\varphi}\mathbf{e}_\varphi) + (r\ddot{\theta} + \dot{r}\dot{\theta})\mathbf{e}_\theta + r\dot{\theta}(-\dot{\theta}\mathbf{e}_r + \cos\theta\dot{\varphi}\mathbf{e}_\varphi) \\ &\quad + (r\sin\theta\ddot{\varphi} + \sin\theta\dot{r}\dot{\varphi} + r\cos\theta\dot{\theta}\dot{\varphi})\mathbf{e}_\varphi - r\sin\theta\dot{\varphi}^2(\sin\theta\mathbf{e}_r + \cos\theta\mathbf{e}_\theta) \\ &= a_r\mathbf{e}_r + a_\theta\mathbf{e}_\theta + a_\varphi\mathbf{e}_\varphi\,,\end{aligned}$$

where the components of the acceleration in spherical coordinates are given by

$$\boxed{\begin{aligned} a_r &= \ddot{r} - r\dot{\theta}^2 - r\sin^2\theta\,\dot{\varphi}^2\,,\\ a_\theta &= r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\sin\theta\cos\theta\,\dot{\varphi}^2\,,\\ a_\varphi &= r\sin\theta\,\ddot{\varphi} + 2\sin\theta\dot{r}\dot{\varphi} + 2r\cos\theta\,\dot{\theta}\dot{\varphi}\,.\end{aligned}} \tag{1.9}$$

1.2.2 Cylindrical coordinates

Cylindrical coordinates (ρ, φ, z) are defined by

$$\boxed{\mathbf{r} = (\rho\cos\varphi, \rho\sin\varphi, z)\,,}$$

where

$$\rho \geqslant 0\,, \qquad 0 \leqslant \varphi < 2\pi\,, \qquad z \in \mathbb{R}$$

(cf. Fig. 1.2). Now

$$\frac{\partial\mathbf{r}}{\partial\rho} = (\cos\varphi, \sin\varphi, 0), \qquad \frac{\partial\mathbf{r}}{\partial\varphi} = \rho(-\sin\varphi, \cos\varphi, 0), \qquad \frac{\partial\mathbf{r}}{\partial z} = (0,0,1)\,,$$

and thus

$$\boxed{h_\rho = 1\,, \qquad h_\varphi = \rho\,, \qquad h_z = 1\,.}$$

Proceeding as before we obtain:

$$\boxed{\begin{aligned}\mathbf{e}_\rho &= \frac{\partial\mathbf{r}}{\partial\rho} = (\cos\varphi, \sin\varphi, 0)\,,\\ \mathbf{e}_\varphi &= \frac{1}{\rho}\frac{\partial\mathbf{r}}{\partial\varphi} = (-\sin\varphi, \cos\varphi, 0)\,,\\ \mathbf{e}_z &= \frac{\partial\mathbf{r}}{\partial z} = (0,0,1)\,.\end{aligned}}$$

Note that, once again,

$$\mathbf{e}_\rho\cdot\mathbf{e}_\varphi = \mathbf{e}_\rho\cdot\mathbf{e}_z = \mathbf{e}_\varphi\cdot\mathbf{e}_z = 0\,, \qquad \mathbf{e}_\rho\times\mathbf{e}_\varphi = \mathbf{e}_z\,,$$

and hence the cylindrical coordinates (ρ, φ, z) are also orthonormal and positively oriented. Differentiating with respect to t the equations for the coordinate vectors we immediately obtain the relations

$$\dot{\mathbf{e}}_\rho = \dot{\varphi}\mathbf{e}_\varphi\,, \qquad \dot{\mathbf{e}}_\varphi = -\dot{\varphi}\mathbf{e}_\rho\,, \qquad \dot{\mathbf{e}}_z = 0\,.$$

Since now

$$\mathbf{r} = \rho\mathbf{e}_\rho + z\mathbf{e}_z\,, \tag{1.10}$$

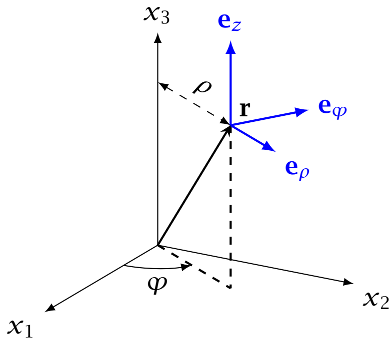


Figure 1.2. Cylindrical coordinates system.

differentiating twice with respect to t and proceeding as before we easily arrive at the following formulas for the components of velocity and acceleration in cylindrical coordinates:

$$v_\rho = \dot\rho\,, \quad v_\varphi = \rho\dot\varphi\,, \quad v_z = \dot z\,; \qquad a_\rho = \ddot\rho - \rho\dot\varphi^2\,, \quad a_\varphi = \rho\ddot\varphi + 2\dot\rho\dot\varphi\,, \quad a_z = \ddot z\,. \tag{1.11}$$

Note that the equations for the components of the velocity could also have been directly obtained using Eq. (1.3). Again, from the orthonormal character of the coordinate vectors $\{\mathbf{e}_\rho, \mathbf{e}_\varphi, \mathbf{e}_z\}$ it follows that

$$v^2 = v_\rho^2 + v_\varphi^2 + v_z^2 = \dot\rho^2 + \rho^2\dot\varphi^2 + \dot z^2\,.$$

Exercise. Prove that in any orthogonal curvilinear system of coordinates the components $a_k := \mathbf{a}\cdot\mathbf{e}_{q_k}$ of the acceleration are given by

$$a_k = h_k\ddot q_k + \sum_{i,j=1}^{3} \Gamma_k^{ij}\dot q_i\dot q_j\,, \qquad \text{with} \quad \Gamma_k^{ij} := \frac{1}{h_k}\frac{\partial\mathbf{r}}{\partial q_k}\cdot\frac{\partial^2\mathbf{r}}{\partial q_i\partial q_j}\,.$$

Exercise. Show that if the vectors $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$ form a positively oriented orthonormal frame then for $i \neq j$ we have

$$\mathbf{n}_i \times \mathbf{n}_j = \operatorname{sgn}(i,j,k)\mathbf{n}_k\,, \tag{1.12}$$

where (i,j,k) is a permutation of $(1,2,3)$ and $\operatorname{sgn}(i,j,k)$ denotes its sign.

Solution. If the frame $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$ is orthonormal then $\mathbf{n}_1\times\mathbf{n}_2$ is of unit length (since $\sin\theta_{\mathbf{n}_1\mathbf{n}_2} = 1$) and perpendicular to both $\mathbf{n}_1$ and $\mathbf{n}_2$, so that $\mathbf{n}_1\times\mathbf{n}_2 = \pm\mathbf{n}_3$. By definition, the frame $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$ is *positively oriented* if $\mathbf{n}_1\times\mathbf{n}_2 = \mathbf{n}_3$, negatively oriented if $\mathbf{n}_1\times\mathbf{n}_2 = -\mathbf{n}_3$. For instance, the canonical frame $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is positively oriented, since $\mathbf{i}\times\mathbf{j} = \mathbf{k}$. Note that

$$(\mathbf{n}_1\times\mathbf{n}_2)\cdot\mathbf{n}_3 = (\pm\mathbf{n}_3)\cdot\mathbf{n}_3 = \pm 1\,,$$

so the frame $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$ is positively oriented if and only if $(\mathbf{n}_1\times\mathbf{n}_2)\cdot\mathbf{n}_3 = 1$. In general, if $i \neq j$ then $\mathbf{n}_i\times\mathbf{n}_j = \varepsilon(i,j,k)\mathbf{n}_k$, with i, j, and k different from one another and $\varepsilon(i,j,k) = \pm1$. To determine the sign $\varepsilon(i,j,k)$, we take the scalar product of $\mathbf{n}_i\times\mathbf{n}_j$ with $\mathbf{n}_k$ and use the elementary properties of determinants to obtain

$$\varepsilon(i,j,k) = (\mathbf{n}_i\times\mathbf{n}_j)\cdot\mathbf{n}_k = \det(\mathbf{n}_i, \mathbf{n}_j, \mathbf{n}_k) = \operatorname{sgn}(i,j,k)\det(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3) = \operatorname{sgn}(i,j,k),$$

since the frame $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$ is positively oriented by hypothesis.

Remark. Introducing *Levi-Civita's completely antisymmetric tensor*

$$\varepsilon_{ijk} = \begin{cases} 1\,, & (i,j,k) \text{ even permutation of } (1,2,3) \\ -1\,, & (i,j,k) \text{ odd permutation of } (1,2,3) \\ 0\,, & \text{otherwise,} \end{cases} \tag{1.13}$$

(i.e., $\varepsilon_{ijk} = \operatorname{sgn}(i,j,k)$ if i, j, k are all different and zero otherwise) we can write

$$\mathbf{n}_i \times \mathbf{n}_j = \sum_{k=1}^{3} \varepsilon_{ijk}\mathbf{n}_k\,. \tag{1.14}$$

Indeed, if $i = j$ both sides of the latter equation vanish. On the other hand, if $i \neq j$ then (1.14) is equivalent to (1.12), since by the definition of ε_{ijk} the only nonzero term in the sum (1.14) is the one with k different from both i and j, and in that case $\varepsilon_{ijk} = \operatorname{sgn}(i,j,k)$ by definition. From Eq. (1.14) it follows that we can express the vector product of two three-dimensional vectors **a** and **b** in terms of the Levi–Civita tensor as follows:

$$\mathbf{a}\times\mathbf{b} = \left(\sum_{i=1}^{3} a_i\mathbf{e}_i\right)\times\left(\sum_{j=1}^{3} b_j\mathbf{e}_j\right) = \sum_{i,j=1}^{3} a_ib_j\mathbf{e}_i\times\mathbf{e}_j = \sum_{i,j=1}^{3} a_ib_j\sum_{k=1}^{3}\varepsilon_{ijk}\mathbf{e}_k = \sum_{k=1}^{3}(\mathbf{a}\times\mathbf{b})_k\mathbf{e}_k\,,$$

with

$$(\mathbf{a}\times\mathbf{b})_k = \sum_{i,j=1}^{3}\varepsilon_{ijk}a_ib_j\,.$$

Note that we can also write the previous formula as

$$(\mathbf{a}\times\mathbf{b})_k = \sum_{i,j=1}^{3}\varepsilon_{kij}a_ib_j\,,$$

since $\varepsilon_{ijk} = \varepsilon_{kij}$. For instance,

$$(\mathbf{a}\times\mathbf{b})_1 = \varepsilon_{123}a_2b_3 + \varepsilon_{132}a_3b_2 = a_2b_3 - a_3b_2,$$

since $\operatorname{sgn}(1,2,3) = +1$ and $\operatorname{sgn}(1,3,2) = -1$. Note that a similar formula can be applied to computed the curl of a vector field **F** in $\mathbb{R}^3$, namely

$$(\nabla\times\mathbf{F})_i = \sum_{j,k=1}^{3}\varepsilon_{ijk}\frac{\partial F_k}{\partial x_j} = \frac{\partial F_k}{\partial x_j} - \frac{\partial F_j}{\partial x_k}\,, \qquad (i,j,k) = \text{cyclic permutation of } (1,2,3)\,. \tag{1.15}$$

Exercise. Consider the system of orthogonal curvilinear coordinates $\mathbf{q} = (q_1,q_2,q_3)$, and let f be a smooth scalar function. Show that

$$\nabla f = \sum_{i=1}^{3}\frac{1}{h_i}\frac{\partial f}{\partial q_i}\mathbf{e}_{q_i} \quad\Longrightarrow\quad (\nabla f)_{q_i} = \frac{1}{h_i}\frac{\partial f}{\partial q_i}\,. \tag{1.16}$$

Solution. Although the gradient has been defined in a *Cartesian* (orthogonal) coordinate system, we can find an identity involving ∇f that is independent of any coordinate system. Indeed,

$$\nabla f\cdot \mathrm{d}\mathbf{r} = \sum_{i=1}^{3}\frac{\partial f}{\partial x_i}\mathrm{d}x_i = \mathrm{d}f\,,$$

where the *differential* $\mathrm{d}f$ of f can be computed in *any* curvilinear coordinate system **q** by the standard formula

$$\mathrm{d}f = \sum_{i=1}^{3} \frac{\partial f}{\partial q_i}\,\mathrm{d}q_i.$$

From the identity

$$\mathrm{d}\mathbf{r} = \sum_{i=1}^{3} \frac{\partial \mathbf{r}}{\partial q_i}\,\mathrm{d}q_i = \sum_{i=1}^{3} h_i \mathbf{e}_{q_i}\,\mathrm{d}q_i$$

we then obtain

$$\mathrm{d}f = \sum_{i=1}^{3} \frac{\partial f}{\partial q_i}\,\mathrm{d}q_i = \nabla f \cdot \mathrm{d}\mathbf{r} = \sum_{i=1}^{3} h_i (\nabla f \cdot \mathbf{e}_{q_i})\,\mathrm{d}q_i = \sum_{i=1}^{3} h_i (\nabla f)_{q_i}\,\mathrm{d}q_i\,.$$

Equating the coefficients of $\mathrm{d}q_i$ in the second and last expression for $\mathrm{d}f$ we deduce that $\frac{\partial f}{\partial q_i} = h_i (\nabla f)_{q_i}$, as claimed.

Remark. Applying Gauss's theorem to the infinitesimal solid whose curvilinear coordinates lie between q_i and $q_i + \mathrm{d}q_i$ ($i = 1, 2, 3$), it can be shown that the *divergence* of a smooth vector field

$$\mathbf{F} = \sum_{i=1}^{3} F_{q_i}(\mathbf{q})\,\mathbf{e}_{q_i}$$

in a curvilinear *orthogonal* system of coordinates $\mathbf{q}$ can be expressed as

$$\nabla \cdot \mathbf{F} = \frac{1}{h_1 h_2 h_3} \sum_{i=1}^{3} \frac{\partial}{\partial q_i} (h_j h_k F_{q_i})\,,$$

where $\{i, j, k\} = \{1, 2, 3\}$. From Eq. (1.16) it then follows that the *Laplacian* of a smooth scalar function f is given by

$$\nabla^2 f := \nabla \cdot (\nabla f) = \frac{1}{h_1 h_2 h_3} \sum_{i=1}^{3} \frac{\partial}{\partial q_i} \left(\frac{h_j h_k}{h_i} \frac{\partial f}{\partial q_i} \right).$$

Similarly, if the orthogonal curvilinear coordinate system $\mathbf{q}$ is *positively oriented*, applying Stokes's theorem to suitable infinitesimal surfaces perpendicular to the unit coordinate vectors $\mathbf{e}_{q_i}$ at an arbitrary point one can show that

$$\nabla \times \mathbf{F} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \mathbf{e}_{q_1} & h_2 \mathbf{e}_{q_2} & h_3 \mathbf{e}_{q_3} \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1 F_{q_1} & h_2 F_{q_2} & h_3 F_{q_3} \end{vmatrix} .$$

For the spherical coordinate system $h_r = 1$, $h_\theta = r$, $h_\varphi = r\sin\theta$, and therefore

$$
\begin{aligned}
\nabla f &= \frac{\partial f}{\partial r}\,\mathbf{e}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\,\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \varphi}\,\mathbf{e}_\varphi\,,\\
\nabla^2 f &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2}\,,\\
\nabla\cdot\mathbf{F} &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2F_r) + \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta F_\theta) + \frac{\partial F_\varphi}{\partial\varphi}\right],\\
\nabla\times\mathbf{F} &= \left[\frac{\partial}{\partial\theta}(\sin\theta\,F_\varphi) - \frac{\partial F_\theta}{\partial\varphi}\right]\frac{\mathbf{e}_r}{r\sin\theta} + \left[\frac{\partial F_r}{\partial\varphi} - \sin\theta\frac{\partial}{\partial r}(rF_\varphi)\right]\frac{\mathbf{e}_\theta}{r\sin\theta}\\
&\quad + \left[\frac{\partial}{\partial r}(rF_\theta) - \frac{\partial F_r}{\partial\theta}\right]\frac{\mathbf{e}_\varphi}{r}\,.
\end{aligned}
\tag{1.17}
$$

Likewise, in cylindrical coordinates $h_\rho = h_z = 1$, $h_\varphi = \rho$, and thus

$$
\begin{aligned}
\nabla f &= \frac{\partial f}{\partial\rho}\,\mathbf{e}_\rho + \frac{1}{\rho}\frac{\partial f}{\partial\varphi}\,\mathbf{e}_\varphi + \frac{\partial f}{\partial z}\,\mathbf{e}_z\,,\\
\nabla^2 f &= \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial f}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial\varphi^2} + \frac{\partial^2 f}{\partial z^2}\,,\\
\nabla\cdot\mathbf{F} &= \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho F_\rho) + \frac{1}{\rho}\frac{\partial F_\varphi}{\partial\varphi} + \frac{\partial F_z}{\partial z}\,,\\
\nabla\times\mathbf{F} &= \left(\frac{\partial F_z}{\partial\varphi} - \rho\frac{\partial F_\varphi}{\partial z}\right)\frac{\mathbf{e}_\rho}{\rho} + \left(\frac{\partial F_\rho}{\partial z} - \frac{\partial F_z}{\partial\rho}\right)\mathbf{e}_\varphi + \left[\frac{\partial}{\partial\rho}(\rho F_\varphi) - \frac{\partial F_\rho}{\partial\varphi}\right]\frac{\mathbf{e}_z}{\rho}\,.
\end{aligned}
\tag{1.18}
$$

Exercise. Show that the volume element in an orthogonal curvilinear coordinate system $\mathbf{q}$ is given by

$$\mathrm{d}^3\mathbf{r} = h_1h_2h_3\,\mathrm{d}q_1\,\mathrm{d}q_2\,\mathrm{d}q_3\,.$$

1.2.3 Motion on a plane in polar coordinates.

Suppose that the particle moves on a plane, which we shall take as the plane $z = 0$, so that $z(t) = 0$ for all t. In this case $\dot z = \ddot z = v_z = a_z = 0$, $\rho = r$ (distance to the origin) and (r,φ) are *polar coordinates* in the plane of motion (cf. Fig. 1.3). The previous formulas (1.11) then reduce to the following:

$$v_r = \dot r\,,\qquad v_\varphi = r\dot\varphi\,,\qquad a_r = \ddot r - r\dot\varphi^2\,,\qquad a_\varphi = r\ddot\varphi + 2\dot r\dot\varphi\,. \tag{1.19}$$

In particular, if $r(t) = R$ for all t (*circular motion*) we obtain the familiar formulas

$$v_r = 0\,,\qquad v_\varphi = R\dot\varphi\,,\qquad a_r = -R\dot\varphi^2\,,\qquad a_\varphi = R\ddot\varphi\,. \tag{1.20}$$

Note, in particular, that, although the radial component of the velocity is identically zero, even when $\ddot\varphi = 0$ there is a negative radial acceleration $-R\dot\varphi^2 = -v^2/R$ (*centripetal acceleration*).

Example 1.1. *Angular velocity.*
Consider next a particle rotating around a fixed axis. Taking the rotation axis as the z axis and the plane of motion as the plane $z = 0$, the particle's motion is described in polar coordinates by Eqs. (1.20). From these equations it follows that

$$\mathbf{v} = R\dot\varphi(t)\,\mathbf{e}_\varphi\,,$$

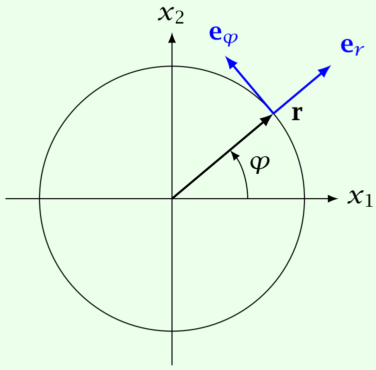


Figure 1.3. Polar coordinates.

and since $\mathbf{e}_\varphi = \mathbf{e}_z \times \mathbf{e}_\rho = \mathbf{e}_z \times \mathbf{e}_r$ we have

$$\mathbf{v} = R\dot{\varphi}(t)\mathbf{e}_z \times \mathbf{e}_r = (\dot{\varphi}(t)\mathbf{e}_z) \times \mathbf{r}\,.$$

Hence in this case we can express the particle's velocity as

$$\boxed{\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}\,,}$$

where

$$\boldsymbol{\omega}(t) = \dot{\varphi}(t)\mathbf{e}_z$$

is called the **angular velocity**. The angular velocity is therefore a *vector* directed along the axis of rotation, whose magnitude is the *absolute value* $|\dot{\varphi}(t)|$ of the angular velocity of rotation. Rotation around the z axis is called "left-handed" if $\dot{\varphi}(t) > 0$ (i.e., if $\boldsymbol{\omega}$ and $\mathbf{e}_z$ have the same direction) and "right-handed" if $\dot{\varphi}(t) < 0$ (i.e., if $\boldsymbol{\omega}$ and $\mathbf{e}_z$ have opposite directions).

1.3 Newton's laws. Inertial frames. Galileo's relativity principle

1.3.1 Newton's laws

In (non-relativistic) classical mechanics, the **linear momentum** (or momentum, for short) of a particle is defined by

$$\boxed{\mathbf{p} = m\mathbf{v} = m\dot{\mathbf{r}}\,,} \tag{1.21}$$

where m is the particle's **mass**. In classical mechanics the mass is a positive *constant* (in particular, velocity independent) parameter characteristic of each particle. In modern notation and terminology, the first two of **Newton's laws** can be stated as follows:

I. *In the absence of external forces, the momentum* (*and, hence, the velocity*) *of a particle remains constant.*

II. *If an external force* $\mathbf{F}$ *acts on a particle, the rate of variation of its momentum is given by*

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{F}\,. \tag{1.22}$$

Since the particle's mass is independent of the velocity, the last equation is equivalent to

$$\boxed{\mathbf{F} = m\mathbf{a} = m\ddot{\mathbf{r}}\,.} \tag{1.23}$$

This is the particle's *equation of motion*.

Remarks.

• As stated above, Newton's first law is a particular case of the second one. Indeed, if $\mathbf{F} = 0$ then $\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = 0$ implies that $\mathbf{p}$, and therefore $\mathbf{v}$, must be constant.

• Newton's first two laws —or, from what we have just remarked, Eqs. (1.22)-(1.23)— are the foundation of classical mechanics. These laws are valid with very high accuracy for motions involving *small velocities* compared to the speed of light, and at *macroscopic scales*[1]. In particular, they do *not* hold for interactions at the atomic and subatomic scales (between elementary particles, atoms, atomic nuclei, molecules, etc.), which are governed by *quantum mechanics*. Nor are they valid for motion in intense gravitational fields, which is governed by Einstein's theory of *general relativity*. Actually, both quantum mechanics (or even *quantum field theory*, which combines quantum mechanics with the special theory of relativity) and general relativity are not universally valid, but rather apply to different physical situations. In fact, at present there is no consistent theory applicable to *all* physical phenomena which unifies quantum mechanics with the general theory of relativity.

• Newton's second law provides an operational definition of mass. Indeed, if we apply the same force $\mathbf{F}$ to two different particles (denoted by 1 and 2), according to Eq. (1.23) their accelerations have the same direction, and the quotient of their magnitudes is given by

$$\frac{|\mathbf{a}_1|}{|\mathbf{a}_2|} = \frac{m_2}{m_1} \,.$$

In this way, the quotient m_2/m_1 can be measured in principle for any pair of particles. From the previous discussion it should also be clear that a particle's mass is a quantitative measure of its *inertia*, i.e., its resistance to being accelerated by an applied force.

• Practically all forces appearing in classical mechanics depend at most on time, position and velocity, and are therefore *independent of acceleration* (and of derivatives of the position vector of order higher than two)[2]. Newton's second law (1.23) can therefore be written in the form

$$\boxed{\ddot{\mathbf{r}} = \frac{1}{m}\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})\,,} \tag{1.24}$$

where $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$ is the force acting on the particle. This vector equation is actually equivalent to the *system of three second-order ordinary differential equations*

$$\ddot{x}_i = \frac{1}{m} F_i(t, x_1, x_2, x_3, \dot{x}_1, \dot{x}_2, \dot{x}_3)\,, \qquad i = 1, 2, 3\,, \tag{1.25}$$

for the three particle coordinates $x_i(t)$. If the function $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$ is of class C^1, the equations (1.25) (or (1.24)) with arbitrary *initial conditions*

$$\mathbf{r}(t_0) = \mathbf{r}_0\,, \qquad \dot{\mathbf{r}}(t_0) = \mathbf{v}_0 \tag{1.26}$$

have (locally) a *unique solution*. In other words, *the position and velocity of the particle at a certain instant t_0 determine its trajectory $\mathbf{r}(t)$ at any other (past or future) time t*. Classical mechanics is thus an essentially *deterministic* theory.

[1] More precisely, if the typical *action* of the system under study, defined as the product of its typical energy and time, is much larger than Planck's constant $h = 6.626\,070\,15 \times 10^{-34}$ J s.

[2] The only exception worth mentioning is the force exerted on an accelerated charge by its own electromagnetic field, the so called *Abraham-Lorentz-Dirac force*, which is proportional to $\dot{\mathbf{a}}$.

- *Newton's third law* (or **law of action and reaction**) states that if particle 2 exerts on particle 1 a force $\mathbf{F}_{12}$ then particle 1 exerts on particle 2 a force $\mathbf{F}_{21}$ of equal magnitude and opposite direction:

$$\mathbf{F}_{21} = -\mathbf{F}_{12} \,. \tag{1.27}$$

A stronger version of Newton's third law states that, in addition, the force $\mathbf{F}_{12}$ (and, hence, $\mathbf{F}_{21}$) must be parallel to the vector $\mathbf{r}_1 - \mathbf{r}_2$, that is to say, to the straight line joining both particles:

$$\mathbf{F}_{12} = -\mathbf{F}_{21} \parallel \mathbf{r}_1 - \mathbf{r}_2 \,. \tag{1.28}$$

It is important to bear in mind that Newton's third law —in either of its two versions (1.27) and (1.28) — does *not* have a fundamental character, since (for example) it does *not* hold in general for the electromagnetic force between two charges in relative motion. It is however verified —in fact, in its most restrictive version (1.28)— by the gravitational and electrostatic forces (see below), as well as by most macroscopic forces that occur in ordinary mechanical problems, as for example the tension of a string. ■

1.3.2 Inertial frames

It is obvious that Newton's first law *cannot* be valid in *all* reference frames. Indeed, let S and S' be two reference frames with parallel axes, and denote by $\mathbf{R}(t)$ the coordinates of the origin of S' with respect to the reference frame S at time t. Let us denote by $\mathbf{r}(t)$ the coordinates of a particle with respect to the reference frame S at each instant t, so that the particle's velocity (with respect to S) is $\mathbf{v}(t) = \dot{\mathbf{r}}(t)$. *In Newtonian mechanics it is assumed that* time *has a universal character*[3], so that (once the unit of time has been set) the relation between the times t and t' of the same event measured in the frames S and S' is simply

$$t' = t - t_0 \,,$$

where t_0 is a constant. From the point of view of S', therefore, the particle's coordinates at the time t' will be given by the vector

$$\mathbf{r}'(t') = \mathbf{r}(t) - \mathbf{R}(t) = \mathbf{r}(t' + t_0) - \mathbf{R}(t' + t_0) \,.$$

The particle's velocity with respect to S' is thus

$$\mathbf{v}'(t') = \frac{\mathrm{d}\mathbf{r}'(t')}{\mathrm{d}t'} = \dot{\mathbf{r}}(t' + t_0) - \dot{\mathbf{R}}(t' + t_0) = \mathbf{v}(t' + t_0) - \dot{\mathbf{R}}(t' + t_0) \,,$$

where as usual the dot denotes differentiation with respect to t. Suppose now that the particle is **free**, i.e., not subject to any force[4]. If Newton's first law holds in S then $\mathbf{v}(t) = \mathbf{v}_0$ for all t. By the previous equation, the particle's velocity relative to S' is

$$\mathbf{v}'(t') = \mathbf{v}_0 - \dot{\mathbf{R}}(t' + t_0) \,,$$

which is *not* constant unless $\dot{\mathbf{R}}$ is. Note that $\dot{\mathbf{R}}$ is constant if and only if $\ddot{\mathbf{R}} = 0$. We conclude that Newton's first law will hold in the reference frame S' (assuming that it holds in S, and that the axes of S and S' are parallel) if and only if its origin moves *without acceleration* with respect to S.

Definition 1.2. A reference frame in which Newton's first law holds is said to be **inertial**.

In view of the above considerations, Newton's first two laws can be formulated in a more accurate and logically satisfactory fashion as follows:

[3] We shall see at the end of the course that this postulate is no longer valid in the special theory of relativity.

[4] Since at the classical level the magnitude of all known forces between two particles tends to zero as the distance between the particles tends to infinity, it is assumed that a particle is free if it is very far away from all other particles.

> I. *There is a class of reference frames (called inertial) with respect to which free particles always move with constant velocity.*
>
> II. *In an* inertial *frame, the force* $\mathbf{F}$ *exerted on a particle is equal to the rate of variation of its momentum* $\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}$.

It should therefore be clear that:

1. Newton's first two laws are logically independent (in particular, the first law *defines* the class of reference frames in which the second law holds).
2. Both laws are not more or less arbitrary *axioms*, but rather *experimentally verifiable* (and, indeed, *verified*) facts (as remarked above, valid only *approximately*, in a certain range of speeds and forces).
3. The relation (1.23) between force and acceleration is (in general) ***only valid in an inertial reference frame***.

• What known reference frames are inertial? Galileo and Newton observed that a reference frame in which distant galaxies are at rest is (to a great approximation) inertial. More recently, it has been found that a reference frame with respect to which the cosmic microwave background radiation (a relic of the *big bang*) appears isotropic is inertial.

1.3.3 Galilean transformations

Let S be an inertial frame, and consider another frame S' whose origin has coordinates $\mathbf{R}(t)$ with respect to S at every instant t (where t denotes the time measured in S). We shall suppose that at any time t the axes of S are related to those of S' by a linear (invertible) transformation $A(t)$, i.e.,

$$\mathbf{e}_i = A(t)\mathbf{e}'_i\,, \qquad i = 1, 2, 3\,.$$

We shall assume from now on that both S and S' are positively oriented orthogonal reference frames (that is, at all times the axes of both S and S' form a positively oriented orthonormal basis of $\mathbb{R}^3$). The matrix A must then be *orthogonal*, i.e, it must verify

$$A^{\mathsf{T}} = A^{-1}\,.$$

The determinant of an orthogonal matrix is equal to ± 1, since

$$A^{\mathsf{T}}A = \mathbb{1} \quad \Longrightarrow \quad (\det A)^2 = 1\,.$$

In fact, we must have $\det A = 1$, since[5]

$$(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 = (\det A)\,(\mathbf{e}'_1 \times \mathbf{e}'_2) \cdot \mathbf{e}'_3\,.$$

We ask ourselves how must $A(t)$ and $\mathbf{R}(t)$ be so that the frame S' is inertial (assuming that S is an inertial frame). To answer this question, note that if we denote by $\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}_0 t$ the

[5]Indeed, since $\mathbf{e}_i = \sum_{k=1}^{3} A_{ki}\mathbf{e}'_i$ (with $A = (A_{ij})_{i,j=1}^{3}$) we have

$$\begin{aligned}(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 &= \sum_{i,j,k=1}^{3} A_{i1}A_{j2}A_{k3}(\mathbf{e}'_i \times \mathbf{e}'_j) \cdot \mathbf{e}'_k = \sum_{i,j,k=1}^{3} A_{i1}A_{j2}A_{k3}\varepsilon_{ijk}(\mathbf{e}'_1 \times \mathbf{e}'_2) \cdot \mathbf{e}'_3 = (\mathbf{e}'_1 \times \mathbf{e}'_2) \cdot \mathbf{e}'_3 \sum_{i,j,k=1}^{3} A_{i1}A_{j2}A_{k3}\varepsilon_{ijk}\\ &= \det A\,(\mathbf{e}'_1 \times \mathbf{e}'_2) \cdot \mathbf{e}'_3\,.\end{aligned}$$

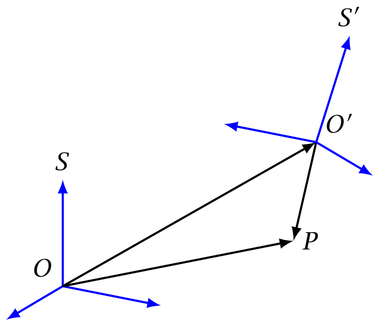


Figure 1.4. Reference frames S and S'. The coordinates of the vectors $\overrightarrow{OO'}$ and $\overrightarrow{OP}$ with respect to S are denoted in the text respectively by $\mathbf{R}$ and $\mathbf{r}$, and thus the coordinates of $\overrightarrow{O'P} = \overrightarrow{OP} - \overrightarrow{OO'}$ with respect to S' are given by $\mathbf{r}' = A \cdot (\mathbf{r} - \mathbf{R})$.

coordinates with respect to S of a free particle at time t, its coordinates with respect to S' at the corresponding time $t' = t - t_0$ are given by[6]

$$\mathbf{r}'(t') = A(t) \cdot \big(\mathbf{r}(t) - \mathbf{R}(t)\big) = A(t) \cdot \big(\mathbf{r}_0 + \mathbf{v}_0 t - \mathbf{R}(t)\big)\,, \qquad t = t' + t_0\,.$$

Differentiating twice with respect to t' we obtain

$$\frac{\mathrm{d}^2\mathbf{r}'(t')}{\mathrm{d}t'^2} = \ddot{A}(t)\mathbf{r}_0 + [t\ddot{A}(t) + 2\dot{A}(t)]\mathbf{v}_0 - [\ddot{A}(t)\mathbf{R}(t) + 2\dot{A}(t)\dot{\mathbf{R}}(t) + A(t)\ddot{\mathbf{R}}(t)]\,.$$

If the reference frame S' is also inertial, the right-hand side (RHS) of this equality must vanish identically for all $t \in \mathbb{R}$ and all $\mathbf{r}_0, \mathbf{v}_0 \in \mathbb{R}^3$. We thus have

$$\ddot{A}(t) = t\ddot{A}(t) + 2\dot{A}(t) = 0\,, \qquad \ddot{A}(t)\mathbf{R}(t) + 2\dot{A}(t)\dot{\mathbf{R}}(t) + A(t)\ddot{\mathbf{R}}(t) = 0\,,$$

or, equivalently,

$$\boxed{\dot{A}(t) = 0\,, \qquad \ddot{\mathbf{R}}(t) = 0\,.}$$

In other words:

The necessary and sufficient conditions in order for S' to be an inertial frame are that the rotation matrix $A(t)$ relating the axes of S and S' be *constant*, and that the origin of S' move with *constant velocity* with respect to S, i.e.,

$$\mathbf{R}(t) = \mathbf{R}_0 + \mathbf{V}_0 t\,,$$

with $\mathbf{R}_0$ and $\mathbf{V}_0$ constant vectors.

Moreover, the transformation relating the space-time coordinates $(t, \mathbf{r})$ and $(t', \mathbf{r}')$ of an event in the inertial reference frames S and S' is given by

$$\boxed{t' = t - t_0\,, \qquad \mathbf{r}' = A \cdot (\mathbf{r} - \mathbf{R}_0 - \mathbf{V}_0 t)\,; \qquad t_0 \in \mathbb{R}\,, \quad \mathbf{R}_0, \mathbf{V}_0 \in \mathbb{R}^3\,, \quad A \in \mathrm{SO}(3, \mathbb{R})\,,} \tag{1.29}$$

where $\mathrm{SO}(3, \mathbb{R})$ denotes the group[7] of 3×3 real orthogonal matrices with unit determinant.

[6]Indeed, let $\mathbf{c} = (c_1, c_2, c_3)$ be the coordinates of a point with respect to the axes $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of S at a certain time t, and denote by $\mathbf{c}' = (c_1', c_2', c_3')$ the coordinates of the *same* point with respect to the axes $\{\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3'\}$ of S' at that instant. We then have

$$\sum_{i=1}^{3} c_i \mathbf{e}_i = \sum_{i=1}^{3} c_i A(t) \cdot \mathbf{e}_i' = \sum_{i,k=1}^{3} c_i A_{ki}(t)\, \mathbf{e}_k' \quad \Longrightarrow \quad c_k' = \sum_{i=1}^{3} A_{ki}(t) c_i\,,$$

or, in matrix notation, $\mathbf{c}' = A(t)\mathbf{c}$.

[7]Recall that a *group* is a set G endowed with an associative product (an application $G \times G \to G$), possessing a unit element and such that every element of G has an inverse.

Definition 1.3. The change of coordinates $(t, \mathbf{r}) \mapsto (t', \mathbf{r}')$ defined by Eq. (1.29) is called a **Galilean transformation.**

Note. A **Galilean boost** is a transformation (1.29) with $t_0 = 0$, $\mathbf{R}_0 = 0$, $A = \mathbb{1}$.

From the previous discussion it then follows that:

The space-time coordinates $(t, \mathbf{r})$ and $(t', \mathbf{r}')$ of the same event in two inertial frames S and S' are related by an appropriate Galilean transformation (1.29).

• It is easy to verify that the composition of two Galilean transformations and the inverse of a Galilean transformations are Galilean transformations (see the exercises at the end of this section). From the mathematical point of view, this means that the set of all Galilean transformations forms a *group*, the so called **Galilean group.**

• From what we have just seen it follows that, given an inertial frame S, any other inertial frame S' is obtained from S by translating its origin with constant velocity and applying a *constant* (i.e., time-independent) rotation to its axes.

1.3.4 Galileo's relativity principle

By Eq. (1.29), the acceleration in the reference frame S' is given by

$$\frac{\mathrm{d}^2\mathbf{r}'}{\mathrm{d}t'^2}(t') = A \cdot \ddot{\mathbf{r}}(t)\,,$$

and from (1.23) it then follows that

$$\boxed{m\,\frac{\mathrm{d}^2\mathbf{r}'}{\mathrm{d}t'^2} = \mathbf{F}'\left(t', \mathbf{r}', \frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}t'}\right),} \tag{1.30}$$

with

$$\mathbf{F}'(t', \mathbf{r}', \dot{\mathbf{r}}') = A \cdot \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})\,, \qquad \dot{\mathbf{r}}' := \frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}t'}\,. \tag{1.31}$$

Thus, if $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$ is the force acting at time t on a particle located at a point $\mathbf{r}$ moving with velocity $\dot{\mathbf{r}}$ as measured in the inertial frame S, the corresponding force measured in the second inertial frame S' is given by Eq. (1.31). The latter equation simply states that the force behaves as a *vector* under a Galilean transformation (1.29). Equivalently, $\mathbf{F}$ and $\mathbf{F}'$ represent the *same* vector in two different frames. In other words, the observers at S and S' measure essentially the *same* force, although of course they assign it different components because their axes do not coincide. Note also that the transformation law (1.31) depends only on the relation between the two inertial frames S and S', and is therefore *independent of the properties of the particle considered* (i.e., its mass, electric charge, etc.).

From Eqs. (1.30)-(1.31) it also follows that Newton's second law —which, as we have seen, is the fundamental law of mechanics— has the same *form* in the inertial frame S' as in the original frame S. In other words:

The laws of mechanics have the *same form* in *all* inertial frames (**Galileo's relativity principle**).

• What happens to Newton's second law in a *non-inertial* frame? We shall see in Chapter 4 that the force measured by a non-inertial observer differs from that measured by an inertial one by several terms *proportional to the mass of the particle considered*, called **fictitious** or **inertial forces**[8]. In other words, *the laws of physics assume their simplest form* (that is, *without* fictitious forces) *only in* inertial *frames.*

[8]An example of such a force is the *centrifugal force* that appears in a frame whose axes are rotating as seen from an inertial frame.

Exercise. Find the parameters of the Galilean transformation obtained by composing (1.29) with a second Galilean transformation

$$t'' = t' - t'_0\,, \qquad \mathbf{r}'' = A' \cdot (\mathbf{r}' - \mathbf{R}'_0 - \mathbf{V}'_0 t')\,; \qquad t'_0 \in \mathbb{R}\,, \quad \mathbf{R}'_0, \mathbf{V}'_0 \in \mathbb{R}^3\,, \quad A' \in \mathrm{SO}(3,\mathbb{R})\,.$$

Solution. Substituting Eq. (1.29) into the previous equations we obtain $t'' = t - t_0 - t'_0$ and

$$\begin{aligned}\mathbf{r}'' &= A' \cdot [A(\mathbf{r} - \mathbf{R}_0 - \mathbf{V}_0 t) - \mathbf{R}'_0 - \mathbf{V}'_0(t - t_0)] = A'A \cdot [\mathbf{r} - \mathbf{R}_0 - \mathbf{V}_0 t - A^{-1}\mathbf{R}'_0 - A^{-1}\mathbf{V}'_0(t - t_0)] \\ &= A'A \cdot [\mathbf{r} - (\mathbf{R}_0 + A^{-1}\mathbf{R}'_0 - t_0 A^{-1}\mathbf{V}'_0) - (\mathbf{V}_0 + A^{-1}\mathbf{V}'_0)t].\end{aligned}$$

These are the equations of a Galilean transformation, with parameters

$$t''_0 = t_0 + t'_0\,, \qquad A'' = A'A\,, \qquad \mathbf{R}''_0 = \mathbf{R}_0 + A^{-1}\mathbf{R}'_0 - t_0 A^{-1}\mathbf{V}'_0, \qquad \mathbf{V}''_0 = \mathbf{V}_0 + A^{-1}\mathbf{V}'_0\,.$$

From the mathematical point of view, the latter equations define the *multiplication law* of the Galilean group.

Exercise. Show that the inverse of (1.29) is a Galilean transformation, and find its parameters.

Solution. From the previous exercise it follows that the inverse of (1.29) is the Galilean transformation with parameters $(t'_0, A', \mathbf{R}'_0, \mathbf{V}'_0)$ satisfying

$$t_0 + t'_0 = 0, \qquad A'A = \mathbb{1}, \qquad \mathbf{R}_0 + A^{-1}\mathbf{R}'_0 - t_0 A^{-1}\mathbf{V}'_0 = \mathbf{V}_0 + A^{-1}\mathbf{V}'_0 = 0.$$

Solving for $(t'_0, A', \mathbf{R}'_0, \mathbf{V}'_0)$ we easily obtain

$$t'_0 = -t_0, \qquad A' = A^{-1}, \qquad \mathbf{R}'_0 = -A(\mathbf{R}_0 + \mathbf{V}_0 t_0), \qquad \mathbf{V}'_0 = -A\mathbf{V}_0.$$

Note that the latter equations could also have been obtained by solving for $(t, \mathbf{r})$ in terms of $(t', \mathbf{r}')$ in Eq. (1.29) (exercise).

1.4 Conservation laws. Conservative forces. Electromagnetic force

1.4.1 Conservation laws

A **conserved quantity** (also called **constant of motion**, **integral of motion** or **first integral**) is any function of $(t, \mathbf{r}, \dot{\mathbf{r}})$ that remains constant as the particle moves. Knowing a conserved quantity is usually very advantageous, since it provides important information on the nature of the motion. For instance, Newton's first law (1.21) immediately yields a **law of conservation of linear momentum:** in the absence of forces, the linear momentum $\mathbf{p}$ of a particle is *conserved*. Let us next define the particle's **angular momentum** with respect to the origin of coordinates by

$$\boxed{\mathbf{L} = \mathbf{r} \times \mathbf{p} = m\mathbf{r} \times \dot{\mathbf{r}}\,,} \tag{1.32}$$

and the **torque** of the force $\mathbf{F}$ (also with respect to the origin) by

$$\boxed{\mathbf{N} = \mathbf{r} \times \mathbf{F}\,.} \tag{1.33}$$

Differentiating with respect to t the definition of angular momentum and applying Newton's second law we easily get the important identity

$$\boxed{\dot{\mathbf{L}} = \mathbf{N}\,.}$$

From this equation it immediately follows the **law of conservation of angular momentum:** if the torque of the force acting on a particle vanishes, its angular momentum is conserved. Note that in this case, since $\mathbf{r}$ is perpendicular to the constant vector $\mathbf{L}$, *the motion takes place in the normal plane to* $\mathbf{L}$ *passing through the origin.*

• From Eq. (1.33) it follows that $\mathbf{N} = 0$ if either $\mathbf{r} = 0$ or the applied force $\mathbf{F}$ is parallel to the particle's position vector $\mathbf{r}$, i.e. (assuming that $\mathbf{F}$ depends only on t, $\mathbf{r}$, and $\dot{\mathbf{r}}$):

$$\mathbf{F} = f(t, \mathbf{r}, \dot{\mathbf{r}})\mathbf{e}_r\,, \tag{1.34}$$

where g is an arbitrary scalar function. This type of force is called **central.**

Consider next the particle's **kinetic energy**, defined by

$$T = \frac{1}{2} m \dot{\mathbf{r}}^2\,. \tag{1.35}$$

Taking the scalar product of Newton's second law with the velocity vector $\dot{\mathbf{r}}$ we obtain

$$\frac{\mathrm{d}T}{\mathrm{d}t} = m\dot{\mathbf{r}}\ddot{\mathbf{r}} = \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})\dot{\mathbf{r}}\,. \tag{1.36}$$

In particular, *kinetic energy is conserved if the force* $\mathbf{F}$ *is perpendicular to the velocity* $\dot{\mathbf{r}}$ *at all times.* This is what happens, for instance, with the *magnetic force* acting on a charged particle (cf. Eq. (1.49) below).

Definition 1.4. We shall say that a force $\mathbf{F}(\mathbf{r})$ is **conservative** if it can be expressed in terms of a scalar **potential** $V(\mathbf{r})$ through the formula

$$\mathbf{F}(\mathbf{r}) = -\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}} \equiv -\nabla V(\mathbf{r}) = -\sum_{i=1}^{3} \frac{\partial V(\mathbf{r})}{\partial x_i}\,\mathbf{e}_i\,. \tag{1.37}$$

Note, in particular, that by its very definition *a conservative force can depend only on the particle's position vector* $\mathbf{r}$ (i.e., it must be independent of t and $\dot{\mathbf{r}}$). If $\mathbf{F}(\mathbf{r}) = -\dfrac{\partial V(\mathbf{r})}{\partial \mathbf{r}}$ is conservative we have

$$\mathbf{F}(\mathbf{r})\dot{\mathbf{r}} = -\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}}\,\dot{\mathbf{r}} = -\frac{\mathrm{d}}{\mathrm{d}t} V(\mathbf{r})\,,$$

and Eq. (1.36) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}(T + V) = 0\,.$$

The previous equation is the **law of conservation of energy:** *if the force acting on a particle is conservative, with potential* $V(\mathbf{r})$*, then the* **total energy**

$$E := T + V = \frac{1}{2} m\dot{\mathbf{r}}^2 + V(\mathbf{r}) \tag{1.38}$$

is conserved.

• More generally, we shall say that a *time-dependent* force $\mathbf{F}(t, \mathbf{r})$ is **irrotational** provided that $\nabla \times \mathbf{F}(t, \mathbf{r}) = 0$ for all $(t, \mathbf{r})$, where $\nabla \times \mathbf{F}$ is the **curl** of $\mathbf{F}$ (cf. Eq. (1.15)). It can be shown (assuming, e.g., that $\mathbf{F}$ is of class C^1 on $\mathbb{R}^4$) that $\mathbf{F}$ is irrotational if and only if there is a *time-dependent* function $V(t,\mathbf{r})$ such that

$$\mathbf{F}(t,\mathbf{r}) = -\frac{\partial V(t,\mathbf{r})}{\partial \mathbf{r}}\,.$$

If the force $\mathbf{F}$ is irrotational, differentiating the definition (1.38) of energy we obtain

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}T}{\mathrm{d}t} + \frac{\mathrm{d}V}{\mathrm{d}t} = m\dot{\mathbf{r}}\ddot{\mathbf{r}} + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial \mathbf{r}}\dot{\mathbf{r}} = \frac{\partial V}{\partial t}\,.$$

Thus if the force is irrotational but depends explicitly on time energy is *not* conserved.

1.4.2 Conservative forces

As we saw in the previous subsection, a force $\mathbf{F}(\mathbf{r})$ is conservative if it is the gradient of a function $-V(\mathbf{r})$. Note that (in a *connected* open subset) the potential $V(\mathbf{r})$ is determined by the force $\mathbf{F}(\mathbf{r})$ up to an arbitrary constant, since

$$\mathbf{F}(\mathbf{r}) = \nabla V_1 = \nabla V_2 \quad \Longleftrightarrow \quad \nabla(V_1 - V_2) = 0 \quad \Longrightarrow \quad V_1 - V_2 = \text{const}\,.$$

It can be shown that the conservative character of a force (independent of time and velocity) $\mathbf{F}(\mathbf{r})$ is *equivalent* to any of the following three conditions:

I. The force $\mathbf{F}$ is **irrotational:**

$$\nabla \times \mathbf{F} = 0\,.$$

II. The **work** done by the force $\mathbf{F}$ along *any* closed curve C vanishes:

$$\int_C \mathbf{F}(\mathbf{r})\cdot \mathrm{d}\mathbf{r} = 0\,.$$

III. The work done by the force $\mathbf{F}$ along *any* curve C with fixed endpoints $\mathbf{r}_1$ and $\mathbf{r}_2$ is *independent of the curve.* In other words,

$$\int_{C_1} \mathbf{F}(\mathbf{r})\cdot \mathrm{d}\mathbf{r} = \int_{C_2} \mathbf{F}(\mathbf{r})\cdot \mathrm{d}\mathbf{r}\,,$$

for any two curves C_1 and C_2 with the same endpoints $\mathbf{r}_1$ and $\mathbf{r}_2$.

The *necessity* of conditions I)–III) above (i.e., that if $\mathbf{F}(\mathbf{r})$ is conservative then I)–III) hold) is straightforward. Indeed, condition I) is a direct consequence of the identity

$$\nabla \times \nabla V(\mathbf{r}) = 0\,.$$

Likewise, the work done by a *conservative* force $\mathbf{F} = -\nabla V$ along any curve C with endpoints $\mathbf{r}_1$ and $\mathbf{r}_2$ is given by

$$\int_C \mathbf{F}(\mathbf{r})\cdot \mathrm{d}\mathbf{r} = -\int_C \frac{\partial V(\mathbf{r})}{\partial \mathbf{r}}\cdot \mathrm{d}\mathbf{r} = -\int_C \mathrm{d}V = V(\mathbf{r}_1) - V(\mathbf{r}_2)\,, \tag{1.39}$$

and is thus independent of the curve considered (condition III). In particular, if C is closed then we can take $\mathbf{r}_1 = \mathbf{r}_2$, and hence $\mathbf{F}$ does no work (condition II). It is shown in advanced calculus courses that the *converse* (i.e., that if any of the conditions I)–III) above hold then $\mathbf{F}(\mathbf{r})$ is conservative) is also true provided that $\mathbf{F}(\mathbf{r})$ is of class C^1 in a *simply connected* open subset[9] of $\mathbb{R}^3$ (in particular, on all of $\mathbb{R}^3$).

[9] By definition, a connected open subset $U \subset \mathbb{R}^3$ is simply connected if any continuous closed curve contained in U can be continuously contracted to a point within U. For example, $\mathbb{R}^3$, $\mathbb{R}^3$ minus one point, the interior of a sphere, a parallelepiped, a cylinder, etc., are simply connected sets, while $\mathbb{R}^3$ minus a line is not.

• From Eq. (1.39) it follows that the work done by a conservative force is equal to the *decrease* in the potential energy as the particle moves from the initial point $\mathbf{r}_1$ to the final one $\mathbf{r}_2$. By the law of conservation of total energy, this coincides with the *increase* in the particle's kinetic energy as it moves from $\mathbf{r}_1$ to $\mathbf{r}_2$.

• More generally, if $\mathbf{F}(t,\mathbf{r},\dot{\mathbf{r}})$ is an arbitrary (not necessarily conservative) force the work W done by $\mathbf{F}$ along a trajectory $C = \{\mathbf{r} = \mathbf{r}(t) \,|\, t \in [t_1, t_2]\}$ starting at a point $\mathbf{r}_1 = \mathbf{r}(t_1)$ with velocity $\dot{\mathbf{r}}_1 = \mathbf{r}(t_1)$ and ending at a point $\dot{\mathbf{r}}_2 = \mathbf{r}(t_2)$ with velocity $\dot{\mathbf{r}}_2 = \dot{\mathbf{r}}(t_2)$ is equal to the increase in the particle's kinetic energy. Indeed,

$$W = \int_C \mathbf{F}(t,\mathbf{r},\dot{\mathbf{r}})\,\mathrm{d}t = \int_{t_1}^{t_2} \mathbf{F}(t,\mathbf{r}(t),\dot{\mathbf{r}}(t))\cdot\dot{\mathbf{r}}(t)\,\mathrm{d}t = \int_{t_1}^{t_2} m\ddot{\mathbf{r}}(t)\cdot\dot{\mathbf{r}}(t)\,\mathrm{d}t = \int_{t_1}^{t_2} \frac{\mathrm{d}T}{\mathrm{d}t}\,\mathrm{d}t$$
$$= T(\dot{\mathbf{r}}_2) - T(\dot{\mathbf{r}}_1),$$

since $T = \frac{1}{2}m\dot{\mathbf{r}}^2$ depends only on $\dot{\mathbf{r}}$. Note, however, that the work along two trajectories with the same endpoints $\mathbf{r}_1$ and $\mathbf{r}_2$ need *not* be the same, since the initial and/or final velocities $\dot{\mathbf{r}}_{1,2}$ will in general be different for both trajectories.

• A particular case of conservative force of great practical interest is that of a *central force* of the form

$$\boxed{\mathbf{F}(\mathbf{r}) = f(r)\,\frac{\mathbf{r}}{r}\,.} \tag{1.40}$$

Indeed, taking into account that

$$\frac{\partial V(r)}{\partial \mathbf{r}} = V'(r)\,\frac{\partial r}{\partial \mathbf{r}} = V'(r)\,\frac{\mathbf{r}}{r}\,,$$

it is obvious that the force (1.40) is generated by the potential

$$\boxed{V(r) = -\int f(r)\,\mathrm{d}r\,,} \tag{1.41}$$

which depends only on the magnitude of the position vector $\mathbf{r}$. Thus *if the force is central and conservative both energy and angular momentum are conserved.*

Exercise. Show that the central force (1.34) is conservative if and only if the function $f(t,\mathbf{r},\dot{\mathbf{r}})$ depends only on r.

Solution. To begin with, f can only depend on $\mathbf{r}$ from the definition of conservative force. If $F_\theta = F_\varphi = 0$ and $F_r = f(\mathbf{r})$ Eq. (1.17) for the curl of $\mathbf{F}$ in spherical coordinates yields

$$\nabla\times\mathbf{F} = \frac{1}{r\sin\theta}\,\frac{\partial f}{\partial\varphi}\,\mathbf{e}_\theta - \frac{1}{r}\,\frac{\partial f}{\partial\theta}\,\mathbf{e}_\varphi = 0 \iff \frac{\partial f}{\partial\theta} = \frac{\partial f}{\partial\varphi} = 0\,,$$

so that f is a function of r only. Alternatively, using Eq. (1.17) for the gradient of $V(\mathbf{r})$ in spherical coordinates we obtain

$$\frac{\partial V}{\partial\mathbf{r}} = \frac{\partial V}{\partial r}\,\mathbf{e}_r + \frac{1}{r}\,\frac{\partial V}{\partial\theta}\,\mathbf{e}_\theta + \frac{1}{r\sin\theta}\,\frac{\partial V}{\partial\varphi}\,\mathbf{e}_\varphi = -\mathbf{F} = -f\mathbf{e}_r \implies \frac{\partial V}{\partial\theta} = \frac{\partial V}{\partial\varphi} = 0\,.$$

Hence V is a function of r only, and so is $f = -\dfrac{\partial V(r)}{\partial r}$.

1.4.3 Gravitational and electrostatic forces

According to Newton's **law of universal gravitation**, the gravitational force exerted by a particle of mass M *fixed* at the origin of coordinates on another particle of mass m located at a point $\mathbf{r}$ is of the form (1.40) with f inversely proportional to the square of the distance to the origin:

$$\boxed{f(r) = -\frac{GMm}{r^2}\,,} \tag{1.42}$$

where[10]

$$G = 6.674\,30(15) \cdot 10^{-11}\ \mathrm{m^3\,Kg^{-1}\,s^{-2}}$$

is the so called *gravitational constant*. By Eq. (1.41), the potential $V(r)$ generating the gravitational force (1.42) is given (up to an arbitrary constant) by

$$\boxed{V(r) = -\frac{GMm}{r}\,.} \tag{1.43}$$

Note that, since $GMm > 0$, the gravitational force is always *attractive*.

- The acceleration caused by the gravitational force (1.40)-(1.42) on the particle of mass m is

$$\boxed{\mathbf{a} = \frac{\mathbf{F}}{m} = -\frac{GM}{r^3}\,\mathbf{r}\,,}$$

independent of m. This nontrivial fact, first observed by Galileo Galilei, is due to the fact that the mass appearing in Newton's law of universal gravitation (the **gravitational mass**) actually coincides[11] with the mass appearing in Newton's second law (the **inertial mass**). The equality between the gravitational and inertial masses —the so called **equivalence principle**, on which Einstein's general theory of relativity is based— has been verified with great accuracy (less than one part in 10^{12}) in different experiments.

Similarly, the electric force exerted by a charge Q fixed at the origin on a point charge q located at point $\mathbf{r}$ is also of the form (1.40), where now

$$\boxed{f(r) = k\,\frac{qQ}{r^2}\,.} \tag{1.44}$$

In the SI system of units, the constant k is given by

$$k = \frac{1}{4\pi\varepsilon_0} \simeq 8.98755 \cdot 10^9\ \mathrm{m\,F^{-1}}\,,$$

where

$$\varepsilon_0 = 8.854\,187\,8128(13) \cdot 10^{-12}\ \mathrm{F\,m^{-1}}$$

is the *vacuum permittivity*. From Eq. (1.44) it follows that the electric force is attractive if the charges q and Q are of opposite signs, and repulsive if they have the same sign. Again, the electric force is obviously conservative, with potential (up to an additive constant)

$$\boxed{V(r) = k\,\frac{qQ}{r}}$$

[10] In these notes we use the CODATA internationally recommended 2018 values of the fundamental physical constants, available at the site `https://physics.nist.gov/cuu/Constants/`.

[11] Obviously, it is only necessary that inertial and gravitational mass differ by a universal (i.e., the same for all particles) proportionality constant.

inversely proportional to the distance between the charges.

More generally, the gravitational force exerted on a particle of mass m located at the point $\mathbf{r}$ by a continuous mass distribution occupying an open subset $U \subset \mathbb{R}^3$ is given by

$$\mathbf{F}(\mathbf{r}) = -Gm \int_U \rho(\mathbf{r}') \frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} \, \mathrm{d}^3\mathbf{r}' =: m\mathbf{g}(\mathbf{r}) = -m \frac{\partial \Phi(\mathbf{r})}{\partial \mathbf{r}}, \tag{1.45}$$

where $\rho(\mathbf{r}')$ is the mass density at $\mathbf{r}' \in U$, $\mathbf{g}(\mathbf{r})$ is the **gravitational field** created by the mass distribution at the point $\mathbf{r}$ and

$$\Phi(\mathbf{r}) = -G \int_U \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \, \mathrm{d}^3\mathbf{r}' \tag{1.46}$$

is the **gravitational potential.** Thus the gravitational force is still conservative in this more general situation, with potential $V(\mathbf{r}) = m\Phi(\mathbf{r})$. Again, the particle's acceleration

$$\mathbf{a} = \frac{\mathbf{F}(\mathbf{r})}{m} = \mathbf{g}(\mathbf{r})$$

is independent of its mass m. Note, however, that in general (unless the mass distribution is spherically symmetric about the origin) the gravitational force (1.45) is *not* central.

Similarly, the force exerted on a point charge q located at a point $\mathbf{r}$ by a *static* charge distribution filling up an open set U is

$$\mathbf{F}(\mathbf{r}) = kq \int_U \rho(\mathbf{r}') \frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3} \, \mathrm{d}^3\mathbf{r}' =: q\mathbf{E}(\mathbf{r}) = -q \frac{\partial \Phi(\mathbf{r})}{\partial \mathbf{r}}, \tag{1.47}$$

where now $\rho(\mathbf{r}')$ is the charge density at a point $\mathbf{r}'$, $\mathbf{E}(\mathbf{r})$ is the **electric field** created by the charge distribution at the point $\mathbf{r}$ and

$$\Phi(\mathbf{r}) = k \int_U \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \, \mathrm{d}^3\mathbf{r}' \tag{1.48}$$

is the **electrostatic potential.** Again, the electrostatic force (1.47) is conservative (with potential $V(\mathbf{r}) = q\Phi(\mathbf{r})$), but *not* central unless the charge distribution is spherically symmetric about the origin.

Exercise. Applying Gauss's theorem to a sphere centered at the origin prove the identity

$$\triangle \left(\frac{1}{r}\right) = -4\pi\delta(\mathbf{r}),$$

where $\delta(\mathbf{r})$ is **Dirac's delta function**. Deduce from Eq. (1.46) that both the gravitational and the electrostatic potential verify **Poisson's equation**

$$\boxed{\triangle\Phi = 4\pi\alpha\,\rho\,,}$$

where $\alpha = G$ for the gravitational potential and $\alpha = -k$ for the electrostatic one. In particular, the gravitational (resp. electrostatic) potential verifies **Laplace's equation** $\triangle\Phi = 0$ in any region of space where there are no masses (resp. charges).

Note. Dirac's delta function $\delta(\mathbf{r})$ is informally defined by the requirements $\delta(\mathbf{r}) = 0$ for all $\mathbf{r} \neq 0$ and $\int_{\mathbb{R}^3} \delta(\mathbf{r})\,\mathrm{d}^3\mathbf{r} = 1$. It can thus be intuitively viewed as the mass density of a point mass located at the origin. In fact, no ordinary function can simultaneously verify the above two requirements, since for an ordinary function the condition $\delta(\mathbf{r}) = 0$ for all $\mathbf{r} \neq 0$ implies that $\int_{\mathbb{R}^3} \delta(\mathbf{r})\,\mathrm{d}^3\mathbf{r} = 0$. We can think of $\delta(\mathbf{r})$ as the "limit" as $\varepsilon \to 0+$ of any family of functions $\delta_\varepsilon(\mathbf{r})$ satisfying $\int_{\mathbb{R}^3} \delta_\varepsilon(\mathbf{r})\,\mathrm{d}^3\mathbf{r} = 1$ for all $\varepsilon > 0$ and such that $\delta_\varepsilon(\mathbf{r})$ is concentrated on a ball centered at the origin whose radius tends to zero as $\varepsilon \to 0+$. (One such family is, for instance, $\delta_\varepsilon(\mathbf{r}) = (\pi\varepsilon)^{-3/2}\mathrm{e}^{-r^2/\varepsilon}$.) From this heuristic definition follows the important property $\int_{\mathbb{R}^3} \delta(\mathbf{r}) f(\mathbf{r})\,\mathrm{d}^3\mathbf{r} = f(0)$, for any sufficiently smooth function $f(\mathbf{r})$. In fact, the latter identity can be taken as a working definition of $\delta(\mathbf{r})$. A rigorous mathematical treatment of Dirac's delta function requires the use of the theory of *distributions* (linear functionals defined on spaces of smooth functions vanishing fast enough at infinity).

Solution. To begin with, let us check that $\triangle(1/r) = 0$ for $\mathbf{r} \neq 0$. Indeed, if $\mathbf{r} \neq 0$ (and hence $r \neq 0$) we have

$$\triangle\left(\frac{1}{r}\right) = \nabla \cdot \left[\nabla\left(\frac{1}{r}\right)\right] = \nabla \cdot \left(-\frac{\mathbf{r}}{r^3}\right) = -\frac{\nabla \cdot \mathbf{r}}{r^3} - \mathbf{r} \cdot \nabla\left(\frac{1}{r^3}\right) = -\frac{3}{r^3} - \mathbf{r} \cdot \left(-\frac{3\mathbf{e}_r}{r^4}\right) = -\frac{3}{r^3} + \frac{3}{r^3} = 0.$$

To show that $\triangle(1/r) = -4\pi\delta(\mathbf{r})$, we only have to prove the equality

$$\int_{\mathbb{R}^3} \triangle\left(\frac{1}{r}\right) \mathrm{d}^3\mathbf{r} = -4\pi\,.$$

As we have just seen that $\triangle(1/r) = 0$ away from the origin, we can integrate over a ball of arbitrary radius R centered at the origin. We thus have

$$\begin{aligned}\int_{\mathbb{R}^3} \triangle\left(\frac{1}{r}\right) \mathrm{d}^3\mathbf{r} &= \int_{|\mathbf{r}|\leqslant R} \triangle\left(\frac{1}{r}\right) \mathrm{d}^3\mathbf{r} = -\int_{|\mathbf{r}|\leqslant R} \nabla \cdot \left(\frac{\mathbf{e}_r}{r^2}\right) \mathrm{d}^3\mathbf{r} = -\int_{|\mathbf{r}|=R} \frac{\mathbf{e}_r}{R^2} \cdot \mathbf{n}\,\mathrm{d}S \\ &= -\int_{|\mathbf{r}|=R} \frac{\mathbf{e}_r}{R^2} \cdot \mathbf{e}_r\,\mathrm{d}S = -\frac{1}{R^2}\int_{|\mathbf{r}|=R} \mathrm{d}S = -\frac{1}{R^2} \cdot 4\pi R^2 = -4\pi\,,\end{aligned}$$

where we have applied Gauss's theorem we obtain the third equality. Taking the Laplacian (with respect to the $\mathbf{r}$ coordinate) of the equation for the gravitational/electrostatic potential,

$$\Phi(\mathbf{r}) = -\alpha \int_U \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\,\mathrm{d}^3\mathbf{r}'\,,$$

we then obtain

$$\begin{aligned}\triangle\Phi(\mathbf{r}) &= -\alpha \int_U \triangle_{\mathbf{r}}\left(\frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\right) \mathrm{d}^3\mathbf{r}' = -\alpha \int_U \rho(\mathbf{r}')\triangle_{\mathbf{r}}\left(\frac{1}{|\mathbf{r} - \mathbf{r}'|}\right) \mathrm{d}^3\mathbf{r}'k = 4\pi\alpha \int_U \rho(\mathbf{r}')\delta(\mathbf{r} - \mathbf{r}')\,\mathrm{d}^3\mathbf{r}' \\ &= 4\pi\alpha\rho(\mathbf{r}).\end{aligned}$$

1.4.4 Electromagnetic force

The **electromagnetic force** (also called **Lorentz force**) acting on a point charge q which moves subject to an electric field $\mathbf{E}(t,\mathbf{r})$ and a magnetic field $\mathbf{B}(t,\mathbf{r})$ is given by

$$\mathbf{F}(t,\mathbf{r},\dot{\mathbf{r}}) = q\big(\mathbf{E}(t,\mathbf{r}) + \dot{\mathbf{r}} \times \mathbf{B}(t,\mathbf{r})\big). \tag{1.49}$$

As is well known, the fields $\mathbf{E}$ and $\mathbf{B}$ verify **Maxwell's equations**

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0}\,, & \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}\,, \\ \nabla \cdot \mathbf{B} &= 0\,, & \nabla \times \mathbf{B} &= \mu_0\mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}\,,\end{aligned}$$

where $\mathbf{J}$ is the current density,

$$c = 2.997\,924\,58 \times 10^8\ \mathrm{m\,s^{-1}}$$

is the speed of light *in vacuo*, and

$$\mu_0 := (c^2\varepsilon_0)^{-1} = 1.256\,637\,062\,12(19) \cdot 10^{-6}\ \mathrm{N\,A^{-2}}$$

is the *vacuum permeability*. From the second and third Maxwell equations it follows that **E** and **B** can be expressed through a **scalar potential** $\Phi(t,\mathbf{r})$ and a **vector potential** $\mathbf{A}(t,\mathbf{r})$ through the equations

$$\boxed{\mathbf{E} = -\frac{\partial \Phi}{\partial \mathbf{r}} - \frac{\partial \mathbf{A}}{\partial t}\,, \qquad \mathbf{B} = \nabla \times \mathbf{A}\,.}$$

Remark. The fields **E** and **B** do *not* uniquely determine the electromagnetic potentials Φ and **A**. Indeed, it is easily verified (exercise) that the potentials

$$\boxed{\hat{\Phi} = \Phi - \frac{\partial f}{\partial t}\,, \qquad \hat{\mathbf{A}} = \mathbf{A} + \frac{\partial f}{\partial \mathbf{r}}\,,} \tag{1.50}$$

where $f(t,\mathbf{r})$ is an arbitrary scalar function[12], generate exactly the same electromagnetic field as Φ and **A**. It can be shown that it is always possible to choose the function f so that the new potentials $\hat{\Phi}$ and $\hat{\mathbf{A}}$ verify the condition

$$\nabla \cdot \hat{\mathbf{A}} + \frac{1}{c^2}\frac{\partial \hat{\Phi}}{\partial t} = 0\,, \tag{1.51}$$

called the *Lorenz gauge*[13]. If the electromagnetic potentials satisfy the Lorenz gauge, it is immediate to check that Maxwell's equations are equivalent to the following two *uncoupled* equations for Φ and **A**:

$$\frac{1}{c^2}\frac{\partial^2 \Phi}{\partial t^2} - \triangle \Phi = \frac{\rho}{\varepsilon_0}\,, \qquad \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} - \triangle \mathbf{A} = \mu_0 \mathbf{J}\,.$$

In particular, *in vacuo* (that is, in any region of space not containing electrical charges or currents), the scalar potential Φ and each component A_i of the vector potential verify the **wave equation**

$$\boxed{\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \triangle u = 0,} \tag{1.52}$$

where c is the velocity of the waves. ■

Exercise. Using the identity (0.13), show that the components of the fields **E** and **B** also verify the wave equation *in vacuo*.

Solution. Indeed,

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \triangle \mathbf{E} = \frac{1}{\varepsilon_0}\nabla\rho - \triangle \mathbf{E} = -\nabla \times \left(\frac{\partial \mathbf{B}}{\partial t}\right) = -\frac{\partial}{\partial t}\nabla \times \mathbf{B} = -\mu_0\frac{\partial \mathbf{J}}{\partial t} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}$$

$$\Longrightarrow \boxed{\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} - \triangle \mathbf{E} = -\frac{1}{\varepsilon_0}\nabla\rho - \mu_0\frac{\partial \mathbf{J}}{\partial t},}$$

$$\nabla \times (\nabla \times \mathbf{B}) = -\triangle \mathbf{B} = \nabla \times \left(\mu_0 \mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}\right) = \mu_0 \nabla \times \mathbf{J} + \frac{1}{c^2}\frac{\partial}{\partial t}\nabla \times \mathbf{E} = \mu_0\nabla \times \mathbf{J} - \frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2}$$

$$\Longrightarrow \boxed{\frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2} - \triangle \mathbf{B} = \mu_0 \times \mathbf{J}.}$$

[12] Equation (1.50) is called a *gauge transformation* of the electromagnetic potentials. It can be shown that if the electromagnetic potentials $(\Phi, \mathbf{A})$ and $(\hat{\Phi}, \hat{\mathbf{A}})$ generate the same electromagnetic field then they are related by a gauge transformation (assuming, for simplicity, that the fields are of class C^2 on $\mathbb{R}^4$).

[13] Indeed, it suffices that the function f be a solution of the partial differential equation

$$\triangle f - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = \nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \Phi}{\partial t}\,.$$

It is shown in differential equations courses that if the potentials **A** and Φ are analytic functions the latter equation has (locally) a solution dependent on two arbitrary functions of the variable **r** (*Cauchy–Kovalevskaya theorem*).

When ρ and $\mathbf{J}$ vanish both framed equations reduce to the wave equation with wave velocity c.

If both the electric and the magnetic field are *static*, i.e., if

$$\mathbf{E} = \mathbf{E}(\mathbf{r}), \qquad \mathbf{B} = \mathbf{B}(\mathbf{r}),$$

Maxwell's second equation reduces to $\nabla \times \mathbf{E}(\mathbf{r}) = 0$, and hence

$$\mathbf{E} = -\frac{\partial \Phi(\mathbf{r})}{\partial \mathbf{r}}.$$

From this equation and the expression (1.49) for the Lorentz force it then follows that

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \mathbf{F}\cdot\dot{\mathbf{r}} = q\mathbf{E}(\mathbf{r})\cdot\dot{\mathbf{r}} = -q\frac{\partial \Phi(\mathbf{r})}{\partial \mathbf{r}}\cdot\dot{\mathbf{r}} = -q\frac{\mathrm{d}\Phi}{\mathrm{d}t} \quad\Longrightarrow\quad \frac{\mathrm{d}}{\mathrm{d}t}(T + q\Phi) = 0.$$

Thus in this case the function

$$\boxed{T + q\Phi(\mathbf{r}),}$$

which can be regarded as the particle's *electromechanical* energy, is conserved, although the Lorentz force is *not* conservative unless $\mathbf{B} = 0$ (cf. the exercise below). To interpret physically this result it suffices to note that the magnetic force does *no* work, since it is perpendicular to the velocity and hence to the infinitesimal displacement $\mathrm{d}\mathbf{r}$, and therefore does not contribute to the particle's energy.

Exercise. Show that the Lorentz force (1.49) is conservative if and only if $\mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} = 0$ (that is, if the electromagnetic field is purely electrostatic).

Exercise. Show that if the potentials $(\Phi, \mathbf{A})$ and $(\hat{\Phi}, \hat{\mathbf{A}})$ generate the same electromagnetic field $(\mathbf{E}, \mathbf{B})$ then (1.50) holds for some scalar function $f(t, \mathbf{r})$.

Solution. If $(\Phi, \mathbf{A})$ and $(\hat{\Phi}, \hat{\mathbf{A}})$ generate the same electromagnetic field then

$$\begin{aligned} &\mathbf{E} = -\frac{\partial \Phi}{\partial \mathbf{r}} - \frac{\partial \mathbf{A}}{\partial t} = -\frac{\partial \hat{\Phi}}{\partial \mathbf{r}} - \frac{\partial \hat{\mathbf{A}}}{\partial t} &&\Longrightarrow\quad \frac{\partial}{\partial \mathbf{r}}(\hat{\Phi} - \Phi) + \frac{\partial}{\partial t}(\hat{\mathbf{A}} - \mathbf{A}) = 0,\\ &\mathbf{B} = \nabla \times \mathbf{A} = \nabla \times \hat{\mathbf{A}} &&\Longrightarrow\quad \nabla \times (\hat{\mathbf{A}} - \mathbf{A}) = 0. \end{aligned}$$

From the second equation we deduce that

$$\hat{\mathbf{A}} = \mathbf{A} + \frac{\partial g}{\partial \mathbf{r}}$$

for some scalar function $g(t, \mathbf{r})$, and substituting into the first equation we then obtain

$$\frac{\partial}{\partial \mathbf{r}}(\hat{\Phi} - \Phi) + \frac{\partial}{\partial t}\frac{\partial g}{\partial \mathbf{r}} = \frac{\partial}{\partial \mathbf{r}}\left(\hat{\Phi} - \Phi + \frac{\partial g}{\partial t}\right) = 0 \quad\Longrightarrow\quad \hat{\Phi} = \Phi - \frac{\partial g}{\partial t} - h(t)$$

for some scalar function of time $h(t)$. Thus Eq. (1.50) holds with

$$f(t, \mathbf{r}) = g(t, \mathbf{r}) + \int h(t)\,\mathrm{d}t.$$

Example 1.5. An electron of mass m and charge $-e < 0$ moves in a uniform electromagnetic field $\mathbf{E} = E\mathbf{e}_2$, $\mathbf{B} = B\mathbf{e}_3$ (with $E > 0$, $B > 0$). Let us compute the particle's trajectory if initially $\mathbf{r}(0) = 0$ and $\mathbf{v}(0) = v_0\mathbf{e}_1$, with $v_0 > 0$. Taking into account Eq. (1.49), the electron's equations of motion are

$$\boxed{m\ddot{x}_1 = -eB\dot{x}_2\,, \qquad m\ddot{x}_2 = -eE + eB\dot{x}_1\,, \qquad m\ddot{x}_3 = 0\,.} \tag{1.53}$$

From the last equation and the initial condition $x_3(0) = \dot{x}_3(0) = 0$ it immediately follows that $x_3(t) = 0$ for all t. Hence the motion takes place in the horizontal plane $x_3 = 0$. The equations for the coordinates x_1 and x_2 can be simplified using the dimensionless variables[a]

$$\boxed{\tau = \frac{eB}{m}\,t\,, \qquad x = \frac{eB^2}{mE}\,x_1\,, \qquad y = \frac{eB^2}{mE}\,x_2\,,}$$

in terms of which

$$\boxed{x'' = -y'\,, \qquad y'' = x' - 1\,,} \tag{1.54}$$

where the prime denotes derivative with respect to τ. In terms of the new variables, the initial conditions are

$$x(0) = y(0) = 0\,, \qquad x'(0) = \frac{eB^2}{mE}\,\dot{x}_1(0)\,\frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{Bv_0}{E} =: 1 + a\,, \qquad y'(0) = 0\,.$$

The equations of motion for the (x, y) variables can be easily solved through standard techniques, since they are a linear system of second-order differential equations with constant coefficients. In this case, however, the simplest course of action is to introduce the complex variable $z = x + \mathrm{i}y$, in terms of which Eqs. (1.54) reduce to the ordinary differential equation

$$z'' = x'' + \mathrm{i}y'' = -y' + \mathrm{i}x' - \mathrm{i} = \mathrm{i}(z' - 1)\,,$$

or equivalently

$$\boxed{w'' = \mathrm{i}\,w'\,, \qquad w := z - \tau.}$$

This is just a linear first-order differential equation in w', with initial condition

$$w'(0) = z'(0) - 1 = x'(0) + \mathrm{i}y'(0) - 1 = a,$$

whose solution is

$$w' = a\mathrm{e}^{\mathrm{i}\tau}\,.$$

Integrating with respect to τ and taking into account the initial condition

$$w(0) = z(0) = 0$$

we obtain

$$w = \mathrm{i}a(1 - \mathrm{e}^{\mathrm{i}\tau}) \qquad \Longrightarrow \qquad z = \tau + \mathrm{i}a(1 - \mathrm{e}^{\mathrm{i}\tau})\,.$$

Taking the real and imaginary parts of z we finally arrive at

$$\boxed{x = \operatorname{Re} z = \tau + a\sin\tau\,, \qquad y = \operatorname{Im} z = a(1 - \cos\tau)\,,} \tag{1.55}$$

or, in terms of the original variables,

$$\boxed{x_1 = \frac{Et}{B} + \frac{1}{\omega}\left(v_0 - \frac{E}{B}\right)\sin(\omega t)\,, \quad x_2 = \frac{1}{\omega}\left(v_0 - \frac{E}{B}\right)[1 - \cos(\omega t)]\,, \qquad \omega := \frac{eB}{m}\,.}$$

Equations (1.55) are the parametric equations of the electron's trajectory. Note that

$$\mathbf{r}(\tau + 2n\pi) = \big(x(2n\pi), y(2n\pi)\big) = \big(2n\pi + x(\tau), y(\tau)\big) = \mathbf{r}(\tau) + 2n\pi\,\mathbf{i} \qquad (n \in \mathbb{Z}), \tag{1.56}$$

so that the whole trajectory can be obtained by translating the arc with $0 \leqslant \tau < 2\pi$ an integer multiple of 2π along the x direction. The qualitative properties of the trajectory depend on the dimensionless parameter

$$\boxed{a = \frac{Bv_0}{E} - 1\,;}$$

note that $a > -1$, since E, B and v_0 are all positive by hypothesis.

i) If $|a| < 1$, i.e.,

$$\boxed{0 < v_0 < \frac{2E}{B}\,,}$$

we have

$$x' = 1 + a\cos\tau > 0\,,$$

and therefore x is an increasing function of τ. In particular, if $a = 0$, or equivalently

$$\boxed{v_0 = \frac{E}{B}\,,}$$

the trajectory is the x axis traversed with constant velocity ($x(\tau) = \tau$ or, in the original variables, $x_1(t) = Et/B = v_0 t$). By Eq. (1.56), it suffices to study the arc of the trajectory with $0 \leqslant \tau \leqslant 2\pi$. In general, $0 \leqslant y \leqslant 2a$ if $a > 0$, or $2a \leqslant y \leqslant 0$ if $a < 0$, for all τ. Moreover, in the interval $0 \leqslant \tau \leqslant 2\pi$ we have

$$y = 0 \iff \tau = 0, 2\pi \quad \implies \quad x = 0, 2\pi\,,$$

while

$$y = 2a \iff \tau = \pi \quad \implies \quad x = \pi\,.$$

At points where y attains its extreme values 0 and $2a$ the electron's velocity is directed along the x axis, since for $\tau = k\pi$ (with $k = 0, 1, 2$) we have

$$x'(k\pi) = 1 + (-1)^k a > 0\,, \qquad y'(k\pi) = a\sin(k\pi) = 0\,.$$

In particular, at such points

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y'}{x'} = 0\,.$$

The electron's trajectory has thus the qualitative shape shown in Fig. 1.5 (red curve).

ii) On the other hand, if $a = 1$, i.e.,

$$\boxed{v_0 = \frac{2E}{B}\,,}$$

then $x'(\tau) = 1 + \cos\tau \geqslant 0$, with (for $0 \leqslant \tau \leqslant 2\pi$)

$$x'(\tau) = 0 \iff \tau = \pi \quad \implies \quad x = \pi\,,\ \ y = 2\,.$$

In fact, when $\tau = \pi$ both velocity components x' and $y' = \sin\tau$ vanish simultaneously. It is easily checked that the trajectory has a cusp at the point $(\pi, 2)$ corresponding to $\tau = \pi$, since the slope to its tangent at this point satisfies

$$\frac{dy}{dx} = \frac{y'}{x'} = \frac{\sin\tau}{1+\cos\tau} = \frac{2\sin(\tau/2)\cos(\tau/2)}{2\cos^2(\tau/2)} = \tan(\tau/2) \underset{\tau\to\pi\mp}{\longrightarrow} \pm\infty\,.$$

The trajectory is in this case a *cycloid* (cf. Fig. 1.5, blue curve).

iii) Finally, if $a > 1$, i.e.,

$$\boxed{v_0 > \frac{2E}{B}\,,}$$

x is no longer a monotonic function of τ in the interval $[0, 2\pi]$. More precisely,

$$x'(\tau) \geqslant 0 \iff \tau \in [0, \arccos(-1/a)) \cup (2\pi - \arccos(-1/a), 2\pi]\,,$$

while

$$x'(\tau) < 0 \iff \tau \in \big(\arccos(-1/a), 2\pi - \arccos(-1/a)\big)\,.$$

Moreover, we have

$$y(2\pi - \tau) = y(\tau), \qquad x(2\pi - \tau) + x(\tau) = 2\pi,$$

so that the points $\mathbf{r}(\tau)$ and $\mathbf{r}(2\pi - \tau)$ are symmetric with respect to the vertical line $x = \pi$. Thus the trajectory is symmetric about the latter line. This can be shown to imply that the trajectory intersects itself at a point on the latter line (cf. Fig. (1.5), green curve).

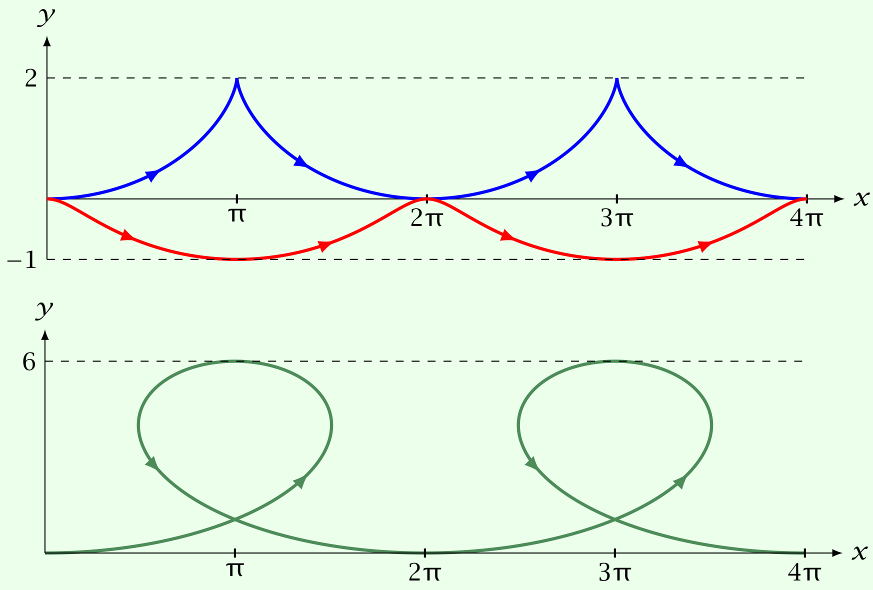


Figure 1.5. Electron's trajectory in Example 1.5 for $v_0 = E/(2B)$ (red line), $v_0 = 2E/B$ (blue line) and $v_0 = 4E/B$ (green line).

[a]From the Lorentz force law (1.49) it follows that evB/m and E/B have dimensions of acceleration and velocity, respectively. Hence eB/m has dimensions of $a/v = t^{-1}$, and $(E/B)(m/eB) = mE/(eB^2)$ has dimensions of $vt = l$.

Exercise. Redo the previous problem assuming that the electric field vanishes ($E = 0$). Show that in this case the particle describes a circle with constant frequency $\omega = eB/m$, called *cyclotron frequency*.

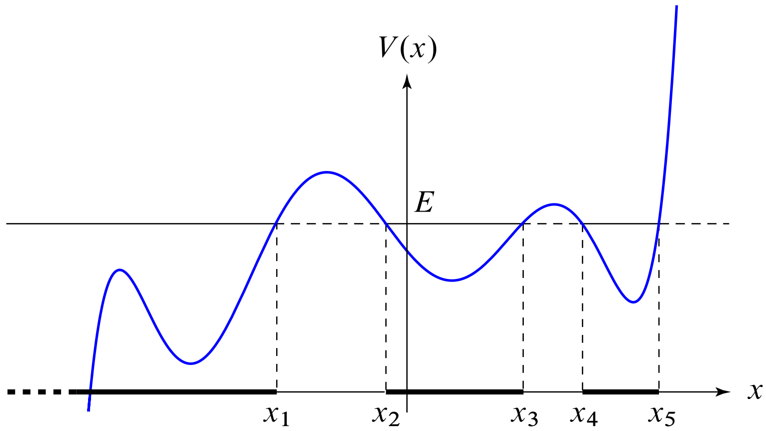


Figure 1.6. One-dimensional potential $V(x)$ with 5 turning points x_i for the energy E shown. The allowed region consists in this case of the three intervals $(-\infty, x_1]$, $[x_2, x_3]$ and $[x_4, x_5]$. Note also that the potential plotted in the figure has exactly 6 equilibria.

1.5 Motion of a particle in a one-dimensional potential

In this section we shall study the motion of a particle in one dimension, subject to a (smooth) force $F(x)$ independent of time and velocity. Such a force is *always conservative*, since $F(x) = -V'(x)$ with

$$V(x) = -\int F(x)\,\mathrm{d}x\,.$$

In this case the law of conservation of energy (1.38) reduces to

$$\boxed{\frac{1}{2}m\dot{x}^2 + V(x) = E\,,} \tag{1.57}$$

where the constant $E \in \mathbb{R}$ is the particle's total energy (which depends on the initial conditions). Conversely, differentiating (1.57) with respect to t we obtain

$$\dot{x}\big(m\ddot{x} - F(x)\big) = 0\,.$$

Hence *if* $\dot{x} \neq 0$ *Eq.* (1.57) *is* equivalent *to the equation of motion* $m\ddot{x} = F(x)$.

- The **equilibrium positions** (or **equilibria**) of the potential $V(x)$ are defined as the points $x_0 \in \mathbb{R}$ for which the equation of motion has the constant solution $x(t) = x_0$. If this is the case $\ddot{x}(t) = 0$ for all t, so that from the equation of motion we obtain

$$F\big(x(t)\big) = F(x_0) = -V'(x_0) = 0\,.$$

Thus *the equilibria are the points at which the force acting on the particle vanishes*. From the mathematical point of view, *the equilibria are the* critical points *of the potential* $V(x)$, that is, the roots of the equation

$$\boxed{V'(x) = 0\,.}$$

Note that, by the *existence and uniqueness theorem* for solutions of ordinary differential equations, if x_0 is an equilibrium the *only* solution of the equation of motion satisfying the initial conditions $x(t_0) = x_0$, $\dot{x}(t_0) = 0$ is the constant solution $x(t) = x_0$. In other words:

If at some instant the particle is at an equilibrium x_0 *with zero velocity* it will remain at x_0 indefinitely.

From Eq. (1.57) it immediately follows that *for a given energy E the motion can only take place in the region defined by the inequality*

$$V(x) \leqslant E\,,$$

that we shall call the **accessible** (or **allowed**) **region** for the energy E. In general (i.e., if the potential is sufficiently smooth), the allowed region is a (countable) *disjoint union of closed intervals*, some of which may be infinite to the right or the left (including the limiting case where the allowed region is the whole real line), or even reduce to isolated points (necessarily equilibria).

By continuity, if at some instant the particle lies on one of the disjoint closed intervals making up the allowed region it will always remain inside that interval.

Of particular interest are the endpoints x_i of the latter intervals, which must satisfy the equation $V(x) = E$. When the particle is at one of these points its velocity vanishes, since

$$x(t) = x_i \iff V(x_i) = E = \frac{1}{2}m\dot{x}(t)^2 + V(x_i) \iff \dot{x}(t) = 0 \tag{1.58}$$

by the law of energy conservation (1.57). We shall say that such a point x_i is a **turning point** of the trajectory if it is not an equilibrium, i.e., if

$$V(x_i) = E\,, \qquad V'(x_i) \neq 0\,.$$

In other words, the turning points are the endpoints of the disjoint closed intervals which make up the allowed region, *excluding the equilibria*.

• The reason for this terminology is the fact that *when the particle reaches a turning point its velocity $\dot{x}$ changes sign, and thus the particle "turns".* For example, if $V'(x_i) > 0$ then $V(x) < V(x_i) = E$ on a sufficiently small interval to the left of x_i, and $V(x) > V(x_i) = E$ on a similar interval to the right of x_i, so that the particle cannot reach the region to the *right* of the turning point. Hence the particle must approach the turning point from its *left*, and therefore $\dot{x}$ changes from positive (right before reaching the turning point) to negative (right afterwards).

If the particle has energy E and x_0 is an equilibrium with $V(x_0) = E$, then *the particle's trajectory cannot cross the equilibrium x_0*. Indeed, since $V(x_0) = E$ if the particle is at x_0 at some instant t_0 its velocity $\dot{x}(t_0)$ must vanish. From the remark on equilibria on p. 35 we conclude that $x(t) = x_0$ for all t, and hence the trajectory in this case consists of the single point x_0. Thus, if (for instance) $x(t_0) < x_0$ then we must have $x(t) < x_0$ for all t, i.e., the whole trajectory lies at the *left* of x_0. From the previous remarks it then follows that:

The trajectory of a particle with energy E is an *interval* (finite or infinite, which might reduce to a single point) on whose interior $V(x) < E$, limited by *turning points* and/or *equilibria* satisfying $V(x) = E$. Moreover, equilibria with $V(x) = E$ limiting the trajectory *cannot be reached in a finite time.*

• The equation of motion $m\ddot{x} = F(x)$ is *invariant* under *time translations* $t \mapsto t + t_0$, for any $t_0 \in \mathbb{R}$, since the time t does not appear explicitly in it. Hence *if $x(t)$ is a solution of the equation of motion so is $x(t + t_0)$, for all $t_0 \in \mathbb{R}$.*

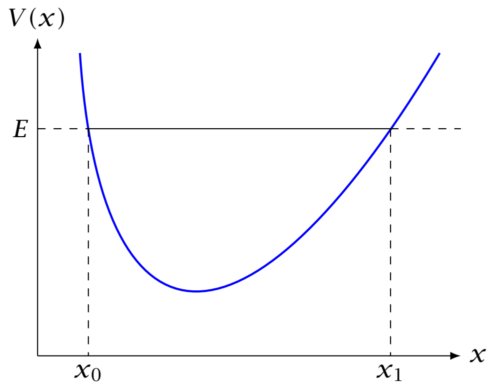


Figure 1.7. One-dimensional potential $V(x)$ with two consecutive turning points x_0, x_1 limiting an allowed interval $[x_0, x_1]$ (for the energy E shown) such that $V(x) < E$ for $x \in (x_0, x_1)$.

The equation of motion is also invariant under the *time reversal* mapping $t \mapsto -t$. Thus *if* $x(t)$ *is a solution of* (1.57) *so is* $x(-t)$. Combining this observation with the previous one it follows that $x(t_0 - t)$ is a solution of the equation of motion if $x(t)$ is.

The law of conservation of energy (1.57) allows us to easily find the general solution of the equation of motion in implicit form. Indeed, solving for $\dot{x}$ in Eq. (1.57) we obtain

$$\dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t} = \pm\sqrt{\frac{2}{m}\,(E - V(x))}\,. \tag{1.59}$$

Each of these *two* equations (corresponding to the two signs before the radical) is a first-order differential equation with *separable variables*, easily solved by separating variables and integrating:

$$t - t_0 = \pm\sqrt{\frac{m}{2}}\int \frac{\mathrm{d}x}{\sqrt{E - V(x)}}\,. \tag{1.60}$$

Here t_0 is an arbitrary integration constant which, without loss of generality, can be taken equal to zero in view of the previous comments. The behavior of the solutions depends crucially on the type of interval inside the allowed region where the motion takes place, as we shall see in more detail below. To simplify the exposition, we shall assume for the time being that *the interval where the motion takes place is limited by turning points* (not by equilibria). Hence on the interior of this interval we must have $V(x) < E$, whereas $V(x) = E$ and $V'(x) \neq 0$ at its endpoints (if any). By the law of conservation of energy (1.57), *the particle's velocity* $\dot{x}(t)$ *can only change sign at the endpoints of the latter interval.*

I) Bounded interval $[x_0, x_1]$

Consider first the case in which the particle's motion takes place in a bounded interval $[x_0, x_1]$ limited by two *consecutive turning points* $x_{0,1}$, so that

$$E = V(x_i) \quad \text{and} \quad V'(x_i) \neq 0\,, \qquad \text{with} \quad i = 0, 1,$$

and $V(x) < E$ for $x_0 < x < x_1$ (cf. Fig. 1.7). Let us suppose, without loss of generality[14], that $x_0 = x(0)$, so that $\dot{x}(0) = 0$. Then $\dot{x} > 0$ for sufficiently small $t > 0$, since otherwise the

[14]Indeed, suppose that the particle is at some point $a \in (x_0, x_1)$ at the initial time $t = t_0$. From Eq. (1.60) we then obtain

$$t = t_0 + \operatorname{sgn}(\dot{x}(a))\sqrt{\frac{m}{2}}\int_a^x \frac{\mathrm{d}s}{\sqrt{E - V(s)}},$$

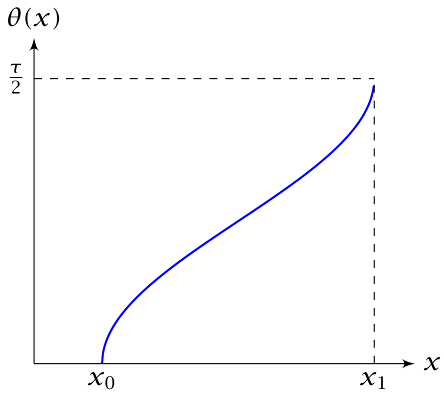


Figure 1.8. Function $\theta(x)$ in Eq. (1.61). Note that $\theta'(x_{0,1}) = +\infty$, by Eq. (1.63).

particle would enter the forbidden region to the left of x_0. We must therefore take the "+" sign in Eq. (1.60), obtaining[15]

$$t = \sqrt{\frac{m}{2}} \int_{x_0}^{x} \frac{\mathrm{d}s}{\sqrt{E - V(s)}} =: \theta(x) \,. \tag{1.61}$$

Hence the particle will reach the point x_1 at time $t = \tau/2$, with[16]

$$\tau = 2\theta(x_1) = \sqrt{2m} \int_{x_0}^{x_1} \frac{\mathrm{d}s}{\sqrt{E - V(s)}} \,. \tag{1.62}$$

Note that, since

$$\theta'(x) = \frac{\sqrt{m/2}}{\sqrt{E - V(x)}} > 0 \,, \qquad x_0 < x < x_1 \,, \tag{1.63}$$

$\theta(x)$ is monotonically increasing, and hence invertible, in the interval $[x_0, x_1]$ (see Fig. 1.8). Thus for $0 \leqslant t \leqslant \tau/2$ the particle's position as a function of time is given by

$$x = \theta^{-1}(t) \,, \qquad 0 \leqslant t \leqslant \frac{\tau}{2} \,.$$

For $t > \tau/2$ (with $t - (\tau/2)$ small enough) $\dot{x}$ becomes negative, since otherwise the particle would reach the forbidden region to the right of x_1. Using again Eq. (1.60), but this time with the "−" sign, and the initial condition $x(\tau/2) = x_1$ we obtain

$$t = \frac{\tau}{2} - \sqrt{\frac{m}{2}} \int_{x_1}^{x} \frac{\mathrm{d}s}{\sqrt{E - V(s)}} = \tau - \sqrt{\frac{m}{2}} \int_{x_0}^{x} \frac{\mathrm{d}s}{\sqrt{E - V(s)}} = \tau - \theta(x) \,. \tag{1.64}$$

In particular, the particle will again reach the point x_0 at time $t = \tau$ (since $\theta(x_0) = 0$). Note also

---

and thus the particle will reach the endpoint x_0 at the time

$$t_1 = t_0 - \operatorname{sgn}(\dot{x}(a)) \sqrt{\frac{m}{2}} \int_{x_0}^{a} \frac{\mathrm{d}s}{\sqrt{E - V(s)}} .$$

This time is *finite*, since the latter integral, which is improper at its lower endpoint x_0, is *convergent*. Indeed, since $V'(x_0) \neq 0$ by hypothesis, the integrand behaves as $(x - x_0)^{-1/2}$ in the vicinity of $s = x_0$. Thus $x(t_1) = x_0$ for some finite time t_1, so that replacing t by $t - t_1$ we have $x(0) = x_0$.

[15] The integral (1.61), which is improper at its lower limit $s = x_0$, is however convergent (see previous footnote).

[16] The integral (1.62) is also improper at its upper limit but certainly convergent, since $V'(x_1) \neq 0$ implies that the integrand behaves as $(x_1 - s)^{-1/2}$ near $s = x_1$.

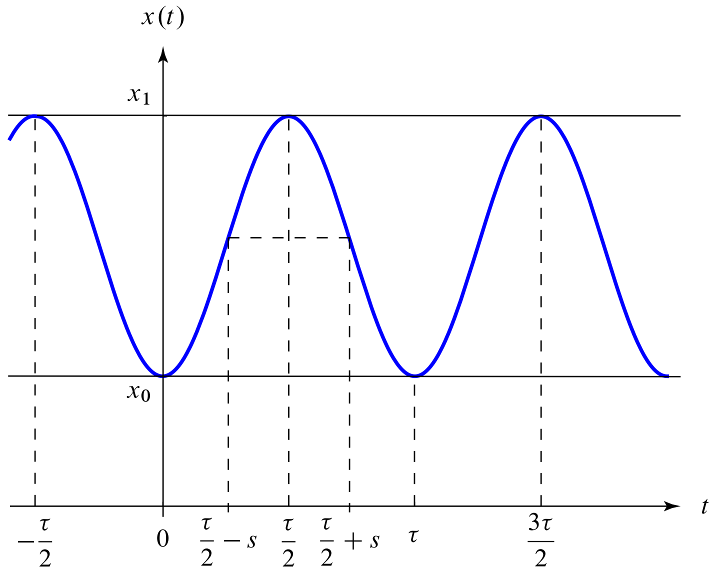


Figure 1.9. Motion of a particle in a one-dimensional potential between two consecutive turning points x_0, x_1.

that from Eq. (1.64) it follows that[17]

$$x = \theta^{-1}(\tau - t)\,, \qquad \frac{\tau}{2} \leqslant t \leqslant \tau\,.$$

Hence the particle's motion for $0 \leqslant t \leqslant \tau$, implicitly given by Eqs. (1.61)-(1.64), can be expressed in terms of the function θ^{-1} by the equations

$$x(t) = \begin{cases} \theta^{-1}(t)\,, & 0 \leqslant t \leqslant \frac{\tau}{2}\,; \\ \theta^{-1}(\tau - t)\,, & \frac{\tau}{2} \leqslant t \leqslant \tau\,. \end{cases} \tag{1.65}$$

Note, in particular, that $x(t)$ is symmetric about $t = \tau/2$, since by the previous equation

$$x\Big(\frac{\tau}{2} - s\Big) = \theta^{-1}\Big(\frac{\tau}{2} - s\Big) = x\Big(\frac{\tau}{2} + s\Big)\,, \qquad 0 \leqslant s \leqslant \frac{\tau}{2}\,.$$

The solution of the equations of motion valid for all t is just the *periodic extension with period* τ of the function $x(t)$ defined in $[0, \tau]$ by Eq. (1.65) (cf. Fig. 1.9). In other words, if $k\tau \leqslant t \leqslant (k+1)\tau$ with $k \in \mathbb{Z}$ then

$$x(t) = x(t - k\tau)\,, \tag{1.66}$$

where the RHS is evaluated using (1.65). Indeed, this function is a solution of the equation of motion due to the invariance of the latter equation under time translations, satisfies the initial conditions $x(0) = x_0$, $x'(0) = 0$ by construction, and is of class C^2 at the junction points $k\tau$ with $k \in \mathbb{Z}$ (exercise). Summarizing:

The motion of a particle between two consecutive turning points $x_{0,1}$ of a one-dimensional potential is *periodic*, with period τ given by Eq. (1.62).

[17] This could have also been proved noting that, by the two remarks on p. 36, if $x = \theta^{-1}(t)$ is a solution to the equation of motion for $t \in [0, \tau/2]$ then $x = \theta^{-1}(\tau - t)$ is a solution for $t \in [\tau/2, \tau]$, which satisfies the initial conditions $x(\tau/2) = \theta^{-1}(\tau/2) = x_1$ and $x'(\tau/2) = 0$ (by Eq. (1.57)). By the existence and uniqueness theorem for second-order ordinary differential equations, $x(t) = \theta^{-1}(\tau - t)$ for $\tau/2 \leqslant t \leqslant \tau$.

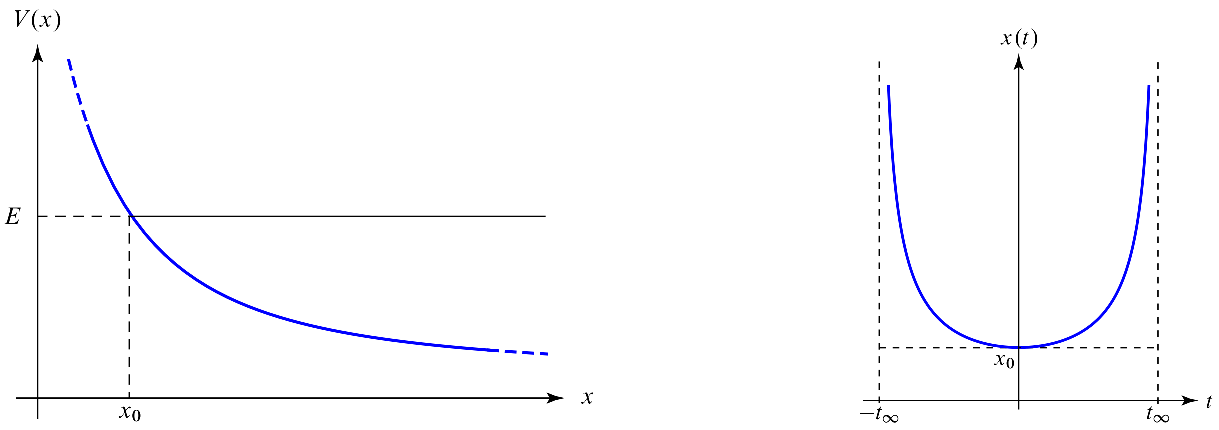


Figure 1.10. Left: one-dimensional potential $V(x)$ with a a turning point x_0 limiting a semiinfinite allowed interval $[x_0, \infty)$ (for the energy E shown) such that $V(x) < E$ for $x > x_0$. Right: corresponding law of motion $x(t)$ (in the case of finite t_∞).

Exercise. Show that the function $x(t)$ defined by Eqs. (1.65)-(1.66) is invariant under time reversal, i.e., that $x(t) = x(-t)$.

Solution. The function $f(t) := x(-t)$ is a solution of the equation of motion, due to the invariance of the latter equation under time reversal. At $t = 0$, the solution $f(t)$ satisfies the *same* initial conditions as $x(t)$, since

$$f(0) = x(0) = x_0 \,, \qquad f'(0) = -x'(0) = 0 \,.$$

By the existence and uniqueness theorem for second-order ordinary differential equations, $f(t) = x(-t) = x(t)$ for all t.

II) Semi-infinite interval $[x_0, \infty)$

Consider next the case in which the particle moves inside a semi-infinite interval[18] $[x_0, \infty)$ limited by a turning point x_0, so that $V(x_0) = E$, $V'(x_0) \neq 0$, and $V(x) < E$ for $x > x_0$ (cf. Fig. 1.10). If the particle is at the point x_0 for $t = 0$ then $\dot{x}(t) > 0$ for $t > 0$, and the relation between the time t and the position x is given by equation (1.61) for all $t > 0$. In particular, the particle reaches (positive) infinity at time

$$\boxed{t_\infty = \theta(\infty) = \sqrt{\frac{m}{2}} \int_{x_0}^{\infty} \frac{\mathrm{d}s}{\sqrt{E - V(s)}} \,,}$$

which is finite or infinite depending on whether the integral on the RHS is convergent or divergent at $+\infty$. For instance, if $V(x) \sim -x^a$ with $a \geqslant 0$ for $x \to \infty$ then t_∞ is finite if $a > 2$, and infinite if $0 \leqslant a \leqslant 2$. Taking into account the definition (1.61) of the function $\theta(x)$, the particle's motion for $0 \leqslant t < t_\infty$ is given by the equation

$$x = \theta^{-1}(t) \,, \qquad 0 \leqslant t < t_\infty \,,$$

with $x(t_\infty) = \infty$. On the other hand, for $-t_\infty < t \leqslant 0$ we have

$$x = \theta^{-1}(-t) \,, \qquad -t_\infty < t \leqslant 0 \,,$$

where again $x(-t_\infty) = \infty$. Indeed, the latter function is a solution to the equation of motion (due to the invariance of this equation under time reversal $t \mapsto -t$), and satisfies the correct initial

[18] The case in which the motion takes place in a semi-infinite interval $(-\infty, x_0]$ limited by a turning point is dealt with analogously.

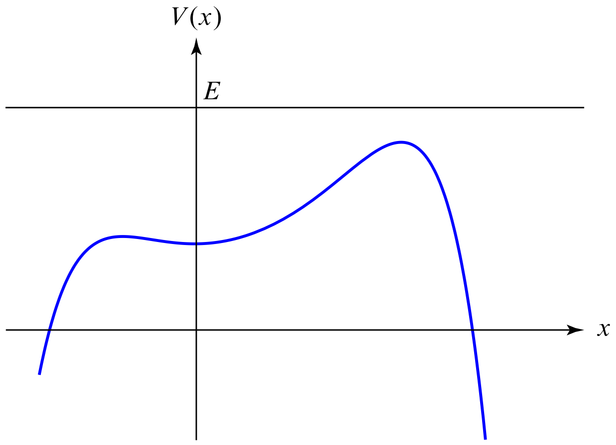


Figure 1.11. One-dimensional potential with $E > V(x)$ for all x (for the energy E shown).

conditions at $t = 0$:

$$x(0) = \theta^{-1}(0) = x_0 \,, \qquad \dot{x}(0) = 0$$

(the last equation is actually a consequence of the first, since x_0 is a turning point). An alternative way of reaching the same conclusion is to observe that if $t < 0$ then $\dot{x}(t) < 0$, since the particle is at the point x_0 for $t = 0$. Therefore we must take the "$-$" sign in Eq. (1.59), which yields the equation

$$t = -\sqrt{\frac{m}{2}} \int_{x_0}^{x} \frac{\mathrm{d}s}{\sqrt{E - V(s)}} = -\theta(x) \quad \Longleftrightarrow \quad x = \theta^{-1}(-t),$$

on account of the initial condition $x(0) = x_0$. In other words, in this case the law of motion is

$$\boxed{x = \theta^{-1}(|t|)\,, \qquad -t_\infty < t < t_\infty\,.}$$

Note, in particular, that (as in Case I) above) $x(t) = x(-t)$.

III) Whole real line $(-\infty, \infty)$

Consider, finally, the case in which for a certain energy E the trajectory is the whole real line, so that $V(x) \leqslant E$ for all x. We must actually have $V(x) < E$ for all $x \in \mathbb{R}$ (cf. Fig. 1.11), since otherwise a point x_0 with $V(x_0) = E$ would be an equilibrium (absolute maximum of V), which cannot be crossed by the trajectory. Let $x(0) = x_0$; then $\dot{x}^2(0)$ is fixed by conservation of energy, namely

$$\dot{x}^2(0) = \frac{2}{m}\,(E - V(x_0)) > 0,$$

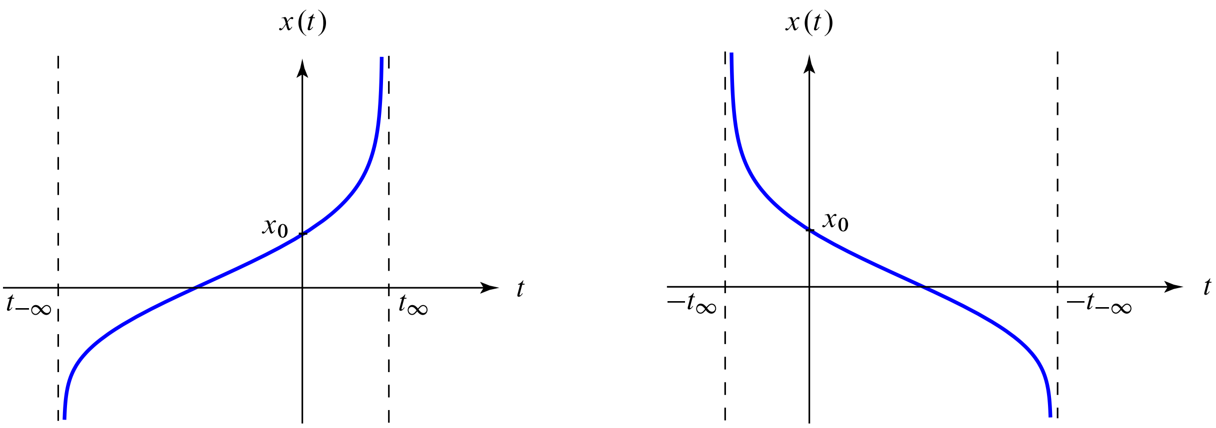


Figure 1.12. Law of motion $x(t)$ for the potential $V(x)$ and the energy E shown in Fig. 1.11 in the cases $\dot{x}(0) > 0$ (left) or $\dot{x}(0) < 0$ (right), assuming that $t_{\pm\infty}$ are finite.

but the sign of $\dot{x}(0)$ is of course undetermined. If (for example) $\dot{x}(0) > 0$, then $\dot{x}(t) > 0$ for all t, since the velocity cannot vanish in this case by the law of conservation of energy. We must therefore take the "+" sign in (1.59) for all t, which yields the relation

$$t = \sqrt{\frac{m}{2}} \int_{x_0}^{x} \frac{\mathrm{d}s}{\sqrt{E - V(s)}} = \theta(x) .$$

In particular, the particle reaches $\pm\infty$ at time

$$\boxed{t_{\pm\infty} = \theta(\pm\infty) = \int_{x_0}^{\pm\infty} \frac{\mathrm{d}s}{\sqrt{E - V(s)}}}$$

(which may again be finite or infinite, according to whether the integral is convergent of divergent at $\pm\infty$), and the particle's motion is governed by the equation

$$\boxed{x = \theta^{-1}(t) , \qquad t_{-\infty} < t < t_{\infty} .}$$

Likewise, if $\dot{x}(0) < 0$ then

$$\boxed{x = \theta^{-1}(-t) , \qquad -t_{\infty} < t < -t_{-\infty} ,}$$

where now $x(-t_{\pm\infty}) = \pm\infty$ (cf. Fig. 1.12).

Example 1.6. Consider the potential

$$\boxed{V(x) = k\left(\frac{x^2}{2} - \frac{x^4}{4a^2}\right) , \qquad k,\ a > 0 ,}$$

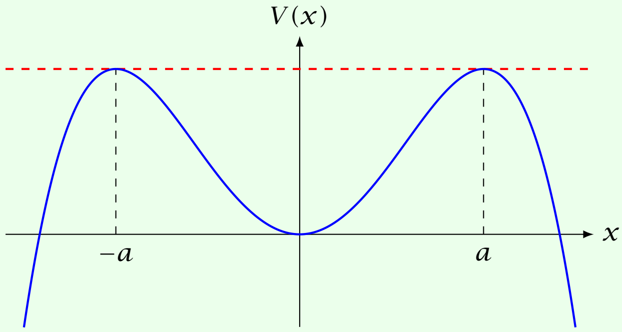


Figure 1.13. Potential in Example 1.6 (blue line) and energy $E = ka^2/4$ (dashed red line).

plotted in Fig. 1.13. Differentiating with respect to x we obtain

$$V'(x) = kx\left(1 - \frac{x^2}{a^2}\right) = 0 \iff x = 0,\ \pm a .$$

Therefore the equilibria are in this case the points $x = 0$ (relative minimum of V) and $x = \pm a$ (global maxima). The allowed region, and therefore the type of trajectory, depends on the value of the energy E as follows:

i) $E < 0$
The allowed region is the union of the two semi-infinite intervals $(-\infty, -c]$ and $[c, \infty)$, c being the only positive root of the equation $V(x) = E$. Thus in this case the trajectory is *unbounded* (to the right if $x(0) > c$, to the left if $x(0) < -c$).

ii) $E = 0$
The allowed region is the union of the semi-infinite intervals $(-\infty, -\sqrt{2}\,a]$ and $[\sqrt{2}\,a, \infty)$ along with the origin, which as we know is an equilibrium. In particular, if $x(0) = 0$ then $x(t) = 0$ for all t (equilibrium solution), while if $|x(0)| \geqslant \sqrt{2}\,a$ the trajectory is *unbounded*.

iii) $0 < E < ka^2/4$
Since $ka^2/4 = V(\pm a)$ is the potential's maximum value, the allowed region is the union of the three intervals $(-\infty, -c_2]$, $[-c_1, c_1]$ and $[c_2, \infty)$, where $c_1 < c_2$ are the two positive roots of the equation $V(x) = E$. Therefore in this case the trajectory is *unbounded* (to the left or right) if $|x(0)| \geqslant c_2$, while if $|x(0)| \leqslant c_1$ the motion is *periodic*, with amplitude c_1.

iv) $E > ka^2/4$
Since $V(x) < E$ for all x, the allowed region (and the trajectory) is *the whole real line*. Note that in this case the time that it takes the particle to reach $\pm\infty$ is *finite*, since for $|x| \to \infty$ the integral

$$\int^{\pm\infty} \frac{\mathrm{d}x}{\sqrt{E - V(x)}} \sim \int^{\pm\infty} \frac{\mathrm{d}x}{x^2}$$

converges.

v) $E = ka^2/4$
We have left for the end the most interesting case, in which $E = ka^2/4$. Since $V(x) \leqslant ka^2/4$ for all x, the allowed region is again the whole real line, and it might therefore superficially seem that the trajectory is also the whole real line. However, this conclusion is *wrong*, since the allowed region now contains the two *equilibria* $x = \pm a$. If the particle starts at $t = 0$ from a point $x_0 \neq \pm a$, it *cannot* reach the points $\pm a$ in a *finite* time. Indeed, if $x(t_0) = \pm a$ for a certain time $t_0 \in \mathbb{R}$, from (1.57) with $V(\pm a) = ka^2/4 = E$ we obtain $\dot{x}(t_0) = 0$. Since the points $\pm a$ are equilibria, this implies that $x(t) = \pm a$ for all t. Therefore in this case the possible trajectories of the particle are the *open* intervals $(-\infty, a)$, $(-a, a)$ and (a, ∞), along with the two equilibria $\pm a$. In particular, if $|x(0)| < a$ the trajectory remains in the interval $(-a, a)$ for all $t \in \mathbb{R}$ and is therefore *bounded*. However, it is *not periodic*, but rather verifies $x(\pm\infty) = \pm a$ if $\dot{x}(0) > 0$ or $x(\pm\infty) = \mp a$ if $\dot{x}(0) < 0$. (Why is $\dot{x}(0) \neq 0$ in this case?)

• In this case it is possible to explicitly integrate the equation of motion when $E = ka^2/4$. Indeed, substituting this value of the energy in Eq. (1.59) we obtain

$$\dot{x} = \pm\sqrt{\frac{k}{2ma^2}}\,(x^2 - a^2)\,.$$

Separating variables and integrating we have

$$\pm\sqrt{\frac{2k}{m}}\,t = \int \frac{2a}{x^2 - a^2}\,\mathrm{d}x = \log\left|\frac{x-a}{x+a}\right| \quad\Longrightarrow\quad \left|\frac{x-a}{x+a}\right| = \mathrm{e}^{\pm 2\omega t}\,, \qquad \omega := \sqrt{\frac{k}{2m}}\,,$$

where without loss of generality we have taken the integration constant equal to zero. If the particle lies initially in one of the intervals $(-\infty, -a)$ or (a, ∞) then

$$\left|\frac{x-a}{x+a}\right| = \frac{x-a}{x+a}\,,$$

and therefore

$$\boxed{x = a\,\frac{1 + \mathrm{e}^{\pm 2\omega t}}{1 - \mathrm{e}^{\pm 2\omega t}} = \mp a \coth(\omega t)\,.}$$

The latter expression actually defines *four* different solutions. Indeed, if initially the particle is in the region $x > a$ with positive (resp. negative) velocity then we must take the "$-$" (resp. "$+$") sign in the previous expression, and the associated solution is therefore defined for $t < 0$ (resp. $t > 0$). This solution corresponds to a motion reaching positive infinity (resp. arriving from positive infinity) in a finite time and tending to the point $x = a$ for $t \to -\infty$ (resp. $t \to +\infty$). Likewise, if $x(0) < -a$ then the solution $x = -a\coth(\omega t)$ with $t > 0$ corresponds to a motion from $x = -\infty$ (for $t \to 0+$) to $x = -a$ (for $t \to \infty$) with positive velocity, while the solution

$x = a\coth(\omega t)$ with $t < 0$ corresponds to a motion from $x = -a$ (for $t \to -\infty$) to $x = -\infty$ (for $t \to 0+$) with negative velocity (cf. Fig. 1.14).

Similarly, if the particle lies initially in the interval $(-a, a)$ then

$$\left|\frac{x-a}{x+a}\right| = \frac{a-x}{a+x},$$

and thus

$$\boxed{x = a\,\frac{1-\mathrm{e}^{\pm 2\omega t}}{1+\mathrm{e}^{\pm 2\omega t}} = \mp a\tanh(\omega t)}$$

(cf. Fig. 1.14). The solution corresponding to the "+" (resp. "−") sign has always positive (resp. negative) velocity, and tends to $\pm a$ for $t \to \pm\infty$ (resp. $t \to \mp\infty$).

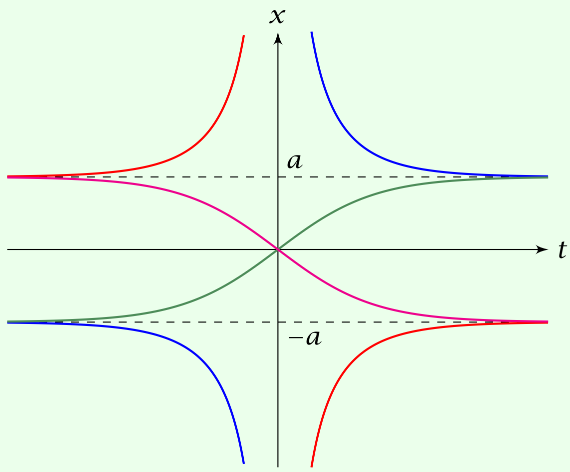


Figure 1.14. Plot of solutions $x = a\coth(\omega t)$ (blue), $x = -a\coth(\omega t)$ (red), $x = a\tanh(\omega t)$ (green), and $x = -a\tanh(\omega t)$ (magenta) in Example 1.6.

To visualize the different trajectories followed by the particle and qualitatively understand their properties, it is useful to plot the momentum $p = m\dot{x}$ as a function of the position x for different values of the energy E. This plot is usually known as the **phase map** of the system. From the law of energy conservation it follows that the equation of the trajectories in the phase map is

$$\frac{p^2}{2m} + V(x) = E\,,$$

which in this case reduces to

$$\frac{p^2}{2m} + k\left(\frac{x^2}{2} - \frac{x^4}{4a^2}\right) = E\,.$$

The corresponding trajectories (obviously symmetric with respect to both axes) are represented in Fig. 1.15. Note that the equation of the trajectories with energy equal to the critical energy $E = ka^2/4$ is

$$p = \pm\sqrt{\frac{mk}{2a^2}}\sqrt{x^4 - 2a^2x^2 + a^4} = \pm\sqrt{\frac{mk}{2a^2}}\,(x^2 - a^2)\,. \tag{1.67}$$

This is the equation of two parabolas whose axis is the vertical line $x = 0$, intersecting at the equilibria $(\pm a, 0)$. These trajectories divide the phase map into 5 disjoint connected regions, in each of which the trajectories have different qualitative properties (they are bounded or unbounded, reach $x = \pm\infty$ or not, etc.). For this reason, the trajectories (1.67) are called *separatrices*.

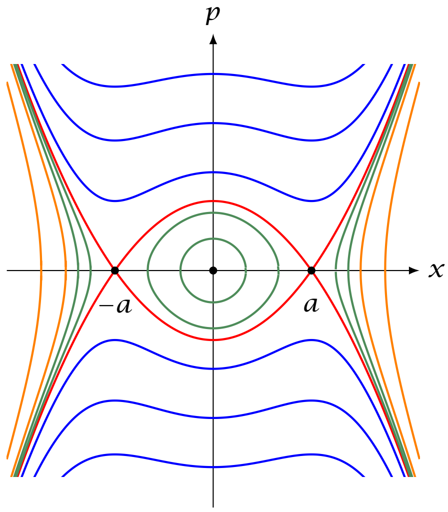


Figure 1.15. Phase map for the potential in Example 1.6. The red line consists of the trajectories with energy $E = ka^2/4$ (including the two equilibria $(\pm a, 0)$), while the orange, green and blue lines represent trajectories with $E < 0$, $0 < E < ka^2/4$ and $E > ka^2/4$, respectively.

1.5.1 Stability of equilibria. Period of the small oscillations

Intuitively speaking, an equilibrium x_0 is **stable** if sufficiently small perturbations of the initial conditions $x(0) = x_0$, $\dot{x}(0) = 0$ lead to solutions $x(t)$ of the equation of motion which remain arbitrarily close to x_0 (and with velocity arbitrarily close to 0) at all times $t > 0$. In the previous example, it is clear that the equilibrium $x = 0$ is *stable*, while $x = \pm a$ are both *unstable* equilibria. Indeed, if we slightly disturb the initial condition $x(0) = \dot{x}(0) = 0$ corresponding to the first of these equilibria, that is, we consider particle motions with $|x(0)|$ and $|\dot{x}(0)|$ small enough, the energy will be slightly positive but much smaller than the critical value $ka^2/4$, and therefore the motion will be periodic and with amplitude close to zero. On the contrary, a perturbation of the initial data $x(0) = \pm a$, $\dot{x}(0) = 0$ such that (for instance) $x(0) = \pm a$ and $|\dot{x}(0)| = \varepsilon > 0$ results in a motion with energy greater than the critical energy $ka^2/4$ no matter how small ε is, and therefore $x(t) \to \pm\infty$ for $t \to \infty$.

In general, an equilibrium is *stable* if and only if it is a *relative minimum* of the potential.

To heuristically justify this statement, suppose that x_0 is a critical point of the potential V, i.e., that $V'(x_0) = 0$. If x_0 is a *relative minimum* of V then in a sufficiently small interval centered at x_0 we have $V'(x) < 0$ for $x < x_0$ and $V'(x) > 0$ for $x > x_0$. Hence $F(x) = -V'(x)$ and $x - x_0$ have *opposite* signs for x sufficiently close to x_0, i.e, *the force acting on the particle always points towards the equilibrium* x_0 in its vicinity. Thus in this case the equilibrium is *stable*. Likewise, if x_0 is a *relative maximum* of V then near x_0 the force $F(x)$ points *away from* x_0, and hence the equilibrium is *unstable*. Finally, if x_0 is an *inflection point* of V then $V'(x)$ has constant sign (positive or negative) for $x \neq x_0$ in the vicinity of x_0. If (for instance) $V'(x) > 0$ then in a neighborhood of x_0 the force $F(x)$ points *away from* x_0 for $x < x_0$, and the equilibrium is again *unstable*.

An alternative proof of the previous result is based solely on energy considerations. Indeed, suppose to begin with that x_0 is a relative minimum of the potential $V(x)$. A solution of the equation of motion with initial conditions $(x(0), \dot{x}(0))$ close to $(x_0, 0)$ will have an energy

$$E = \frac{1}{2} m\dot{x}(0)^2 + V(x(0))$$

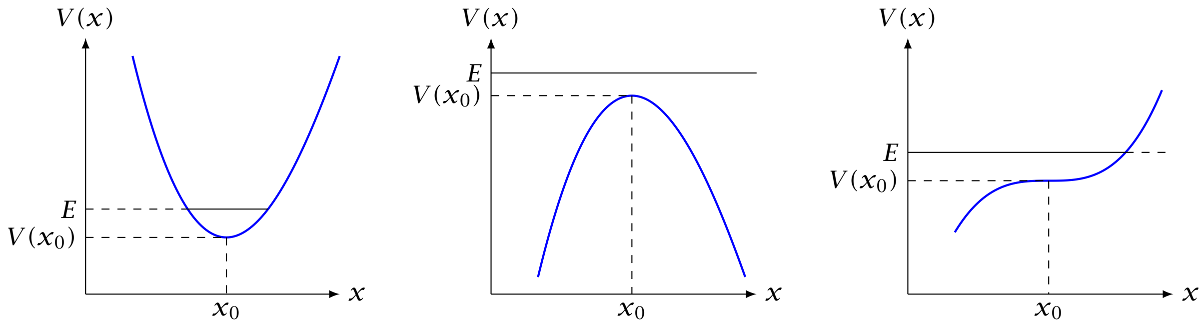


Figure 1.16. Potential $V(x)$ with a local minimum (left), a local maximum (center) or an inflection point (right) at x_0.

slightly greater than $V(x_0)$, as $|\dot{x}(0)| \gtrsim 0$ and (for $x(0)$ close enough to x_0) $V(x(0)) \gtrsim V(x_0)$, since by hypothesis x_0 is a local minimum of V. Hence the motion will consist in oscillations about x_0 with amplitude decreasing as E approaches $V(x_0)$ —see Fig. 1.16 (left). On the contrary, if x_0 is a relative maximum of V then a solution with (for instance) $x(0) = x_0$ and $|\dot{x}(0)|$ small but non-vanishing will have energy

$$E = \frac{1}{2} m\dot{x}(0)^2 + V(x_0)$$

slightly larger than $V(x_0)$. Hence the particle will move away from x_0 by a finite amount (to the left or right) no matter how small $|\dot{x}(0)|$ is —see Fig. 1.16 (center). Finally, suppose that x_0 is an inflection point of V, with (for instance) $V'(x)$ strictly increasing near x_0. In this case a solution with, e.g., initial conditions $x(0) = x_0$ and $|\dot{x}(0)|$ small but non-vanishing will again have energy $E \gtrsim V(x_0)$, and thus the particle will move away from x_0 by a finite amount (to the *left*) no matter how small $|\dot{x}(0)|$ is —see Fig. 1.16 (right). (In fact, it is apparent from Fig. 1.16 that when x_0 is either a local maximum or an inflection point that any solution with initial conditions close to equilibrium will have energy close to $V(x_0)$, and will move away by a finite amount from equilibrium either to the left or to the right.)

Suppose that the potential $V(x)$ is smooth (say, of class C^2) at an equilibrium x_0, and that furthermore $V''(x_0) \neq 0$ (which is the generic case). If the equilibrium is *stable* we must then have

$$\boxed{V'(x_0) = 0\,, \qquad V''(x_0) > 0\,.}$$

For initial conditions $(x(0), \dot{x}(0))$ sufficiently close to $(x_0, 0)$ the motion is periodic, as $E \gtrsim V(x_0)$ and the particle oscillates between two consecutive turning points close to x_0; see Fig. 1.16 (left). To find an approximation to the period of these small amplitude oscillations, taking into account that $|x - x_0| \ll 1$ we Taylor expand the force $F(x)$ about x_0 to first order in $x - x_0$:

$$F(x) = -V'(x) = -V''(x_0)(x - x_0) + \mathrm{O}\left((x - x_0)^2\right).$$

Hence the particle's equation of motion is approximately

$$\ddot{x} = \frac{F(x)}{m} \simeq -\frac{V''(x_0)}{m}(x - x_0)\,,$$

which can be written as

$$\ddot{\xi} + \omega^2 \xi = 0$$

with

$$\xi := x - x_0\,, \qquad \omega := \sqrt{\frac{V''(x_0)}{m}}\,.$$

As is well known, the general solution of this equation can be written as

$$\xi = A\cos(\omega t + \alpha)\,,$$

where $\alpha \in [0, 2\pi)$ and $A \geqslant 0$ are arbitrary constants (with $A \ll 1$ to be consistent with the hypothesis $|x - x_0| = |\xi| \ll 1$). Hence the particle's motion near the equilibrium x_0 is approximately described by the equation

$$\boxed{x(t) \simeq x_0 + A\cos(\omega t + \alpha)\,.}$$

In other words, the period τ of the *small oscillations* about the equilibrium x_0 is approximately given by

$$\boxed{\tau \simeq \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{m}{V''(x_0)}}\,.} \tag{1.68}$$

In particular, from the above formula it follows that the period of the *small* oscillations about a stable equilibrium is approximately *independent of the amplitude* (or, equivalently, the energy). Obviously, this is exactly true for *any* amplitude for the harmonic potential $V(x) = k(x-x_0)^2/2$. For a general potential, the period of the oscillations of *arbitrary* amplitude is given *exactly* by Eq. (1.62), and is thus in general *dependent on the amplitude*. More precisely,

$$\boxed{\tau = \sqrt{2m}\int_{x_1}^{x_2}\frac{\mathrm{d}x}{\sqrt{E - V(x)}} = \sqrt{2m}\int_{x_1}^{x_2}\frac{\mathrm{d}x}{\sqrt{V(x_{1,2}) - V(x)}}\,,} \tag{1.69}$$

where $x_1 < x_2 = x_1 + 2A$ are the two turning points (roots of the equation $V(x) = E$) closest to the equilibrium x_0 and A is the amplitude.

Example 1.7. The period of the small oscillations about the origin for the potential in Example 1.6 is *approximately*

$$\tau \simeq 2\pi\sqrt{\frac{m}{k}}\,. \tag{1.70}$$

Note, however, that this approximation is only correct for *small* amplitudes $A \ll a$. For *arbitrary* amplitude $0 < A < a$, the period is *exactly* given by the formula

$$\tau = \sqrt{2m}\int_{-A}^{A}\frac{\mathrm{d}x}{\sqrt{E - V(x)}} = 2\sqrt{2m}\int_{0}^{A}\frac{\mathrm{d}x}{\sqrt{E - V(x)}}\,,$$

with

$$E = V(A) = \frac{kA^2}{4a^2}\,(2a^2 - A^2)\,.$$

Substituting this value of E in the previous formula for τ, setting $x = As$ and operating we obtain

$$\tau = 4\sqrt{\frac{m}{k}}\int_0^1\frac{\mathrm{d}s}{\sqrt{(1-s^2)\left(1 - \frac{\varepsilon^2}{2}(1+s^2)\right)}}\,, \qquad \varepsilon := \frac{A}{a} \in (0,1)\,. \tag{1.71}$$

Note that when ε tends to 1, that is A tends to a, the integral tends to infinity, which is consistent with the fact that for $A = a$ the particle takes an infinite time to reach the equilibria $x = \pm a$. If ε is small, taking into account that

$$\left(1 - \tfrac{\varepsilon^2}{2}(1+s^2)\right)^{-1/2} = 1 + \frac{\varepsilon^2}{4}(1+s^2) + \mathrm{O}(\varepsilon^4)$$

and using the above formula for the period we obtain the more accurate expansion

$$\tau = 4\sqrt{\frac{m}{k}}\left(\int_0^1 \frac{ds}{\sqrt{1-s^2}} + \frac{\varepsilon^2}{4}\int_0^1 \frac{1+s^2}{\sqrt{1-s^2}}\,ds + O(\varepsilon^4)\right)$$
$$= \sqrt{\frac{m}{k}}\left[2\pi + \varepsilon^2\int_0^{\pi/2}(1+\cos^2\theta)\,d\theta + O(\varepsilon^4)\right] = 2\pi\sqrt{\frac{m}{k}}\left(1+\frac{3}{8}\varepsilon^2 + O(\varepsilon^4)\right) \quad (1.72)$$

(cf. Fig. 1.17).

Note. In this case, the *exact* value of the period can be expressed in terms of the *complete elliptic integral of the first kind*

$$K(\alpha) := \int_0^{\pi/2}(1-\alpha^2\sin^2 s)^{-1/2}\,ds = \int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-\alpha^2x^2)}}, \qquad 0 \leqslant \alpha < 1, \quad (1.73)$$

by performing the change of variables $x = \sin t$ in the integral (1.71), namely

$$\tau = 4\sqrt{\frac{m}{k}}\left(1-\frac{\varepsilon^2}{2}\right)^{-1/2} K\left(\frac{\varepsilon}{\sqrt{2-\varepsilon^2}}\right).$$

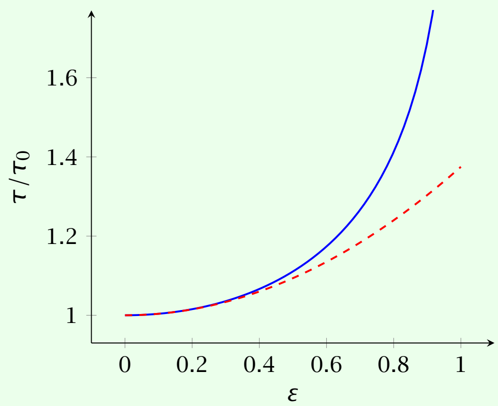


Figure 1.17. Period of the oscillations about the origin for the potential in Example 1.6 (in units of the approximate period of the small oscillations $\tau_0 = 2\pi\sqrt{m/k}$) as a function of the parameter $\varepsilon = A/a$ (solid blue line) compared to its approximation (1.72) (dashed red line).

Exercise. Find the relation between the energy E and the amplitude A of the small oscillations about a stable equilibrium x_0 with $V''(x_0) > 0$.

Solution. The turning points (points on the trajectory at a maximum and minimum distance of the equilibrium x_0) are the two roots $x_{1,2}$ of the equation $E = V(x)$ to the left and right of x_0. Using the approximation

$$V(x) \simeq V(x_0) + \frac{1}{2}V''(x_0)(x-x_0)^2$$

we obtain

$$E = V(x_{1,2}) \simeq V(x_0) + \frac{1}{2}V''(x_0)(x_{1,2}-x_0)^2$$
$$\implies A := \frac{1}{2}(x_2 - x_1) \simeq |x_{1,2} - x_0| \simeq \sqrt{\frac{2(E-V(x_0))}{V''(x_0)}}\,.$$

Exercise. Find the dependence on the amplitude and energy of the period of the oscillations around $x = 0$ of a particle of mass m moving subject to the potential $V(x) = k|x|^n$, with $k > 0$ and $n \in \mathbb{N}$.

Exercise. Redo the discussion in Example 1.6 for the potential

$$V(x) = k\left(\frac{x^4}{4a^2} - \frac{x^2}{2}\right), \qquad k,\ a > 0\,.$$

In particular, determine the stable equilibria and compute the period of small oscillations about them. Show that the period of the oscillations of amplitude $A \gg a$ about the origin is approximately proportional to a/A, and find the proportionality constant.

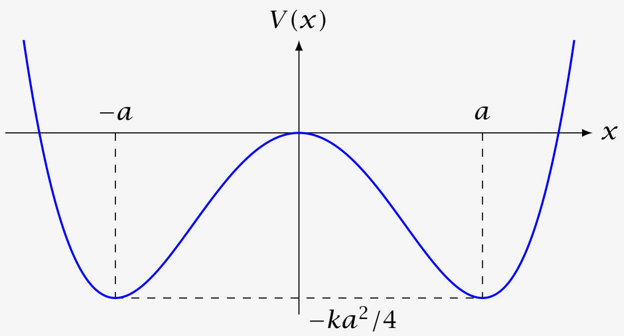


Figure 1.18. Potential $V(x) = k\left(\frac{x^4}{4a^2} - \frac{x^2}{2}\right)$.

Solution. The potential in this exercise differs only by a sign from that of Example 1.6, and thus its graph is as shown in Fig. 1.18. In particular, the equilibria are again 0 (unstable) and $\pm a$ (both stable). In this case E must be greater than the absolute minimum $V(\pm a) = -ka^2/4$ of V, and therefore we can have the following types of trajectories:

i) $E = -ka^2/4$
In this case either $x(t) = -a$ or $x(t) = a$ for all t (equilibrium solutions).

ii) $-ka^2/4 < E < 0$
The particle oscillates about the equilibrium $x = -a$ if $x(0) < 0$, or about $x = a$ if $x(0) > 0$, so that the motion is periodic and bounded. The period of the small oscillations about either equilibrium is approximately given by

$$\tau \simeq 2\pi\sqrt{\frac{m}{V''(\pm a)}} = 2\pi\sqrt{\frac{m}{2k}}\,.$$

iii) $E = 0$
If $x(0) > 0$ then the motion is bounded ($0 < x \leqslant \sqrt{2}\,a$, where $\sqrt{2}\,a$ is the positive root of $V(x) = 0$) but *not periodic*, since the unstable equilibrium at the origin cannot be reached in a finite time. Similarly, if $x(0) < 0$ the motion is again bounded (with $-\sqrt{2}\,a \leqslant x < 0$) but not periodic. Finally, if $x(0) = 0$ then $x(t) = 0$ for all t (equilibrium solution).

iv) $E > 0$
In this case the motion is bounded and periodic, since the two roots $\pm A$ of the equation $V(x) = 0$ are turning points. Note that $A > \sqrt{2}\,a$ is not infinitesimally small, so that Eq. (1.68) does not

apply in this case. The period of the oscillations as a function of their amplitude A is given, however, by the *exact* formula

$$\tau = 2\sqrt{2m}\int_0^A \frac{\mathrm{d}x}{\sqrt{V(A)-V(x)}} = 4a\sqrt{\frac{2m}{k}}\int_0^A \frac{\mathrm{d}x}{\sqrt{(A^2-x^2)(A^2-2a^2+x^2)}}\,.$$

Calling $\varepsilon := a/A < 1/\sqrt{2}$ and performing the change of variable $x = As$ (so that s is a dimensionless variable) we obtain

$$\tau = 4\varepsilon\sqrt{\frac{2m}{k}}\int_0^1 \frac{\mathrm{d}s}{\sqrt{(1-s^2)(1-2\varepsilon^2+s^2)}}\,.$$

When $A \gg a$ the dimensionless parameter ε is very small, and we can thus approximate the period by

$$\tau \simeq 4C\sqrt{\frac{2m}{k}}\,\varepsilon\,,$$

where the constant C is given by[a]

$$C := \int_0^1 \frac{\mathrm{d}s}{\sqrt{1-s^4}} \overset{s=\cos\theta}{=} \int_0^{\pi/2} \frac{\mathrm{d}\theta}{\sqrt{2-\sin^2\theta}} = \frac{1}{\sqrt{2}}K(1/\sqrt{2}) = 1.31103\ldots\,,$$

where $K(\alpha)$ is the complete elliptic integral of the first kind defined by Eq. (1.73). Thus the period of the oscillations with *large* amplitude $A \gg a$ is (approximately) *inversely proportional* to the amplitude.

Note. The exact value of the period of the oscillations of amplitude $A > \sqrt{2}\,a$ can also be determined in terms of a complete elliptic integral of the first kind. Indeed, performing the change of variable $s = \cos\theta$ in the integral for τ we obtain

$$\tau = 4\varepsilon\sqrt{\frac{2m}{k}}\int_0^{\pi/2} \frac{\mathrm{d}\theta}{\sqrt{2(1-\varepsilon^2)-\sin^2\theta}} = 4\varepsilon\sqrt{\frac{m}{k(1-\varepsilon^2)}}\,K\Big(\frac{1}{\sqrt{2(1-\varepsilon^2)}}\Big).$$

[a] The constant C can be also be expressed in terms of Euler's *gamma function*

$$\Gamma(x) := \int_0^\infty t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t, \qquad x > 0,$$

as $C = \dfrac{\sqrt{\pi}}{4}\,\dfrac{\Gamma(1/4)}{\Gamma(3/4)}$.

1.6 Dynamics of a system of particles. Conservation laws

1.6.1 Dynamics of a system of particles

We shall study in this section the motion of a system of N particles of mass m_i ($i = 1,\ldots,N$). Let us denote by $\mathbf{r}_i$ the coordinates of the i-th particle in a certain inertial system, and define $\mathbf{F}_{ij}$ as the force exerted by particle j on particle i (in particular, $\mathbf{F}_{ii} = 0$) and $\mathbf{F}_i^{(\mathrm{e})}$ as the external force acting on the i-th particle. Newton's second law of motion applied to the i-th particle of the system then states that

$$m_i\ddot{\mathbf{r}}_i = \sum_{j=1}^N \mathbf{F}_{ij} + \mathbf{F}_i^{(\mathrm{e})}\,, \qquad i = 1,\ldots,N\,, \tag{1.74}$$

where the sum over j on the RHS represents the internal force exerted on particle i by the remaining particles of the system. These N *vector* equations are actually a system of $3N$ *scalar* second-order ordinary differential equations in the unknowns $\mathbf{r}_1, \dots, \mathbf{r}_N$. By the existence and uniqueness theorem for such systems, if the RHS of Eq. (1.74) is a function of class C^1 in the variables $(t, \mathbf{r}_1, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N)$ in a certain open subset $A \subset \mathbb{R}^{6N+1}$ then the system's equations of motion (1.74) have (locally) a *unique* solution verifying any initial condition of the form

$$\mathbf{r}_1(t_0) = \mathbf{r}_{10}\,, \dots, \mathbf{r}_N(t_0) = \mathbf{r}_{N0}\,; \quad \dot{\mathbf{r}}_1(t_0) = \mathbf{v}_{10}\,, \dots, \dot{\mathbf{r}}_N(t_0) = \mathbf{v}_{N0}$$

with $(t_0, \mathbf{r}_{10}, \dots, \mathbf{r}_{N0}, \mathbf{v}_{10}, \dots, \mathbf{v}_{N0}) \in A$. In other words, *the trajectories of all the particles in the system are determined by their positions and velocities at any instant.* In this sense, Newtonian mechanics is a *completely deterministic* theory.

Summing over i in Eq. (1.74) we obtain

$$\sum_{i=1}^{N} m_i \ddot{\mathbf{r}}_i = \sum_{i,j=1}^{N} \mathbf{F}_{ij} + \sum_{i=1}^{N} \mathbf{F}_i^{(e)}\,. \tag{1.75}$$

If —as we shall assume throughout this section— *Newton's third law* holds, the internal forces verify the condition

$$\mathbf{F}_{ij} + \mathbf{F}_{ji} = 0\,,$$

which summed over i, j immediately yields

$$0 = \sum_{i,j=1}^{N} \mathbf{F}_{ij} + \sum_{i,j=1}^{N} \mathbf{F}_{ji} = 2 \sum_{i,j=1}^{N} \mathbf{F}_{ij}\,,$$

where in the last step we have used the fact that the summation indices (i, j) are dummy (i.e., $\sum_{i,j=1}^{N} \mathbf{F}_{ji} = \sum_{i,j=1}^{N} \mathbf{F}_{ij}$). Denoting by

$$\boxed{\mathbf{F}^{(e)} := \sum_{i=1}^{N} \mathbf{F}_i^{(e)}}$$

the **total external force** acting on the system, Eq. (1.75) can be written more concisely as

$$\boxed{\sum_{i=1}^{N} m_i \ddot{\mathbf{r}}_i = \mathbf{F}^{(e)}\,.} \tag{1.76}$$

Let us next define the system's **center of mass** as the point with coordinates

$$\boxed{\mathbf{R} := \frac{1}{M} \sum_{i=1}^{N} m_i \mathbf{r}_i\,,} \tag{1.77}$$

where

$$M = \sum_{i=1}^{N} m_i$$

is the total mass of the system. In other words, the center of mass (which is generally abbreviated by CM) is the average of the particles' coordinates weighted by their masses. In terms of the CM, Eq. (1.76) adopts the simple form

$$\boxed{M\ddot{\mathbf{R}} = \mathbf{F}^{(e)}\,.} \tag{1.78}$$

In other words:

The center of mass moves as a single particle of mass M on which the total *external* force acting on the system is exerted.

As a consequence, *the motion of the center of mass is not affected by the internal forces acting on the system.* In particular, if the total external force vanishes then $\ddot{\mathbf{R}} = 0$. Hence:

In the absence of external forces the center of mass moves with constant velocity.

1.6.2 Conservation laws

Consider first the system's total **momentum** $\mathbf{P}$, defined as the sum of the momenta of its constituent particles:

$$\mathbf{P} := \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i = M\dot{\mathbf{R}} \,. \tag{1.79}$$

Thus *the total momentum of the system coincides with the momentum of its center of mass regarded as a single particle of mass M.* It also follows from the latter equation that Eq. (1.78) is equivalent to

$$\dot{\mathbf{P}} = \mathbf{F}^{(e)} \,.$$

In particular, *in the absence of* external *forces the system's total momentum is* conserved.

Consider next the system's **angular momentum** with respect to the origin of coordinates, defined as the sum of the angular momenta of its N constituent particles:

$$\mathbf{L} := \sum_{i=1}^{N} m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i \,. \tag{1.80}$$

Let us denote by $\mathbf{r}_i'$ the position vector of the i-th particle with respect to the CM, so that

$$\mathbf{r}_i = \mathbf{R} + \mathbf{r}_i' \,. \tag{1.81}$$

Substituting the latter expression for $\mathbf{r}_i$ in Eq. (1.80) we obtain

$$\mathbf{L} = \sum_{i=1}^{N} m_i (\mathbf{R} + \mathbf{r}_i') \times (\dot{\mathbf{R}} + \dot{\mathbf{r}}_i') = M\mathbf{R} \times \dot{\mathbf{R}} + \mathbf{R} \times \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i' + \left(\sum_{i=1}^{N} m_i \mathbf{r}_i' \right) \times \dot{\mathbf{R}} + \sum_{i=1}^{N} m_i \mathbf{r}_i' \times \dot{\mathbf{r}}_i' \,. \tag{1.82}$$

On the other hand, from Eq. (1.81) and the definition (1.77) of the CM it easily follows that

$$\sum_{i=1}^{N} m_i \mathbf{r}_i = M\mathbf{R} = \sum_{i=1}^{N} m_i \mathbf{R} + \sum_{i=1}^{N} m_i \mathbf{r}_i' = M\mathbf{R} + \sum_{i=1}^{N} m_i \mathbf{r}_i' \quad \Longrightarrow \quad \boxed{\sum_{i=1}^{N} m_i \mathbf{r}_i' = 0 \,,} \tag{1.83}$$

and therefore the second and third terms in the RHS of Eq. (1.82) vanish identically. We thus have

$$\mathbf{L} = M\mathbf{R} \times \dot{\mathbf{R}} + \sum_{i=1}^{N} m_i \mathbf{r}_i' \times \dot{\mathbf{r}}_i' \,. \tag{1.84}$$

In other words, *the angular momentum of the system is the sum of the angular momentum of its center of mass and the* **internal angular momentum** (last term in the RHS of Eq. (1.84)) due to the motion of the particles around the CM.

Example 1.8. Suppose that the system moves as a whole with velocity (not necessarily uniform) $\mathbf{v}(t)$, i.e.,

$$\dot{\mathbf{r}}_i = \mathbf{v}(t)\,, \qquad i = 1, \dots, N\,.$$

In this case

$$\dot{\mathbf{R}} = \frac{1}{M}\sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i = \mathbf{v}(t)\,,$$

and thus the internal angular momentum vanishes:

$$\dot{\mathbf{r}}'_i = \dot{\mathbf{r}}_i - \dot{\mathbf{R}} = 0\,, \quad i = 1, \dots, N \quad \Longrightarrow \quad \mathbf{L} = M\mathbf{R}\times\dot{\mathbf{R}} = M\mathbf{R}\times\mathbf{v}\,.$$

By Eq. (1.80), the time derivative of the angular momentum is given by

$$\dot{\mathbf{L}} = \sum_{i=1}^{N} m_i \mathbf{r}_i \times \ddot{\mathbf{r}}_i = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{F}_i^{(\mathrm{e})} + \sum_{i,j=1}^{N} \mathbf{r}_i \times \mathbf{F}_{ij}\,.$$

Again, it is easy to check that the last term vanishes if Newton's third law holds in its *stronger version*, that is if

$$\mathbf{F}_{ji} = -\mathbf{F}_{ij} \parallel \mathbf{r}_i - \mathbf{r}_j\,, \qquad i \neq j\,. \tag{1.85}$$

Indeed,

$$0 = \sum_{i,j=1}^{N} (\mathbf{r}_i - \mathbf{r}_j)\times\mathbf{F}_{ij} = \sum_{i,j=1}^{N} \mathbf{r}_i\times\mathbf{F}_{ij} - \sum_{i,j=1}^{N} \mathbf{r}_j\times\mathbf{F}_{ij} = \sum_{i,j=1}^{N} \mathbf{r}_i\times\mathbf{F}_{ij} + \sum_{i,j=1}^{N} \mathbf{r}_j\times\mathbf{F}_{ji} = 2\sum_{i,j=1}^{N} \mathbf{r}_i\times\mathbf{F}_{ij}\,.$$

Thus in this case we have

$$\dot{\mathbf{L}} = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{F}_i^{(\mathrm{e})} =: \mathbf{N}^{(\mathrm{e})}\,, \tag{1.86}$$

where by definition $\mathbf{N}^{(\mathrm{e})}$ is the **total torque of the external forces** acting on the system. In other words:

If Newton's third law holds in its stronger version (1.85) then the time derivative of the system's angular momentum is equal to the total torque of the *external* forces acting on it. In particular if the total torque of the *external* forces acting on the system vanishes its angular momentum is conserved.

Exercise. Show that in general the total torque of the external forces is different from the torque of the total external force with respect to the CM.

Solution. By definition , the torque of the total external force with respect to the CM is the vector $\mathbf{R}\times\mathbf{F}^{(\mathrm{e})}$. Taking into account the definitions of $\mathbf{R}$ and $\mathbf{F}^{\mathrm{e}}$ we easily obtain

$$\mathbf{R}\times\mathbf{F}^{(\mathrm{e})} = \frac{1}{M}\left(\sum_{i=1}^{N} m_i\mathbf{r}_i\right)\times\left(\sum_{j=1}^{N} F_j^{(\mathrm{e})}\right) = \sum_{i,j=1}^{N} \frac{m_i}{M}\mathbf{r}_i\times\mathbf{F}_j^{(\mathrm{e})},$$

which is different in general than $\sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{F}_i^{(e)}$. For example, for two particles we have

$$\mathbf{R} \times \mathbf{F}^{(e)} - \sum_{i=1}^{2} \mathbf{r}_i \times \mathbf{F}_i^{(e)} = \sum_{i,j=1}^{2} \mathbf{r}_i \times \left(\frac{m_i}{M} \sum_{j=1}^{2} \mathbf{F}_j^{(e)} - \mathbf{F}_i^{(e)} \right) = \frac{\mathbf{r}_1}{M} \times \left(m_1 \mathbf{F}_2^{(e)} - m_2 \mathbf{F}_1^{(e)} \right) + \frac{\mathbf{r}_2}{M} \times \left(m_2 \mathbf{F}_1^{(e)} - m_1 \mathbf{F}_2^{(e)} \right) = \frac{\mathbf{r}_1 - \mathbf{r}_2}{M} \times \left(m_1 \mathbf{F}_2^{(e)} - m_2 \mathbf{F}_1^{(e)} \right),$$

which does not vanish unless

$$m_1 \mathbf{F}_2^{(e)} - m_2 \mathbf{F}_1^{(e)} \parallel \mathbf{r}_1 - \mathbf{r}_2 \,.$$

A common situation, however, in which $\mathbf{R} \times \mathbf{F}^{(e)}$ is equal to the total torque of the external forces arises when all the particles have the same mass and the total external force $\mathbf{F}_i^{(e)}$ is independent of i. For example, this will be the case if the particles move in a constant gravitational field, or if they move in a constant electric field and they have the same charge. Indeed, if $m_i = m$ and $\mathbf{F}_i^{(e)} = \mathbf{F}$ for all i we have $M = Nm$, $\mathbf{F}^{(e)} = N\mathbf{F}$, and therefore

$$\mathbf{R} \times \mathbf{F}^{(e)} = \sum_{i=1}^{N} \frac{m}{Nm} \mathbf{r}_i \times (N\mathbf{F}) = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{F} = \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{F}_i^{(e)}.$$

Let us next study the **kinetic energy** of the system, defined as

$$T = \frac{1}{2} \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i^2 \,. \tag{1.87}$$

Using again the decomposition (1.81) and the identity (1.83) we easily obtain:

$$T = \frac{1}{2} M \dot{\mathbf{R}}^2 + \frac{1}{2} \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i'^2 \,, \tag{1.88}$$

where the last term is the **internal kinetic energy**, due to the motion of the particles with respect to the CM. Hence *the kinetic energy of the system is the sum of the kinetic energy of its CM and the internal kinetic energy.*

We shall say that the forces acting on the system are **conservative** —or, equivalently, that the system itself is conservative— if there is a *single* scalar function $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$ such that

$$\mathbf{F}_i := \mathbf{F}_i^{(e)} + \sum_{j=1}^{N} \mathbf{F}_{ij} = -\frac{\partial V}{\partial \mathbf{r}_i} \,, \qquad i = 1, \dots, N \,. \tag{1.89}$$

As in the case of a single particle addressed in Section 1.4.2, *if the forces acting on the system are conservative the system's* ***total energy***

$$E = T + V(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

is conserved. Indeed, if the forces acting on the system are conservative we have

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i \cdot \ddot{\mathbf{r}}_i + \frac{\mathrm{d}V}{\mathrm{d}t} = \sum_{i=1}^{N} \dot{\mathbf{r}}_i \cdot \mathbf{F}_i + \sum_{i=1}^{N} \frac{\partial V}{\partial \mathbf{r}_i} \dot{\mathbf{r}}_i = \sum_{i=1}^{N} \left(\mathbf{F}_i + \frac{\partial V}{\partial \mathbf{r}_i} \right) \dot{\mathbf{r}}_i = 0 \,.$$

Let us assume that there exist certain functions $V_i(\mathbf{r}_i)$, $V_{ij}(\mathbf{r}_i, \mathbf{r}_j)$ (with $i \neq j$, $1 \leqslant i, j \leqslant N$) such that $V_{ij}(\mathbf{r}_i, \mathbf{r}_j) = V_{ji}(\mathbf{r}_j, \mathbf{r}_i)$ (i.e., $V_{ij}(\mathbf{r}_i, \mathbf{r}_j)$ is symmetric under exchange of the particles i and j) and

$$\mathbf{F}_i^{(e)} = -\frac{\partial V_i(\mathbf{r}_i)}{\partial \mathbf{r}_i}, \qquad \mathbf{F}_{ij} = -\frac{\partial V_{ij}(\mathbf{r}_i, \mathbf{r}_j)}{\partial \mathbf{r}_i}, \qquad 1 \leqslant i \neq j \leqslant N. \tag{1.90a}$$

We shall then show that the system is conservative, with potential (up to an additive constant)

$$V = \sum_{i=1}^{N} V_i(\mathbf{r}_i) + \sum_{1 \leqslant i < j \leqslant N} V_{ij}(\mathbf{r}_i, \mathbf{r}_j). \tag{1.90b}$$

Indeed, it is easily verified that Eqs. (1.90) imply the more general relation (1.89):

$$\begin{aligned}\mathbf{F}_i + \frac{\partial V}{\partial \mathbf{r}_i} &= \sum_{j=1}^{N} \mathbf{F}_{ij} + \frac{\partial}{\partial \mathbf{r}_i} \sum_{1 \leqslant j < k \leqslant N} V_{jk}(\mathbf{r}_j, \mathbf{r}_k) = \sum_{j=1}^{N} \mathbf{F}_{ij} + \sum_{k=i+1}^{N} \frac{\partial}{\partial \mathbf{r}_i} V_{ik}(\mathbf{r}_i, \mathbf{r}_k) + \sum_{j=1}^{i-1} \frac{\partial}{\partial \mathbf{r}_i} V_{ji}(\mathbf{r}_j, \mathbf{r}_i) \\ &= \sum_{j=1}^{N} \mathbf{F}_{ij} + \sum_{j=i+1}^{N} \frac{\partial}{\partial \mathbf{r}_i} V_{ij}(\mathbf{r}_i, \mathbf{r}_j) + \sum_{j=1}^{i-1} \frac{\partial}{\partial \mathbf{r}_i} V_{ij}(\mathbf{r}_i, \mathbf{r}_j) = \sum_{\substack{j=1 \\ j \neq i}}^{N} \left(\mathbf{F}_{ij} + \frac{\partial V_{ij}}{\partial \mathbf{r}_i} \right) = 0,\end{aligned}$$

where we have used the identities $\mathbf{F}_{ii} = 0$ and $V_{ji}(\mathbf{r}_j, \mathbf{r}_i) = V_{ij}(\mathbf{r}_i, \mathbf{r}_j)$.

- By Newton's third law we must have

$$\mathbf{F}_{ji} = -\frac{\partial V_{ji}(\mathbf{r}_j, \mathbf{r}_i)}{\partial \mathbf{r}_j} = -\frac{\partial V_{ij}(\mathbf{r}_i, \mathbf{r}_j)}{\partial \mathbf{r}_j} = -\mathbf{F}_{ij} = \frac{\partial V_{ij}(\mathbf{r}_i, \mathbf{r}_j)}{\partial \mathbf{r}_i},$$

so that the function $V_{ij}(\mathbf{r}_i, \mathbf{r}_j)$ should verify the system of partial differential equations

$$\frac{\partial V_{ij}}{\partial \mathbf{r}_i} + \frac{\partial V_{ij}}{\partial \mathbf{r}_j} = 0.$$

It can be shown (exercise) that the general solution of this system is an arbitrary function of the difference $\mathbf{r}_i - \mathbf{r}_j$. We thus have

$$V_{ij}(\mathbf{r}_i, \mathbf{r}_j) = U_{ij}(\mathbf{r}_i - \mathbf{r}_j), \qquad \text{with } U_{ji}(\mathbf{r}) = U_{ij}(-\mathbf{r}). \tag{1.91}$$

Substituting in (1.90b) we obtain the following more explicit formula for the potential V:

$$V = \sum_{i=1}^{N} V_i(\mathbf{r}_i) + \sum_{1 \leqslant i < j \leqslant N} U_{ij}(\mathbf{r}_i - \mathbf{r}_j). \tag{1.92}$$

In fact, in most conservative physical systems of interest the potential is of the form (1.92).

Exercise. Study under what conditions on the functions U_{ij} Newton's third law holds in its stronger version (1.85).

Solution. If $\mathbf{r} := \mathbf{r}_i - \mathbf{r}_j$, Eq. (1.85) will hold if and only if

$$0 = \mathbf{r} \times \mathbf{F}_{ij} = -\mathbf{r} \times \frac{\partial U_{ij}(\mathbf{r}_i - \mathbf{r}_j)}{\partial \mathbf{r}_i} = -\mathbf{r} \times \frac{\partial U_{ij}(\mathbf{r})}{\partial \mathbf{r}},$$

i.e., if the gradient of $U_{ij}(\mathbf{r})$ has only a radial component. By the formula for the gradient in spherical coordinates

$$\frac{\partial U_{ij}}{\partial \mathbf{r}} = \frac{\partial U_{ij}}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial U_{ij}}{\partial \theta}\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial U_{ij}}{\partial \varphi}\mathbf{e}_\varphi,$$

this will be the case if and only if U_{ij} depends only on $r = |\mathbf{r}_i - \mathbf{r}_j|$, i.e., if there is a function of one variable u_{ij} such that $U_{ij} = u_{ij}(|\mathbf{r}_i - \mathbf{r}_j|)$ (with $u_{ij} = u_{ji}$, in view of Eq. (1.91)). This condition is in fact satisfied by most conservative forces, like the gravitational or the electrostatic ones (see next exercise). When it holds, Eq. (1.92) reads

$$V = \sum_{i=1}^{N} V_i(\mathbf{r}_i) + \sum_{1\leqslant i<j\leqslant N} u_{ij}(|\mathbf{r}_i - \mathbf{r}_j|). \tag{1.93}$$

Exercise. Write down the potential for a system of charged particles of mass m_i and charge q_i ($i = 1, \dots, N$) moving in an external electric field generated by an electrostatic potential $\Phi(\mathbf{r})$.

Solution. The external force acting on the i-th particle is due to its interaction with the electric field $\mathbf{E}(\mathbf{r}) = -\dfrac{\partial \Phi(\mathbf{r})}{\partial \mathbf{r}}$ generated by the electrostatic potential $\Phi(\mathbf{r})$, namely

$$\mathbf{F}_i^{(\mathrm{e})} = q_i \mathbf{E}(\mathbf{r}_i) = -q_i \frac{\partial \Phi}{\partial \mathbf{r}}(\mathbf{r}_i) = -q_i \frac{\partial \Phi(\mathbf{r}_i)}{\partial \mathbf{r}_i} = -\frac{\partial V_i(\mathbf{r}_i)}{\partial \mathbf{r}_i}, \qquad \text{with} \quad V_i(\mathbf{r}_i) = q_i \Phi(\mathbf{r}_i).$$

On the other hand, the force exerted by particle j on particle i is the sum of the electric and gravitational forces between both particles, given by

$$\mathbf{F}_{ij} = (kq_iq_j - Gm_im_j)\frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} = -\frac{\partial}{\partial \mathbf{r}_i}\frac{kq_iq_j - Gm_im_j}{|\mathbf{r}_i - \mathbf{r}_j|} = -\frac{\partial u_{ij}(|\mathbf{r}_i - \mathbf{r}_j|)}{\partial \mathbf{r}_i},$$

with $k = 1/(4\pi\varepsilon_0)$ and

$$u_{ij}(r) = \frac{kq_iq_j - Gm_im_j}{r} = u_{ji}(r).$$

The system is thus conservative, with potential

$$V = \sum_{i=1}^{N} q_i \Phi(\mathbf{r}_i) + \sum_{1\leqslant i<j\leqslant N} \frac{kq_iq_j - Gm_im_j}{|\mathbf{r}_i - \mathbf{r}_j|}.$$

In practice, the electrostatic coupling constant kq_iq_j is usually much greater than the gravitational one Gm_im_j. For instance, for protons

$$q_i = q_j = 1.602\,176\,634 \cdot 10^{-19}\,\mathrm{C}, \qquad m_i = m_j = 1.672\,621\,923\,69(51) \cdot 10^{-27}\,\mathrm{kg},$$

so that

$$\frac{Gm_i m_j}{kq_i q_j} \simeq 8.09355 \cdot 10^{-37}.$$

For this reason, the gravitational interaction between the charges is usually neglected, and the potential reduces accordingly to

$$\boxed{V = \sum_{i=1}^{N} q_i \Phi(\mathbf{r}_i) + \frac{1}{4\pi\varepsilon_0} \sum_{1\leqslant i<j\leqslant N} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|}}.$$

2 Motion in a central potential

2.1 Two-body problem. Reduction to the equivalent one-body problem

We shall study in this section the motion of two point masses m_1 and m_2 not subject to any external forces. If $\mathbf{F}_{12}$ denotes the force exerted by the second particle on the first, by Newton's third law the first particle exerts a force $-\mathbf{F}_{12}$ on the second one, and the system's equations of motion are thus

$$\begin{aligned} m_1\ddot{\mathbf{r}}_1 &= \mathbf{F}_{12}(t,\mathbf{r}_1,\mathbf{r}_2,\dot{\mathbf{r}}_1,\dot{\mathbf{r}}_2) \\ m_2\ddot{\mathbf{r}}_2 &= -\mathbf{F}_{12}(t,\mathbf{r}_1,\mathbf{r}_2,\dot{\mathbf{r}}_1,\dot{\mathbf{r}}_2)\,. \end{aligned} \tag{2.1}$$

It is convenient to rewrite these equations in terms of the variables

$$\mathbf{R} = \frac{1}{M}(m_1\mathbf{r}_1 + m_2\mathbf{r}_2)\,, \qquad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{2.2}$$

(cf. Fig. 2.1), where $M := m_1 + m_2$ denotes the system's total mass. Solving for $\mathbf{r}_1$ and $\mathbf{r}_2$ in the previous equations we easily obtain the inverse relations

$$\mathbf{r}_1 = \mathbf{R} + \frac{m_2}{M}\mathbf{r}\,, \qquad \mathbf{r}_2 = \mathbf{R} - \frac{m_1}{M}\mathbf{r}\,. \tag{2.3}$$

As we saw in Section 1.6.1, since there are no external forces the center of mass $\mathbf{R}$ moves without acceleration, i.e., $\ddot{\mathbf{R}} = 0$. As to the **relative coordinate** $\mathbf{r}$, using Eqs. (2.1) and (2.3) we immediately obtain

$$\mu\ddot{\mathbf{r}} = \mathbf{F}_{12}\left(t,\mathbf{R} + \frac{m_2}{M}\mathbf{r},\mathbf{R} - \frac{m_1}{M}\mathbf{r},\dot{\mathbf{R}} + \frac{m_2}{M}\dot{\mathbf{r}},\dot{\mathbf{R}} - \frac{m_1}{M}\dot{\mathbf{r}}\right), \tag{2.4}$$

where

$$\mu := \frac{m_1 m_2}{m_1 + m_2} \tag{2.5}$$

is the so called **reduced mass** of the system. Note that this is a second-order differential equation in the single (vector) variable $\mathbf{r}$, since the motion of the center of mass is known ($\mathbf{R} = \mathbf{R}_0 + \mathbf{V}_0 t$,

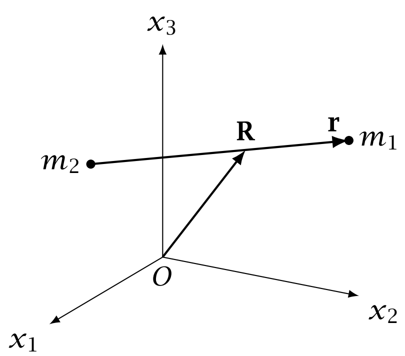


Figure 2.1. Coordinates $\mathbf{R}$ and $\mathbf{r}$ in the two-body problem.

with $\mathbf{R}_0$ and $\mathbf{V}_0$ constant vectors determined by the initial conditions). Hence:

> The two-body problem (2.1) is always equivalent to the one-body problem (2.4).

If (as is usually the case) the internal force satisfies the condition[1]

$$\mathbf{F}_{12} = \mathbf{F}(t, \mathbf{r}_1 - \mathbf{r}_2, \dot{\mathbf{r}}_1 - \dot{\mathbf{r}}_2)\,, \tag{2.6}$$

that is, if it depends only on the particles' *relative* coordinates and velocities, the equation of motion of the relative coordinate $\mathbf{r}$ reduces to

$$\boxed{\mu\ddot{\mathbf{r}} = \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})\,.} \tag{2.7}$$

In other words, if the force $\mathbf{F}_{12}$ is of the form (2.6) the relative coordinate $\mathbf{r}$ moves as a particle of mass μ under the force $\mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$, and the two-body problem (2.1) reduces to the one-body problem (2.7).

Once Eq. (2.7) is solved, the motion of the coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$ is easily found using Eqs. (2.3). Since $\ddot{\mathbf{R}} = 0$, if we move the origin to the center of mass the resulting reference frame, called **center of mass frame**, remains inertial. In the CM frame, Eqs. (2.3) simplify to

$$\mathbf{r}_1 = \frac{m_2}{M}\mathbf{r}\,, \qquad \mathbf{r}_2 = -\frac{m_1}{M}\mathbf{r}\,.$$

In many applications, the mass m_2 is much larger than m_1. In this case $m_1/M \simeq 0$, $m_2/M \simeq 1$, and thus (in the CM frame)

$$\mathbf{r}_1 \simeq \mathbf{r}\,, \qquad \mathbf{r}_2 \simeq 0\,.$$

In other words, in this case the heavy particle is approximately fixed at the origin (i.e., the CM), and the relative coordinate $\mathbf{r}$ is approximately equal to the radius vector of the light particle.

2.2 Constants of motion. Law of motion and equation of the trajectory. Bounded Orbits

2.2.1 Constants of motion

The most important example of force satisfying condition (2.6) is that of a *central force* of the form

$$\boxed{\mathbf{F}_{12} = f(|\mathbf{r}_1 - \mathbf{r}_2|)\,\frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|}\,.}$$

In this case Eq. (2.7) reduces to

$$\boxed{\mu\,\ddot{\mathbf{r}} = f(r)\,\frac{\mathbf{r}}{r}\,,} \tag{2.8}$$

which is the equation of motion of a particle of mass μ subject to the central force

$$\mathbf{F}(\mathbf{r}) = f(r)\,\frac{\mathbf{r}}{r}\,. \tag{2.9}$$

We shall study in this section how to find the general solution of Eq. (2.8), and analyze the qualitative behavior of its trajectories.

---

[1]The dependence of $\mathbf{F}$ only on the *relative* coordinates and velocities is very natural, since it is clearly related to the *homogeneity* of space. Note that the the homogeneity of time would also require that $\mathbf{F}$ be time-independent, which is almost always the case.

As we saw in Section 1.4.2, the force (2.9) is *conservative*, since

$$\mathbf{F}(\mathbf{r}) = -\frac{\partial V(r)}{\partial \mathbf{r}}\,, \qquad \text{with} \qquad \boxed{V(r) = -\int f(r)\,\mathrm{d}r\,.}$$

Thus the total energy is conserved, i.e.,

$$\boxed{\frac{1}{2}\mu\dot{\mathbf{r}}^2 + V(r) = E}$$

remains constant throughout the motion. Moreover, since the force (2.9) is *central* the angular momentum $\mathbf{L} = \mu\mathbf{r}\times\dot{\mathbf{r}}$ is also conserved. If, as we shall assume from now on,

$$\mathbf{L} \neq 0,$$

the motion takes place in the plane perpendicular to $\mathbf{L}$ *passing through the origin* (i.e., the center of force)[2]. We shall choose the z axis in the direction of the constant vector $\mathbf{L}$, so that

$$\boxed{\mathbf{L} = L\mathbf{e}_z\,, \qquad \text{with} \quad L = |\mathbf{L}| > 0.} \tag{2.10}$$

In particular, with this choice of axes *the motion takes place in the* $z = 0$ *plane.* Let us introduce polar coordinates (r, φ) in the latter plane through the usual equations

$$\mathbf{r} = r(\cos\varphi, \sin\varphi, 0) \qquad (r > 0\,, \quad 0 \leqslant \varphi < 2\pi)\,.$$

The unit coordinate vectors are

$$\mathbf{e}_r = (\cos\varphi, \sin\varphi, 0) = \frac{\mathbf{r}}{r}\,, \qquad \mathbf{e}_\varphi = (-\sin\varphi, \cos\varphi, 0)\,,$$

and hence

$$\dot{\mathbf{e}}_r = \dot{\varphi}\mathbf{e}_\varphi\,, \qquad \dot{\mathbf{e}}_\varphi = -\dot{\varphi}\mathbf{e}_r\,,$$

whence we easily obtain the formulas (cf. Example 1.2.3)

$$v_r = \dot{r}\,, \quad v_\varphi = r\dot{\varphi}\,; \qquad a_r = \ddot{r} - r\dot{\varphi}^2\,, \quad a_\varphi = r\ddot{\varphi} + 2\dot{r}\dot{\varphi}\,.$$

The equations of motion in polar coordinates are therefore

$$\boxed{\begin{aligned} \ddot{r} - r\dot{\varphi}^2 &= \frac{f(r)}{\mu} \\ r\ddot{\varphi} + 2\dot{r}\dot{\varphi} &= 0\,. \end{aligned}} \tag{2.11}$$

[2] If $\mathbf{L} = 0$ the trajectory is a straight line passing through the origin, which is actually a particular (degenerate) case of motion on a plane through the origin. Indeed, if $\mathbf{r}$ is not identically zero (degenerate case of motion along a line through the origin) the velocity vector must be parallel to $\mathbf{r}$ at all times, and consequently $v_\theta = r\dot{\theta}$ and $v_\varphi = r\sin\theta\,\dot{\varphi}$ vanish identically. Therefore the angles θ and φ are constant, and the trajectory lies on the straight line passing through the origin in the direction of the *constant* vector $\mathbf{e}_r$ (including the degenerate case in which $\dot{\mathbf{r}}$ is identically zero and the particle is at rest on a point of the latter line). Taking the line of the motion as the x axis we have $r = |x|$ and $\mathbf{e}_r = \operatorname{sgn} x\,\mathbf{i}$, so that $\mathbf{F}(\mathbf{r}) = \operatorname{sgn} x\, f(|x|)\,\mathbf{i}$ and $V(r) = V(|x|)$. Thus when $\mathbf{L} = 0$ the motion takes place in one dimension under the conservative force $F(x) = \operatorname{sgn} x\, f(|x|)$ —or, equivalently, the potential $V(|x|)$—, a problem studied in the previous chapter.

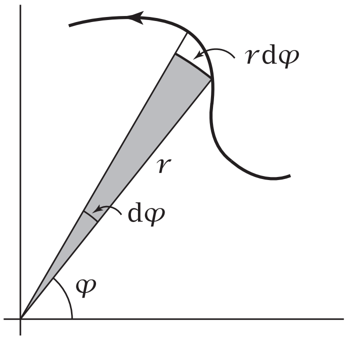


Figure 2.2. Infinitesimal area swept by the position vector **r**.

2.2.2 Law of motion and equation of the trajectory

In order to determine the **trajectory** described by the particle (i.e., r as a function of φ or vice versa) and the law of motion (i.e., r and φ as functions of t) it is easier to use the laws of conservation of energy and angular momentum, as we shall next see. Indeed, the angular momentum of the particle is given by

$$\mathbf{L} = L\mathbf{e}_z = \mu\mathbf{r}\times\dot{\mathbf{r}} = \mu r\mathbf{e}_r \times (\dot{r}\mathbf{e}_r + r\dot{\varphi}\mathbf{e}_\varphi) = \mu r^2\dot{\varphi}\mathbf{e}_z\,,$$

so that

$$\boxed{\mu r^2\dot{\varphi} = L > 0\,.} \tag{2.12}$$

From the previous equation (or, more precisely, our choice of the z axis in the direction of **L**) it follows that

$$\boxed{\dot{\varphi}(t) > 0\,, \qquad \forall t\,;}$$

in particular, if the trajectory surrounds the origin it must be traversed in an *anticlockwise* direction. Note also that the second equation of motion (2.11) is just the time derivative of Eq. (2.12) (divided by μr).

An immediate consequence of the conservation of angular momentum is the so-called *law of areas*, first formulated by Johannes Kepler in the early 17th century. Indeed, note that the area $A(\varphi)$ swept by the particle's position vector when moving between two points on its trajectory with polar coordinates $(r(\varphi_0), \varphi_0)$ and $(r(\varphi), \varphi)$ is given by

$$A(\varphi) = \frac{1}{2}\int_{\varphi_0}^{\varphi} r^2(\alpha)\,\mathrm{d}\alpha$$

(cf. Fig. 2.2). Since

$$\dot{A} = \frac{\mathrm{d}A}{\mathrm{d}\varphi}\,\dot{\varphi} = \frac{1}{2}\,r^2\dot{\varphi} = \frac{L}{2\mu} \tag{2.13}$$

is constant, the area ΔA swept over a time Δt is simply

$$\boxed{\Delta A = \frac{L}{2\mu}\,\Delta t\,.}$$

In other words, *the particle sweeps out equal areas in equal times* as it moves along its trajectory (**law of areas**). Note that this property is valid for *any* central force $f(t,\mathbf{r},\dot{\mathbf{r}})\,\mathbf{e}_r$, more general than (2.9), since it only requires the conservation of angular momentum.

We can use the conservation law of angular momentum to express $\dot{\varphi}$ in terms of r and the angular momentum L as

$$\boxed{\dot{\varphi} = \frac{L}{\mu r^2}\,.} \tag{2.14}$$

Substituting this expression into the energy conservation equation we immediately obtain

$$E = \frac{1}{2}\mu\dot{\mathbf{r}}^2 + V(r) = \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\mu r^2\dot{\varphi}^2 + V(r) = \frac{1}{2}\mu\dot{r}^2 + \frac{L^2}{2\mu r^2} + V(r)\,, \tag{2.15}$$

or equivalently,

$$\boxed{\frac{1}{2}\mu\dot{r}^2 + U_L(r) = E\,,} \tag{2.16}$$

where the **effective potential** $U_L(r)$ is defined by

$$\boxed{U_L(r) := V(r) + \frac{L^2}{2\mu r^2}\,.} \tag{2.17}$$

Although U_L *depends on the* (*constant*) *value of the angular momentum*, we shall from now on adhere to the customary practice of dropping the subindex L and simply writing U instead of U_L. Note also that the effective force generated by the last term in $U(r)$, namely

$$-\frac{\partial}{\partial\mathbf{r}}\left(\frac{L^2}{2\mu r^2}\right) = -\frac{\partial}{\partial r}\left(\frac{L^2}{2\mu r^2}\right)\mathbf{e}_r = \frac{L^2}{\mu r^3}\,\mathbf{e}_r = \mu r\dot{\varphi}^2\,\mathbf{e}_r = \frac{\mu v_\varphi^2}{r}\,\mathbf{e}_r\,,$$

can be interpreted as a *centrifugal force*. Again, it is easy to see that the first equation of motion (2.11) is the time derivative of Eq. (2.15) (divided by $\mu\dot{r}$). In other words, *the laws of conservation of energy and angular momentum are obtained by integrating once with respect to time t the equations of motion* (2.11).

Equations (2.12)-(2.16) easily yield the law of motion and the equation of the trajectory *in implicit form*. Indeed, the law of motion is directly obtained by integrating Eq. (2.16) (after separating variables) and using the conservation of angular momentum:

$$\boxed{t = \pm\sqrt{\frac{\mu}{2}}\int\frac{\mathrm{d}r}{\sqrt{E-U(r)}}\,, \qquad \varphi = \frac{L}{\mu}\int\frac{\mathrm{d}t}{r^2(t)}\,,} \tag{2.18}$$

where it is understood that in the second equation we must substitute the value of $r(t)$ obtained from the first one. (We shall see later, however, that these equations are almost never the easiest way of finding the law of motion.)

As to the equation of the trajectory, from Eq. (2.16) it follows that

$$\frac{\mu}{2}\,\dot{\varphi}^2 r'^2(\varphi) + U(r) = E\,,$$

where (as we shall do throughout this section) we have denoted by a prime the derivative with respect to the angle φ. Using Eq. (2.14) to express $\dot{\varphi}$ as a function of r we obtain

$$\frac{L^2}{2\mu}\left(\frac{r'}{r^2}\right)^2 + U(r) = E\,.$$

Introducing the dependent variable

$$\boxed{u = \frac{1}{r}\,,}$$

we can rewrite the previous equation as

$$u'^2 = \frac{2\mu}{L^2}\,(E - U(1/u))\,, \tag{2.19}$$

whose integration yields the equation of the trajectory:

$$\varphi = \pm\frac{L}{\sqrt{2\mu}}\int^{1/r}\frac{\mathrm{d}u}{\sqrt{E - U(1/u)}}\,. \tag{2.20}$$

In practice, to find the equation of the trajectory it is sometimes convenient to differentiate Eq. (2.19) with respect to φ, thus obtaining a second-order equation that is often easier to integrate than (2.19). Indeed, proceeding in this way we obtain

$$2u'u'' = \frac{2\mu}{L^2u^2}\frac{\mathrm{d}U}{\mathrm{d}r}\,u'\,, \tag{2.21}$$

and therefore, taking into account the definition (2.17) of U,

$$u'' = \frac{\mu}{L^2u^2}\frac{\mathrm{d}U}{\mathrm{d}r} = \frac{\mu}{L^2u^2}\left(\frac{\mathrm{d}V}{\mathrm{d}r} - \frac{L^2}{\mu}u^3\right),$$

or equivalently,

$$u'' + u = -\frac{\mu}{L^2u^2}\,f(1/u)\,. \tag{2.22}$$

The latter equation is known as **Binet's equation**. Binet's equation, written as

$$f(r) = -\frac{L^2}{\mu r^2}(u'' + u)\,,$$

is often used to compute the central force law $f(r)$ if the equation of the trajectory $r = r(\varphi)$ is known.

- In general, *if the equation of the trajectory is known it is possible to find the law of motion* implicitly. Indeed, suppose that the equation of the trajectory is $r = r(\varphi)$. From the conservation of angular momentum we obtain

$$t = \frac{\mu}{L}\int r^2(\varphi)\,\mathrm{d}\varphi\,, \tag{2.23}$$

which yields t as a function of φ. Inverting this relation we obtain $\varphi(t)$, while the motion of the radial coordinate can of course be found by substituting $\varphi(t)$ in the equation of the trajectory:

$$r = r(\varphi(t))\,.$$

Remark. Binet's equation actually holds for a general central force $f(r,\varphi)\mathbf{e}_r$ in two dimensions, even if energy is not conserved unless f is independent of the polar angle φ. To derive Binet's equation in this more general setting we start from the equation of motion in the radial direction

$$\ddot{r} - r\dot{\varphi}^2 = \frac{f(r,\varphi)}{\mu}\,.$$

Since angular momentum is still conserved, we can replace $r\dot{\varphi}^2$ in the latter equation by $L^2/(\mu^2r^3)$. Using this fact and the identity

$$\frac{\mathrm{d}}{\mathrm{d}t} = \dot{\varphi}\,\frac{\mathrm{d}}{\mathrm{d}\varphi} = \frac{L}{\mu r^2}\,\frac{\mathrm{d}}{\mathrm{d}\varphi}$$

we obtain

$$\frac{L}{\mu r^2}\frac{\mathrm{d}}{\mathrm{d}\varphi}\left(\frac{L}{\mu r^2}r'\right) - \frac{L^2}{\mu^2r^3} = -\frac{L^2u^2}{\mu^2}\,u'' - \frac{L^2u^3}{\mu^2} = \frac{f(r,\varphi)}{\mu} \quad\Longrightarrow\quad u'' + u = -\frac{\mu}{L^2u^2}\,f(1/u,\varphi)\,.$$

Example 2.1. Let us find the equation of the trajectories of a particle of mass μ subject to the central force

$$\boxed{\mathbf{F} = \frac{k}{r^3}\,\mathbf{e}_r\,.}$$

In this case $f(1/u) = ku^3$, and thus Binet's equation reduces to

$$\boxed{u'' + Cu = 0\,, \qquad C := 1 + \frac{k\mu}{L^2}\,.}$$

The solutions of this equation depend on the sign of the dimensionless constant C. More precisely:

I) $C < 0$

In this case —which can only happen if $k < 0$, i.e., if the force is *attractive*— the general solution of Binet's equation is

$$u = a\,\mathrm{e}^{\gamma\varphi} + b\,\mathrm{e}^{-\gamma\varphi} \quad \iff \quad r = \left(a\,\mathrm{e}^{\gamma\varphi} + b\,\mathrm{e}^{-\gamma\varphi}\right)^{-1} \qquad (a, b \in \mathbb{R})\,.$$

with

$$\gamma := \sqrt{|C|} > 0\,.$$

It is easy to see that if a and b are both positive the solution can be expressed as

$$\boxed{r = A\,\mathrm{sech}\big(\gamma(\varphi - \varphi_0)\big) \qquad (A > 0)\,,} \tag{2.24}$$

if either a or b vanish then

$$\boxed{r = \mathrm{e}^{\pm\gamma(\varphi - \varphi_0)}\,,} \tag{2.25}$$

whereas if a and b have opposite signs we have[a]

$$\boxed{r = A\,\mathrm{csch}\big(\gamma(\varphi - \varphi_0)\big) \qquad (A \neq 0)\,.} \tag{2.26}$$

The trajectories (2.24) are *bounded* ($r \leqslant A$), while (2.25) and (2.26) are not (in the former case $r \to \infty$ for $\varphi \to \pm\infty$, while in the latter $r \to \infty$ for $\varphi \to \varphi_0$). It is also easy to check that all of these trajectories are of *spiral* type, since the angle φ can take arbitrarily large (positive or negative) values and r tends to 0 as $\varphi \to \pm\infty$.

More precisely, for (2.24) $\varphi \in \mathbb{R}$ and

$$\lim_{\varphi \to \pm\infty} r(\varphi) = 0\,.$$

Let $t_{\pm\infty}$ denote the times corresponding to $\varphi = \pm\infty$ (i.e., such that $\varphi(t_{\pm\infty}) = \pm\infty$). Then the trajectory (2.24) spirals away from the origin as $t \to t_{-\infty}$, reaches its point of minimum distance to the origin ($r = A$), and spirals back into the origin as $t \to t_\infty$. Likewise, for the trajectory (2.25) $\varphi \in \mathbb{R}$ and $r \to 0$ as $\varphi \to \infty$ (for the "$-$" sign in the exponential) or $\varphi \to -\infty$ (for the "$+$" sign). Hence in the first case the trajectory spirals into the origin as $t \to t_\infty$, while in the second one it spirals away from the origin as $t \to t_{-\infty}$. Finally, for the trajectory (2.26) we have $\varphi > \varphi_0$ for $A > 0$ and $\varphi < \varphi_0$ for $A < 0$, and $r \to 0$ as $\varphi \to \infty$ or $\varphi \to -\infty$, respectively. Thus for $A > 0$ the trajectory spirals into the origin for $t \to t_\infty$, while for $A < 0$ it spirals away from the origin as $t \to t_{-\infty}$. Moreover, in both cases $r \to \infty$ as $\varphi \to \varphi_0$, which is a necessary condition for the trajectory to have an asymptote making an angle φ_0 with the positive x axis. In fact, it can be shown that in this case the trajectory does have such an asymptote (see next exercise).

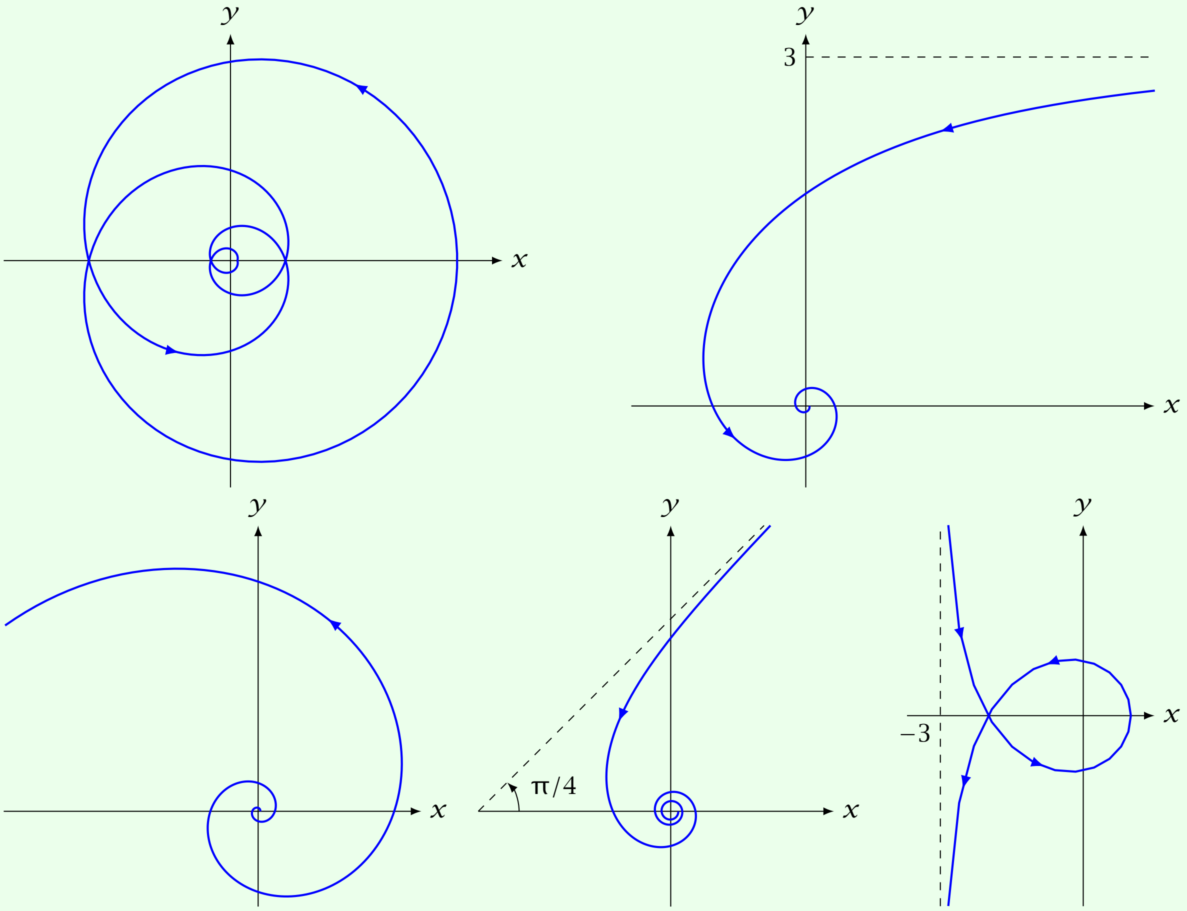


Figure 2.3. Trajectories of a particle in the central force field $\mathbf{F} = k\mathbf{e}_r/r^3$. From left to right and top to bottom, $r = \operatorname{sech}(\varphi/3)$, $r = \operatorname{csch}(\varphi/3)$, $r = e^{\varphi/3}$, $r = (\varphi - \frac{\pi}{4})^{-1}$, and $r = \sec(\varphi/3)$. (The orientation of the trajectories agrees with the convention used in this chapter, according to which $L_z > 0$ and hence $\dot{\varphi} > 0$.)

II) $C = 0$

In this case $k = -L^2/\mu < 0$, and the general solution of Binet's equation is

$$u = a + b\varphi \iff \boxed{r = \frac{1}{a + b\varphi}} \qquad (a, b \in \mathbb{R}), \tag{2.27}$$

so that $\varphi > -a/b$ for $b > 0$ and $\varphi < -a/b$ for $b < 0$. If $b = 0$ the trajectory is simply a *circle* centered at the origin. On the other hand, for $b \neq 0$ the trajectories are *unbounded* ($r \to \infty$ for $\varphi \to -a/b$) and spiraling towards the origin (when $b > 0$) or away from it (when $b < 0$), as $r \to 0$ when φ tends to ∞ or $-\infty$. Moreover, it can be shown that in this case the trajectory does have an asymptote making an angle $-a/b$ with the positive real axis (see next exercise).

III) $C > 0$

In this case (which takes place, in particular, if $k > 0$) the equation of the trajectories is

$$u = \frac{1}{A}\cos\big(\gamma(\varphi - \varphi_0)\big) \iff \boxed{r = A\sec\big(\gamma(\varphi - \varphi_0)\big)} \qquad (A > 0\,,\ 0 \leqslant \varphi_0 < 2\pi)\,, \tag{2.28}$$

so that (since $r > 0$) the angle φ can be taken (for instance) in the range $(\varphi_0 - \frac{\pi}{2\gamma}, \varphi_0 + \frac{\pi}{2\gamma})$. The trajectories are *not spiraling*, since $r \geqslant A > 0$, but are *unbounded* ($r \to \infty$ as $\varphi \to \varphi_0 \pm \frac{\pi}{2\gamma}$). In fact, it can be shown that in this case the trajectory has two asymptotes making an angle $\varphi_0 \pm \frac{\pi}{2\gamma}$ with the positive real axis.

Note. If $C \leqslant 0$ (and, in particular, $k < 0$), the trajectories in the previous example are generically called *Cotes spirals*.

[a] More precisely, in cases (2.24) and (2.26) the parameters A and φ_0 are given by

$$\varphi_0 = \frac{\log|b/a|}{2\gamma}, \qquad A = \frac{\operatorname{sgn} a}{2\sqrt{|ab|}}.$$

Exercise. Show that a plane curve with polar equation $r = f(\varphi)$ has an asymptote making an angle α with the positive x axis provided that

$$\lim_{\varphi\to\alpha} f(\varphi) = \infty, \qquad c := \lim_{\varphi\to\alpha} f(\varphi)\sin(\varphi - \alpha) < \infty,$$

and in that case the Cartesian equation of the asymptote is $\cos\alpha\, y - \sin\alpha\, x = c$. Using this result, prove that:
i) The trajectory (2.26) has an asymptote of equation

$$\cos\varphi_0\, y - \sin\varphi_0\, x = \frac{A}{\gamma}.$$

ii) The trajectory (2.27) with $b \neq 0$ has an asymptote of equation

$$\cos\varphi_0\, y - \sin\varphi_0\, x = \frac{1}{b}, \qquad \varphi_0 := -a/b.$$

iii) The trajectory (2.28) has two asymptotes of equations

$$\cos\left(\varphi_0 \pm \tfrac{\pi}{2\gamma}\right) y - \sin\left(\varphi_0 \pm \tfrac{\pi}{2\gamma}\right) x = \mp\frac{A}{\gamma}.$$

The energy of a trajectory $r = r(\varphi)$ can be computed from Eq. (2.19) and the definition (2.17) of U, namely

$$E = \frac{L^2}{2\mu}(u'^2 + u^2) + V(1/u). \tag{2.29}$$

Since $E = \mu v^2/2 + V(1/u)$, the particle's speed as a function of its distance r to the origin is given by

$$v = \frac{L}{\mu}\sqrt{u'^2 + u^2}. \tag{2.30}$$

Both formulas can also be obtained directly, taking into account that

$$v^2 = \dot r^2 + r^2\dot\varphi^2 = \dot\varphi^2(r'^2 + r^2) = \frac{L^2}{\mu^2 r^4}\left(\frac{u'^2}{u^4} + \frac{1}{u^2}\right) = \frac{L^2}{\mu^2}(u'^2 + u^2).$$

In Example 2.1 we can take

$$V(r) = -\int f(r)\,\mathrm{d}r = -k\int\frac{\mathrm{d}r}{r^3} = \frac{k}{2r^2},$$

and thus

$$E = \frac{L^2}{2\mu}\,(u'^2 + u^2) + \frac{k}{2}\,u^2 = \frac{L^2}{2\mu}\,(u'^2 + Cu^2).$$

For instance, for the trajectories (2.24) it is easily checked that

$$E = -\frac{L^2|C|}{2\mu A^2} < 0\,,$$

while for (2.26) and (2.28) we have

$$E = \frac{L^2|C|}{2\mu A^2} > 0\,.$$

It is also immediate to verify that the energy of the trajectories (2.27) is

$$E = \frac{L^2 b^2}{2\mu} \geqslant 0\,,$$

while the trajectories (2.25) have energy $E = 0$. These results agree with the fact that the trajectories (2.24) are *bounded* while (2.26), (2.27) (if $b \neq 0$) and (2.28) are *unbounded*. Indeed, note that, as in this case

$$\lim_{r\to\infty} V(r) = 0\,,$$

if the particle reaches infinity we must necessarily have

$$E = \frac{1}{2}\,\mu v_\infty^2 \geqslant 0\,.$$

Example 2.2. Let us find the central force causing a particle to describe the spiral $r = a\varphi$. To this end, it suffices to substitute $u = 1/(a\varphi)$ into Binet's equation, which yields

$$f(r) = -\frac{L^2}{\mu r^2}\,(u'' + u) = -\frac{L^2}{\mu a r^2}\left(\frac{2}{\varphi^3} + \frac{1}{\varphi}\right) = -\frac{L^2}{\mu a r^2}\left(\frac{2a^3}{r^3} + \frac{a}{r}\right) = \boxed{-\frac{L^2}{\mu a^3}\left(\frac{2a^5}{r^5} + \frac{a^3}{r^3}\right)}\,.$$

The motion of the angular coordinate φ is easily determined from Eq. (2.23):

$$t = \frac{\mu}{L}\int a^2\varphi^2\,\mathrm{d}\varphi = \frac{\mu a^2}{3L}\,(\varphi^3 - \varphi_0^3) \quad\Longrightarrow\quad \boxed{\varphi = \left(\frac{3L}{\mu a^2}\,t + \varphi_0^3\right)^{1/3}}\,,$$

with $\varphi_0 = \varphi(0)$, whence it follows that

$$\boxed{r = a\varphi = a\left(\frac{3L}{\mu a^2}\,t + \varphi_0^3\right)^{1/3}.}$$

The energy of this trajectory is computed without difficulty using Eq. (2.29). To this end, we first need to find the potential, which is given by

$$V(r) = \frac{L^2}{\mu a^3}\int\left(\frac{2a^5}{r^5} + \frac{a^3}{r^3}\right)\mathrm{d}r = -\frac{L^2}{2\mu a^2}\left(\frac{a^4}{r^4} + \frac{a^2}{r^2}\right)$$

up to an arbitrary constant that has been taken equal to zero so that

$$\lim_{r\to\infty} V(r) = 0\,.$$

Substituting into Eq. (2.29) we obtain

$$E = \frac{L^2}{2\mu}\,(u'^2 + u^2) - \frac{L^2}{2\mu}\,(a^2u^4 + u^2) = \frac{L^2}{2\mu}\,(u'^2 - a^2u^4) = 0\,.$$

By the law of conservation of energy, the particle reaches infinity with zero speed. The speed of the particle at any point on the trajectory can be computed using Eq. (2.30), but since we know the potential V and the energy E it can be more directly obtained from the energy equation:

$$E = 0 = \frac{1}{2}\mu v^2 + V(r) = \frac{1}{2}\mu v^2 - \frac{L^2}{2\mu a^2}\left(\frac{a^4}{r^4} + \frac{a^2}{r^2}\right) \quad \Longrightarrow \quad \boxed{v = \frac{L}{\mu r^2}\sqrt{r^2 + a^2}}\,.$$

From this equation it also follows that, as we already knew, $v \to 0$ for $r \to \infty$.

Exercise. If we apply Binet's equation to a circle of radius a centered at the origin we apparently obtain a force inversely proportional to the square of the distance from origin:

$$f(r) = -\frac{L^2}{a\mu}\frac{1}{r^2}\,. \tag{2.31}$$

Is this result correct?

Solution. The result is clearly *false* as stated, since *any* potential V whose corresponding effective potential U has some critical point r_0 admits a circular orbit $r = r_0$ with energy $U(r_0)$ (cf. Eq. (2.16)). Hence Eq. (2.31) is *not* correct. This fact is not totally surprising, since to obtain Binet's equation from Eq. (2.21) it is necessary to divide by u', which is not allowed if $u = 1/r$ is constant. Equation (2.31) is however *true* if properly interpreted. Indeed, since $r = a$ along the circular orbit, what this equation actually states is that

$$f(a) = -\frac{L^2}{\mu a^3} = -\mu a\dot{\varphi}^2 = -\frac{\mu v^2}{a}\,, \tag{2.32}$$

i.e., that the central force $f(a)$ along the orbit generates the centripetal acceleration $-\mu v^2/a$. Equivalently, (2.32) is the necessary and sufficient condition for the potential $U(r)$ to have a critical point at $r = a$. Thus in the case of a circular orbit centered at the origin Binet's equation is simply the condition for the existence of such an orbit, and does not provide any information about the force law at distances from the origin different from the orbit's radius.

2.2.3 Bounded orbits

As we have just shown in the previous section, the motion of the radial coordinate r is determined by the law of conservation of energy (2.16)-(2.17). Formally, this is the equation of motion of a particle of mass μ in the one-dimensional potential $U(r)$. For this reason, most of the results from Section 1.5 are also valid in the present situation. For instance, from Eq. (2.16) it follows that the motion can only take place in the region defined by the inequality

$$U(r) = V(r) + \frac{L^2}{2\mu r^2} \leqslant E\,. \tag{2.33}$$

It is however important to note that, unlike the variable x in Section 1.5, the coordinate r *can only take non-negative values.*

Example 2.3. For the **Kepler potential**

$$V(r) = -k/r\,, \qquad \text{with } k > 0\,,$$

the effective potential

$$U(r) = -\frac{k}{r} + \frac{L^2}{2\mu r^2}$$

behaves as shown in Fig. 2.4. Indeed, $U(r)$ diverges as $L^2/(2\mu r^2)$ for $r \to 0$ and tends to zero as $-k/r$ for $r \to \infty$. Moreover,

$$U'(r) = \frac{k}{r^2} - \frac{L^2}{\mu r^3} = \frac{k}{r^3}\left(r - \frac{L^2}{k\mu}\right),$$

so that U decreases for $0 < r < a := L^2/(k\mu)$ and increases for $r > a$, attaining its minimum value at $r = a$. The particle's energy must be greater than or equal to this minimum value, given by

$$U(a) = -\frac{k}{2a} = -\frac{k^2\mu}{2L^2} =: E_{\min}\,.$$

From Fig. 2.4 it follows that the trajectories with energy $E \geqslant 0$ are *unbounded*, since in this case the inequality (2.33) implies that $r \in [r_0, \infty)$, where $r_0 > 0$ is the only root of the equation $U(r) = E$. Therefore in this case the particle "arrives" from infinity, reaches a minimum distance from the origin equal to r_0 (which is a turning point of U) and goes back to infinity. On the contrary, if $E_{\min} < E < 0$ then $r_1 \leqslant r \leqslant r_2$, where $r_1 < r_2$ are the two roots of the equation $U(r) = E$ (cf. Fig. 2.4), which are again turning points. Therefore in this case the trajectory is *bounded* and stays away from the origin. Finally, if $E = E_{\min}$ then the trajectory is the *circle* $r = a$ (cf. Fig. 2.4).

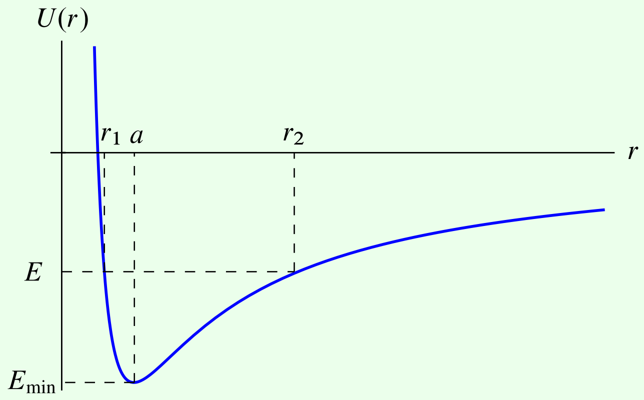


Figure 2.4. Effective potential $U(r)$ for the Kepler potential $V(r) = -k/r$ (with $k > 0$).

Exercise. Repeat the previous discussion for the potential $V(r) = k/(2r^2)$ of Example 2.1.

Solution. The effective potential is in this case given by

$$U(r) = \frac{k}{2r^2} + \frac{L^2}{2\mu r^2} = \frac{L^2}{2\mu}\frac{C}{r^2}\,, \tag{2.34}$$

with $C = 1 + (k\mu/L^2)$. Since $L^2/(2\mu) > 0$, the behavior of U depends on the sign of C (cf. Fig. 2.5).

i) For $C < 0$, the trajectories with energy $E \geqslant 0$ are unbounded and can reach the origin, since the allowed region for such energies is the whole semiaxis $r \geqslant 0$. On the other hand, for $E < 0$ the allowed region is an interval of the form $[0, r_0]$, where $r_0 > 0$ is a turning point. Thus in this case the trajectories with negative energy are bounded but fall into the origin. Note that this is consistent with our previous analysis, since for $C < 0$ the trajectories with positive energy $L^2|C|/(2\mu A^2)$ are given by Eq. (2.26) and those with zero energy by Eq. (2.25), whereas the trajectories with negative energy $-L^2|C|/(2\mu A^2)$ obey Eq. (2.24).

ii) For $C = 0$ the effective potential vanishes identically. Thus in this case the trajectories with positive energy are unbounded and fall into the origin, since the allowed region is again the semiaxis $r \geqslant 0$. These trajectories are the curves (2.27) with $b \neq 0$, whose energy is indeed $L^2 b^2/(2\mu) > 0$. On the other hand, for $E = 0$ *all* the points on the semiaxis $r \geqslant 0$ are equilibrium solutions of $U(r)$ (since $U'(r) = 0$ and $U(r) = E = 0$ for all r). Hence in this case the possible trajectories are the circles $r = r_0$ with arbitrary $r_0 \geqslant 0$ (i.e., the curves (2.27) with $b = 0$).

iii) Finally, for $C > 0$ we must have $E > 0$, and all the trajectories reach infinity but stay away from the origin (indeed, the allowed region is an interval of the form $[r_0, \infty]$, where r_0 is a turning point). As we saw above, these trajectories have equation (2.28) and their energy is indeed $L^2 C/(2\mu A^2) > 0$.

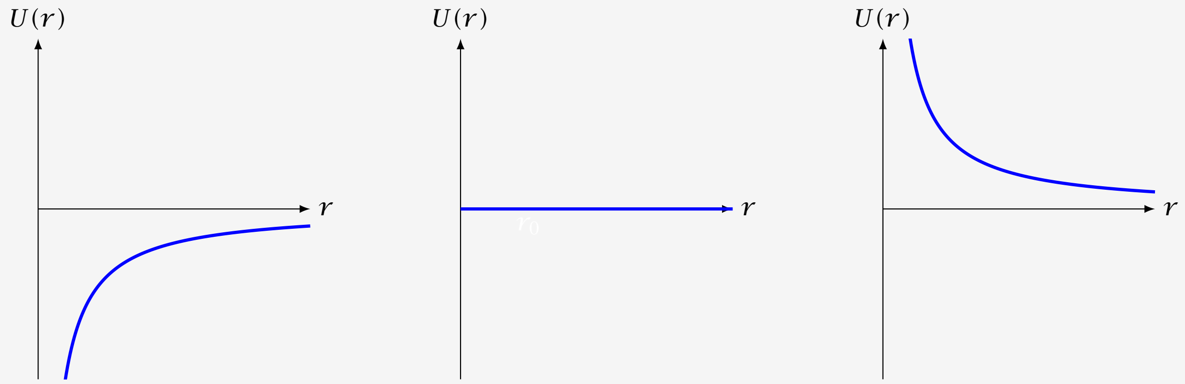


Figure 2.5. Plot of the effective potential (2.34) for $C < 0$ (left), $C = 0$ (middle) and $C > 0$ (right).

Of particular interest are *bounded orbits*, in which the radial coordinate moves between two consecutive turning points $0 < r_1 < r_2$ of the effective potential $U(r)$ (cf. Fig. 2.6). In this case, the points on the trajectory at the minimum distance r_1 from the origin are called **periapsides** or **pericenters** (**perigees**, **perihelia** or **periastra**[3] if the center of force is respectively Earth, the Sun or a star), while those at the maximum distance r_2 are called **apoapsides** or **apocenters** (**apogees**, **aphelia** or **apoastra**[4]). Both of these types of points are jointly referred to as **apsides**[5] (or *apsidal points*).

As we saw in Section 1.5, the motion of the radial coordinate is in this case *periodic* in time,

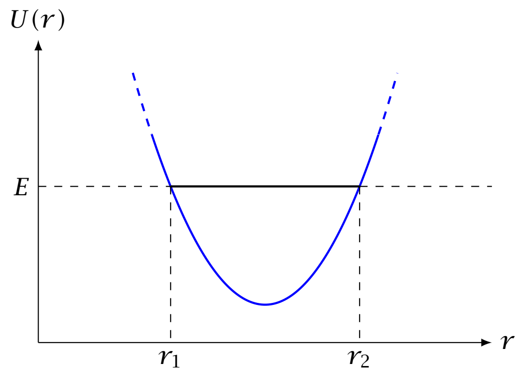


Figure 2.6. Effective potential $U(r)$ admitting a bounded orbit of energy E with $r_1 \leqslant r \leqslant r_2$, where $r_1 < r_2$ are two consecutive turning points of U for the energy E.

[3] Singular *periapsis*, *perigee*, *perihelion* and *periastron*.
[4] Singular *apoapsis*, *apogee*, *aphelion* and *apoastron*.
[5] Singular *apsis*.

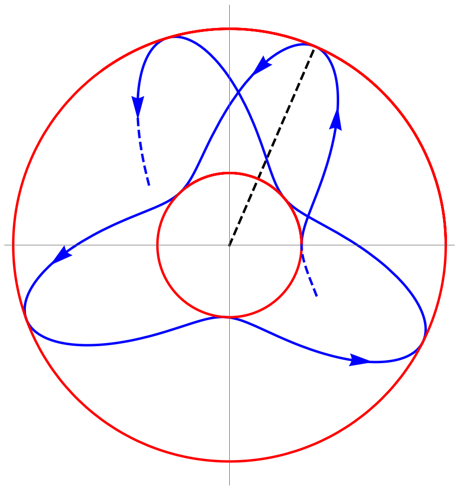


Figure 2.7. Precession of the periapsis in a bounded orbit. The circles $r = r_1$ and $r = r_2$ (minimum and maximum values of the radial coordinate) have been drawn in red, while the dashed black segment represents the line $\varphi = \Delta\varphi_{12}$ joining the origin with an apoapsis.

with period (depending in general on the energy and the angular momentum)

$$\tau_r = \sqrt{2\mu} \int_{r_1}^{r_2} \frac{\mathrm{d}r}{\sqrt{E - U(r)}} \,. \tag{2.35}$$

However, *this does not mean that the particle's motion is periodic*. Indeed, the necessary and sufficient condition for the motion to be periodic is that when the radial coordinate r completes a certain integer number of its periods the angle φ increases by an integer multiple of 2π, so that the particle returns to its starting point. Obviously, this is equivalent to requiring that the bounded orbit be *closed*. Let us next examine under what conditions this will be the case.

To begin with, when the radial coordinate r increases from r_1 to r_2 the variable $u = 1/r$ decreases, and hence $\frac{\mathrm{d}\varphi}{\mathrm{d}u} < 0$ (since $\dot\varphi = \frac{\mathrm{d}\varphi}{\mathrm{d}u}\,\dot u > 0$), so that we must take the minus sign in Eq. (2.20). Setting (without loss of generality) $\varphi(r_1) = 0$ we obtain

$$\varphi = -\frac{L}{\sqrt{2\mu}} \int_{1/r_1}^{1/r} \frac{\mathrm{d}u}{\sqrt{E - U(1/u)}} = \frac{L}{\sqrt{2\mu}} \int_{1/r}^{1/r_1} \frac{\mathrm{d}u}{\sqrt{E - U(1/u)}} =: \varphi_1(r) \,. \tag{2.36}$$

In particular, when r reaches its maximum value r_2 the angle φ increases by

$$\Delta\varphi_{12} = \varphi_1(r_2) = \frac{L}{\sqrt{2\mu}} \int_{1/r_2}^{1/r_1} \frac{\mathrm{d}u}{\sqrt{E - U(1/u)}} \,. \tag{2.37}$$

On the other hand, when r decreases from r_2 to r_1 the variable u increases, so that we should take the "+" in Eq. (2.20). In other words, we have

$$\varphi = \Delta\varphi_{12} + \frac{L}{\sqrt{2\mu}} \int_{1/r_2}^{1/r} \frac{\mathrm{d}u}{\sqrt{E - U(1/u)}} =: \varphi_2(r) \tag{2.38}$$

Thus when r takes once again its minimum value r_1 the angle φ has increased by

$$\Delta\varphi = \varphi_2(r_1) = 2\Delta\varphi_{12} \,.$$

The **displacement of the periapsis** $\Delta\varphi$, defined as the increase in the azimuthal angle φ between two consecutive periapsides, is thus given by

$$\Delta\varphi = \sqrt{\frac{2L^2}{\mu}} \int_{1/r_2}^{1/r_1} \frac{\mathrm{d}u}{\sqrt{E - U(1/u)}} = \sqrt{\frac{2L^2}{\mu}} \int_{r_1}^{r_2} \frac{\mathrm{d}r}{r^2\sqrt{E - U(r)}} \,. \tag{2.39}$$

In general, $\Delta\varphi$ is *not* an integer multiple of 2π, and thus the particles does *not* return to its initial position (i.e., the point with polar coordinates $r = r_1$, $\varphi = 0$) when the coordinate r takes the value r_1 for the second time (cf. Fig. (2.7)). The necessary and sufficient condition for the motion to be *periodic* —or, equivalently, the orbit to be *closed*— is that after n periods of the r coordinate the increase of the angle φ, which is obviously equal to $n\Delta\varphi$, be an integer multiple $2m\pi$ of 2π. We have thus proved the following:

The motion on a bounded orbit in which r varies between two consecutive turning points $r_1 < r_2$ of the effective potential U is *periodic* if and only if the orbit is *closed*. The necessary and sufficient condition for this to happen is that the displacement of the periapsis (2.39) be a *rational multiple* of 2π.

Remark. According to *Bertrand's theorem*, the only central potentials for which *all* bounded orbits are closed (and, hence, periodic) are the harmonic potential ($V(r) = \frac{1}{2}kr^2$ with $k > 0$) and the Kepler potential ($V(r) = -k/r$ with $k > 0$). ■

Exercise. Prove that the orbits in a central force field are symmetric about the line joining the origin with an apsis.

Solution. The function $u(\varphi)$ is a solution to Binet's equation, which is invariant under the transformations $\varphi \mapsto -\varphi$ and $\varphi \mapsto \varphi + \varphi_0$ with φ_0 an arbitrary constant. Hence if $u(\varphi)$ is a solution so are $u(-\varphi)$, $u(\varphi + \varphi_0)$, and $u(\varphi_0 - \varphi)$, for every $\varphi_0 \in \mathbb{R}$. Since the angle φ does not appear explicitly in Binet's equation, we can assume without loss of generality that the apsis considered has polar angle $\varphi = 0$. The orbit will then be symmetric about this apsis provided that $u(\varphi) = u(-\varphi)$. Since $u(-\varphi)$ is also a solution of Binet's equation, in order to prove the latter equality it suffices to show that $u(\varphi)$ and $g(\varphi) = u(-\varphi)$ satisfy the same initial conditions at $\varphi = 0$, i.e., that $u(0) = g(0)$ and $u'(0) = g'(0)$. The first of these equalities is obvious, while the second one easily follows from the fact that $u'(0) = 0$, since by definition of apsis u has a maximum or a minimum at $\varphi = 0$.

Exercise. Show that the *displacement of the apoapsis* (increase in the angle φ between two consecutive apoapsides) is also given by Eq. (2.39).

Solution. Since $u(-\varphi) = u(\varphi)$ by the previous exercise, if the polar angle of a periapsis is $\varphi = 0$ there are two consecutive apoapsides of the orbit at angles $\varphi = \pm\Delta\varphi_{12}$, and thus their angular displacement is again $2\Delta\varphi_{12} = \Delta\varphi$.

Remark. From the previous considerations it also follows that $u(\varphi + \Delta\varphi) = u(\varphi)$ (indeed, by definition of $\Delta\varphi$ both $u(\varphi)$ and $u(\varphi + \Delta\varphi)$ are solutions of Binet's equation with the same initial conditions $u(0) = 1/r_1$, $u'(0) = 0$ at $\varphi = 0$.) Thus $u(\varphi)$ is an even periodic function of period $\Delta\varphi$. ■

Exercise. Find the period of a circular orbit $r = a$ in a central potential.

Solution. The conservation of angular momentum implies that

$$\varphi(t) = \varphi(0) + \frac{Lt}{\mu a^2}.$$

Hence the motion is periodic, with period $\tau = 2\pi\mu a^2/L$. In general, the period of a closed orbit around the origin in a central potential can be found from the law of areas through the formula $\tau = 2\mu A/L$, where A is the area enclosed by the curve. (We are actually assuming that, as is usually the case, $r(\varphi)$ is a one-valued function of φ).

Example 2.4. For the harmonic potential $V(r) = \frac{1}{2}kr^2$, with $k > 0$, the effective potential $U(r)$ has the appearance of Fig. 2.8. Therefore in this case all orbits are bounded. The periapsis displacement of any of these orbits is given by

$$\Delta\varphi = \sqrt{\frac{2L^2}{\mu}} \int_{1/r_2}^{1/r_1} \frac{\mathrm{d}u}{\sqrt{E - \frac{L^2}{2\mu}u^2 - \frac{k}{2u^2}}} = \sqrt{\frac{2L^2}{\mu}} \int_{u_2}^{u_1} \frac{u\,\mathrm{d}u}{\sqrt{-\frac{L^2}{2\mu}u^4 + Eu^2 - \frac{k}{2}}},$$

where $u_1 > u_2$ are the two roots of the equation $-\frac{L^2}{2\mu}u^4 + Eu^2 - \frac{k}{2} = 0$. Performing the change of variable $s = u^2$ the last integral becomes

$$\Delta\varphi = \int_{s_2}^{s_1} \frac{\mathrm{d}s}{\sqrt{p(s)}},$$

where

$$p(s) = -s^2 + \frac{2\mu E}{L^2}s - \frac{k\mu}{L^2} = -\left(s - \frac{\mu E}{L^2}\right)^2 + \frac{\mu^2E^2}{L^4}\left(1 - \frac{kL^2}{\mu E^2}\right) \tag{2.40}$$

and $s_1 > s_2$ are the two roots of the equation $p(s) = 0$ (obviously, $s_i = u_i^2$). Performing then the change of variable

$$s = \frac{\mu E}{L^2} + \frac{\mu E}{L^2}\left(1 - \frac{kL^2}{\mu E^2}\right)^{1/2} \sin\theta, \qquad -\frac{\pi}{2} \leqslant \theta \leqslant \frac{\pi}{2}, \tag{2.41}$$

we finally obtain

$$\Delta\varphi = \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta = \pi.$$

Since $\Delta\varphi$ is a rational multiple of 2π, *all* the orbits are periodic (in agreement with Bertrand's theorem).

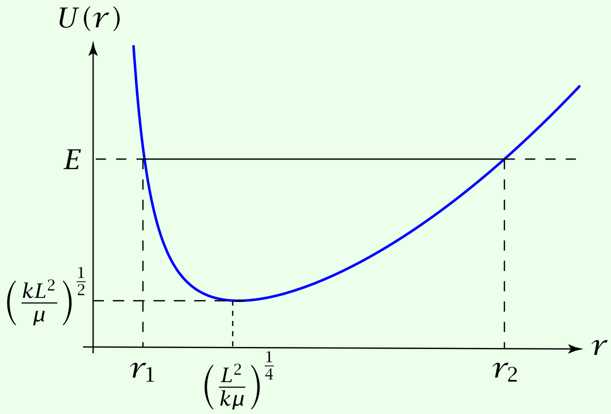


Figure 2.8. Effective potential $U(r)$ for the harmonic potential $V(r) = kr^2/2$ (with $k > 0$).

Exercise. Show that all the orbits of the harmonic potential $V(r) = kr^2/2$ (with $k > 0$) are ellipses centered at the origin, and compute the period of the motion.

Solution. The equation of the orbits is in this case

$$\varphi = \sqrt{\frac{L^2}{2\mu}} \int^{1/r} \frac{\mathrm{d}u}{\sqrt{E - \frac{L^2}{2\mu}u^2 - \frac{k}{2u^2}}} = \int^{1/r} \frac{u\,\mathrm{d}u}{\sqrt{-u^4 + \frac{2\mu E}{L^2}u^2 - \frac{k\mu}{L^2}}} = \frac{1}{2}\int^{1/r^2} \frac{\mathrm{d}s}{\sqrt{p(s)}},$$

with $p(s)$ defined by Eq. (2.40). Performing the change of variable (2.41) (note that, as seen in Example 2.4, in this case $E^2 \geqslant kL^2/\mu$) we obtain

$$\varphi = \varphi_0 + \frac{1}{2}\arcsin\left(\frac{\frac{L^2}{\mu E r^2} - 1}{\sqrt{1 - \frac{L^2 k}{\mu E^2}}}\right) \quad \Longrightarrow \quad \frac{L^2}{\mu E r^2} = 1 + \sqrt{1 - \frac{L^2 k}{\mu E^2}}\,\sin\left(2(\varphi - \varphi_0)\right).$$

Taking, without loss of generality, $\varphi_0 = \pi/4$, the above equation can be written as

$$r^2 = \frac{\alpha}{1 - e\cos 2\varphi}, \qquad \text{with} \quad \alpha = \frac{L^2}{\mu E}, \quad e = \sqrt{1 - \frac{L^2 k}{\mu E^2}} \leqslant 1$$

(and $e = 1$ if and only if $L = 0$). In Cartesian coordinates,

$$r^2 - er^2\cos 2\varphi = x^2 + y^2 - e(x^2 - y^2) = \boxed{(1-e)x^2 + (1+e)y^2 = \alpha},$$

which is the equation of an *ellipse centered at the origin* with semiaxes

$$a = \sqrt{\frac{\alpha}{1-e}}, \qquad b = \sqrt{\frac{\alpha}{1+e}}.$$

The period of the motion τ is easily computed using the law of areas:

$$\frac{L\tau}{2\mu} = \pi ab = \frac{\pi\alpha}{\sqrt{1-e^2}} = \frac{\pi L^2/\mu E}{\sqrt{L^2 k/\mu E^2}} = \frac{\pi L}{\sqrt{k\mu}} \quad \Longrightarrow \quad \boxed{\tau = 2\pi\sqrt{\frac{\mu}{k}}}.$$

Note, in particular, that in this case the period is *independent* of E and L, and is therefore the same for all orbits.

Note. In this case, the equation of the orbits can be obtained more easily by solving the equations of motion in Cartesian coordinates, namely

$$\ddot{x} + \frac{k}{\mu}x = 0, \qquad \ddot{y} + \frac{k}{\mu}y = 0.$$

Indeed, calling $\omega = \sqrt{k/\mu}$ and setting, without loss of generality, $\dot{x}(0) = 0$, $y(0) = 0$ (i.e., taking the x axis in the direction of an apsis[a]) we obtain

$$x = a\,\cos(\omega t), \qquad y = b\,\sin(\omega t),$$

with a and b real constants, which are the parametric equations of an ellipse centered at the origin with semiaxes $|a|$ and $|b|$. (In fact, from the choice of the x axis it follows that $a > 0$, and the condition $\dot{\varphi} > 0$ implies that $\dot{y}(0) = b\omega > 0$, i.e., that $b > 0$.) From the above equations it also follows that the period of the motion is

$$\tau = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{\mu}{k}}.$$

[a] Indeed, at an apsis we have $\dot{r} = v_r = 0$, and hence $\mathbf{r}$ is perpendicular to $\dot{\mathbf{r}}$.

Exercise. Find the equation of the orbits for the *repulsive* harmonic potential $V(r) = -kr^2/2$ (with $k > 0$).

Example 2.5. According to Einstein's theory of general relativity, the potential felt by a particle of mass m at a distance r from a mass $M \gg m$ fixed at the origin is effectively given by

$$V(r) = -\frac{GMm}{r} - \frac{GML^2}{mc^2r^3},$$

where $L > 0$ is the particle's angular momentum. In planetary motion the general relativity correction is much smaller than the Kepler potential term. Indeed, the quotient of the two terms in $V(r)$ is given by $L^2/(m^2c^2r^2)$, which for nearly circular orbits can be estimated by taking $L = mrv$:

$$\frac{(mrv)^2}{m^2c^2r^2} = \frac{v^2}{c^2}.$$

The velocity v of a planet in the solar system is at most 59 Km/s (Mercury's maximum velocity), so that $v^2/c^2 = O(10^{-8})$. In spite of this fact, the general relativity correction causes a displacement of the periapsis of the planetary orbits slightly different from 2π, so that these orbits are in general *not* closed as is the case for the Kepler potential. The periapsis displacement, given by[a]

$$\Delta\varphi = \sqrt{\frac{2L^2}{m}} \int_{u_2}^{u_1} \frac{du}{\sqrt{E + GMmu - \frac{L^2}{2m} u^2\left(1 - \frac{2GM}{c^2}u\right)}}, \tag{2.42}$$

where $u_2 < u_1$ are the two positive roots of the cubic polynomial under the square root, cannot be expressed in terms of elementary functions. However, since the general relativity correction is very small it is possible to compute it approximately to order c^{-2} or, more accurately, to order 2 in the small *dimensionless* parameter[b] $GMm/(Lc)$, as follows. We start by writing

$$\begin{aligned}
\Delta\varphi &= 2\int_{u_2}^{u_1}\left(1 - \frac{2GM}{c^2}u\right)^{-1/2}\left[\frac{2m}{L^2}(E + GMmu)\left(1 - \frac{2GM}{c^2}u\right)^{-1} - u^2\right]^{-1/2} du \\
&\simeq 2\int_{u_2}^{u_1}\left(1 + \frac{GM}{c^2}u\right)\left[\frac{2m}{L^2}(E + GMmu)\left(1 + \frac{2GM}{c^2}u\right) - u^2\right]^{-1/2} du \\
&= 2\left(1 - \frac{4G^2M^2m^2}{L^2c^2}\right)^{-1/2}\int_{u_2}^{u_1}\left(1 + \frac{GM}{c^2}u\right)P(u)^{-1/2}\,du,
\end{aligned}$$

where $P(u)$ is the cubic polynomial

$$P(u) = -u^2 + \frac{2m}{L^2}\left(1 - \frac{4G^2M^2m^2}{L^2c^2}\right)^{-1}\left[E + GMm\left(1 + \frac{2E}{mc^2}\right)u\right]$$

and we have taken into account that $GMu/c^2 \sim GM/(c^2a) = O(v^2/c^2)$ (see footnote b) to approximate the factor $\left(1 - \frac{2GM}{c^2}u\right)^{-1/2}$. Likewise,

$$\left(1 - \frac{4G^2M^2m^2}{L^2c^2}\right)^{-1/2}\left(1 + \frac{GM}{c^2}u\right) \simeq \left(1 + \frac{2G^2M^2m^2}{L^2c^2}\right)\left(1 + \frac{GM}{c^2}u\right) \simeq 1 + \frac{2G^2M^2m^2}{L^2c^2} + \frac{GM}{c^2}u,$$

and hence

$$\Delta\varphi \simeq 2\int_{u_2}^{u_1}\left(1 + \frac{2G^2M^2m^2}{L^2c^2} + \frac{GM}{c^2}u\right)P(u)^{-1/2}\,du.$$

The *quadratic* polynomial $P(u)$ differs from the cubic polynomial appearing under the square root in Eq. (2.42) by terms of order c^{-2}, so the positive roots $u_2 < u_1$ of the latter polynomial are approximately equal to the two roots $u_2^* < u_1^*$ of $P(u)$. Since $P(u) = (u_1^* - u)(u - u_2^*)$ we thus have

$$\Delta\varphi \simeq 2\int_{u_2^*}^{u_1^*}\left(1 + \frac{2G^2M^2m^2}{L^2c^2} + \frac{GM}{c^2}u\right)[(u_1^* - u)(u - u_2^*)]^{-1/2}\,\mathrm{d}u = \left(1 + \frac{2G^2M^2m^2}{L^2c^2}\right)I_0 + \frac{GM}{c^2}I_1,$$

with

$$I_k := 2\int_{u_2^*}^{u_1^*} u^k[(u_1^* - u)(u - u_2^*)]^{-1/2}\,\mathrm{d}u\,.$$

The latter integrals are easily computed by the standard change of variable

$$u = \frac{1}{2}(u_1^* + u_2^*) + \frac{1}{2}(u_1^* - u_2^*)\sin\theta\,, \qquad -\frac{\pi}{2} \leqslant \theta \leqslant \frac{\pi}{2}\,,$$

with the result

$$I_0 = 2\int_{-\pi/2}^{\pi/2}\mathrm{d}\theta = 2\pi\,, \qquad I_1 = \int_{-\pi/2}^{\pi/2}[u_1^* + u_2^* + (u_1^* - u_2^*)\sin\theta]\,\mathrm{d}\theta = \pi(u_1^* + u_2^*)\,.$$

Inserting these values into the last formula for $\Delta\varphi$ we obtain

$$\Delta\varphi - 2\pi \simeq \frac{4\pi G^2M^2m^2}{L^2c^2} + \frac{\pi GM}{c^2}(u_1^* + u_2^*)\,.$$

Since the constant multiplying $u_1^* + u_2^*$ in the last term is already of order c^{-2}, we can compute the roots u_i^* to order 1, i.e., as the roots of the polynomial

$$\lim_{c\to\infty} P(u) = -u^2 + \frac{2GMm^2}{L^2}u + \frac{2m}{L^2}E\,.$$

We thus obtain

$$u_1^* + u_2^* \simeq \frac{2GMm^2}{L^2}\,,$$

which yields the following formula for the *advance* of the periapsis after one period of the radial coordinate r:

$$\Delta\varphi - 2\pi \simeq 6\pi\left(\frac{GMm}{Lc}\right)^2.$$

Due to the smallness of the dimensionless parameter GMm/LC in the previous expression, we can use the formula we shall derive in the next section relating the angular momentum to the semi-major axis a and the eccentricity e of a Keplerian orbit, namely

$$L^2 = GMm^2a(1 - e^2)\,.$$

We thus have

$$\boxed{\Delta\varphi - 2\pi \simeq \frac{6\pi GM}{c^2a(1 - e^2)}\,.}$$

For planetary motion this effect is very small (of order v^2/c^2), but it *accumulates* with each period. The *rate of advance* of the periapsis is given by

$$\frac{\Delta\varphi - 2\pi}{\tau} \simeq \frac{6\pi GM}{c^2a(1 - e^2)\tau}\,,$$

where τ is the period of the radial coordinate r. To order c^{-2}, we can use the formula for τ derived in the next section for the Kepler potential, namely

$$\tau = \frac{2\pi a^{3/2}}{\sqrt{GM}}.$$

We thus finally obtain

$$\frac{\Delta\varphi - 2\pi}{\tau} \simeq \frac{3(GM)^{3/2}}{c^2 a^{5/2}(1-e^2)}.$$

If we measure lengths in astronomical units (AU) and times in years we have

$$GM = 4\pi^2 \text{ AU}^3/\text{year}^2$$

and the previous formula reads

$$\frac{\Delta\varphi - 2\pi}{\tau} \simeq \frac{24\pi^3}{c^2 a^{5/2}(1-e^2)}.$$

In the solar system the rate of advance of the perihelion is maximum for Mercury, since its orbit has the smallest semi-major axis ($a = 0.38709893$ AU) and one of the largest eccentricities ($e = 0.20563069$). Taking into account that

$$c = 2.99792458 \cdot 10^8 \,\frac{\text{m}}{\text{s}} = 2.99792458 \cdot 10^8 \,\frac{3.1558149504 \cdot 10^7}{1.495978707 \cdot 10^{11}} \,\frac{\text{AU}}{\text{year}} = 6.3241077 \times 10^4 \,\frac{\text{AU}}{\text{year}},$$

we obtain the following value for the rate of advance of Mercury's perihelion:

$$\boxed{\frac{\Delta\varphi - 2\pi}{\tau} \simeq 2.08387 \cdot 10^{-6} \text{ rad/year} = 42.9829''/\text{century}.}$$

---

[a]More precisely, in the previous formula the energy E should be replaced by

$$\frac{mc^2}{2}\left[\left(1 + \frac{E}{mc^2}\right)^2 - 1\right] = E + \frac{E^2}{2mc^2}.$$

However, the last term is much smaller than the first one, since in planetary motion $E \ll mc^2$ (see next footnote).

[b]To determine the order of magnitude of $GMm/(Lc)$ in planetary motion, we can use the value of L we shall obtain from the analysis of the Kepler problem in the next section,

$$L = m\sqrt{GMa(1-e^2)},$$

where a and e are respectively the semi-major axis and the eccentricity of the Keplerian orbit. We thus obtain

$$\frac{GMm}{Lc} = \frac{1}{c}\sqrt{\frac{GM}{a(1-e^2)}}.$$

Using the formula for the period τ from the next section,

$$\tau = \frac{2\pi a^{3/2}}{\sqrt{GM}},$$

we have

$$\frac{GMm}{Lc} = \frac{2\pi a}{\tau c\sqrt{1-e^2}} = \frac{v/c}{\sqrt{1-e^2}}, \qquad v := \frac{2\pi a}{\tau}.$$

In planetary motion the orbital velocity $v = 2\pi a/\tau$ is at most 48.9 Km/s (Mercury's orbital velocity), while the factor of $(1-e^2)^{-1/2}$ is only slightly larger than one even for relatively eccentric orbits (1.02 for Mercury). Hence $GMm/(Lc)$ is typically of order 10^{-4} in planetary motion. Similarly, using the formula in the next section for the energy of a Keplerian orbit as a good estimate for E we obtain

$$\frac{E}{mc^2} \simeq \frac{GM}{2ac^2} = \frac{2\pi^2 a^2}{c^2\tau^2} = \frac{1}{2}\left(\frac{v}{c}\right)^2 = O(10^{-8})\,.$$

2.3 Kepler's problem. Planetary motion

2.3.1 Kepler's problem

We shall study in this section **Kepler's problem**, i.e., the motion of two bodies of masses m_1 and m_2 subject only to their mutual gravitational attraction

$$\mathbf{F}_{12} = -\mathbf{F}_{21} = -\frac{k(\mathbf{r}_1 - \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|^3}\,,$$

where the constant k is given by

$$k = Gm_1m_2 = GM\mu > 0\,.$$

Therefore in this case

$$\mathbf{F}(\mathbf{r}) = -\frac{k}{r^2}\,\mathbf{e}_r \quad \Longrightarrow \quad f(r) = -\frac{k}{r^2}\,, \quad V(r) = -\frac{k}{r}\,,$$

and the associated one-body problem is

$$\mu\ddot{\mathbf{r}} = -k\frac{\mathbf{r}}{r^3}\,,$$

or equivalently

$$\ddot{\mathbf{r}} = -GM\frac{\mathbf{r}}{r^3}\,.$$

The equation of the orbits is easily obtained from Binet's equation, which for this potential is particularly simple:

$$u'' + u = \frac{\mu k}{L^2}\,.$$

The general solution of this equation can be expressed in the form

$$u = \frac{\mu k}{L^2}\big(1 + e\cos(\varphi - \varphi_0)\big)\,,$$

with e and φ_0 integration constants. Note that we can assume without loss of generality that $e \geqslant 0$, since if $e < 0$ it suffices to replace φ_0 with $\pi + \varphi_0$ in the previous equation. Clearly φ_0 is the polar angle of the orbit's periapsis, so that taking the x axis as the line joining the origin to the periapsis we can set $\varphi_0 = 0$. The equation of the orbits reduces then to

$$r = \frac{\alpha}{1 + e\cos\varphi}\,, \qquad \text{with} \quad \alpha := \frac{L^2}{\mu k}\,. \tag{2.43}$$

The parameter e can be related to the energy and angular momentum of the orbit using Eq. (2.29):

$$E = \frac{L^2}{2\mu}(u'^2 + u^2) - ku = \frac{\mu k^2}{2L^2}\left[e^2 \sin^2\varphi + (1 + e\cos\varphi)^2 - 2(1 + e\cos\varphi)\right] = \frac{\mu k^2}{2L^2}(e^2 - 1)\,.$$

Since $e \geqslant 0$, we have

$$e = \sqrt{1 + \frac{2EL^2}{\mu k^2}}\,. \tag{2.44}$$

Note that, as we saw in Example 2.3, for the Kepler potential $E \geqslant E_{\min} = -\mu k^2/(2L^2)$, and thus the quantity under the radical is nonnegative.

- The orbits of Kepler's potential are **conic sections**. Indeed, from Eq. (2.43) we obtain

$$r = \alpha - ex \quad \Longrightarrow \quad x^2 + y^2 = \alpha^2 - 2\alpha e x + e^2 x^2 \quad \Longrightarrow \quad (1 - e^2)x^2 + y^2 + 2\alpha e x = \alpha^2\,,$$

which is a second-degree polynomial equation in (x, y). The type of conic depends on the sign of $1 - e^2$ as follows:

$$\begin{aligned} e > 1 \quad &\Longrightarrow \quad \textbf{hyperbola} \\ e = 1 \quad &\Longrightarrow \quad \textbf{parabola} \\ 0 < e < 1 \quad &\Longrightarrow \quad \textbf{ellipse} \\ e = 0 \quad &\Longrightarrow \quad \textbf{circle}\,. \end{aligned}$$

In terms of the energy (cf. Eq. (2.44)),

$$\begin{aligned} E > 0 \quad &\Longrightarrow \quad \textbf{hyperbola} \\ E = 0 \quad &\Longrightarrow \quad \textbf{parabola} \\ -\frac{\mu k^2}{2L^2} < E < 0 \quad &\Longrightarrow \quad \textbf{ellipse} \\ E = -\frac{\mu k^2}{2L^2} \quad &\Longrightarrow \quad \textbf{circle}\,, \end{aligned}$$

where $-\mu k^2/(2L^2)$ is the minimum energy that a particle of mass μ and angular momentum L can have. Note, in particular, that this result is consistent with the qualitative discussion of Example 2.3. Note also that in the Kepler potential *all bounded orbits are closed* (and hence *periodic*), confirming once again Bertrand's theorem.

2.3.2 Planetary motion

The most interesting case is that of elliptical orbits (including, in particular, circular ones), in which $0 \leqslant e < 1$ or $E < 0$, since it is the relevant case when studying the motion of the planets around the Sun. The Cartesian equation of the orbits can be rewritten as

$$(1 - e^2)\left(x + \frac{\alpha e}{1 - e^2}\right)^2 + y^2 = \alpha^2 + \frac{\alpha^2 e^2}{1 - e^2} = \frac{\alpha^2}{1 - e^2}\,,$$

which is the equation of an ellipse centered at the point

$$\left(-\frac{\alpha e}{1 - e^2}, 0\right) \tag{2.45}$$

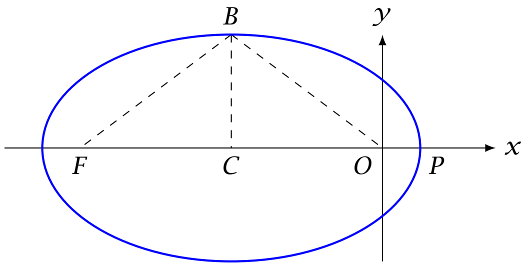


Figure 2.9. Geometry of an elliptical orbit in planetary motion. The point C is the center of the ellipse, F and O (the origin) its foci and P the periapsis. The distances $\overline{CP} = a$ and $\overline{CB} = b$ are respectively the ellipse's semi-major and semi-minor axes, and $\overline{OC} = \overline{FC} = c$ is its focal distance. By the defining property of the ellipse we have $\overline{BO} + \overline{BF} = 2\sqrt{b^2+c^2} = \overline{PO} + \overline{PF} = (a-c) + (a+c) = 2a \implies a = \sqrt{b^2+c^2}$.

with semi-major and semi-minor axes respectively given by

$$\boxed{a = \frac{\alpha}{1-e^2}\,, \qquad b = \frac{\alpha}{\sqrt{1-e^2}}\,.} \tag{2.46}$$

Recall that the *focal distance* c (defined as the distance of the center of the ellipse to either of its foci) and the *eccentricity* ε of an ellipse with semiaxes $a \geqslant b$ are given by

$$\boxed{c = \sqrt{a^2-b^2}\,, \qquad \varepsilon = \frac{c}{a}} \tag{2.47}$$

(cf. Fig. 2.9). Using the previous expressions for a and b we easily obtain

$$c = \frac{\alpha}{1-e^2}\sqrt{1-(1-e^2)} = \frac{\alpha e}{1-e^2} = ea \quad \Longrightarrow \quad \boxed{e = \varepsilon\,.}$$

Therefore the constant e appearing in the equation of the orbits is the *eccentricity* of the ellipse, and Eq. (2.44) relates the particle's energy and angular momentum to the eccentricity of its orbit. The above equations also determine the position of the *foci* of the ellipse, which by definition are the two points on the major axis (in this case, the x axis) at a distance c from the center of the ellipse. Indeed, from Eqs. (2.45)–(2.47) it follows that the center of the ellipse has coordinates $(-c, 0)$, and thus the foci are the points $(-2c, 0)$ and $(0, 0)$. In particular, this shows that one of the foci is the origin of coordinates, that is, the center of gravitational attraction. Hence *the bounded orbits in planetary motion are ellipses, one of whose foci is the Sun* (Kepler's first law).

- From Eq. (2.46) and the expression (2.44) for the eccentricity it follows that the energy of an elliptical orbit is

$$\boxed{E = -\frac{\mu k^2}{2L^2}(1-e^2) = -\frac{\mu k^2 \alpha}{2aL^2} = -\frac{k}{2a}\,.} \tag{2.48}$$

We see, therefore, that *the energy depends only on the semi-major axis of the orbit* (i.e., it is independent of its eccentricity).

- The period τ of elliptic orbits in planetary motion is easily determined from the law of areas (2.13), taking into account that the area of an ellipse is equal to πab:

$$\frac{L\tau}{2\mu} = \pi ab = \pi\sqrt{\alpha}\,a^{3/2} \quad \Longrightarrow \quad \boxed{\tau = \frac{2\pi\mu}{L}\sqrt{\alpha}\,a^{3/2} = 2\pi\sqrt{\frac{\mu}{k}}\,a^{3/2}\,,} \tag{2.49}$$

or in terms of the energy

$$\boxed{\tau = \pi k \sqrt{\frac{\mu}{2}}\,|E|^{-3/2}\,.}$$

In particular, *the period depends only on the semi-major axis, and is independent of the eccentricity*. In planetary motion the formula for the period is usually expressed in the form

$$\boxed{\tau = \frac{2\pi a^{3/2}}{\sqrt{GM}} \simeq \frac{2\pi a^{3/2}}{\sqrt{GM_\odot}}\,,}$$

$M_\odot$ being the Sun's mass. Thus (with great approximation) *the ratio τ^2/a^3 is the same for all planets* (Kepler's third law).

• Let us denote by p and p' respectively the distance of the perihelion and aphelion of the ellipse to the origin. From the equation of the orbit (2.43) it easily follows that the particle is in the perihelion (resp. in the aphelion) when $\varphi = 0$ (resp. $\varphi = \pi$), and therefore

$$p = \frac{\alpha}{1+e} = a(1-e)\,, \qquad p' = \frac{\alpha}{1-e} = a(1+e)\,.$$

This is also apparent from the geometry of the ellipse (cf. Fig. 2.9), since $a(1-e) = a - c$, $= a(1+e) = a + c$.

• It is also straightforward to compute the speed at any point in the orbit from the law of conservation of energy:

$$v^2 = \frac{2}{\mu}\left(E - V(r)\right) = \frac{k}{\mu}\left(\frac{2}{r} - \frac{1}{a}\right) = \frac{k}{\mu\alpha}\left[2(1+e\cos\varphi) - (1-e^2)\right] = \boxed{\frac{k^2}{L^2}\left(1 + e^2 + 2e\cos\varphi\right).}$$

Note that v is maximal at the perihelion ($\varphi = 0$) and minimal at the aphelion ($\varphi = \pi$), with respective values

$$v_p = \frac{k}{L}(1+e)\,, \qquad v_{p'} = \frac{k}{L}(1-e)\,.$$

In particular, the quotient

$$\frac{v_p}{v_{p'}} = \frac{1+e}{1-e}$$

depends only on the eccentricity of the orbit. It is sometimes of interest to express the speeds v_p and $v_{p'}$ as a function of p and p', instead of L. To this end, it suffices to note that

$$\alpha = \frac{L^2}{k\mu} = p(1+e) = p'(1-e) \quad \Longrightarrow \quad L = \sqrt{k\mu p(1+e)} = \sqrt{k\mu p'(1-e)}\,,$$

and therefore

$$v_p = \sqrt{\frac{k}{\mu p}(1+e)}\,, \qquad v_{p'} = \sqrt{\frac{k}{\mu p'}(1-e)}.$$

Example 2.6. The mean value over a period of a planetary orbit of any quantity $f(r)$ is defined by

$$\langle f(r)\rangle := \frac{1}{\tau}\int_0^\tau f(r)\,\mathrm{d}\tau\,.$$

Taking into account that

$$\mathrm{d}t = \frac{\mathrm{d}\varphi}{\dot\varphi} = \frac{\mu}{L}r^2\,\mathrm{d}\varphi\,,$$

the time integral can be transformed into the following integral over the polar angle φ:

$$\langle f(r)\rangle = \frac{\mu}{\tau L}\int_0^{2\pi} r^2 f(r)\,\mathrm{d}\varphi = \frac{\mu\alpha^2}{\tau L}\int_0^{2\pi} f\left(\frac{\alpha}{1+e\cos\varphi}\right)\frac{\mathrm{d}\varphi}{(1+e\cos\varphi)^2}\,.$$

Using Eqs. (2.43), (2.46) and (2.49) we obtain

$$\frac{\mu\alpha^2}{\tau L} = \frac{\mu\alpha^2}{\tau\sqrt{\alpha k\mu}} = \frac{1}{\tau}\sqrt{\frac{\mu}{k}}\,a^{3/2}(1-e^2)^{3/2} = \frac{(1-e^2)^{3/2}}{2\pi}\,,$$

and thus

$$\langle f(r)\rangle = \frac{(1-e^2)^{3/2}}{2\pi}\int_0^{2\pi} f\left(\frac{a(1-e^2)}{1+e\cos\varphi}\right)\frac{\mathrm{d}\varphi}{(1+e\cos\varphi)^2}\,.$$

For instance, the mean distance of a planet to the Sun is given by

$$\langle r\rangle = a(1-e^2)^{5/2}I(e)\,,\qquad I(e) := \frac{1}{2\pi}\int_0^{2\pi}\frac{\mathrm{d}\varphi}{(1+e\cos\varphi)^3}\,.$$

The integral $I(e)$ can be computed using the residue theorem taught in complex analysis courses, with the result

$$I(e) = \frac{e^2+2}{2(1-e^2)^{5/2}}\,.$$

We finally obtain

$$\boxed{\langle r\rangle = \left(1+\frac{e^2}{2}\right)a\,.}$$

Exercise. Integrate Eq. (2.23) to find the relation between t and φ in planetary motion.

Solution. Using Eq. (2.43) for the Kepler orbits and the first relation (2.46) we obtain

$$t = \frac{L^3}{\mu k^2}\int\frac{\mathrm{d}\varphi}{(1+e\cos\varphi)^2} = \sqrt{\frac{\mu}{k}}\,a^{3/2}(1-e^2)^{3/2}\int\frac{\mathrm{d}\varphi}{(1+e\cos\varphi)^2}\,.$$

To compute the integral, we start by making the change of variable

$$\begin{cases} u = \tan(\varphi/2) \quad\Longrightarrow\quad \cos\varphi = 2\cos^2(\varphi/2)-1 = \dfrac{2}{\sec^2(\varphi/2)}-1 = \dfrac{2}{1+u^2}-1 = \dfrac{1-u^2}{1+u^2}\,,\\[2mm] \mathrm{d}u = \dfrac{1}{2}\sec^2(\varphi/2)\,\mathrm{d}\varphi = \dfrac{1}{2}(1+u^2)\,\mathrm{d}\varphi, \end{cases}$$

and hence

$$\int\frac{\mathrm{d}\varphi}{(1+e\cos\varphi)^2} = 2\int\frac{\mathrm{d}u}{(1+u^2)\left[1+\frac{e(1-u^2)}{1+u^2}\right]^2} = 2\int\frac{(1+u^2)\,\mathrm{d}u}{[1+e+(1-e)u^2]^2}\,.$$

Setting now

$$u = \sqrt{\frac{1+e}{1-e}}\,v$$

we obtain

$$\begin{aligned}\int\frac{\mathrm{d}\varphi}{(1+e\cos\varphi)^2} &= \frac{2}{(1+e)^2}\sqrt{\frac{1+e}{1-e}}\int\frac{1+\frac{1+e}{1-e}v^2}{(1+v^2)^2} = 2(1-e^2)^{-3/2}\int\frac{1-e+(1+e)v^2}{(1+v^2)^2}\\ &= 2(1-e^2)^{-3/2}\left[(1+e)\arctan v - 2e\int\frac{\mathrm{d}v}{(1+v^2)^2}\right].\end{aligned}$$

The last integral is computed integrating by parts in the integral of $(1+v^2)^{-1}$:

$$\arctan v = \int \frac{\mathrm{d}v}{1+v^2} = \frac{v}{1+v^2} + \int \frac{2v^2\,\mathrm{d}v}{(1+v^2)^2} = \frac{v}{1+v^2} + 2\arctan v - 2\int \frac{\mathrm{d}v}{(1+v^2)^2}$$
$$\implies \quad 2\int \frac{\mathrm{d}v}{(1+v^2)^2} = \frac{v}{1+v^2} + \arctan v\,.$$

Putting everything together we obtain:

$$\int \frac{\mathrm{d}\varphi}{(1+e\cos\varphi)^2} = 2(1-e^2)^{-3/2}\left(\arctan v - \frac{ev}{1+v^2}\right).$$

Since

$$\frac{v}{1+v^2} = \sqrt{\frac{1-e}{1+e}}\,\frac{u}{\frac{1-e}{1+e}u^2+1} = \frac{\sqrt{1-e^2}\,u}{(1-e)u^2+1+e} = \frac{\sqrt{1-e^2}\,\tan(\frac{\varphi}{2})}{2e+(1-e)\sec^2(\frac{\varphi}{2})}$$
$$= \sqrt{1-e^2}\,\frac{\sin(\frac{\varphi}{2})\cos(\frac{\varphi}{2})}{2e\cos^2(\frac{\varphi}{2})+1-e} = \frac{1}{2}\sqrt{1-e^2}\,\frac{\sin\varphi}{1+e\cos\varphi}$$

we finally arrive at the formula

$$\boxed{t = \sqrt{\frac{\mu}{k}}\,a^{3/2}\left[2\arctan\left(\sqrt{\frac{1-e}{1+e}}\tan(\tfrac{\varphi}{2})\right) - e\sqrt{1-e^2}\,\frac{\sin\varphi}{1+e\cos\varphi}\right]},$$

where we have discarded the integration constant so that $t=0$ at the periapsis $\varphi=0$. This expression is too unwieldy in practice, and the time dependence of r (and hence φ) in the Kepler problem is usually computed inverting Kepler's equation introduced in the next exercise.

Exercise. Repeat the previous calculation for hyperbolic and parabolic orbits.

Exercise. Given an elliptic orbit of eccentricity e and major semiaxis a, define the *eccentric anomaly* $\psi(t)$ by the equation

$$\boxed{\omega t = \psi - e\sin\psi,} \tag{2.50}$$

where $\omega = \dfrac{2\pi}{\tau} = \sqrt{\dfrac{k}{\mu}}\,a^{-3/2}$ is the mean orbital frequency. (Note that

$$\frac{\mathrm{d}}{\mathrm{d}\psi}(\psi - e\sin\psi) = 1 - e\cos\psi \geqslant 1-e > 0\,,$$

so that (2.50) uniquely determines ψ as a function of t by the inverse function theorem.) Show that the radius vector $r(t)$ can be expressed in terms of the eccentric anomaly through the equation

$$\boxed{r = a(1-e\cos\psi)\,.} \tag{2.51}$$

Solution. From the first Eq. (2.18) with

$$U(r) = \frac{L^2}{2\mu r^2} - \frac{k}{r} = \frac{ak(1-e^2)}{2r^2} - \frac{k}{r}$$

and Eq. (2.48) we easily obtain

$$t = \sqrt{\frac{\mu}{2k}} \int_p^r \frac{s\,\mathrm{d}s}{\sqrt{-\frac{s^2}{2a} + s - \frac{a}{2}(1 - e^2)}},$$

where we have chosen as lower limit in the integral the orbit's perihelion p so that $\mathrm{d}t/\mathrm{d}r > 0$ till r reaches the next aphelion (i.e., for $0 \leqslant t \leqslant \tau/2$). Since

$$P(s) := -\frac{s^2}{2a} + s - \frac{a}{2}(1 - e^2) = \frac{1}{2a}\left[a^2e^2 - (s - a)^2\right]$$

we perform the natural change of variable $s = a(1 - e\cos\beta)$ in the integral, so that $\mathrm{d}s = ae\sin\beta\,\mathrm{d}\beta$ and

$$P(s) = \frac{ae^2}{2}\sin^2\beta\,.$$

Taking into account that $p = a(1 - e)$ implies that $\beta = 0$ when $r = p$ we finally obtain

$$t = \sqrt{\frac{\mu a^3}{k}} \int_0^\psi (1 - e\cos\beta)\,\mathrm{d}\beta = \frac{1}{\omega}(\psi - e\sin\psi)$$

with $r = a(1 - e\cos\psi)$, as claimed.

Equation (2.50) is usually called in the literature *Kepler's equation*. The geometric meaning of ψ can be understood from Fig. 2.10. Indeed, since the point P on the elliptic orbit and the point P' on the circle of radius a centered at C lie on the same vertical they must have the same abscissa (measured from the focus F at the origin), namely

$$a\cos\psi - c = a(\cos\psi - e) = r\cos\varphi\,.$$

On the other hand, from the equation of the ellipse in polar coordinates

$$r = \frac{a(1 - e^2)}{1 + e\cos\varphi}$$

we obtain

$$r = a(1 - e^2) - er\cos\varphi = a(1 - e^2) - ea(\cos\psi - e) = a(1 - e\cos\psi),$$

namely Eq. (2.51).

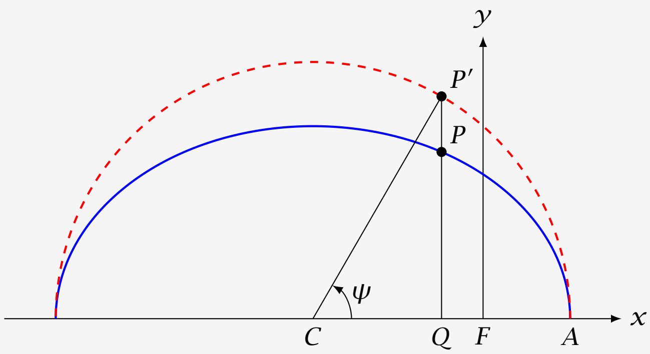


Figure 2.10. Eccentric anomaly ψ of point P on an elliptic orbit (solid blue line). Note that $\overline{CF} = c = ea$, $\overline{CA} = a$, and the dashed red curve represents the circle of radius a centered at C.

Note. Strictly speaking, we have proved (2.50)-(2.51) only for one half-period, i.e., for $0 \leqslant t \leqslant \tau/2$ or equivalently $0 \leqslant \psi \leqslant \pi$. However, using the identities $r(t) = r(-t) = r(t + k\tau)$ (with $k \in \mathbb{Z}$) it is straightforward to show that the latter equations are in fact valid for *all* values of t and ψ. Indeed, if $\psi \mapsto -\psi$ then r does not change in Kepler's equation and $t \mapsto -t$, which is again consistent with the identity $r(t) = r(-t)$. Thus Kepler's equation can be extended to the interval $-\tau/2 \leqslant t \leqslant \tau/2$, i.e., to a whole period of the motion. Likewise, when ψ changes by $2k\pi$ (with $k \in \mathbb{Z}$) r does not change in Kepler's equation and t changes by $2k\pi/\omega = k\tau$, which is consistent with the identity $r(t + k\tau) = r(\tau)$. This establishes Kepler's equation for an arbitrary time $t \in \mathbb{R}$.

Exercise. If ψ is the angle ACP' in Fig. (2.10), derive Kepler's equation using the law of areas.

Solution. According to the law of areas, if t is the time taken by the planet to travel from the periapsis A to the point P in Fig. 2.10 we have

$$\frac{PFA}{\pi ab} = \frac{t}{\tau} = \frac{\omega t}{2\pi} \quad \Longrightarrow \quad PFA = \frac{1}{2}ab \cdot \omega t\,,$$

where PFA denotes the area swept by the planet's position vector as it travels from A to P along its orbit. On the other hand, from the latter figure it follows that

$$PFA = PQA - PQF, \tag{2.52}$$

where PQF and PQA respectively denote the areas of the triangle PQF and the sector delimited by the elliptic arc AP and the segments PQ and QA. Since the the ellipse in Fig. (2.10) is obtained dilating the circle of radius a and center C (dashed red line in Fig. 2.10) by b/a in the vertical direction, we have

$$\begin{aligned} PQA &= \frac{b}{a}P'QA = \frac{b}{a}(P'AC - P'QC) = \frac{b}{a}\left(\frac{a^2\psi}{2} - \frac{a^2}{2}\sin\psi\cos\psi\right) \\ &= \frac{1}{2}ab(\psi - \sin\psi\cos\psi). \end{aligned} \tag{2.53}$$

Here $P'QA$ is the area of the circular sector delimited by the arc AP' and the segments $P'Q$ and QA, $P'AC$ is the area of the circular sector determined by the points P', A and C, and $P'QC$ is the area of the triangle $P'QC$. On the other hand,

$$\begin{aligned} &QP = \frac{b}{a}QP' = b\sin\psi, \quad QF = CF - CQ = ae - a\cos\psi \\ &\qquad \Longrightarrow \quad PQF = \frac{1}{2}b\sin\psi \cdot a(e - \cos\psi). \end{aligned} \tag{2.54}$$

Combining Eqs. (2.52)-(2.54) we finally obtain

$$\frac{1}{2}ab \cdot \omega t = \frac{1}{2}ab(\psi - \sin\psi\cos\psi) - \frac{1}{2}ab\sin\psi(e - \cos\psi) = \frac{1}{2}ab(\psi - e\sin\psi),$$

which yields Kepler's equation (2.50).

Exercise. Find the analogue of Kepler's equation for *hyperbolic* orbits of the Kepler potential.

Solution. In this case the energy is positive, and can be expressed as $k/(2a)$ if we define $a = \alpha/(e^2 - 1)$ (cf. Eq. (2.29)). Proceeding as for elliptic orbits and taking into account that

$$L^2 = k\mu\alpha = k\mu a(e^2 - 1)$$

we arrive at the formula

$$t = \sqrt{\frac{\mu}{2k}} \int_p^r \frac{s\,\mathrm{d}s}{\sqrt{P(s)}}, \tag{2.55}$$

where now

$$P(s) := \frac{s^2}{2a} + s - \frac{a}{2}(e^2 - 1) = \frac{1}{2a}\left[(s+a)^2 - a^2e^2\right]$$

and $s = p = \alpha/(e+1) = a(e-1)$ is obtained for $\beta = 0$. This suggests performing the change of variable $s = a(e\cosh\beta - 1)$, so that $\mathrm{d}s = ae\sinh\beta\,\mathrm{d}\beta$, $P(s) = a^2e^2\sinh^2\beta$ and

$$t = \sqrt{\frac{\mu}{k}}\, a^{3/2} \int_0^\psi (e\cosh\beta - 1)\,\mathrm{d}\beta = \frac{1}{\omega}(e\sinh\psi - \psi),$$

with $\omega = \sqrt{k/\mu}\,a^{-3/2}$ and $r = a(e\cosh\psi - 1)$. The latter is the analogue of Kepler's equation for hyperbolic orbits. The motion of the radial coordinate is obtained inverting the relation $\omega t = (e\sinh\psi - \psi)$ for t as a function of ψ. This is possible, since the RHS of the previous equation has derivative $e\cosh\psi - 1 \geqslant e - 1 > 0$, and is therefore a monotonically increasing function of ψ. Strictly speaking, we have established the equations

$$\boxed{\omega t = e\sinh\psi - \psi, \qquad r = a(e\cosh\psi - 1)} \tag{2.56}$$

for $t \geqslant 0$, or equivalently $\psi \geqslant 0$. However, from the identity $r(t) = r(-t)$ and the fact that $t \mapsto -t$ implies $\psi \mapsto -\psi$ and $r \mapsto r$ in Kepler's equation we deduce that Eqs. (2.56) hold for all real values of t and β.

Exercise. Find the equation of the orbits in the *repulsive* $1/r$ potential $V(r) = k/r$ (with $k > 0$).

Solution. Binet's equation is in this case

$$u'' + u = -\frac{k\mu}{L^2},$$

whose general solution can be taken as

$$u = \frac{k\mu}{L^2}\left(e\cos(\varphi - \varphi_0) - 1\right).$$

Again, we can assume w.l.o.g. that $e > 0$ and $\varphi_0 = 0$ (by an appropriate choice of the x axis). In fact, since $r > 0$ we must have $e > 1$. We can thus write

$$\boxed{r = \frac{\alpha}{e\cos\varphi - 1}, \qquad \alpha := \frac{L^2}{k\mu}.}$$

All the orbits in this case are clearly *unbounded*, since $r \to \infty$ for $\varphi \to \pm\arccos(1/e)$. It is also clear that the polar angle of the periapsis is $\varphi = 0$, and its distance to the origin is equal to $\alpha/(e-1)$. To find the Cartesian equation of the orbit we multiply both sides of the previous equation by $e\cos\varphi - 1$, thus obtaining

$$r = ex - \alpha \quad \Longrightarrow \quad x^2 + y^2 = e^2x^2 - 2\alpha ex + \alpha^2 \quad \Longleftrightarrow \quad (e^2-1)x^2 - 2\alpha ex - y^2 = -\alpha^2,$$

or equivalently

$$(e^2 - 1)\left(x - \frac{\alpha e}{e^2 - 1}\right)^2 - y^2 = -\alpha^2 + \frac{\alpha^2 e^2}{e^2 - 1} = \frac{\alpha^2}{e^2 - 1}.$$

This is the equation of a *hyperbola* with center $(\alpha e/(e^2-1), 0)$ and semiaxes

$$a = \frac{\alpha}{e^2-1}, \qquad b = \frac{\alpha}{\sqrt{e^2-1}}$$

In fact, since $ex = \alpha + r > 0$ the orbit is the branch of the latter hyperbola in the half-plane $x > 0$. The focal distance of the hyperbola is given by

$$c = \sqrt{a^2+b^2} = \frac{\alpha^2}{e^2-1}\sqrt{1+(e^2-1)} = \frac{\alpha e}{e^2-1} = ea,$$

and hence its eccentricity c/a is equal to the parameter e. This implies that the center of the hyperbola is the point $(c, 0)$, and hence the foci are the points $(0,0)$ and $(0, 2c)$. The energy is now given by

$$E = \frac{L^2}{2\mu}(u'^2 + u^2) + ku = \frac{\mu k^2}{2L^2}\left[e^2\sin^2\varphi + (e\cos\varphi - 1)^2 + 2(e\cos\varphi - 1)\right] = \frac{\mu k^2}{2L^2}(e^2-1),$$

or, using the equation for the semi-major axis,

$$E = \frac{\mu k^2 \alpha}{2aL^2} = \frac{k}{2a}.$$

Finally, the eccentricity can be related to the energy and angular momentum of the orbit through the equation

$$e = \sqrt{1 + \frac{2EL^2}{\mu k^2}},$$

which is the same as for the Kepler potential.

Exercise. Find the analogue of Kepler's equation for the orbits of the repulsive Kepler potential.

Solution. Proceeding as for the Kepler problem and using the formulas

$$E = \frac{k}{2a}, \qquad L^2 = k\mu\alpha = k\mu a(e^2-1)$$

derived in the previous exercise we again arrive at Eq. (2.55), where now

$$P(s) := \frac{s^2}{2a} - s - \frac{a}{2}(e^2-1) = \frac{1}{2a}\left[(s-a)^2 - a^2e^2\right].$$

We therefore perform the change of variable $s = a(1 + e\cosh\psi)$, obtaining

$$t = \sqrt{\frac{\mu}{k}}\, a^{3/2}\int_0^{\psi}(e\cosh\beta + 1)\,\mathrm{d}\beta = \frac{1}{\omega}(\psi + e\sinh\psi).$$

This is the analogue of Kepler's equation for this potential.

3 Lagrangian and Hamiltonian mechanics

3.1 Introduction to the calculus of variations

3.1.1 Fundamental problem of the calculus of variations

The fundamental problem of the **calculus of variations** (in its simplest version) is that of finding the extrema (i.e., maxima or minima) of a function of the form

$$\boxed{F[y] = \int_{x_1}^{x_2} f\big(x, y(x), y'(x)\big)\,dx} \qquad \left(' = \frac{d}{dx}\right), \tag{3.1}$$

where $f : \mathbb{R}^3 \to \mathbb{R}$ is of class C^2 (i.e., twice continuously differentiable). The domain of F consists of the functions $y : [x_1, x_2] \to \mathbb{R}$ (also assumed to be of class C^2) that satisfy the **boundary conditions**

$$\boxed{y(x_1) = y_1\,, \qquad y(x_2) = y_2} \tag{3.2}$$

with $y_1, y_2 \in \mathbb{R}$ *fixed*. From the mathematical viewpoint,

$$F : C_0^2([x_1, x_2]) \to \mathbb{R}$$

is therefore a function whose domain is the space $C_0^2([x_1, x_2])$ of scalar functions $y : [x_1, x_2] \to \mathbb{R}$ of class C^2 in the interval $[x_1, x_2]$ satisfying the conditions (3.2). In other words, the application F assigns to each function $y : [x_1, x_2] \to \mathbb{R}$, which can be identified with its **graph**

$$\{(x, y(x)) : x_1 \leqslant x \leqslant x_2\} \subset \mathbb{R}^2\,,$$

the number given by the RHS of Eq. (3.1). A function like F, whose domain is a set of functions, is usually called a **functional**. The function f appearing in Eq. (3.1) is called the **density** of the corresponding functional F.

Many interesting problems in mathematics and physics reduce to finding the extrema (maxima or minima) of an appropriate functional of the form (3.1)-(3.2). We shall list below a few of the most noted ones.

Example 3.1. What is the shortest curve joining two fixed points on a plane?

Let us denote the two fixed points by (x_i, y_i) $(i = 1, 2)$, with $x_1 \neq x_2$ (this can always be arranged by suitably choosing the axes). If we restrict ourselves, for the sake of simplicity, to plane curves that are graphs of functions $y : [x_1, x_2] \to \mathbb{R}$, the problem considered is equivalent to finding the minimum of the *length functional*

$$\boxed{F[y] = \int_{x_1}^{x_2} \sqrt{1 + y'(x)^2}\,dx}$$

with the condition (3.2).

Example 3.2. *The brachistochrone problem.* A particle of mass m is forced to move in a vertical plane along a curve with fixed endpoints (x_1, y_1) and (x_2, y_2), with $x_1 < x_2$ and $y_1 > y_2$. For which curve is the time taken by the particle to travel between both endpoints a *minimum*?

If we neglect friction, the reaction force exerted by the curve on the particle is normal to the curve at each point. Hence the reaction force does no work, and consequently energy is conserved:

$$\frac{1}{2}mv^2 + mgy = E\,.$$

If the curve in question is the graph of a function $y(x)$, the differential of time along the curve is given by

$$\mathrm{d}t = \frac{\mathrm{d}s}{v} = \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g\left(\frac{E}{mg} - y(x)\right)}}\,\mathrm{d}x\,.$$

Thus the problem proposed is equivalent to finding the minimum of the functional (proportional to the travel time)

$$\boxed{F[y] = \int_{x_1}^{x_2} \sqrt{\frac{1 + y'(x)^2}{y_0 - y(x)}}\,\mathrm{d}x} \qquad \left(y_0 := \frac{E}{mg}\right) \tag{3.3}$$

with the condition (3.2). Note that from energy conservation it follows that $y \leqslant y_0$, and that the energy (and hence y_0) depends only on the particle's initial velocity v_0. Indeed, $E = \frac{1}{2}mv_0^2 + mgy_1$, or equivalently $y_0 = y_1 + \frac{v_0^2}{2g}$; in particular, $y_0 = y_1$ if the particle is initially at rest.

Example 3.3. *Fermat's principle.* What is the trajectory followed by a light ray traveling from a point (x_1, y_1) to a second point (x_2, y_2) in a flat optical medium with refractive index $n(x, y)$?

According to **Fermat's principle** (in the approximation of geometric optics), the trajectory of the light ray joining the points (x_1, y_1) and (x_2, y_2) is the curve for which the time taken by light to cover the distance between both points is *minimum*. Suppose, again, that $x_1 \neq x_2$, and that the trajectory is the graph of a function $y(x)$. By definition of index of refraction, the speed of light at a point (x, y) of the medium is given by

$$v(x, y) = \frac{c}{n(x, y)}\,,$$

where c is the speed of light *in vacuo*. Since

$$\mathrm{d}t = \frac{\mathrm{d}s}{v(x, y)} = n(x, y)\,\frac{\mathrm{d}s}{c}\,,$$

the problem proposed is equivalent to determining the minimum of the functional (proportional to light's travel time)

$$\boxed{F[y] = \int_{x_1}^{x_2} n(x, y(x))\sqrt{1 + y'(x)^2}\,\mathrm{d}x\,,} \tag{3.4}$$

again with the condition (3.2). The functional (3.4), which has dimensions of length, is called *optical length*. Note that if the refractive index is *constant* F is proportional to the length functional of the first example, and thus the path followed by light rays in this case is the shortest curve joining the points (x_1, y_1) and (x_2, y_2). Likewise, if the refractive index is proportional to $(y_0 - y)^{-1/2}$ the path followed by light is the brachistochrone of the previous example.

3.1.2 Euler–Lagrange equations

In order to solve the fundamental problem of the calculus of variations formulated in the previous subsection, we shall proceed in essentially the same way as in real analysis when determining the extrema of an ordinary function $F : \mathbb{R} \to \mathbb{R}$. The key idea is that in both cases the extrema are points (functions, in this case) for which the *variation* of the function when we infinitesimally increase its argument vanishes at first order.

More precisely, suppose that $y(x)$ is an extremum (maximum or minimum) of the functional (3.1) with the condition (3.2). Let $\eta(x)$ be an arbitrary function (of class C^2) satisfying the conditions

$$\eta(x_1) = \eta(x_2) = 0\,, \tag{3.5}$$

so that for all $\varepsilon \in \mathbb{R}$ the function $y_\varepsilon := y + \varepsilon\eta$ satisfies the boundary conditions (3.2). The functions $y_\varepsilon(x)$ (with $\varepsilon \in \mathbb{R}$) form a **one-parameter family** containing the extremum $y(x)$ for $\varepsilon = 0$. More informally, if ε is small we can think of $y_\varepsilon(x)$ as a small *variation* of the extremum $y(x)$. In any case, if we restrict the functional F to these functions we obtain the *scalar function of one variable*

$$g(\varepsilon) := F[y_\varepsilon] = \int_{x_1}^{x_2} f\big(x, y(x) + \varepsilon\eta(x), y'(x) + \varepsilon\eta'(x)\big)\,\mathrm{d}x,$$

which by construction has an extremum at $\varepsilon = 0$. We know that the *necessary* (although in general not sufficient) condition for this to happen is that $g'(0) = 0$. Since

$$\begin{aligned} g'(\varepsilon) &= \int_{x_1}^{x_2} \frac{\partial}{\partial\varepsilon} f\big(x, y(x) + \varepsilon\eta(x), y'(x) + \varepsilon\eta'(x)\big)\,\mathrm{d}x \\ &= \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\big(x, y(x) + \varepsilon\eta(x), y'(x) + \varepsilon\eta(x)\big)\,\eta(x) + \frac{\partial f}{\partial y'}\big(x, y(x) + \varepsilon\eta(x), y'(x) + \varepsilon\eta(x)\big)\eta'(x)\right]\mathrm{d}x \end{aligned}$$

we have

$$g'(0) = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\big(x, y(x), y'(x)\big)\,\eta(x) + \frac{\partial f}{\partial y'}\big(x, y(x), y'(x)\big)\,\eta'(x)\right]\mathrm{d}x\,.$$

Hence, if the function $y(x)$ is an extremum of the functional F with the boundary conditions (3.2) it must satisfy

$$\int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}\big(x, y(x), y'(x)\big)\,\eta(x) + \frac{\partial f}{\partial y'}\big(x, y(x), y'(x)\big)\,\eta'(x)\right]\mathrm{d}x = 0 \tag{3.6}$$

for *any* function $\eta(x)$ satisfying (3.5). Equation (3.6) can be simplified integrating by parts the last term, since

$$\begin{aligned} &\int_{x_1}^{x_2} \frac{\partial f}{\partial y'}\big(x, y(x), y'(x)\big)\,\eta'(x)\,\mathrm{d}x \\ &\quad = \frac{\partial f}{\partial y'}\big(x, y(x), y'(x)\big)\,\eta(x)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta(x)\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\big(x, y(x), y'(x)\big)\right)\mathrm{d}x \\ &\quad = -\int_{x_1}^{x_2} \eta(x)\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\big(x, y(x), y'(x)\big)\right)\mathrm{d}x\,, \end{aligned}$$

where we have taken into account condition (3.5). Substituting back into Eq. (3.6) we finally obtain

$$g'(0) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} F[y_\varepsilon] = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y}(x, y(x), y'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}(x, y(x), y'(x))\right)\right] \eta(x)\,\mathrm{d}x = 0\,. \tag{3.7}$$

Since this condition must be verified for *any* function η satisfying (3.5), the term in square brackets must vanish identically in the interval $[x_1, x_2]$, and hence *the extremum $y(x)$ must satisfy the* **Euler–Lagrange equation**

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}(x, y(x), y'(x))\right) - \frac{\partial f}{\partial y}(x, y(x), y'(x)) = 0\,, \qquad \forall x \in [x_1, x_2]\,. \tag{3.8}$$

- Clearly, the argument leading to the Euler–Lagrange equation (3.8) is still valid if $y(x)$ is only a *local* extremum of the functional F.

- It is important to remember that the Euler–Lagrange equation (3.8) is a *necessary*, but in general *not sufficient*, condition for the function $y(x)$ to be an extremum of the functional F. In fact, the solutions of this differential equation can be regarded as the *critical points* of F (in the same way as the points at which the derivative of a function $F : \mathbb{R} \to \mathbb{R}$ vanishes are the critical points of the function). Indeed, what the previous argument shows is that the functional $F[y]$ is *stationary* (i.e., approximately constant) when $y(x)$ is a solution of the Euler–Lagrange equations. For this reason, the functions $y(x)$ satisfying the latter equation are usually called *stationary points* of the functional (3.1).

- If (as we are assuming throughout) the function f is of class C^2, Eq. (3.8) can be written in the equivalent form
$$\frac{\partial^2 f}{\partial y'^2}\, y'' + \frac{\partial^2 f}{\partial y \partial y'}\, y' + \frac{\partial^2 f}{\partial x \partial y'} - \frac{\partial f}{\partial y} = 0\,.$$
In particular, if the condition
$$\frac{\partial^2 f}{\partial y'^2} \neq 0$$
is satisfied the previous equation is a *second-order ordinary differential equation* in the unknown function y. To find the stationary points of the functional F, we must supplement this equation with the *boundary conditions* (3.2).

- Multiplying the left-hand side (LHS) of the Euler–Lagrange equation (3.8) by y' we obtain
$$y'\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} - y'\frac{\partial f}{\partial y} = \frac{\mathrm{d}}{\mathrm{d}x}\left(y'\frac{\partial f}{\partial y'}\right) - y'\frac{\partial f}{\partial y} - y''\frac{\partial f}{\partial y'} = \frac{\mathrm{d}}{\mathrm{d}x}\left(y'\frac{\partial f}{\partial y'} - f\right) + \frac{\partial f}{\partial x}\,.$$
Hence if $y' \neq 0$ the Euler–Lagrange equation can be written in the equivalent form
$$\frac{\mathrm{d}}{\mathrm{d}x}\left(y'\frac{\partial f}{\partial y'} - f\right) + \frac{\partial f}{\partial x} = 0\,. \tag{3.9}$$
In particular, if f does not explicitly depend on x (that is, if it is a function of y and y' only), the function in parentheses in the LHS of Eq. (3.9) is conserved:
$$\frac{\partial f}{\partial x} = 0 \quad \Longrightarrow \quad h := y'\frac{\partial f}{\partial y'} - f = \text{const.} \tag{3.10}$$
It is said in this case that the function h is a **first integral** of the Euler–Lagrange equation (3.8), since when h is conserved the *second-order* equation (3.8) is equivalent to the *first-order* equation (3.10). We shall call h the **energy integral**, since in many mechanical problems it is equal

to the system's mechanical energy[1]. Likewise, if f does not depend on y it follows from the Euler–Lagrange equation that the partial derivative of f with respect to y' is conserved:

$$\boxed{\frac{\partial f}{\partial y} = 0 \quad \Longrightarrow \quad \frac{\partial f}{\partial y'} = \text{const.}} \tag{3.11}$$

Example 3.4. The problems set forth in Examples 3.1-3.2 are easily solved using the Euler–Lagrange equation and its conservation laws (3.10)-(3.11). Indeed, for the length functional Eq. (3.11) yields

$$\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1+y'^2}} = \text{const.} \quad \Longrightarrow \quad y' = \text{const.}$$

Therefore $y(x) = ax + b$, where the constants a and b must be chosen so that conditions (3.2) are satisfied. Hence the curve of minimum[a] length is the line segment joining the given points.

As to the brachistochrone functional (3.3), Eq. (3.10) reads

$$\left(\frac{y'^2}{\sqrt{1+y'^2}} - \sqrt{1+y'^2}\right)(y_0 - y)^{-1/2} = -(1+y'^2)^{-1/2}(y_0-y)^{-1/2} = \text{const.}$$
$$\Longrightarrow \quad (y_0 - y)(1+y'^2) = 2a\,,$$

with $a > 0$ constant. Hence

$$y' = \pm\sqrt{\frac{2a}{y_0 - y} - 1} = \pm\sqrt{\frac{2a - y_0 + y}{y_0 - y}}\,,$$

and thus

$$x - x_0 = \pm\int\sqrt{\frac{y_0 - y}{2a - y_0 + y}}\,\mathrm{d}y\,,$$

with x_0 constant. Performing the change of variable

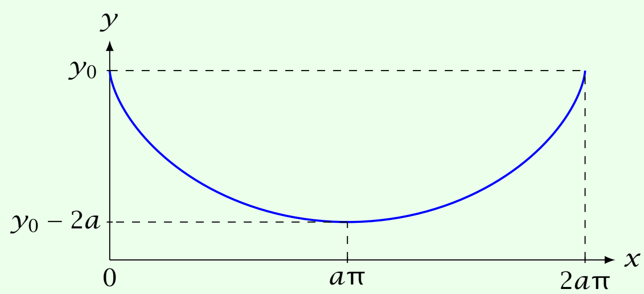
$$y_0 - y = 2a\sin^2\theta$$

we obtain

$$x - x_0 = \mp 4a\int\frac{\sin^2\theta\cos\theta}{\cos\theta}\,\mathrm{d}\theta = \mp 4a\int\sin^2\theta\,\mathrm{d}\theta = \mp 2a\int(1-\cos 2\theta)\,\mathrm{d}\theta = \mp a\,(2\theta - \sin 2\theta).$$

The parametric equations of the sought for curve are therefore

$$\boxed{x = x_0 \mp a(2\theta - \sin 2\theta)\,, \qquad y = y_0 - 2a\sin^2\theta = y_0 - a\,(1 - \cos 2\theta)\,.} \tag{3.12}$$



[1]The energy integral is also called *Jacobi integral* by some authors.

Figure 3.1. Arc of the cycloid (3.12) with $x_0 = 0$ and "+" sign in a period $0 \leqslant \theta \leqslant \pi$.

These are the equations of an inverted *cycloid*[b] traced out by a circle of radius a (cf. Fig. 3.1), where the constants x_0 and a must again be determined imposing the boundary conditions (3.2).

[a] *Stricto sensu*, we have only shown that the straight line is a *stationary point* of the length functional.

[b] Note that the double sign can actually be omitted, since the points on the curve corresponding to the "−" sign can be obtained from those with the "+" sign changing θ by $-\theta$.

Example 3.5. The Euler–Lagrange equation for the optical length functional (3.4) reads

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{n(x,y)\,y'}{\sqrt{1+y'^2}}\right) - \sqrt{1+y'^2}\,\frac{\partial n(x,y)}{\partial y} = 0\,.$$

This equation can be expressed in a more compact form taking into account that if s is the arc length along the path of the light ray then

$$\frac{\mathrm{d}}{\mathrm{d}s} = \left(\frac{\mathrm{d}s}{\mathrm{d}x}\right)^{-1}\frac{\mathrm{d}}{\mathrm{d}x} = (1+y'^2)^{-1/2}\,\frac{\mathrm{d}}{\mathrm{d}x}\,.$$

In this way we obtain the equation

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}s}\left(n(x,y)\,\frac{\mathrm{d}y}{\mathrm{d}s}\right) = \frac{\partial n(x,y)}{\partial y}\,.}$$

For instance, if the refractive index does not depend on the y coordinate the previous equation yields

$$n(x)\,\frac{\mathrm{d}y}{\mathrm{d}s} = k \quad\Longrightarrow\quad \frac{n^2(x)\,y'^2}{1+y'^2} = k^2 \quad\Longrightarrow\quad y' = \pm\frac{k}{\sqrt{n^2(x)-k^2}}\,,$$

where k is a constant. Thus in this case the equation of the light rays is

$$\boxed{y = y_0 \pm k\int\frac{\mathrm{d}x}{\sqrt{n^2(x)-k^2}}\,.}$$

In particular, if

$$n(x) = \frac{n_0}{x} \qquad (n_0 > 0\,,\quad x > 0)$$

we have

$$y - y_0 = \pm k\int\frac{x\,\mathrm{d}x}{\sqrt{n_0^2-k^2x^2}} = \mp\frac{1}{k}\sqrt{n_0^2-k^2x^2} \quad\Longrightarrow\quad x^2+(y-y_0)^2 = \frac{n_0^2}{k^2}\,.$$

Therefore the paths of the light rays are in this case arcs of *circles* whose centers lie on the y axis.

We shall next consider a more general version of the fundamental problem of the calculus of variations, in which the functional F depends on n scalar functions $y_1,\dots,y_n$ of one real variable x. Equivalently (and more advantageously from the notational point of view), we can regard F as a function of a single *vector-valued* function $\mathbf{y} := (y_1,\dots,y_n) : \mathbb{R} \to \mathbb{R}^n$. More precisely, consider the functional

$$F[\mathbf{y}] = \int_{x_1}^{x_2} f\big(x,\mathbf{y}(x),\mathbf{y}'(x)\big)\,\mathrm{d}x\,, \tag{3.13}$$

whose domain is the space of functions $\mathbf{y} : [x_1, x_2] \to \mathbb{R}^n$ of class C^2 on the interval $[x_1, x_2]$ satisfying conditions similar to (3.2):

$$\boxed{\mathbf{y}(x_1) = \mathbf{y}_1\,, \qquad \mathbf{y}(x_2) = \mathbf{y}_2\,,} \tag{3.14}$$

for certain fixed vectors $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^n$.

As before, in order to find the extrema of the functional (3.13) subject to the boundary conditions (3.14) we consider a variation

$$\mathbf{y}_\varepsilon(x) = \mathbf{y}(x) + \varepsilon\boldsymbol{\eta}(x)$$

about a hypothetical extremum $\mathbf{y}(x)$, where the vector-valued function $\boldsymbol{\eta} =: (\eta_1, \dots, \eta_n)$ must satisfy

$$\boldsymbol{\eta}(x_1) = \boldsymbol{\eta}(x_2) = 0 \tag{3.15}$$

so that $\mathbf{y}_\varepsilon$ verifies conditions (3.14) for all ε. Restricting the functional F to the modified extremum $\mathbf{y}_\varepsilon$ we obtain, as before, the scalar function of one variable

$$g(\varepsilon) := F[\mathbf{y}_\varepsilon] = \int_{x_1}^{x_2} f(x, \mathbf{y}_\varepsilon(x), \mathbf{y}'_\varepsilon(x))\,\mathrm{d}x\,, \tag{3.16}$$

whose derivative at $\varepsilon = 0$ must vanish. Computing this derivative and integrating by parts, taking into account conditions (3.15), we easily obtain

$$g'(0) = \int_{x_1}^{x_2} \sum_{i=1}^{n} \left[\frac{\partial f}{\partial y_i}(x, \mathbf{y}(x), \mathbf{y}'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'_i}(x, \mathbf{y}(x), \mathbf{y}'(x)) \right] \eta_i(x)\,. \tag{3.17}$$

Since this expression must vanish identically for all functions η_i satisfying conditions (3.15), we conclude that the extrema of the functional (3.13) must verify the n **Euler–Lagrange equations**

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'_i} - \frac{\partial f}{\partial y_i} = 0\,, \qquad i = 1, \dots, n\,. \tag{3.18}$$

Again, the latter equations are only *necessary* for the function $\mathbf{y}(x)$ to be an extremum of the functional F. Indeed, the solutions of Eqs. (3.18) are actually the *critical* or *stationary points* of the functional (3.13).

- Expanding Eqs. (3.18) we obtain

$$\sum_{j=1}^{n} \frac{\partial^2 f}{\partial y'_i \partial y'_j} y''_j + \sum_{j=1}^{n} \frac{\partial^2 f}{\partial y_j \partial y'_i} y'_j + \frac{\partial^2 f}{\partial x \partial y'_i} - \frac{\partial f}{\partial y_i} = 0\,, \qquad i = 1, \dots, n\,.$$

Hence if the *Hessian* of the density f with respect to the variables y'_i does not vanish identically, i.e., if

$$\det\left(\frac{\partial^2 f}{\partial y'_i \partial y'_j}\right)_{1\leqslant i,j\leqslant n} \not\equiv 0\,,$$

the Euler–Lagrange (3.18) equations are a *system of* n *second-order ordinary differential equations in the* n *unknown scalar functions* $y_i(x)$ ($i = 1, \dots, n$), which must be supplemented by the $2n$ boundary conditions (3.14).

- Multiplying the LHS of the Euler–Lagrange (3.18) by y'_i and summing over i we obtain

$$\sum_{i=1}^{n} y'_i \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'_i} - \sum_{i=1}^{n} y'_i \frac{\partial f}{\partial y_i} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{i=1}^{n} y'_i \frac{\partial f}{\partial y'_i} - f\right) + \frac{\partial f}{\partial x} = 0\,.$$

Hence *if f does not depend explicitly on the variable x the function*

$$h := \sum_{i=1}^{n} y_i' \frac{\partial f}{\partial y_i'} - f = \mathbf{y}' \frac{\partial f}{\partial \mathbf{y}'} - f$$

is conserved. As in the scalar case, h is usually called the **energy** (or Jacobi) **integral.** It is also evident that if the density f is independent of the variable y_i, the derivative of f with respect to y_i' is conserved:

$$\frac{\partial f}{\partial y_i} = 0 \quad \Longrightarrow \quad \frac{\partial f}{\partial y_i'} = \text{const.}$$

Example 3.6. Let us find the equation of the paths followed by light rays in an optical (three-dimensional) medium with refractive index $n(\mathbf{r})$.

According to Fermat's principle, the trajectory $\mathbf{r} = \mathbf{r}(u)$ of the light ray joining two points $\mathbf{r}_1, \mathbf{r}_2 \in \mathbb{R}^3$ (where $u \in [u_1, u_2]$ is any parameter along the path) must minimize the time taken by light to cover the distance between both points. Since

$$\frac{\mathrm{d}s}{\mathrm{d}u} = \sqrt{\mathbf{r}'^2(u)}\,,$$

where the prime denotes derivative with respect to u, we have

$$\mathrm{d}t = \frac{\mathrm{d}t}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}u}\,\mathrm{d}u = \frac{1}{v}\sqrt{\mathbf{r}'^2(u)}\,\mathrm{d}u = \frac{n(\mathbf{r}(u))}{c}\sqrt{\mathbf{r}'^2(u)}\,\mathrm{d}u\,.$$

Thus the sought for trajectory must minimize the optical length functional (proportional to the travel time)

$$F[\mathbf{r}] = \int_{u_1}^{u_2} n(\mathbf{r}(u))\sqrt{\mathbf{r}'^2(u)}\,\mathrm{d}u$$

with the boundary conditions

$$\mathbf{r}(u_1) = \mathbf{r}_1\,, \qquad \mathbf{r}(u_2) = \mathbf{r}_2\,.$$

(Note that in this example u plays the role of x and $\mathbf{r}$ that of $\mathbf{y}$.) The Euler–Lagrange equations for this functional are

$$\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\partial}{\partial x_i'}\left(n\sqrt{\mathbf{r}'^2}\right)\right) - \sqrt{\mathbf{r}'^2}\,\frac{\partial n}{\partial x_i} = \frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{n\,x_i'}{\sqrt{\mathbf{r}'^2}}\right) - \sqrt{\mathbf{r}'^2}\,\frac{\partial n}{\partial x_i} = 0\,, \qquad i = 1, 2, 3\,,$$

where $\mathbf{r} = (x_1, x_2, x_3)$. Taking into account that

$$\frac{1}{\sqrt{\mathbf{r}'^2}}\frac{\mathrm{d}}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}s}\,,$$

the previous equations can be written in the following more geometric fashion:

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right) = \frac{\partial n}{\partial \mathbf{r}}\,.$$

This is the fundamental equation of geometric optics. For example, if the index of refraction depends only on r (that is, if the optical medium is *spherically symmetric*) then

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\mathbf{r}\times n\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right) = \mathbf{r}\times\frac{\partial n}{\partial \mathbf{r}} = n'(r)\,\mathbf{r}\times\frac{\mathbf{r}}{r} = 0\,.$$

Hence in this case *the path of the light ray is contained in a plane passing through the origin of coordinates* (perpendicular to the constant vector $n\mathbf{r}\times\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}$).

3.1.3 Variation and variational derivative

Let uus denote by

$$\boxed{\delta\mathbf{y}(x) := \mathbf{y}_\varepsilon(x) - \mathbf{y}(x) = \varepsilon\boldsymbol{\eta}(x)}$$

the **variation** of the function $\mathbf{y}(x)$, where the vector-valued function $\boldsymbol{\eta}$ satisfies the boundary conditions (3.15). The change in the functional F when its argument $\mathbf{y}$ is incremented by $\delta\mathbf{y}$ is then (using the notation of Eq. (3.16))

$$F[\mathbf{y}+\delta\mathbf{y}] - F[\mathbf{y}] = F[\mathbf{y}_\varepsilon] - F[\mathbf{y}] = g(\varepsilon) - g(0)\,.$$

To first order in the small parameter ε, this change is given by

$$\varepsilon g'(0) = \varepsilon\,\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} F[\mathbf{y}_\varepsilon] =: \delta F[\mathbf{y}]\,,$$

so that (by definition of derivative) we have

$$F[\mathbf{y}+\delta\mathbf{y}] - F[\mathbf{y}] = \varepsilon\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\bigg|_{\varepsilon=0} F[\mathbf{y}_\varepsilon]\right) + o(\varepsilon) = \delta F(\mathbf{y}) + o(\varepsilon)\,.$$

The functional $\delta F[\mathbf{y}]$ is called the **variation** of F at $\mathbf{y}$. From Eq. (3.17) it follows that we can write this variation as

$$\boxed{\delta F[\mathbf{y}] = \int_{x_1}^{x_2}\frac{\delta f}{\delta\mathbf{y}}(x,\mathbf{y}(x),\mathbf{y}'(x))\cdot\delta\mathbf{y}(x)\,\mathrm{d}x\,,} \tag{3.19}$$

where

$$\boxed{\frac{\delta f}{\delta\mathbf{y}} := \frac{\partial f}{\partial\mathbf{y}} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial\mathbf{y}'}\,.} \tag{3.20}$$

The (n-component) vector-valued function $\frac{\delta f}{\delta\mathbf{y}}(x,\mathbf{y},\mathbf{y}')$ is called the **variational derivative** of the density f with respect to the function $\mathbf{y}(x)$. In particular, with this notation the Euler–Lagrange equations (3.18) of the functional (3.13) simply express the *vanishing of the variational derivative of its density* f:

$$\boxed{\frac{\delta f}{\delta\mathbf{y}} = 0 \iff \delta F[\mathbf{y}] = 0\,.} \tag{3.21}$$

In other words, if $\mathbf{y}(x)$ satisfies the Euler–Lagrange equations for a density $f(x,\mathbf{y},\mathbf{y}')$ the variation $F(\mathbf{y}+\delta\mathbf{y}) - F(\mathbf{y})$ of the corresponding functional F is $o(\varepsilon)$, so that F is “stationary” at $\mathbf{y}(x)$ (i.e., does not vary appreciably near $\mathbf{y}(x)$).

Consider two functionals of the form (3.13)-(3.14) with densities f_1 and f_2 differing by the **total derivative** with respect to x of a function $g(x,\mathbf{y})$:

$$f_2(x,\mathbf{y},\mathbf{y}') = f_1(x,\mathbf{y},\mathbf{y}') + \frac{\mathrm{d}}{\mathrm{d}x}g(x,\mathbf{y})\,,\qquad \frac{\mathrm{d}}{\mathrm{d}x}g(x,\mathbf{y}) := \frac{\partial g(x,\mathbf{y})}{\partial x} + \frac{\partial g(x,\mathbf{y})}{\partial\mathbf{y}}\mathbf{y}'\,.$$

We then have

$$F_2[\mathbf{y}] - F_1[\mathbf{y}] = \int_{x_1}^{x_2} \frac{\mathrm{d}}{\mathrm{d}x} g(x,\mathbf{y}(x))\,\mathrm{d}x = g(x_1,\mathbf{y}(x_1)) - g(x_2,\mathbf{y}(x_2)) = g(x_1,\mathbf{y}_1) - g(x_2,\mathbf{y}_2)\,,$$

on account of the boundary conditions (3.14) satisfied by the functions $\mathbf{y}(x)$ in the domain of the functionals F_1 and F_2. Hence the latter functionals *differ by a constant*, and therefore *they have the same variational derivative* (as they have the same variation). It follows that the Euler–Lagrange equations of the densities f_1 and f_2 must be exactly the *same*, as can be also checked by direct differentiation (exercise). In other words:

Two densities differing by a total derivative give rise to the *same* Euler–Lagrange equations.

Exercise. Show that if the variational derivative of a function $f(x,\mathbf{y},\mathbf{y}')$ vanishes identically then f is the total derivative of a function $g(x,\mathbf{y})$. This implies the converse of the previous result: *if two densities f_1 and f_2 give rise to the* same *Euler–Lagrange equations they must necessarily differ by a total derivative.*

Solution. Indeed, if the variational derivative of $f(x,\mathbf{y},\mathbf{y}')$ vanishes identically we have

$$\frac{\delta f}{\delta y_i} := \frac{\partial f}{\partial y_i} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y_i'} = -\sum_{j=1}^{n} \frac{\partial^2 f}{\partial y_i' \partial y_j'}\, y_j'' - \sum_{j=1}^{n} \frac{\partial^2 f}{\partial y_i' \partial y_j}\, y_j' - \frac{\partial^2 f}{\partial x \partial y_i'} + \frac{\partial f}{\partial y_i} = 0 \tag{3.22}$$

for $i = 1,\dots,n$ and *all* $(x,\mathbf{y},\mathbf{y}',\mathbf{y}'')$. Since none of the partial derivatives appearing in the latter identity depend on $\mathbf{y}''$ the coefficient of y_j'' must vanish identically. We thus obtain

$$\frac{\partial^2 f}{\partial y_i' y_j'} = 0, \qquad i,j = 1,\dots,n\,,$$

i.e, $\dfrac{\partial f}{\partial y_i'}$ is independent of $\mathbf{y}'$ for all i. Hence

$$\frac{\partial f}{\partial y_i'} = g_i(x,\mathbf{y}) \quad \Longrightarrow \quad f = \sum_{i=1}^{n} g_i(x,\mathbf{y})\, y_i' + g_0(x,\mathbf{y})$$

for certain functions $g_i(x,\mathbf{y})$, $g_0(x,\mathbf{y})$. Substituting into Eq. (3.22) we then obtain

$$-\sum_{j=1}^{n} \frac{\partial g_i}{\partial y_j}\, y_j' - \frac{\partial g_i}{\partial x} + \sum_{j=1}^{n} \frac{\partial g_j}{\partial y_i}\, y_j' + \frac{\partial g_0}{\partial y_i} = 0\,. \tag{3.23}$$

Since none of the partial derivatives in Eq. (3.23) depend on $\mathbf{y}'$, equating to zero the coefficient of y_j' in the latter identity we deduce that

$$\frac{\partial g_i}{\partial y_j} = \frac{\partial g_j}{\partial y_i}\,, \qquad i,j = 1,\dots,n\,.$$

It can be shown that the latter equations imply that there is a function $k(x,\mathbf{y})$ such that

$$g_i = \frac{\partial k}{\partial y_i}\,, \qquad i = 1,\dots,n\,.$$

Equation (3.23) then reduces to

$$\frac{\partial}{\partial y_i}\left(g_0 - \frac{\partial k}{\partial x}\right) = 0 \quad \Longrightarrow \quad g_0 - \frac{\partial k}{\partial x} = l(x)$$

for some function $l(x)$. We then have

$$f = \sum_{i=1}^{n} \frac{\partial k(x,\mathbf{y})}{\partial y_i} y_i' + \frac{\partial k(x,\mathbf{y})}{\partial x} + l(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(k(x,\mathbf{y}) + \int l(x)\,\mathrm{d}x\right).$$

3.2 Hamilton's principle for unconstrained systems

3.2.1 Hamilton's principle for a single particle

Consider, first, the motion of a particle of mass m subject to an *irrotational* force

$$\mathbf{F}(t,\mathbf{r}) = -\frac{\partial V(t,\mathbf{r})}{\partial \mathbf{r}}. \tag{3.24}$$

Newton's equations of motion are in this case

$$m\ddot{x}_i = -\frac{\partial V}{\partial x_i}, \qquad i = 1,2,3, \tag{3.25}$$

where we have again denoted by x_i the i-th component of the particle's position vector $\mathbf{r}$. We ask ourselves if Eqs. (3.25) are the Euler–Lagrange equations of some functional

$$\int_{t_1}^{t_2} L\big(t,\mathbf{r}(t),\dot{\mathbf{r}}(t)\big)\,\mathrm{d}t.$$

(Note, again, that in this case t, $\mathbf{r}$ and L respectively play the roles of x, $\mathbf{y}$, and f.) Although it is not difficult to answer this question in the affirmative simply by inspection, we can proceed more systematically as follows. Writing Eqs. (3.25) in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x}_i) + \frac{\partial V}{\partial x_i} = 0, \qquad i = 1,2,3,$$

we see that it suffices to find a function $L(t,\mathbf{r},\dot{\mathbf{r}})$ verifying the equations

$$\frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i, \qquad \frac{\partial L}{\partial x_i} = -\frac{\partial V}{\partial x_i}; \qquad i = 1,2,3.$$

Integrating first the three equations for $\dfrac{\partial L}{\partial x_i}$ we obtain

$$L = -V + g(t,\dot{\mathbf{r}}),$$

and substituting back into the remaining equations we have

$$\frac{\partial g}{\partial \dot{x}_i} = m\dot{x}_i, \qquad i = 1,2,3,$$

which determines the function g:

$$g = \frac{1}{2}m\dot{\mathbf{r}}^2 + h(t).$$

Thus the simplest function with the desired property is[2]

$$\boxed{L = \frac{1}{2}m\dot{\mathbf{r}}^2 - V(t,\mathbf{r}) = T - V(t,\mathbf{r}).} \tag{3.26}$$

[2]Note that $h(t)$ is obviously a total derivative, since

$$h(t) = \frac{\mathrm{d}}{\mathrm{d}t}\int h(s)\,\mathrm{d}s.$$

Therefore adding it to the Lagrangian (3.26) does not change its Euler–Lagrange equations (3.25), as we saw in Section (3.1.3).

The function L is called the system's **Lagrangian**. We have therefore proved the so-called **Hamilton's principle** in its simplest form:

The trajectory followed by a particle of mass m subject to the irrotational force (3.24) as it moves from a point $\mathbf{r}_1$ at time t_1 to another point $\mathbf{r}_2$ at time t_2 is a *stationary point* of the action functional

$$S[\mathbf{r}] = \int_{t_1}^{t_2} L\big(t, \mathbf{r}(t), \dot{\mathbf{r}}(t)\big)\,\mathrm{d}t \qquad (\text{with} \quad \mathbf{r}(t_1) = \mathbf{r}_1, \ \ \mathbf{r}(t_2) = \mathbf{r}_2) \tag{3.27}$$

where $L = T - V(t, \mathbf{r})$. In other words, Newton's equations of motion are equivalent to the *Euler–Lagrange equations*

$$\frac{\delta L}{\delta \mathbf{r}} = 0,$$

i.e., to the *vanishing of the variation of the action:*

$$\delta S[\mathbf{r}] = 0.$$

The functional (3.27) is called the **action**. Note that the action has dimensions of energy × time or length × momentum, since L (like T or V) has dimensions of energy.

3.2.2 Hamilton's principle for a system of particles

Hamilton's principle is extended without difficulty to a system of N particles, provided that the total forces $\mathbf{F}_i$ acting on each particle are *irrotational*, i.e.,

$$\mathbf{F}_i(t, \mathbf{r}_1, \dots, \mathbf{r}_N) = -\frac{\partial V(t, \mathbf{r}_1, \dots, \mathbf{r}_N)}{\partial \mathbf{r}_i}, \qquad i = 1, \dots, N. \tag{3.28}$$

Indeed, it is easy to check that Newton's equations of motion for the system:

$$m_i \ddot{\mathbf{r}}_i = -\frac{\partial V(t, \mathbf{r}_1, \dots, \mathbf{r}_N)}{\partial \mathbf{r}_i}, \qquad i = 1, \dots, N,$$

are the Euler–Lagrange equations of the action

$$\boxed{S[\mathbf{r}_1, \dots, \mathbf{r}_N] = \int_{t_1}^{t_2} L\big(t, \mathbf{r}_1(t), \dots, \mathbf{r}_N(t), \dot{\mathbf{r}}_1(t), \dots, \dot{\mathbf{r}}_N(t)\big)\,\mathrm{d}t\,,} \tag{3.29}$$

where in this case the Lagrangian L is given by

$$L(t, \mathbf{r}_1, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N) = T - V(t, \mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i^2 - V(t, \mathbf{r}_1, \dots, \mathbf{r}_N). \tag{3.30}$$

To see this, we write in vector form the three Euler–Lagrange equations for the i-th particle (i.e., one for each of the components of its position vector $\mathbf{r}_i$):

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\mathbf{r}}_i} - \frac{\partial L}{\partial \mathbf{r}_i} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{\mathbf{r}}_i} + \frac{\partial V}{\partial \mathbf{r}_i} = \frac{\mathrm{d}}{\mathrm{d}t}(m_i \dot{\mathbf{r}}_i) + \frac{\partial V}{\partial \mathbf{r}_i} = m_i \ddot{\mathbf{r}}_i - \mathbf{F}_i\,,$$

and observe that this is precisely the equation of motion of $\mathbf{r}_i$. In other words, the following more general version of Hamilton's principle holds:

The trajectory followed by a system of N particles subject to the irrotational forces (3.28) is a *stationary point* of the action (3.29) with Lagrangian $L = T - V(t, \mathbf{r}_1, \dots, \mathbf{r}_N)$. In other words, the system's equations of motion are again the *Euler–Lagrange equations*

$$\frac{\delta L}{\delta \mathbf{r}_i} = 0, \qquad i = 1, \dots, N,$$

which express the *vanishing of the variation of the action*:

$$\delta S[\mathbf{r}_1, \dots, \mathbf{r}_N] = 0.$$

• Hamilton's variational principle is sometimes called **principle of least action**, since in many cases of interest the trajectories of a mechanical system turn out to be *minima* of the action (at least *locally*). More properly, this principle should be called **principle of stationary action**, since as we know the Euler–Lagrange equations only guarantee the stationary character of the action.

• From Hamilton's principle and the conservation laws of the Euler–Lagrange equations derived in the previous subsection, it follows that if the Lagrangian (3.30) does not depend explicitly on time the energy integral

$$h = \sum_{i=1}^{N} \dot{\mathbf{r}}_i \frac{\partial L}{\partial \dot{\mathbf{r}}_i} - L = \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i^2 - L = 2T - (T - V) = T + V,$$

which in this case coincides with the system's *total energy*, is conserved. This result is consistent with the one obtained in Section 1.6.2, since L is independent of time if and only if the potential V does not depend on t, in which case the forces (3.28) acting on the system are not only irrotational but *conservative*.

• Likewise, if the Lagrangian L does not depend (for instance) on the x coordinate of the i-th particle, i.e., if $\frac{\partial L}{\partial x_i} = 0$, then $\frac{\partial L}{\partial \dot{x}_i}$ is conserved:

$$\frac{\partial L}{\partial x_i} = 0 \quad \Longrightarrow \quad \frac{\partial L}{\partial \dot{x}_i} = \text{const.}$$

Note that this result is nothing more than the conservation law of the x component of the i-th particle's momentum, since

$$\frac{\partial L}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i} = m_i \dot{x}_i = (\mathbf{p}_i)_x.$$

This is in agreement with the discussion in Section 1.6.2, since

$$\frac{\partial L}{\partial x_i} = -\frac{\partial V}{\partial x_i} = (\mathbf{F}_i)_x.$$

Obviously, the same result holds for the coordinates y_i or z_i.

3.2.3 Covariance of the Lagrangian formulation

One of the great advantages of the Lagrangian formulation of mechanics is its ***covariance** under coordinate changes*, i.e., that it treats all systems of curvilinear coordinates on the same footing. More specifically, consider a particle of mass m subject to an irrotational force, whose trajectories are the critical points of the action (3.27). Let $(q_1, q_2, q_3) =: \mathbf{q}$ be a system of *curvilinear coordinates*, and denote (with a slight abuse of notation) by $\mathbf{r}(\mathbf{q})$ the function expressing the

Cartesian coordinates $\mathbf{r}$ in terms of the curvilinear ones $\mathbf{q}$. Suppose that the particle's trajectory in the coordinates q_i is given by a certain function $\mathbf{q}(t) = (q_1(t), q_2(t), q_3(t))$. In Cartesian coordinates the trajectory is then $\mathbf{r} = \mathbf{r}(\mathbf{q}(t))$, which (with a slight abuse of notation) we shall denote by $\mathbf{r}(t)$. We can then express the Lagrangian $L(t, \mathbf{r}(t), \dot{\mathbf{r}}(t))$ in terms of $\mathbf{q}(t)$ and its time derivatives $\dot{\mathbf{q}}(t)$ using the change of coordinates formula $\mathbf{r} = \mathbf{r}(\mathbf{q})$ and its time derivative

$$\dot{\mathbf{r}} = \sum_{i=1}^{3} \frac{\partial \mathbf{r}(\mathbf{q})}{\partial q_i} \dot{q}_i =: \frac{\partial \mathbf{r}(\mathbf{q})}{\partial \mathbf{q}} \dot{\mathbf{q}},$$

so that

$$L(t, \mathbf{r}, \dot{\mathbf{r}}) = L\left(t, \mathbf{r}(\mathbf{q}), \frac{\partial \mathbf{r}(\mathbf{q})}{\partial \mathbf{q}} \dot{\mathbf{q}}\right) =: \widetilde{L}(t, \mathbf{q}, \dot{\mathbf{q}}). \tag{3.31}$$

The *action* of the trajectory $\mathbf{r}(t)$ is then given by

$$S[\mathbf{r}] = \int_{t_1}^{t_2} L(t, \mathbf{r}(t), \dot{\mathbf{r}}(t))\,\mathrm{d}t = \int_{t_1}^{t_2} \widetilde{L}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))\,\mathrm{d}t =: \widetilde{S}[\mathbf{q}].$$

By Hamilton's principle, the equations of motion are obtained from the condition $\delta S[\mathbf{r}] = 0$, i.e., $\delta \widetilde{S}[\mathbf{q}] = 0$, which is in turn equivalent to the Euler–Lagrange equations $\frac{\delta \widetilde{L}}{\delta \mathbf{q}} = 0$. Hence:

The equations of motion of the particle in curvilinear coordinates (q_1, q_2, q_3) are the Euler–Lagrange equations of the Lagrangian $\widetilde{L}$ in Eq. (3.31), that is

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \widetilde{L}}{\partial \dot{q}_i} - \frac{\partial \widetilde{L}}{\partial q_i} = 0, \qquad i = 1, 2, 3. \tag{3.32}$$

Note that $\widetilde{L}$ is nothing but the expression of the Lagrangian L in terms of the curvilinear coordinates $\mathbf{q}$ and their time derivatives. With this understanding the tilde can be dropped, and Eqs. (3.32) can be simply written as

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, 2, 3. \tag{3.33}$$

Example 3.7. *Equations of motion in spherical coordinates.*

The kinetic energy of a particle of mass m in spherical coordinates (r, θ, φ) is given by

$$\frac{1}{2} m \dot{\mathbf{r}}^2 = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta\, \dot{\varphi}^2)$$

(cf. Eq. (1.8)). Thus the Lagrangian in these coordinates is given by

$$L = T - V(t, r, \theta, \varphi) = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta\, \dot{\varphi}^2) - V(t, r, \theta, \varphi). \tag{3.34}$$

The particle's *equations of motion in spherical coordinates* are thus

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} &= m\ddot{r} - mr(\dot{\theta}^2 + \sin^2 \theta\, \dot{\varphi}^2) + \frac{\partial V}{\partial r} = 0, \\
\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} &= m \frac{\mathrm{d}}{\mathrm{d}t} (r^2 \dot{\theta}) - mr^2 \sin \theta \cos \theta\, \dot{\varphi}^2 + \frac{\partial V}{\partial \theta} = 0, \\
\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{\varphi}} - \frac{\partial L}{\partial \varphi} &= m \frac{\mathrm{d}}{\mathrm{d}t} (r^2 \sin^2 \theta\, \dot{\varphi}) + \frac{\partial V}{\partial \varphi} = 0.
\end{aligned} \tag{3.35}$$

If the potential V does not depend on the azimuthal angle φ then the quantity

$$\boxed{\frac{\partial L}{\partial \dot{\varphi}} = mr^2 \sin^2\theta\, \dot{\varphi}\,,}$$

is conserved. This conserved quantity is nothing but the z component of the angular momentum $\mathbf{J}$, since[a],

$$\mathbf{J} = r\mathbf{e}_r \times m(\dot{r}\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta + r\sin\theta\dot{\varphi}\,\mathbf{e}_\varphi) = mr^2\dot{\theta}\mathbf{e}_\varphi - mr^2\sin\theta\dot{\varphi}\,\mathbf{e}_\theta\,, \quad \mathbf{e}_z = \cos\theta\,\mathbf{e}_r - \sin\theta\mathbf{e}_\theta$$
$$\Longrightarrow \quad J_z = \mathbf{J}\cdot\mathbf{e}_z = mr^2\sin^2\theta\,\dot{\varphi}\,.$$

Likewise, if the potential V is independent of t, i.e., if V is a function of (r,θ,φ) only, then $\frac{\partial L}{\partial t} = 0$ and hence the quantity

$$h = \dot{r}\frac{\partial L}{\partial \dot{r}} + \dot{\theta}\frac{\partial L}{\partial \dot{\theta}} + \dot{\varphi}\frac{\partial L}{\partial \dot{\varphi}} - L = m\,(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\varphi}^2) - L = 2T - (T-V) = T + V\,,$$

is conserved. This is, of course, the law of *conservation of energy* discussed in the previous chapters. Note, finally, that if the potential is independent of θ the function $\frac{\partial L}{\partial \dot{\theta}}$ is *not* conserved, since the angle θ appears explicitly in the kinetic energy and as a consequence $\frac{\partial L}{\partial \theta}$ never vanishes.

[a]Throughout this section, to avoid confusion with the Lagrangian L we shall denote by $\mathbf{J}$ the angular momentum.

Example 3.8. *Equations of motion in polar coordinates.*

If a particle moves on a plane subject to an irrotational force with potential V, its Lagrangian in polar coordinates (r,φ) is given by

$$\boxed{L = T - V(r,\varphi) = \frac{m}{2}\,(\dot{r}^2 + r^2\dot{\varphi}^2) - V(t,r,\varphi).}$$

Its corresponding Euler–Lagrange equations are

$$\boxed{\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} &= m\ddot{r} - mr\dot{\varphi}^2 + \frac{\partial V}{\partial r} = 0\,,\\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\varphi}} - \frac{\partial L}{\partial \varphi} &= m\frac{\mathrm{d}}{\mathrm{d}t}(r^2\dot{\varphi}) + \frac{\partial V}{\partial \varphi} = 0\,.\end{aligned}} \tag{3.36}$$

Taking into account that

$$\nabla V = \frac{\partial V}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial V}{\partial \varphi}\mathbf{e}_\varphi = -\mathbf{F} = -F_r\mathbf{e}_r - F_\varphi\mathbf{e}_\varphi\,,$$

we can write the previous equations as

$$\ddot{r} - r\dot{\varphi}^2 = \frac{F_r}{m}\,, \qquad \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}t}(r^2\dot{\varphi}) = \frac{F_\varphi}{m}\,.$$

Thus the left-hand sides of the previous equations are nothing but the radial and angular components of the acceleration, a_r and a_φ. When the potential V is independent of φ (in which case the force is *central*) the second equation of motion yields the law of conservation of angular momentum

$$mr^2\dot{\varphi} = \text{const.} = J\,,$$

while the equation of motion for the radial coordinate can be written as

$$m\ddot{r} - \frac{J^2}{mr^3} + \frac{\partial V(t,r)}{\partial r} = 0\,.$$

Finally, when V does not depend explicitly on time (i.e., when the force is not just irrotational but conservative) we have $\frac{\partial L}{\partial t} = 0$, which implies that the energy integral

$$h = \dot{r}\frac{\partial L}{\partial \dot{r}} + \dot{\varphi}\frac{\partial L}{\partial \dot{\varphi}} - L = m\,(\dot{r}^2 + r^2\dot{\varphi}^2) - L = 2T - (T - V) = T + V\,.$$

is conserved. This is nothing but the law of conservation of energy discussed in the previous two chapters.

Exercise. Compare Eqs. (3.35) with Newton's second law written in spherical coordinates (see Eqs. (1.9)).

Hint. It suffices to note that

$$\frac{\partial V}{\partial \mathbf{r}} = \frac{\partial V}{\partial r}\mathbf{e}_r + \frac{1}{r}\frac{\partial V}{\partial \theta}\mathbf{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial V}{\partial \varphi}\mathbf{e}_\varphi\,.$$

Exercise. i) Write down the equations of motion of a particle subject to an irrotational force in an arbitrary system of orthogonal curvilinear coordinates $\mathbf{q} = (q_1, q_2, q_3)$. ii) Use the latter equations to find an expression for the components of the acceleration in the curvilinear coordinates $\mathbf{q}$.

Hint. In any orthogonal curvilinear coordinate system we have

$$\dot{\mathbf{r}}^2 = \sum_{i=1}^{3} h_i(\mathbf{q})^2\dot{q}_i^2\,, \qquad \frac{\partial V}{\partial \mathbf{r}} = \sum_{i=1}^{3}\frac{1}{h_i(\mathbf{q})}\frac{\partial V}{\partial q_i}\mathbf{e}_{q_i}\,,$$

where the scale factors $h_i(\mathbf{q})$ where defined in Eq. (1.2). From the expression for $\dot{\mathbf{r}}^2$ it is straightforward to compute the Lagrangian in the curvilinear coordinates $\mathbf{q}$. The components of the acceleration can then be found from the Euler–Lagrange equations of motion using Newton's second law:

$$a_{q_i} = \frac{F_{q_i}}{m} = -\frac{1}{mh_i(\mathbf{q})}\frac{\partial V}{\partial q_i}\,, \qquad 1 \leqslant i \leqslant 3\,.$$

The final result is

$$a_{q_i} = h_i\ddot{q}_i + 2\dot{q}_i\sum_{j=1}^{3}\frac{\partial h_i}{\partial q_j}\dot{q}_j - \sum_{j=1}^{3}\frac{h_j}{h_i}\frac{\partial h_j}{\partial q_i}\dot{q}_j^2\,, \qquad 1 \leqslant i \leqslant 3\,.$$

3.2.4 Lagrangian of a charged particle in an electromagnetic field

We shall next derive the Lagrangian formulation of the equations of motion of a particle of mass m and charge e in the electromagnetic field generated by the potentials $\Phi(t,\mathbf{r})$ and $\mathbf{A}(t,\mathbf{r})$. As we saw in Chapter 1, the equations of motion are

$$m\ddot{\mathbf{r}} = -e\left(\frac{\partial \Phi}{\partial \mathbf{r}} + \frac{\partial \mathbf{A}}{\partial t}\right) + e\dot{\mathbf{r}} \times (\nabla \times \mathbf{A})\,.$$

To simplify these equations, we recall the identity

$$\mathbf{a} \times (\nabla \times \mathbf{b}) = \nabla(\mathbf{a}\cdot\mathbf{b}) - (\mathbf{a}\cdot\nabla)\mathbf{b},$$

where $\mathbf{a}(\mathbf{r})$, $\mathbf{b}(\mathbf{r})$ are vector fields and $\mathbf{a}\cdot\nabla$ is the differential operator

$$\mathbf{a}\cdot\nabla = \sum_{i=1}^{3} a_i \frac{\partial}{\partial x_i}\,.$$

We shall also use in what follows the alternative notation

$$(\mathbf{a}\cdot\nabla)\mathbf{b} \equiv \sum_{i=1}^{3} a_i\,\frac{\partial \mathbf{b}}{\partial x_i} =: \frac{\partial \mathbf{b}}{\partial \mathbf{r}}\,\mathbf{a}\,.$$

From the previous identity it follows that

$$-\frac{\partial \mathbf{A}}{\partial t} + \dot{\mathbf{r}} \times (\nabla \times \mathbf{A}) = -\frac{\partial \mathbf{A}}{\partial t} + \frac{\partial}{\partial \mathbf{r}}\,(\dot{\mathbf{r}}\cdot\mathbf{A}) - \frac{\partial \mathbf{A}}{\partial \mathbf{r}}\cdot\dot{\mathbf{r}} = \frac{\partial}{\partial \mathbf{r}}\,(\dot{\mathbf{r}}\cdot\mathbf{A}) - \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\,,$$

and hence we can rewrite the equation of motion as

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(m\dot{\mathbf{r}} + e\mathbf{A}) + e\frac{\partial}{\partial \mathbf{r}}\,(\Phi - \dot{\mathbf{r}}\cdot\mathbf{A}) = 0\,.$$

These are the Euler–Lagrange equations of a Lagrangian $L(t, \mathbf{r}, \dot{\mathbf{r}})$ provided that

$$\frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} + e\mathbf{A}\,, \qquad \frac{\partial L}{\partial \mathbf{r}} = -e\frac{\partial}{\partial \mathbf{r}}\,(\Phi - \dot{\mathbf{r}}\cdot\mathbf{A})\,.$$

Integrating the second equation we obtain

$$L = e(\dot{\mathbf{r}}\cdot\mathbf{A} - \Phi) + g(t, \dot{\mathbf{r}})\,,$$

and substituting back into the first one we have

$$\frac{\partial g}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} \quad \Longrightarrow \quad g = \frac{1}{2}m\dot{\mathbf{r}}^2\,,$$

up to an arbitrary function of t. We have thus obtained the following result:

The equations of motion of a particle of mass m and charge e in the electromagnetic field generated by the potentials $\Phi(t, \mathbf{r})$ and $\mathbf{A}(t, \mathbf{r})$ are the Euler–Lagrange equations of the Lagrangian

$$L(t, \mathbf{r}, \dot{\mathbf{r}}) = \frac{1}{2}m\dot{\mathbf{r}}^2 - e\Phi(t, \mathbf{r}) + e\dot{\mathbf{r}}\cdot\mathbf{A}(t, \mathbf{r})\,. \tag{3.37}$$

Note that we can express the latter Lagrangian as

$$L = T - U\,,$$

where the potential U is given by

$$U(t, \mathbf{r}, \dot{\mathbf{r}}) = e\big[\Phi(t, \mathbf{r}) - \dot{\mathbf{r}}\cdot\mathbf{A}(t, \mathbf{r})\big] \tag{3.38}$$

and is thus *velocity dependent.* If the fields are *static*, i.e., if

$$\frac{\partial \Phi}{\partial t} = 0\,, \qquad \frac{\partial \mathbf{A}}{\partial t} = 0\,,$$

then L is independent of t and therefore

$$h = \dot{\mathbf{r}}\cdot\frac{\partial L}{\partial \dot{\mathbf{r}}} - L = m\dot{\mathbf{r}}^2 + e\dot{\mathbf{r}}\cdot\mathbf{A} - L = \frac{1}{2}\,m\dot{\mathbf{r}}^2 + e\Phi(\mathbf{r})\,,$$

is conserved. This is the conservation law of the particle's electromechanical energy.

3.3 Systems with constraints

3.3.1 Motion of a particle on a smooth surface

The simplest case of a mechanical system with constraints is that of a particle of mass m subject to an external irrotational force with potential $V(t,\mathbf{r})$, whose coordinates $\mathbf{r}$ satisfy at every instant t the **constraint** (restriction)

$$\boxed{\phi(t,\mathbf{r}) = 0\,.} \tag{3.39}$$

In particular, if ϕ does not depend on t then the particle is forced to move on the *surface* of equation $\phi(\mathbf{r}) = 0$. (In general, (3.39) is the equation of a moving surface.) Although the external force is irrotational, it is essential to take also into account the **reaction** (or **constraint**) **force** $\mathbf{F}^{(c)}(t,\mathbf{r},\dot{\mathbf{r}})$ exerted by the constraint surface (3.39) on the particle at each instant, so that Newton's equations of motion are in this case

$$\boxed{m\ddot{\mathbf{r}} + \frac{\partial V(t,\mathbf{r})}{\partial \mathbf{r}} = \mathbf{F}^{(c)}(t,\mathbf{r},\dot{\mathbf{r}})\,.} \tag{3.40}$$

We ask ourselves whether Eqs. (3.40) are the Euler–Lagrange equations of some action functional. In order to answer this question, let us introduce two *independent coordinates* $(q_1,q_2) = \mathbf{q}$ parametrizing the surface (3.39). For instance, if

$$\phi(t,\mathbf{r}) = \mathbf{r}^2 - a(t)^2\,, \tag{3.41}$$

which is the equation of a sphere centered at the origin with variable radius $a(t) \geqslant 0$, we can use spherical coordinates $q_1 = \theta$, $q_2 = \varphi$. We shall express (with a slight abuse of notation) the relation between the generalized coordinates $\mathbf{q}$ and the Cartesian ones $\mathbf{r}$ in the general form

$$\mathbf{r} = \mathbf{r}(t,\mathbf{q})\,, \tag{3.42}$$

where for each fixed t the mapping $\mathbf{q} \mapsto \mathbf{r}$ must be *bijective* (from an open subset of $\mathbb{R}^2$ to an open subset of the surface at time t). For instance, for the constraint (3.41) the function $\mathbf{r}(t,\mathbf{q}) = \mathbf{r}(t,\theta,\varphi)$ is given by

$$\mathbf{r}(t,\theta,\varphi) = a(t)(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)\,.$$

In general, we can specify the position of the particle at each instant t using the value $\mathbf{q}(t)$ taken by its **generalized coordinates** q_i at that time: indeed, $\mathbf{r} = \mathbf{r}(t,\mathbf{q}(t))$. It is important to note that, while the three Cartesian coordinates x_i are *not* independent, since they are connected by the relation (3.39), the two generalized coordinates q_i are by construction independent variables (i.e., can take *arbitrary* values in some open subset of $\mathbb{R}^2$). For this reason, it is easy to convince oneself that only two of the three (scalar) equations of motion (3.40) can actually be independent.

We shall assume that the constraint surface (3.39) is *smooth*, i.e., that there is no *friction*. If this is the case *the constraint force at each instant t is* perpendicular *to the corresponding* instantaneous *constraint surface* $\phi(t,\mathbf{r}) = 0$. When this happens we shall say that the constraint (3.39) is **ideal**. To formulate analytically the condition of ideal constraints, note that for each t the two vectors

$$\frac{\partial \mathbf{r}(t,\mathbf{q})}{\partial q_i}\,, \qquad i = 1,2\,, \tag{3.43}$$

are *tangent to the constraint surface* at the point with generalized coordinates $\mathbf{q}$, and in fact are a *basis* of the tangent plane to the instantaneous surface $\phi(t,\mathbf{r}) = 0$ at this point. Hence *the constraint is ideal if the constraint force verifies the condition*

$$\boxed{\mathbf{F}^{(c)} \cdot \frac{\partial \mathbf{r}}{\partial q_i} = 0\,, \qquad i = 1,2\,,} \tag{3.44}$$

at each point. Projecting the equation of motion (3.40) onto the tangent plane to the constraint surface (i.e., multiplying scalarly by each of the two vectors (3.43)) we obtain the two *independent* equations

$$m\ddot{\mathbf{r}}\,\frac{\partial \mathbf{r}}{\partial q_i}+\frac{\partial V}{\partial \mathbf{r}}\,\frac{\partial \mathbf{r}}{\partial q_i}=0\,,\qquad i=1,2\,,$$

or equivalently

$$\boxed{m\ddot{\mathbf{r}}\,\frac{\partial \mathbf{r}}{\partial q_i}+\frac{\partial V}{\partial q_i}=0\,,\qquad i=1,2\,.}\tag{3.45}$$

Differentiating the relation (3.42) with respect to t, q_i and $\dot{q}_i$ we obtain the identities[3]

$$\boxed{\dot{\mathbf{r}}=\frac{\partial \mathbf{r}}{\partial t}+\frac{\partial \mathbf{r}}{\partial \mathbf{q}}\,\dot{\mathbf{q}}=\dot{\mathbf{r}}(t,\mathbf{q},\dot{\mathbf{q}})\quad\Longrightarrow\quad\frac{\partial \dot{\mathbf{r}}}{\partial \dot{q}_i}=\frac{\partial \mathbf{r}}{\partial q_i}\,,\qquad\frac{\partial \dot{\mathbf{r}}}{\partial q_i}=\frac{\partial^2 \mathbf{r}}{\partial t\,\partial q_i}+\frac{\partial^2 \mathbf{r}}{\partial q_i\,\partial \mathbf{q}}\,\dot{\mathbf{q}}=\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial \mathbf{r}}{\partial q_i}}\,,\tag{3.46}$$

and hence

$$\ddot{\mathbf{r}}\,\frac{\partial \mathbf{r}}{\partial q_i}=\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{\mathbf{r}}\,\frac{\partial \mathbf{r}}{\partial q_i}\right)-\dot{\mathbf{r}}\,\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbf{r}}{\partial q_i}\right)=\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{\mathbf{r}}\,\frac{\partial \dot{\mathbf{r}}}{\partial \dot{q}_i}\right)-\dot{\mathbf{r}}\,\frac{\partial \dot{\mathbf{r}}}{\partial q_i}=\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial}{\partial \dot{q}_i}\left(\frac{1}{2}\,\dot{\mathbf{r}}^2\right)-\frac{\partial}{\partial q_i}\left(\frac{1}{2}\,\dot{\mathbf{r}}^2\right).\tag{3.47}$$

Thus Eqs. (3.45) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial T}{\partial \dot{q}_i}-\frac{\partial}{\partial q_i}(T-V)=0\,,\qquad i=1,2\,,$$

or, taking into account that V does not depend on $\dot{\mathbf{q}}$,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial}{\partial \dot{q}_i}(T-V)-\frac{\partial}{\partial q_i}(T-V)=0\,,\qquad i=1,2\,.}\tag{3.48}$$

These are the Euler–Lagrange equations of the Lagrangian $L=T-V$, where *it is understood that the kinetic energy T and the potential V must be expressed in terms of the independent variables $(t,\mathbf{q},\dot{\mathbf{q}})$ using Eq.* (3.42) *and its derivative with respect to t* (i.e., the first Eq. (3.46)). We have thus proved the following fundamental result:

> The trajectory $\mathbf{q}(t)$ followed by a particle as it moves from a point with generalized coordinates $\mathbf{q}_1$ (at $t=t_1$) to a second point with generalized coordinates $\mathbf{q}_2$ (at $t=t_2$) obeying the constraint (3.39) at all times is a *stationary point* of the *action*
>
> $$S[\mathbf{q}]=\int_{t_1}^{t_2}L\big(t,\mathbf{q}(t),\dot{\mathbf{q}}(t)\big)\,\mathrm{d}t\qquad(\text{with}\quad\mathbf{q}(t_1)=\mathbf{q}_1,\ \ \mathbf{q}(t_2)=\mathbf{q}_2)\,,$$
>
> where the Lagrangian L equals $T-V$ expressed in terms of the independent variables $(t,\mathbf{q},\dot{\mathbf{q}})$. The equations of motion are therefore the Euler–Lagrange equations of L,
>
> $$\frac{\delta L}{\delta \mathbf{q}}=0\,,$$
>
> expressing the vanishing of the variation of the action functional:
>
> $$\delta S[\mathbf{q}]=0\,.$$

In other words:

[3]We are using again the notation

$$\frac{\partial \mathbf{r}}{\partial \mathbf{q}}\,\dot{\mathbf{q}}:=\sum_{j=1,2}\frac{\partial \mathbf{r}}{\partial q_j}\,\dot{q}_j\,,\qquad\frac{\partial^2 \mathbf{r}}{\partial q_i\partial \mathbf{q}}\,\dot{\mathbf{q}}:=\sum_{j=1,2}\frac{\partial^2 \mathbf{r}}{\partial q_i\partial q_j}\,\dot{q}_j\,.$$

Hamilton's principle remains valid in this case if the constraint is *ideal*, i.e., if the constraint force $\mathbf{F}^{(c)}$ satisfies condition (3.44). Moreover, when this is the case the Lagrangian L is equal to $T - V$ expressed in terms of the generalized coordinates q_i and their time derivatives $\dot{q}_i$.

- A **virtual displacement** is a curve $\mathbf{r} = \mathbf{r}(u)$ ($u \in [u_1, u_2]$) entirely contained in the *instantaneous* constraint surface $\phi(t, \mathbf{r}) = 0$ at a certain *fixed* time t, i.e., such that

$$\phi(t, \mathbf{r}(u)) = 0\,, \qquad \forall u \in [u_1, u_2].$$

In other words, a virtual displacement is a succession of *possible* particle positions at the *same* instant t. Note that if the constraint is time-dependent the particle's trajectories are *not* virtual displacements, since for $t' \neq t$ the vector $\mathbf{r}(t')$ belongs to the surface $\phi(t', \mathbf{r}) = 0$, in general different from $\phi(t, \mathbf{r}) = 0$. On the other hand, since a virtual displacement $\mathbf{r}(u)$ is contained in the instantaneous constraint surface $\phi(t, \mathbf{r}) = 0$ for all $u \in [u_1, u_2]$, its tangent vector $\mathbf{r}'(u)$ is *tangent* to the latter surface at the point $\mathbf{r}(u)$. Hence the ideal constraint condition implies in this case that

$$\mathbf{F}^{(c)}(t, \mathbf{r}(u), \mathbf{v}(u)) \cdot \mathbf{r}'(u) = 0\,, \qquad \forall u \in [u_1, u_2]\,, \tag{3.49}$$

where the prime denotes derivative with respect to u and $\mathbf{v}(u)$ is a possible velocity[4] for the particle at the point $\mathbf{r}(u)$. Thus *the work W_{12} done by the constraint force along the virtual displacement* $\mathbf{r}(u)$ *vanishes:*

$$W_{12} = \int_{u_1}^{u_2} \mathbf{F}^{(c)}(t, \mathbf{r}(u), \mathbf{v}(u)) \cdot \mathbf{r}'(u)\, \mathrm{d}u = 0\,, \tag{3.50}$$

for arbitrary $\mathbf{v}(u) \in \mathbb{R}^3$. Conversely, if the constraint force satisfies (3.50) for *any* virtual displacement $\mathbf{r}(u)$ then Eq. (3.49) holds, which implies that the constraint force is perpendicular to the surface $\phi(t, \mathbf{r}) = 0$ at each point (since any vector tangent to the latter surface can be obtained as the tangent vector to a curve contained in it, i.e., to a virtual displacement). We have thus proved the following result, known as the **principle of virtual work**:

The constraint is ideal —and, thus, *Hamilton's principle* holds— if and only if the constraint force does no work along any *virtual* displacement of the particle.

If the constraint equation (3.39) is *independent of* t (which is the most common case in practice), then the particle's trajectories are virtual displacements, and the principle of virtual work simply states that *the constraint is ideal if and only if the constraint force does no work along* any *trajectory.*

[4] Differentiating the constraint equation $\phi(t, \mathbf{r}) = 0$ with respect to time we obtain

$$\frac{\partial \phi}{\partial t}(t, \mathbf{r}) + \frac{\partial \phi}{\partial \mathbf{r}}(t, \mathbf{r})\dot{\mathbf{r}} = 0\,,$$

which is the condition satisfied by the particle's velocity if the particle is at the point $\mathbf{r}$ at time t. Thus the vector field $\mathbf{v}(u)$ must satisfy the condition

$$\frac{\partial \phi}{\partial t}(t, \mathbf{r}(u)) + \frac{\partial \phi}{\partial \mathbf{r}}(t, \mathbf{r}(u))\mathbf{v}(u) = 0\,, \qquad \forall u \in [u_1, u_2].$$

If the constraint is time-independent then $\mathbf{v}(u)$ must simply be orthogonal to the gradient $\dfrac{\partial \phi}{\partial \mathbf{r}}(t, \mathbf{r}(u))$, and thus tangent to the constraint surface at each point $\mathbf{r}(u)$.

3.3.2 System of N particles with constraints

Consider next the most general case of a system of N particles subject to the irrotational forces (3.28) and to the $l < 3N$ *independent* constraints[5]

$$\boxed{\phi_i(t, \mathbf{r}_1, \dots, \mathbf{r}_N) = 0\,, \qquad i = 1, \dots, l\,.} \tag{3.51}$$

Constraints of this type, which are independent of the particles' *velocities*, are called *holonomic*. The vector

$$\mathbf{x} := (\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{3N}$$

representing the state of the system must belong at each instant t to the surface in $\mathbb{R}^{3N}$ — or *manifold*, in a more mathematical language— specified by Eqs. (3.51). Since this manifold has dimension $3N - l = n$, in general it can be parametrized by n independent coordinates $(q_1, \dots, q_n) =: \mathbf{q}$, in terms of which the vector $\mathbf{x}$ will be expressed by a certain function $\mathbf{x}(t, \mathbf{q})$:

$$\mathbf{x} = \mathbf{x}(t, \mathbf{q})\,. \tag{3.52}$$

In other words, *the state of the system at each instant is uniquely determined by the value of the n **generalized coordinates** q_i at that instant.* We shall accordingly say that the system possesses n **degrees of freedom**. In particular, the system's trajectory in the space $\mathbb{R}^{3N}$ can be specified by a curve $\mathbf{q}(t)$ in the open subset of $\mathbb{R}^n$ in which the generalized coordinates q_i vary, called **configuration space**, through the equation

$$\mathbf{x} = \mathbf{x}(t, \mathbf{q}(t)).$$

It is important to note that, while the *Cartesian* coordinates $\mathbf{x}$ are *not* independent (since they are related by the constraint equations (3.51)), the *generalized* coordinates $\mathbf{q}$ are by construction independent variables.

Again, *we shall suppose that the constraints are **ideal**,* in the sense that the constraint force acting on the point $\mathbf{x}$ representing the state of the system, i.e., the vector

$$\mathbf{F}^{(c)}(t, \mathbf{x}, \dot{\mathbf{x}}) := \left(\mathbf{F}_1^{(c)}(t, \mathbf{x}, \dot{\mathbf{x}}), \dots, \mathbf{F}_N^{(c)}(t, \mathbf{x}, \dot{\mathbf{x}})\right) \in \mathbb{R}^{3N},$$

is *orthogonal to the constraint manifold* defined by Eqs. (3.51) at all times. Since the n vectors

$$\frac{\partial \mathbf{x}(t, \mathbf{q})}{\partial q_i}\,, \qquad i = 1, \dots, n\,,$$

are a basis of the tangent space to the constraint manifold at each point, the previous condition is equivalent to the relations

$$\boxed{\mathbf{F}^{(c)} \cdot \frac{\partial \mathbf{x}}{\partial q_i} = \left(\mathbf{F}_1^{(c)}, \dots, \mathbf{F}_N^{(c)}\right) \cdot \left(\frac{\partial \mathbf{r}_1}{\partial q_i}, \dots, \frac{\partial \mathbf{r}_N}{\partial q_i}\right) = \sum_{j=1}^{N} \mathbf{F}_j^{(c)} \cdot \frac{\partial \mathbf{r}_j}{\partial q_i} = 0\,, \qquad i = 1, \dots, n\,.} \tag{3.53}$$

As in the case of a single particle treated above, this condition is equivalent to the *principle of virtual work*, according to which *the constraint forces do no work along any virtual displacement of the system,* which by definition is any curve $\mathbf{x}(u)$ (with $u \in [u_1, u_2]$) entirely contained in an *instantaneous* constraint surface $\phi_i(t, \mathbf{x}) = 0$ ($i = 1, \dots, l$) at a *fixed* instant t. Indeed, if

[5] Mathematically, the independence of the constraints (3.51) is equivalent to the condition that the Jacobian matrix of the vector-valued function $\boldsymbol{\phi} := (\phi_1, \dots, \phi_l)$ with respect to the $3N$ variables $\mathbf{x} := (\mathbf{r}_1, \dots, \mathbf{r}_N)$ be of maximal rank (equal to l) at all points:

$$\operatorname{rank}\left(\frac{\partial \phi_i}{\partial x_j}\right)_{\substack{1 \leqslant i \leqslant l \\ 1 \leqslant j \leqslant 3N}} = l\,.$$

$\mathbf{x}(u) = (\mathbf{r}_1(u), \dots, \mathbf{r}_N(u))$ the work W_{12} done by the constraint forces acting on the system along the virtual displacement $\mathbf{x}(u)$ is given by

$$W_{12} = \sum_{i=1}^{N} \int_{u_1}^{u_2} F_i^{(c)} \cdot \mathbf{r}_i'(u)\,\mathrm{d}u = \int_{u_1}^{u_2} \mathbf{F}^{(c)} \cdot \mathbf{x}'(u)\,\mathrm{d}u\,.$$

Hence the principle of virtual work, i.e., the requirement that $W_{12} = 0$ for an *arbitrary* virtual displacement $\mathbf{x}(u)$, is equivalent to requiring that

$$\mathbf{F}^{(c)} \cdot \mathbf{x}'(u) = 0$$

for every tangent vector $\mathbf{x}'(u)$ to the instantaneous constraint manifold at time t.

Under these conditions —that is, if the constraints are *ideal* and the applied forces acting on the system are *irrotational* —, proceeding as in the previous subsection one can prove that *Hamilton's principle is still valid*:

The trajectory $\mathbf{q}(t)$ joining two states $\mathbf{q}_1$ (at $t = t_1$) and $\mathbf{q}_2$ (at $t = t_2$) of a system of particles subject to irrotational forces and *ideal* holonomic constraints is a *stationary point* of the action

$$S[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))\,\mathrm{d}t\,,$$

where the Lagrangian L equals $T - V$ expressed in terms of the independent variables $(t, \mathbf{q}, \dot{\mathbf{q}})$. The equations of motion are thus the Euler–Lagrange equations

$$\frac{\delta L}{\delta \mathbf{q}} = 0\,,$$

expressing the vanishing of the variation of the action functional:

$$\delta S[\mathbf{q}] = 0\,.$$

Exercise. Prove in detail the latter result.

Solution. The system's equations of motion can be written in vector form as

$$(m_1\ddot{\mathbf{r}}_1, \dots, m_N\ddot{\mathbf{r}}_N) + \frac{\partial V}{\partial \mathbf{x}} = \mathbf{F}^{(c)}\,.$$

Projecting onto the direction of the vector $\dfrac{\partial \mathbf{x}}{\partial q_i}$ and taking into account Eq. (3.53) we obtain

$$(m_1\ddot{\mathbf{r}}_1, \dots, m_N\ddot{\mathbf{r}}_N) \cdot \frac{\partial \mathbf{x}}{\partial q_i} + \frac{\partial V}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial q_i} = \sum_{j=1}^{N} m_j\ddot{\mathbf{r}}_j\frac{\partial \mathbf{r}_j}{\partial q_i} + \frac{\partial V}{\partial q_i} = 0\,, \qquad i = 1, \dots, n\,.$$

From Eqs. (3.47) (with $\mathbf{r}_j$ instead of $\mathbf{r}$) it then follows that

$$\sum_{j=1}^{N} m_j\left[\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{q}_i}\left(\frac{1}{2}\dot{\mathbf{r}}_j^2\right) - \frac{\partial}{\partial q_i}\left(\frac{1}{2}\dot{\mathbf{r}}_j^2\right)\right] + \frac{\partial V}{\partial q_i} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial}{\partial q_i}(T - V) = 0\,, \qquad i = 1, \dots, n\,.$$

Since V is independent of $\dot{\mathbf{q}}$, the latter equations can be written in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0\,, \qquad i = 1, \dots, n\,,$$

which are the Euler–Lagrange equations of the Lagrangian $L = T - V$.

From what we have just seen, to write down the equations of motion of a mechanical system of N particles subject to l independent ideal holonomic constraints, the remaining (applied) forces being irrotational, we can proceed as follows:

1. Introduce $n = 3N - l$ independent generalized coordinates $(q_1, \dots, q_n) = \mathbf{q}$ parametrizing the constraint manifold (3.51).
2. Express the kinetic energy
$$T = \frac{1}{2} \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i^2$$
and the potential V of the irrotational forces in terms of $(t, \mathbf{q}, \dot{\mathbf{q}})$, thus obtaining the two functions $T(t, \mathbf{q}, \dot{\mathbf{q}})$ and $V(t, \mathbf{q})$.
3. The system's equations of motion in the generalized coordinates q_i are the Euler–Lagrange equations of the Lagrangian
$$L(t, \mathbf{q}, \dot{\mathbf{q}}) = T(t, \mathbf{q}, \dot{\mathbf{q}}) - V(t, \mathbf{q}),$$
i.e.,
$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, \dots, n. \tag{3.54}$$

Notation. In classical mechanics textbooks, equations (3.54) are often referred to simply as **Lagrange's equations** for the Lagrangian L.

• One of the advantages of the Lagrangian formulation for systems with constraints is that, as we have just seen, *in order to find the equations of motion it is not necessary to know the constraint forces* (all that is needed is to check that the constraints are *ideal*). In fact, *once these equations have been found* the constraint forces can always be computed using the formula[6]

$$\boxed{\mathbf{F}_i^{(c)} = m_i \ddot{\mathbf{r}}_i + \frac{\partial V}{\partial \mathbf{r}_i}, \qquad 1 \leqslant i \leqslant N,} \tag{3.55}$$

which is nothing but Newton's second law applied to the i-th particle.

Remark. Hamilton's principle is key to understanding in what sense classical mechanics is the $\hbar \to 0$ limit of quantum mechanics, with the help of Feynman's *path integral* formulation of the latter theory. According to this formulation, the probability $P(t_1, \mathbf{q}_1; t_2, \mathbf{q}_2)$ that a mechanical system with classical Lagrangian $L(t, \mathbf{q}, \dot{\mathbf{q}})$ whose generalized coordinates take the value $\mathbf{q}_1$ at a certain time t_1 is found to have generalized coordinates $\mathbf{q}_2$ at a later time t_2 is given by

$$P(t_1, \mathbf{q}_1; t_2, \mathbf{q}_2) = \left| \Phi(t_1, \mathbf{q}_1; t_2, \mathbf{q}_2) \right|^2,$$

where the *probability amplitude* $\Phi(t_1, \mathbf{q}_1; t_2, \mathbf{q}_2)$ (in general complex) is given by

$$\Phi(t_1, \mathbf{q}_1; t_2, \mathbf{q}_2) = \text{const.} \sum_{\mathbf{q}} \mathrm{e}^{\frac{\mathrm{i}}{\hbar} S[\mathbf{q}]}. \tag{3.56}$$

[6] Indeed, $\ddot{\mathbf{r}}_i$ can be computed in terms of $(t, \mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$ by differentiating $\mathbf{r}_i(t, \mathbf{q})$ twice with respect to time. Once the equations of motion have been found using the Lagrangian formalism, the generalized accelerations $\ddot{\mathbf{q}}$, and hence the accelerations $\ddot{\mathbf{r}}_i$ and the constraint forces $\mathbf{F}_i^{(c)}$, can be expressed in terms of $(t, \mathbf{q}, \dot{\mathbf{q}})$. Note that, in general, the constraint force will depend (usually in a complicated way) on the velocity of the particles.

The latter "sum" —technically an integral, usually called the *path integral*— is extended to *all* paths $\mathbf{q}(t)$ satisfying the boundary conditions $\mathbf{q}(t_i) = \mathbf{q}_i$, $i = 1, 2$, and

$$S[\mathbf{q}] = \int_{t_1}^{t_2} L(t, \mathbf{q}(t), \dot{\mathbf{q}}(t))\, \mathrm{d}t$$

is the classical action of the path $\mathbf{q}(t)$. Thus all paths contribute to the probability amplitude $\Phi(t_1, \mathbf{q}_1; t_2, \mathbf{q}_2)$ with the same absolute magnitude, but with different *phases* proportional to their classical action. In the classical limit $\hbar \to 0$ we have $S[\mathbf{q}] \gg \hbar$, and thus the term $\mathrm{e}^{\mathrm{i}S[\mathbf{q}]/\hbar}$ is highly oscillatory near paths satisfying $\delta S[\mathbf{q}] \neq 0$. As a consequence the contributions to the sum coming from such paths is vanishingly small as $\hbar \to 0$, since very close to a path $\mathbf{q}$ with $\delta S[\mathbf{q}] \neq 0$ there is a neighboring path whose phase differs by an odd multiple of π from that of $\mathbf{q}$. Thus in the limit $\hbar \to 0$ the overwhelming contribution to the sum (3.56) comes from the path[7] satisfying $\delta S[\mathbf{q}] = 0$, i.e., from the classical trajectory. In other words, the validity of Hamilton's principle (when $\hbar \to 0$, i.e., in the classical limit) hinges on the fact that in this limit the path with the largest contribution to the probability amplitude $\Phi(t_1, \mathbf{q}_1; t_2, \mathbf{q}_2)$ is the one making the classical action stationary. ■

Example 3.9. *The spherical pendulum.* A spherical pendulum consists of a particle of mass m attached to a rigid massless rod of length l and negligible mass whose other end is fixed, subject only to Earth's gravitational field $\mathbf{g} = -g\mathbf{e}_z$. In this case there is only one (time-independent) constraint

$$\phi(t, \mathbf{r}) = \mathbf{r}^2 - l^2 = 0\,, \tag{3.57}$$

(if we place the origin at the pendulum's anchor point), and there are therefore $3 - 1 = 2$ degrees of freedom. We shall take as generalized coordinates the polar and azimuthal angles $\theta \in [0, \pi]$, $\varphi \in [0, 2\pi)$ of the spherical coordinate system[a], in terms of which

$$\mathbf{r}(\theta, \varphi) = l\,(\sin\theta\cos\varphi\,\mathbf{e}_1 + \sin\theta\sin\varphi\,\mathbf{e}_2 + \cos\theta\,\mathbf{e}_z).$$

The constraint force $\mathbf{F}^{(c)}$ (in this case, the rod's reaction) is directed along the rod (towards the origin), and is thus perpendicular to the constraint surface (3.57). Hence the constraint is ideal, and we can apply the Lagrangian formalism. The potential of the external force $-mg\mathbf{e}_z$ is simply

$$V = mgz = mgl\cos\theta\,,$$

and the kinetic energy is given by

$$T = \frac{1}{2}m\dot{\mathbf{r}}^2 = \frac{1}{2}ml^2(\dot\theta^2 + \sin^2\theta\,\dot\varphi^2).$$

We thus have

$$\boxed{L = ml^2\left[\frac{1}{2}(\dot\theta^2 + \sin^2\theta\,\dot\varphi^2) - k\cos\theta\right], \qquad k := \frac{g}{l},}$$

and Lagrange's equations read

$$\boxed{\ddot\theta = \sin\theta\cos\theta\,\dot\varphi^2 + k\sin\theta\,, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(\sin^2\theta\,\dot\varphi) = 0\,.}$$

[7] We are assuming for the sake of simplicity that, as is usually the case, there is a unique classical trajectory satisfying the boundary conditions $\mathbf{q}(t_i) = \mathbf{q}_i$, $i = 1, 2$.

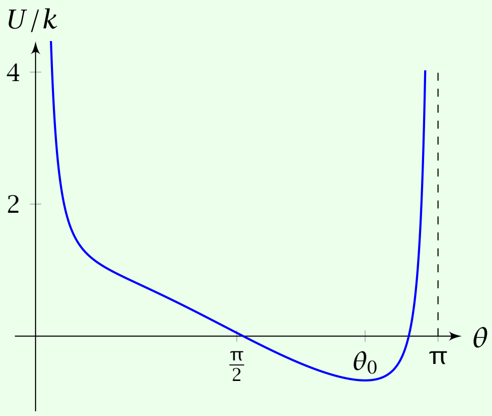


Figure 3.2. Effective potential $U(\theta)$ for $k = 10c^2$.

From the second Lagrange equation we obtain

$$\sin^2\theta\,\dot{\varphi} = \frac{J_z}{ml^2} =: c = \text{const.}\,,$$

where $\mathbf{J}$ is the particle's angular momentum. This was to be expected, since

$$\dot{\mathbf{J}} = \mathbf{N} = \mathbf{r}\times(\mathbf{F}+\mathbf{F}^{(c)}) = \mathbf{r}\times\mathbf{F} = -mg\mathbf{r}\times\mathbf{e}_z \quad\Longrightarrow\quad \dot{J}_z = 0\,.$$

Substituting into the first Lagrange equation we obtain the following second-order differential equation for the angle θ:

$$\boxed{\ddot{\theta} = c^2\,\frac{\cos\theta}{\sin^3\theta} + k\,\sin\theta\,.} \tag{3.58}$$

If $c = 0$ (i.e., $J_z = 0$) then the particle moves along a meridian[b] $\varphi = \text{const.}$ (since $\dot{\varphi} = 0$) and Eq. (3.58) becomes the equation of motion of the simple pendulum $\ddot{\alpha} + k\sin\alpha = 0$, where $\alpha = \pi - \theta$. Let us see next what happens in the more interesting case $c \neq 0$. Equation (3.58) is formally the equation of motion of a particle of unit mass moving in the effective one-dimensional potential

$$\boxed{U(\theta) = -\int\left(c^2\,\frac{\cos\theta}{\sin^3\theta} + k\,\sin\theta\right)\mathrm{d}\theta = k\,\cos\theta + \frac{c^2}{2\sin^2\theta}}$$

plotted in Fig. 3.2. The shape of the potential $U(\theta)$ can be determined by taking into account the following facts:

i. $U(\theta)$ diverges as $(\sin\theta)^{-2}$ as $\theta \to 0, \pi$.

ii. The derivative $U'(\theta)$ has the sign of $\theta - \theta_0$, for some $\theta_0 \in (\pi/2, \pi)$.

To prove the last statement note that

$$U'(\theta) = -(\sin\theta)^{-3}(c^2\cos\theta + k\sin^4\theta)$$

has the sign of $f(\theta) := -(c^2\cos\theta + k\sin^4\theta)$. The function f, and hence U', is clearly negative for $\theta \leqslant \pi/2$. On the other hand,

$$f'(\theta) = c^2\sin\theta - 4k\sin^3\theta\cos\theta$$

is positive over the interval $[\pi/2, \pi)$, so that f is increasing on $[\pi/2, \pi]$ from $f(\pi/2) = -k < 0$ to $f(\pi) = c^2 > 0$. It follows that there is a unique $\theta_0 \in (\pi/2, \pi)$ such that $f(\theta) = 0$, with $f(\theta) < 0$ for $\pi/2 \leqslant \theta < \theta_0$ and $f(\theta) > 0$ for $\theta_0 < \theta \leqslant \pi$. Thus $f(\theta)$, and hence $U'(\theta)$, has the sign of $\theta - \theta_0$, as stated.

In fact, since L does not depend explicitly on time the energy integral

$$h = \dot{\theta}\frac{\partial L}{\partial \dot{\theta}} + \dot{\varphi}\frac{\partial L}{\partial \dot{\varphi}} - L = T + V = ml^2\left[\frac{1}{2}(\dot{\theta}^2 + \sin^2\theta\,\dot{\varphi}^2) + k\cos\theta\right],$$

is conserved. Thus h is the particle's total energy E, which using the conservation of J_z can be expressed as

$$ml^2\left(\frac{1}{2}\dot{\theta}^2 + U(\theta)\right) = E\,.$$

The motion of the angular coordinate θ is easily determined integrating the latter equation:

$$\boxed{t = \pm\int\frac{\mathrm{d}\theta}{\sqrt{2\left(\frac{E}{ml^2} - U(\theta)\right)}}},$$

while the azimuthal angle φ then follows from the conservation of J_z:

$$\boxed{\varphi = c\int\frac{\mathrm{d}t}{\sin^2\theta(t)}}\,.$$

Finally, the equation of the trajectory (θ as a function of φ, or vice versa) is obtained combining the previous equations:

$$\dot{\theta} = \frac{\mathrm{d}\theta}{\mathrm{d}\varphi}\,\dot{\varphi} = \frac{c}{\sin^2\theta}\frac{\mathrm{d}\theta}{\mathrm{d}\varphi} = \pm\sqrt{2\left(\frac{E}{ml^2} - U(\theta)\right)} \quad\Longrightarrow\quad \boxed{\varphi = \pm c\int\frac{\mathrm{d}\theta}{\sin^2\theta\sqrt{2\left(\frac{E}{ml^2} - U(\theta)\right)}}}\,.$$

From the form of the effective potential $U(\theta)$ it follows that the motion of the coordinate θ is always *periodic*. Indeed, the period of this motion is given by

$$\tau_\theta = 2\int_{\theta_1}^{\theta_2}\frac{\mathrm{d}\theta}{\sqrt{2\left(\frac{E}{ml^2} - U(\theta)\right)}},$$

where $\theta_1 < \theta_2$ are the two roots of the equation $E/(ml^2) - U(\theta) = 0$ in the interval $(0, \pi)$. The pendulum's motion, however, is *not* periodic in general, since when the coordinate θ returns to its initial value after a period the azimuthal angle φ does not necessarily increase by a multiple of 2π. More precisely, from the equation of the trajectory it follows that in a period of θ the angle φ increases by

$$\Delta\varphi = 2c\int_{\theta_1}^{\theta_2}\frac{\mathrm{d}\theta}{\sin^2\theta\sqrt{2\left(\frac{E}{ml^2} - U(\theta)\right)}} = \sqrt{2}\int_{\theta_1}^{\theta_2}\frac{\mathrm{d}\theta}{\sin^2\theta\sqrt{\frac{E}{mlc^2} - \frac{k}{c^2}\cos\theta - \frac{1}{2\sin^2\theta}}}\,.$$

Hence *the motion is periodic if $\Delta\varphi$ is a rational multiple of* 2π.

Equation (3.58) possesses the *constant solution* $\theta = \theta_0$, with $\theta_0 \in (\pi/2, \pi)$ the unique solution of the equation

$$c^2\cos\theta_0 + k\sin^4\theta_0 = 0\,,$$

corresponding to a rotation around the z axis with constant angular velocity $\dot{\varphi} = c/\sin^2\theta_0$. The frequency ω of the small oscillations of the angle θ about the solution $\theta = \theta_0$ is given by

$$\begin{aligned}\omega^2 = U''(\theta_0) &= -k\cos\theta_0 + \frac{c^2}{\sin^2\theta_0} + 3c^2\,\frac{\cos^2\theta_0}{\sin^4\theta_0} = -k\cos\theta_0 - k\,\frac{1-\cos^2\theta_0 + 3\cos^2\theta_0}{\cos\theta_0}\\ &= k\,\frac{1+3\cos^2\theta_0}{|\cos\theta_0|}\,,\end{aligned}$$

where we have taken into account that $c^2/\sin^4\theta_0 = -k/\cos\theta_0$.

Finally, the constraint force is easily computed using Eq. (3.55) and noting that in this case $\mathbf{F}^{(c)}$ is directed along $\mathbf{e}_r$ (perpendicular to the constraint surface). Therefore $\mathbf{F}^{(c)} = R\mathbf{e}_r$, with

$$R = m\ddot{\mathbf{r}}\cdot\mathbf{e}_r - m\mathbf{g}\cdot\mathbf{e}_r = ma_r + mg\mathbf{e}_z\cdot\mathbf{e}_r = mg\cos\theta - ml(\dot{\theta}^2 + \sin^2\theta\,\dot{\varphi}^2)\,,$$

where we have applied Eq. (1.9) with $r = l$ and $\dot{r} = \ddot{r}0$. Using the law of conservation of energy we finally obtain

$$\boxed{R = 3mg\cos\theta - \frac{2E}{l}\,.}$$

[a]Of course, if the pendulum's pivot is fixed to the ceiling the angle θ must be restricted to the range $[\pi/2, \pi]$.

[b]In fact, $c = 0$ is also possible if either $\theta = 0$ or $\theta = \pi$, but these are just the two equilibria of the motion in a meridian $\varphi = \text{const}$.

Exercise. Show that $R < 0$ for $\theta \geqslant \pi/2$.

Solution. Indeed, we have

$$E = ml^2U(\theta_{1,2}) = mgl\cos\theta_{1,2} + \frac{ml^2c^2}{2\sin^2\theta_{1,2}} \quad\Longrightarrow\quad R = mg\cos\theta + 2mg(\cos\theta - \cos\theta_1) - \frac{mlc^2}{2\sin^2\theta_1}\,,$$

where $\cos\theta \leqslant \cos\theta_1$ (since $\theta \geqslant \theta_1$) and $\cos\theta \leqslant 0$ for $\theta \in [\pi/2, \pi]$.

3.4 Noether's theorem

Consider a mechanical system with Lagrangian $L(t, \mathbf{q}, \dot{\mathbf{q}})$, where $\mathbf{q} = (q_1, \dots, q_n)$ are the n generalized coordinates. We define the **canonical momentum** associated with the generalized coordinate q_i as the partial derivative of L with respect to the corresponding **generalized velocity** $\dot{q}_i$:

$$\boxed{p_i := \frac{\partial L}{\partial \dot{q}_i}\,.} \tag{3.59}$$

Lagrange's equation of motion for the coordinate q_i is then

$$\dot{p}_i = \frac{\partial L}{\partial q_i}\,. \tag{3.60}$$

We shall say that the coordinate q_i is **cyclic** (or **ignorable**) if L is independent of q_i, i.e.,

$$\frac{\partial L}{\partial q_i} = 0\,.$$

From Eq. (3.60) we then obtain the following conservation law:

If the coordinate q_i is cyclic, its corresponding canonical momentum p_i is conserved.

Likewise, if L does not explicitly depend on t we saw in Section 3.1.2 that the energy integral

$$h = \sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = \sum_{i=1}^{n} p_i \dot{q}_i - L \tag{3.61}$$

is conserved. In many mechanical systems the kinetic energy is a *quadratic form* in the generalized velocities, i.e., is of the form

$$T = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(t, \mathbf{q})\, \dot{q}_i \dot{q}_j\,, \qquad \text{with } a_{ij} = a_{ji}\,,$$

and $L = T - V(t, \mathbf{q})$. A mechanical system of this type is called **natural.** In fact, most of the systems considered so far —with the important exception of the Lagrangian of a charged particle in an electromagnetic field (3.37)— are natural. In a natural mechanical system, the generalized momenta are given by

$$p_i = \frac{\partial T}{\partial \dot{q}_i} = \sum_{j=1}^{n} a_{ij}(t, \mathbf{q})\, \dot{q}_j\,, \qquad i = 1, \dots, n\,,$$

are linear in the generalized velocities $\dot{q}_i$, and the energy integral is simply

$$h = \sum_{i,j=1}^{n} a_{ij}(t, \mathbf{q})\, \dot{q}_i \dot{q}_j - L = 2T - (T - V) = T + V\,.$$

Hence:

In a natural mechanical system the energy integral is equal to the total energy. In particular, in natural mechanical systems the conservation of h, -which will occur if the coefficients a_{ij} and V are both independent of t, is nothing but the law of conservation of energy.

Example 3.10. Consider, first, the Lagrangian of a particle of mass m in Cartesian coordinates $\mathbf{r} = (x_1, x_2, x_3)$, given by

$$L = \frac{1}{2} m \dot{\mathbf{r}}^2 - V(t, \mathbf{r})\,.$$

In this case

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m \dot{x}_i\,,$$

and thus the canonical momentum corresponding to the coordinate x_i is the i-th component of the linear momentum. Moreover, L is clearly natural and therefore the energy integral h coincides with the particle's energy.

Consider next the Lagrangian of a particle of mass m in spherical coordinates (r, θ, φ):

$$L = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2\theta\, \dot{\varphi}^2) - V(t, r, \theta, \varphi)\,, \tag{3.62}$$

for which

$$p_r = m\dot{r}\,, \qquad p_\theta = m r^2 \dot{\theta}\,, \qquad p_\varphi = m r^2 \sin^2\theta\, \dot{\varphi}\,. \tag{3.63}$$

In this case the kinetic energy (the term in parentheses in the Lagrangian) depends on r and θ, so that p_r and p_θ are *not* conserved even if V is independent of r or θ. On the other hand, if V does not depend on φ then L is independent of the latter coordinate, and hence p_φ is conserved:

$$\frac{\partial V}{\partial \varphi} = 0 \quad \Longrightarrow \quad p_\varphi = \text{const.}$$

As we know, p_φ is the z component of the particle's angular momentum. Moreover, since the kinetic energy is quadratic in the generalized velocities the Lagrangian is natural, and hence the energy integral coincides with the energy $T + V$, as we saw in Example 3.7. Hence if L does not depend on t —i.e., if the potential V is independent of time— energy is conserved.

Consider, finally, the Lagrangian (3.37) of a particle of mass m and charge e moving in an electromagnetic field with potentials $\Phi(t, \mathbf{r})$ and $\mathbf{A}(t, \mathbf{r})$. The canonical momentum corresponding to the coordinate x_i is now

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i + eA_i(t, \mathbf{r}) .$$

Thus in this case *the canonical and the linear momenta are in general different.* In particular, if L does not depend on the coordinate x_i, i.e., if

$$\frac{\partial \Phi}{\partial x_i} = 0 , \qquad \frac{\partial \mathbf{A}}{\partial x_i} = 0 ,$$

p_i is conserved but $m\dot{x}_i$ is not conserved in general. The energy integral is given by

$$h = \sum_{i=1}^{3} p_i \dot{x}_i - L = (m\dot{\mathbf{r}} + e\mathbf{A}) \cdot \dot{\mathbf{r}} - L = \frac{1}{2} m \dot{\mathbf{r}}^2 + e\Phi .$$

Therefore in this case h is the sum of the particle's kinetic and electrostatic energies. If L does not depend on t, that is if

$$\frac{\partial \Phi}{\partial t} = 0 , \qquad \frac{\partial \mathbf{A}}{\partial t} = 0 ,$$

then h is conserved. Although the system is *not* natural, we can also interpret h in this case as the total energy. Indeed, if Φ and $\mathbf{A}$ do not depend on t the electric force is *conservative* with potential $e\Phi(\mathbf{r})$, and therefore h is the sum of the kinetic energy and the potential energy of the electric force. But this is the total energy of the particle, since the magnetic force does no work as it is always perpendicular to the particle's velocity.

The conservation of the canonical momentum p_i and the energy integral h are clearly a consequence of the *invariance* of the Lagrangian under translations in the coordinate q_i ($q_i \mapsto q_i + \varepsilon$) or the time t ($t \mapsto t + \varepsilon$), respectively, where $\varepsilon \in \mathbb{R}$ is a *continuous parameter*. In fact, one of the fundamental principles of modern physics is the fact that continuous transformations leaving invariant the Lagrangian —or, more generally, the action— give rise to conserved quantities. This is precisely the import of **Noether's theorem**:

Suppose that the *action* of a mechanical system with Lagrangian $L(t, \mathbf{q}, \dot{\mathbf{q}})$ is *invariant* under a *one-parameter family* of invertible transformations

$$\tilde{t} = t + \varepsilon\tau(t, \mathbf{q}) + O(\varepsilon^2) , \qquad \tilde{\mathbf{q}} = \mathbf{q} + \varepsilon\boldsymbol{\eta}(t, \mathbf{q}) + O(\varepsilon^2) , \tag{3.64}$$

i.e., that

$$\int_{\tilde{t}_1}^{\tilde{t}_2} L\left(\tilde{t}, \tilde{\mathbf{q}}, \frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}\tilde{t}}\right) \mathrm{d}\tilde{t} = \int_{t_1}^{t_2} L(t, \mathbf{q}, \dot{\mathbf{q}})\, \mathrm{d}t , \qquad \forall t_1, t_2 . \tag{3.65}$$

Then the function

$$I(t, \mathbf{q}, \dot{\mathbf{q}}) := \mathbf{p}\,\boldsymbol{\eta} - h\tau ,$$

where $\mathbf{p} = \dfrac{\partial L}{\partial \dot{\mathbf{q}}}$ and $h = \mathbf{p}\dot{\mathbf{q}} - L$, is conserved.

Proof. We begin by computing the derivatives of $\dfrac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}t}$ and $\dfrac{\mathrm{d}\tilde{t}}{\mathrm{d}t}$ with respect to ε at $\varepsilon = 0$, that we shall need in the sequel:

$$\frac{\mathrm{d}\tilde{t}}{\mathrm{d}t} = 1 + \varepsilon\,\dot{\tau} + O(\varepsilon^2) \qquad \Longrightarrow \qquad \frac{\mathrm{d}\tilde{t}}{\mathrm{d}t}\Big|_{\varepsilon=0} = 1\,, \qquad \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0}\frac{\mathrm{d}\tilde{t}}{\mathrm{d}t} = \dot{\tau}\,.$$

$$\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}\tilde{t}} = \left(\frac{\mathrm{d}\tilde{t}}{\mathrm{d}t}\right)^{-1}\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}t} = \Big(1 + \varepsilon\,\dot{\tau} + O(\varepsilon^2)\Big)^{-1}\Big(\dot{\mathbf{q}} + \varepsilon\,\dot{\boldsymbol{\eta}} + O(\varepsilon^2)\Big) = \Big(1 - \varepsilon\,\dot{\tau} + O(\varepsilon^2)\Big)\Big(\dot{\mathbf{q}} + \varepsilon\,\dot{\boldsymbol{\eta}} + O(\varepsilon^2)\Big)$$

$$= \dot{\mathbf{q}} + \varepsilon(\dot{\boldsymbol{\eta}} - \dot{\mathbf{q}}\dot{\tau}) + O(\varepsilon^2) \qquad \Longrightarrow \qquad \frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0}\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}\tilde{t}} = \dot{\boldsymbol{\eta}} - \dot{\mathbf{q}}\dot{\tau}\,.$$

The invariance of the action under the transformation $(t,\mathbf{q}) \mapsto (\tilde{t},\tilde{\mathbf{q}})$ can be expressed in the equivalent form[8]

$$L\left(\tilde{t},\tilde{\mathbf{q}},\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}\tilde{t}}\right)\frac{\mathrm{d}\tilde{t}}{\mathrm{d}t} = L(t,\mathbf{q},\dot{\mathbf{q}})\,. \tag{3.66}$$

Differentiating (3.66) with respect to ε and setting $\varepsilon = 0$ we then obtain

$$\frac{\mathrm{d}\tilde{t}}{\mathrm{d}t}\Big|_{\varepsilon=0}\frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} L\left(\tilde{t},\tilde{\mathbf{q}},\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}\tilde{t}}\right) + L\,\frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0}\frac{\mathrm{d}\tilde{t}}{\mathrm{d}t} = \frac{\partial L}{\partial t}\tau + \frac{\partial L}{\partial\mathbf{q}}\boldsymbol{\eta} + \frac{\partial L}{\partial\dot{\mathbf{q}}}(\dot{\boldsymbol{\eta}} - \dot{\mathbf{q}}\dot{\tau}) + L\dot{\tau} = 0\,. \tag{3.67}$$

Using Lagrange's equations we can rewrite the previous equation as follows:

$$0 = \frac{\partial L}{\partial t}\tau + \dot{\mathbf{p}}\,\boldsymbol{\eta} + \mathbf{p}\,\dot{\boldsymbol{\eta}} - \mathbf{p}\,\dot{\mathbf{q}}\,\dot{\tau} + L\dot{\tau} = \frac{\partial L}{\partial t}\tau + \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{p}\,\boldsymbol{\eta}) - h\dot{\tau} = \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{p}\,\boldsymbol{\eta} - h\tau) + \left(\dot{h} + \frac{\partial L}{\partial t}\right)\tau\,.$$

It is straightforward to check that the last term vanishes identically on account of Lagrange's equations:

$$\dot{h} = \dot{\mathbf{p}}\,\dot{\mathbf{q}} + \mathbf{p}\,\ddot{\mathbf{q}} - \frac{\partial L}{\partial t} - \frac{\partial L}{\partial\mathbf{q}}\dot{\mathbf{q}} - \mathbf{p}\,\ddot{\mathbf{q}} = \left(\dot{\mathbf{p}} - \frac{\partial L}{\partial\mathbf{q}}\right)\dot{\mathbf{q}} - \frac{\partial L}{\partial t} = -\frac{\partial L}{\partial t}\,.$$ ■

Remark. Generally speaking, a *symmetry* of an object is any transformation leaving the object invariant. The set of all symmetries of an object is a *group* (with composition as group multiplication), since i) the composition of two symmetries is clearly a symmetry, ii) the inverse of a symmetry is also a symmetry (why?), and iii) the identity transformation is obviously a symmetry. Thus the family of transformations (3.64) are a *one-parameter group* of symmetries of the *action*. Families of symmetries of an object depending on one or more continuous parameters —like the transformations (3.64)— are usually called *continuous symmetries*. Thus the import of Noether's theorem is that *every continuous symmetry of the action is yields a conservation law*. As mentioned before, this is in fact one of the most fundamental principles in modern physics, which actually holds in much more general settings like classical or quantum field theory. ■

Example 3.11. Consider a system of N particles subject only to irrotational forces generated by a potential $V(t,\mathbf{r}_1,\dots,\mathbf{r}_N)$. We can then take the Cartesian coordinates $\mathbf{q} = (\mathbf{r}_1,\dots,\mathbf{r}_N)$ as generalized coordinates, and $L = T - V$ as the system's Lagrangian. The kinetic energy

$$T = \frac{1}{2}\sum_{i=1}^{N} m_i\dot{\mathbf{r}}_i^2$$

is then invariant under two types of transformations:

i. *Translations* of the particles' coordinates in the direction of a unit vector $\mathbf{n}$:

$$\tilde{t} = t\,, \qquad \tilde{\mathbf{r}}_i = \mathbf{r}_i = \mathbf{r}_i + \varepsilon\mathbf{n} \qquad (1 \leqslant i \leqslant N, \quad \varepsilon \in \mathbb{R})\,; \tag{3.68}$$

indeed, $\dot{\tilde{\mathbf{r}}}_i = \dot{\mathbf{r}}_i$.

[8] Indeed, integrating (3.66) between t_1 and t_2 we obtain Eq. (3.65). Conversely, Eq. (3.65) implies (3.66), since the times t_1 and t_2 are arbitrary.

ii. *Rotations* of the particles' coordinates around an axis **n**:

$$\tilde{t} = t\,, \qquad \tilde{\mathbf{r}}_i = R(\varepsilon)\mathbf{r}_i \qquad (1 \leqslant i \leqslant N, \quad \varepsilon \in \mathbb{R}); \tag{3.69}$$

indeed, $\dot{\tilde{\mathbf{r}}}_i = R(\varepsilon)\dot{\mathbf{r}}_i$ and hence $\dot{\tilde{\mathbf{r}}}_i^2 = \dot{\mathbf{r}}_i^2$.

Obviously, the Lagrangian —and hence the action, since $\tilde{t} = t$— will be invariant under the latter transformations if and only if the potential $V(t, \mathbf{r}_1, \dots, \mathbf{r}_N)$ is, i.e., provided that

$$V(t, \tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_N) = V(t, \mathbf{r}_1, \dots, \mathbf{r}_N)\,.$$

Suppose, first, that the potential is invariant under the translations (3.68), for which

$$\tau = 0\,, \qquad \boldsymbol{\eta}_i = \mathbf{n} \qquad (1 \leqslant i \leqslant N)\,.$$

The corresponding *conserved quantity* is then

$$I = \sum_{i=1}^{N} \frac{\partial L}{\partial \dot{\mathbf{r}}_i}\boldsymbol{\eta}_i - 0 \cdot h = \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i \cdot \mathbf{n} = \mathbf{n} \cdot \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i = \boxed{\mathbf{P} \cdot \mathbf{n}}\,,$$

i.e., the component of the system's *total linear momentum* along the direction of the vector **n**.

Suppose next that the potential is invariant under the rotations (3.69). What is the conserved quantity associated with this invariance of the action? To answer this question, let us take the z axis in the direction of the vector **n**, so that

$$R(\varepsilon) = \begin{pmatrix} \cos\varepsilon & -\sin\varepsilon & 0 \\ \sin\varepsilon & \cos\varepsilon & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Expanding $R(\varepsilon)$ in powers of ε we obtain

$$\tilde{\mathbf{r}}_i = R(\varepsilon)\mathbf{r}_i = \mathbf{r}_i + \varepsilon A\mathbf{r}_i + O(\varepsilon^2)\,, \qquad 1 \leqslant i \leqslant N\,, \tag{3.70}$$

with

$$A = R'(0) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Since

$$A\mathbf{r}_i = (-y_i, x_i, 0) = \mathbf{e}_3 \times \mathbf{r}_i = \mathbf{n} \times \mathbf{r}_i\,,$$

we can rewrite Eq. (3.70) in vector form as

$$\tilde{\mathbf{r}}_i = \mathbf{r}_i + \varepsilon\, \mathbf{n} \times \mathbf{r}_i + O(\varepsilon^2)\,, \qquad 1 \leqslant i \leqslant N\,.$$

Hence in this case

$$\tau = 0\,, \qquad \boldsymbol{\eta}_i = \mathbf{n} \times \mathbf{r}_i\,, \qquad 1 \leqslant i \leqslant N\,,$$

and the conserved quantity associated with the invariance of the action under rotations around the **n** axis is therefore

$$I = \sum_{i=1}^{N} \frac{\partial L}{\partial \dot{\mathbf{r}}_i}\boldsymbol{\eta}_i - 0 \cdot h = \sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i \cdot (\mathbf{n} \times \mathbf{r}_i) = \mathbf{n} \cdot \sum_{i=1}^{N} m_i \mathbf{r}_i \times \dot{\mathbf{r}}_i = \boxed{\mathbf{J} \cdot \mathbf{n}\,,}$$

where **J** is the system's total angular momentum.

Exercise. Determine the conserved quantity $I(t,\mathbf{q},\dot{\mathbf{q}})$ associated with the invariance of the action under the *space-time dilations*

$$\tilde{t} = \lambda^{\alpha} t\,, \qquad \tilde{\mathbf{q}} = \lambda\mathbf{q} \qquad (\lambda > 0,\ \ \alpha \in \mathbb{R})\,. \tag{3.71}$$

Solution. As in the formulation of Noether's theorem the parameter $\varepsilon = 0$ corresponds to the identity transformation, we set $\lambda = e^{\varepsilon}$ in Eqs. (3.71). Expanding to first order in ε we then obtain

$$\tilde{t} = e^{\alpha\varepsilon} t = t + \varepsilon\alpha t + O(\varepsilon^2)\,, \qquad \tilde{\mathbf{q}} = e^{\varepsilon}\mathbf{q} = \mathbf{q} + \varepsilon\mathbf{q} + O(\varepsilon)^2\,.$$

Hence

$$\tau = \alpha t\,, \qquad \boldsymbol{\eta} = \mathbf{q}\,,$$

and thus the conserved quantity associated with the invariance of the action under the transformations (3.71) is given by

$$I(t,\mathbf{q},\dot{\mathbf{q}}) = \mathbf{q}\,\frac{\partial L}{\partial \dot{\mathbf{q}}} - \alpha t h\,.$$

Note that the action is invariant under the dilations (3.71) if the Lagrangian L verifies the condition

$$L\Big(\tilde{t},\tilde{\mathbf{q}},\frac{d\tilde{\mathbf{q}}}{d\tilde{t}}\Big)\,d\tilde{t} = L(\lambda^{\alpha}t,\lambda\mathbf{q},\lambda^{1-\alpha}\dot{\mathbf{q}})\lambda^{\alpha}\,dt = L(t,\mathbf{q},\dot{\mathbf{q}})\,dt\,,$$

i.e., if L transforms under dilations as

$$L(\lambda^{\alpha}t,\lambda\mathbf{q},\lambda^{1-\alpha}\dot{\mathbf{q}}) = \lambda^{-\alpha}L(t,\mathbf{q},\dot{\mathbf{q}})\,.$$

Suppose, for instance, that the system is *natural*. In this case the previous condition becomes

$$\frac{1}{2}\lambda^{2-2\alpha}\sum_{i,j=1}^{n} a_{ij}(\lambda^{\alpha}t,\lambda\mathbf{q})\dot{q}_i\dot{q}_j - V(\lambda^{\alpha}t,\lambda\mathbf{q}) = \frac{1}{2}\lambda^{-\alpha}\sum_{i,j=1}^{n} a_{ij}(t,\mathbf{q})\dot{q}_i\dot{q}_j - \lambda^{-\alpha}V(t,\mathbf{q})\,.$$

Equating the coefficient of $\dot{q}_i\dot{q}_j$ in both sides of this equality we obtain

$$\lambda^{2-2\alpha}a_{ij}(\lambda^{\alpha}t,\lambda\mathbf{q}) = \lambda^{-\alpha}a_{ij}(t,\mathbf{q}) \iff a_{ij}(\lambda^{\alpha}t,\lambda\mathbf{q}) = \lambda^{\alpha-2}a_{ij}(t,\mathbf{q})\,,$$

and hence

$$V(\lambda^{\alpha}t,\lambda\mathbf{q}) = \lambda^{-\alpha}V(t,\mathbf{q})\,.$$

For example, if the matrix a_{ij} is constant then we must have $\alpha = 2$, and therefore

$$V(\lambda^{2}t,\lambda\mathbf{q}) = \lambda^{-2}V(t,\mathbf{q})\,.$$

Consider, for instance, the case of a particle of mass m that moves subject to the central potential $V(r) = k/(2r^2)$, with $k \neq 0$. From the previous discussion it easily follows that in this case the action is invariant under the transformation (3.71) with $\alpha = 2$. In this case the energy $T + V = E$ and the function

$$I = m\mathbf{r}\dot{\mathbf{r}} - 2ht = mr\dot{r} - 2Et = \frac{d}{dt}\Big(\frac{1}{2}mr^2 - Et^2\Big) = \text{const.} \tag{3.72}$$

is conserved. Note that the value of the conserved quantity I can be easily expressed in terms of the initial data $r_0 := r(0)$ and $\dot{r}_0 := \dot{r}(0)$ by evaluating it at $t = 0$:

$$I = mr\dot{r} - 2Et\,\big|_{t=0} = mr_0\dot{r}_0\,.$$

Integrating Eq. (3.72) we can easily determine the motion of the r coordinate:

$$\frac{1}{2}mr^2 = \frac{1}{2}mr_0^2 + mr_0\dot{r}_0 t + Et^2 \quad \Longrightarrow \quad \boxed{r = \sqrt{r_0^2 + 2r_0\dot{r}_0 t + \frac{2E}{m}t^2}\,.}$$

The motion of the angular coordinate φ (in the plane of motion) is obtained integrating the law of conservation of angular momentum $mr^2\dot{\varphi} = J$:

$$\boxed{\varphi = \varphi_0 + \frac{J}{m}\int_0^t \frac{\mathrm{d}s}{r^2(s)} = \varphi_0 + \frac{J}{m}\int_0^t \frac{\mathrm{d}s}{r_0^2 + 2r_0\dot{r}_0 s + \frac{2E}{m}s^2}\,, \qquad \varphi_0 := \varphi(0)\,.}$$

If $E = 0$ the integral is elementary:

$$\varphi = \begin{cases} \varphi_0 + \frac{Jt}{mr_0^2}\,, & \dot{r}_0 = 0 \\ \varphi_0 + \frac{J}{2mr_0\dot{r}_0}\log\left(1 + \frac{2\dot{r}_0 t}{r_0}\right), & \dot{r}_0 \neq 0\,. \end{cases}$$

In the more general case $E \neq 0$ the integral can be evaluated in terms of hyperbolic, rational or trigonometric functions depending on whether the discriminant of the polynomial in the denominator (equal to $4r_0^2(\dot{r}_0^2 - 2E/m) = -4(J^2 + km)/m^2$) is respectively positive, zero or negative (exercise).

3.5 Small oscillations

In this section we shall analyze the motion of a conservative mechanical system near a stable equilibrium. We shall assume that the constraints are *holonomic* and *time-independent*, so that the position vector of each particle is a function only of the generalized coordinates $\mathbf{q} = (q_1, \ldots, q_n)$ (not explicitly depending on time):

$$\mathbf{r}_k = \mathbf{r}_k(\mathbf{q})\,, \qquad k = 1, \ldots, N\,.$$

Since

$$\dot{\mathbf{r}}_k = \sum_{j=1}^{n} \frac{\partial \mathbf{r}_k(\mathbf{q})}{\partial q_j}\,\dot{q}_j\,, \tag{3.73}$$

the system's kinetic energy is of the form

$$T(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\sum_{i,j=1}^{n} t_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j$$

with

$$t_{ij}(\mathbf{q}) = \sum_{k=1}^{N} m_k \frac{\partial \mathbf{r}_k(\mathbf{q})}{\partial q_i} \cdot \frac{\partial \mathbf{r}_k(\mathbf{q})}{\partial q_j} = t_{ji}(\mathbf{q})\,.$$

If the system is conservative, with potential energy $V(\mathbf{q})$, its Lagrangian is given by

$$\boxed{L = \frac{1}{2}\sum_{i,j=1}^{n} t_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j - V(\mathbf{q})\,,}$$

and Lagrange's equations of motion read:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\sum_{j=1}^{n} t_{ij}(\mathbf{q})\dot{q}_j\Big) + \frac{\partial V}{\partial q_i}(\mathbf{q})$$
$$= \sum_{j=1}^{n} t_{ij}(\mathbf{q})\ddot{q}_j + \sum_{j,k=1}^{n} \frac{\partial t_{ij}}{\partial q_k}(\mathbf{q})\,\dot{q}_j\dot{q}_k + \frac{\partial V}{\partial q_i}(\mathbf{q}) = 0\,, \qquad i = 1,\dots,n\,. \quad (3.74)$$

Thus an equilibrium solution $\mathbf{q}(t) = \mathbf{q}_0$ exists if and only if

$$\frac{\partial V}{\partial q_i}(\mathbf{q}_0) = 0\,, \qquad i = 1,\dots,n\,,$$

i.e., if $\mathbf{q}_0$ is a *critical point* of the potential $V(\mathbf{q})$. As in the one-dimensional case (cf. Section 1.5.1), it can be shown that the equilibrium $\mathbf{q}_0$ is *stable* if and only if $\mathbf{q}_0$ is a *local minimum* of V.

We wish to describe the motion of the system near a stable equilibrium $\mathbf{q}_0$. To this end, let us assume without loss of generality that $\mathbf{q}_0 = 0$ (choosing $\mathbf{q} - \mathbf{q}_0$ as new generalized coordinates if necessary) and normalize the potential so that $V(0) = 0$. The energy of the equilibrium solution $\mathbf{q} = 0$ is then $E_0 = 0$. Consider now a motion of the system close to the equilibrium solution $\mathbf{q} = 0$, i.e., with $|\mathbf{q}(0)|$ and $|\dot{\mathbf{q}}(0)|$ small. The energy of such a motion is then close to $E_0 = 0$ and verifies

$$E = T + V(\mathbf{q}) \geqslant V(\mathbf{q}) \geqslant V(0) = 0$$

(since $\mathbf{q} = 0$ is by hypothesis a local minimum of V), i.e., E is positive and small. Taking into account that the first-order partial derivatives of V vanish at the origin (since $\mathbf{q} = 0$ is by hypothesis a critical point of V), we can write its Taylor expansion about 0 as[9]

$$V(\mathbf{q}) = \frac{1}{2}\sum_{i,j=1}^{n} b_{ij}q_iq_j + o(|\mathbf{q}|^2)\,, \qquad \text{with} \quad b_{ij} = \frac{\partial^2 V}{\partial q_i q_j}(0) = b_{ji}\,.$$

Furthermore, since 0 is a local minimum of V the quadratic form $\sum_{i,j=1}^{n} b_{ij}q_iq_j$ is positive semidefinite, i.e., the eigenvalues of the symmetric $n\times n$ matrix

$$B = (b_{ij})_{1\leqslant i,j\leqslant n}$$

are nonnegative. We shall actually assume that B is *positive definite* (i.e., all its eigenvalues are strictly positive), so that

$$V(\mathbf{q}) \simeq \frac{1}{2}\sum_{i,j=1}^{n} b_{ij}q_iq_j$$

near the origin. Similarly, the system's kinetic energy can be expanded near the equilibrium $\mathbf{q} = \dot{\mathbf{q}} = 0$ as

$$T = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\dot{q}_i\dot{q}_j + o(|\mathbf{q}|^2 + |\dot{\mathbf{q}}|^2)\,, \qquad \text{with} \quad a_{ij} = t_{ij}(0) = a_{ji}\,.$$

Note also that the the quadratic form

$$T_0 := \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\dot{q}_i\dot{q}_j$$

[9]In the previous formula we are using the standard notation $o(t)$ to denote a function verifying $\lim_{t\to 0}\frac{o(t)}{t} = 0$.

is *positive definite*. Indeed, the kinetic energy $T(\mathbf{q}, \dot{\mathbf{q}})$ is nonnegative and vanishes only if $\dot{\mathbf{r}}_k = 0$ for all $k = 1, \ldots, N$. In particular, $T_0 = T(0, \dot{\mathbf{q}})$ is also nonnegative, and hence positive semidefinite, and by Eq. (3.73) it can only vanish when

$$\sum_{j=1}^{n} \frac{\partial \mathbf{r}_k}{\partial q_j}(0)\, \dot{q}_j = 0\,, \qquad k = 1, \ldots, N\,. \tag{3.75}$$

Recall that the vectors

$$\frac{\partial \mathbf{x}}{\partial q_j} := \left(\frac{\partial \mathbf{r}_1}{\partial q_j}, \ldots, \frac{\partial \mathbf{r}_N}{\partial q_j} \right), \qquad j = 1, \ldots, n\,,$$

are linearly independent at each point, since they are a basis of the tangent space to the system's constraint manifold (see Section 3.3.2). Thus equations (3.75), which are equivalent to the single relation

$$\sum_{j=1}^{n} \frac{\partial \mathbf{x}}{\partial q_j}(0)\, \dot{q}_j = 0,$$

imply that $\dot{q}_j = 0$ for all j. This shows that $T_0 = 0$ if and only if $\dot{\mathbf{q}} = 0$, and hence T_0 is positive definite, as claimed. Thus near the equilibrium solution $\mathbf{q} = \dot{\mathbf{q}} = 0$ we also have $T \simeq T_0$. It follows that for small displacements $\mathbf{q}$ and small velocities $\dot{\mathbf{q}}$ the system's Lagrangian can be approximated by

$$L_0 := \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \dot{q}_i \dot{q}_j - \frac{1}{2} \sum_{i,j=1}^{n} b_{ij} q_i q_j\,. \tag{3.76}$$

The motion of the system near its stable equilibrium $\mathbf{q} = 0$ is thus approximately governed by the Euler–Lagrange equations of the Lagrangian L_0, namely

$$\sum_{j=1}^{n} (a_{ij}\ddot{q}_j + b_{ij} q_j) = 0\,, \qquad i = 1, \ldots, n\,,$$

or in matrix form

$$A\ddot{\mathbf{q}} + B\mathbf{q} = 0\,, \tag{3.77}$$

where A is the positive definite $n \times n$ symmetric matrix with matrix elements a_{ij}.

Equations (3.77) are a system of n second-order linear homogeneous differential equations with constant coefficients, which can also be derived by linearizing the exact equations of motion obtained from the Lagrangian L (exercise). One of the easiest ways of solving the linearized equations of motion (3.77) is by transforming the Lagrangian L_0 to a suitable canonical form. Indeed, since T_0 is positive definite there is a non-singular (in general non-orthogonal) linear change of variables

$$q_i = \sum_{j=1}^{n} M_{ij} \tilde{q}_j\,, \qquad i = 1, \ldots, n\,,$$

or in matrix form

$$\mathbf{q} = M\tilde{\mathbf{q}}$$

(and consequently $\dot{\mathbf{q}} = M\dot{\tilde{\mathbf{q}}}$), transforming the positive definite quadratic form T_0 into $\dot{\tilde{\mathbf{q}}}^2/2$. In other words, there exists a non-singular matrix M such that

$$M^{\mathsf{T}} A M = \mathbb{1}\,.$$

Since

$$L_0 = T_0 - \frac{1}{2} \mathbf{q}^{\mathsf{T}} \cdot B\mathbf{q}\,,$$

in the new generalized coordinates $\tilde{\mathbf{q}}$ the Lagrangian L_0 can then be written as

$$L_0 = \frac{1}{2}\dot{\tilde{\mathbf{q}}}^2 - \frac{1}{2}\sum_{i,j=1}^{n} \tilde{b}_{ij}\tilde{q}_i\tilde{q}_j = \frac{1}{2}\dot{\tilde{\mathbf{q}}}^2 - \frac{1}{2}\tilde{\mathbf{q}}^{\mathsf{T}} \cdot \tilde{B}\tilde{\mathbf{q}},$$

where $\tilde{b}_{ij}$ is the matrix element of the matrix

$$\tilde{B} = M^{\mathsf{T}}BM .$$

Since the matrix $\tilde{B}$ is still symmetric and positive definite (exercise), it can be diagonalized by an orthogonal transformation, i.e., there is a real *orthogonal* matrix O such that

$$O^{\mathsf{T}}\tilde{B}O = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} =: \Lambda \tag{3.78}$$

with Λ diagonal. Note that the numbers λ_i are all *positive*, since $\tilde{B}$ is a positive definite matrix. Defining new generalized coordinates $\mathbf{Q}$ by the linear transformation $\tilde{\mathbf{q}} = O\mathbf{Q}$, and taking into account that $\dot{\tilde{\mathbf{q}}}^2 = (O\dot{\mathbf{Q}})^2 = \dot{\mathbf{Q}}^2$ since O is orthogonal, we easily find

$$\boxed{L_0 = \frac{1}{2}\dot{\mathbf{Q}}^2 - \frac{1}{2}\mathbf{Q} \cdot \Lambda\mathbf{Q} = \frac{1}{2}\sum_{i=1}^{n}(\dot{Q}_i^2 - \lambda_i Q_i^2) .}$$

By the *covariance* of Lagrange's equations (cf. Section 3.2.3), the linearized equations of motion in the generalized coordinates $\mathbf{Q}$ are simply the Euler–Lagrange equations of the latter Lagrangian with respect to the variables $(\mathbf{Q}, \dot{\mathbf{Q}})$, namely the *decoupled* system

$$\boxed{\ddot{Q}_i + \lambda_i Q_i = 0 , \qquad i = 1, \dots, n .} \tag{3.79}$$

In other words, in the generalized coordinates $\mathbf{Q}$ the system is equivalent to a collection of n *decoupled harmonic oscillators* with frequencies

$$\omega_i := \sqrt{\lambda_i}.$$

The general solution of the system (3.79) is therefore

$$\boxed{Q_i = A_i \cos(\omega_i t + \alpha_i) , \qquad i = 1, \dots, n ,}$$

or in vector form

$$\mathbf{Q} = \sum_{i=1}^{n} A_i \cos(\omega_i t + \alpha_i)\mathbf{e}_i,$$

with $A_i \geqslant 0$ and $\alpha_i \in [0, 2\pi)$ arbitrary constants. In other words, equations (3.79) possess a fundamental system of solutions of the form

$$\boxed{\mathbf{Q}^{(i)}(t) = \cos(\omega_i t + \alpha_i)\mathbf{e}_i , \qquad i = 1, \dots, n ,} \tag{3.80}$$

where $\mathbf{e}_i = (0, \dots, 1, \dots, 0)$ is the i-th canonical basis vector. The generalized coordinates $\mathbf{Q}$ and the n fundamental solutions (3.80) are respectively called the system's **normal coordinates** and **normal modes**. Likewise, the n frequencies ω_i ($i = 1, \dots, n$) are called the system's **normal frequencies**. In terms of the original generalized coordinates

$$\boxed{\mathbf{q} = M\tilde{\mathbf{q}} = MO\mathbf{Q}} \tag{3.81}$$

the normal modes (3.80) become

$$\boxed{\mathbf{q}^{(i)}(t) = \mathbf{c}_i \cos(\omega_i t + \alpha_i)\,, \qquad i = 1, \dots, n\,,} \tag{3.82}$$

where the n-dimensional vectors $\mathbf{c}_i$ are given by

$$\boxed{\mathbf{c}_i = (MO)\mathbf{e}_i\,.} \tag{3.83}$$

In other words, the vectors $\mathbf{c}_i$ are the *columns* of the matrix MO satisfying

$$(MO)^\mathsf{T} A(MO) = \mathbb{1}\,, \qquad (MO)^\mathsf{T} B(MO) = \Lambda \tag{3.84}$$

(exercise). In particular, the n vectors $\mathbf{c}_i$ are *linearly independent*. The general solution of the system's linearized equations of motion (3.77) —which, by the previous argument, is an approximate solution of its exact equations of motion (3.74) near the stable equilibrium $\mathbf{q} = 0$— is an arbitrary linear combination

$$\boxed{\mathbf{q}(t) = \sum_{i=1}^{n} a_i \mathbf{q}^{(i)}(t)\,,}$$

with $a_i \in \mathbb{R}$ constant, of the n normal modes $\mathbf{q}^{(i)}(t)$.

Remarks.

- The matrices M and O, and therefore the vectors $\mathbf{c}_i$, are not unique.

- The vectors $\mathbf{c}_i$ defined by Eq. (3.83) are in general not mutually orthogonal nor of unit length. However, since the n vectors $O\mathbf{e}_i$ are orthonormal (being the columns of an orthogonal matrix), the vectors $\mathbf{c}_i$ satisfy the relations

$$\mathbf{c}_i \cdot A\mathbf{c}_j = \delta_{ij}\,, \qquad i, j = 1, \dots, n\,.$$

Indeed, taking into account that $M^\mathsf{T} AM = \mathbb{1}$ we have

$$\mathbf{c}_i \cdot A\mathbf{c}_j = (MO\mathbf{e}_i) \cdot (AMO\mathbf{e}_j) = (O\mathbf{e}_i) \cdot (M^\mathsf{T} AMO\mathbf{e}_j) = (O\mathbf{e}_i) \cdot (O\mathbf{e}_j) = \delta_{ij}\,. \qquad \blacksquare$$

How does one find in practice the frequencies ω_i and the corresponding vectors $\mathbf{c}_i$ determining the system's normal modes (3.82) in the original coordinates $\mathbf{q}$? To answer this question, it suffices to note that since $\mathbf{q}^{(i)}(t)$ is a solution of the linearized equations (3.77) the vector $\mathbf{c}_i$ must satisfy the linear system

$$\boxed{(B - \omega_i^2 A)\mathbf{c}_i = 0\,, \qquad i = 1, \dots, n\,.} \tag{3.85}$$

Since $\mathbf{c}_i$ is nonzero we must therefore have $\det(B - \omega_i^2 A) = 0$. In other words, the normal frequencies $\omega_i = \sqrt{\lambda_i}$ are the square roots of the n solutions λ_i (counting multiplicities) of the *characteristic equation*

$$\boxed{\det(B - \lambda A) = 0\,.} \tag{3.86}$$

The numbers λ_i are called the eigenvalues of the matrix B relative to the positive definite matrix A (in particular, when $A = \mathbb{1}$ the λ_i's are the ordinary eigenvalue of B). Note that the λ_i's are the eigenvalues of the matrix $\widetilde{B}$ in the previous discussion (cf. Eq. (3.78)). For each such eigenvalue $\lambda_i = \omega_i^2$, the corresponding (eigen)vector $\mathbf{c}_i$ is then found solving the linear system (3.85). Note that the previous argument guarantees that there is a basis $\{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ of $\mathbb{R}^n$ whose elements $\mathbf{c}_i$ satisfy Eqs. (3.85) (just take as $\mathbf{c}_i$ the i-th column of the matrix MO constructed as explained above).

Remarks.

- If ω_i^2 is a simple root of the characteristic equation it can be shown that the corresponding vector $\mathbf{c}_i$ is determined by Eq. (3.85) up to a multiplicative constant. In general, if the frequency ω_i is m times *degenerate*, i.e., if ω_i^2 is a root of the characteristic equation with multiplicity $m > 1$, it can be shown that there are exactly m linearly independent solutions of Eq. (3.85). Both of these statements are easily proved by noting that from Eqs. (3.78) and (3.84) it follows that the system (3.85) is equivalent to

$$(\Lambda - \omega_i^2)(MO)^{-1}\mathbf{c}_i = 0\,, \qquad i = 1, \dots, n\,.$$

- Comparison of Eqs. (3.81) and (3.83) shows that once the vectors $\mathbf{c}_i$ have been obtained the normal coordinates $\mathbf{Q}$ can be found through the formula

$$\mathbf{Q} = C^{-1}\mathbf{q}\,, \tag{3.87}$$

where C is the matrix whose columns are the components of the vectors $\mathbf{c}_1, \dots, \mathbf{c}_n$ —i.e., the change of basis matrix from the canonical basis of $\mathbb{R}^n$ to the basis $\{\mathbf{c}_1, \dots, \mathbf{c}_n\}$. ■

Example 3.12. *Double pendulum.*

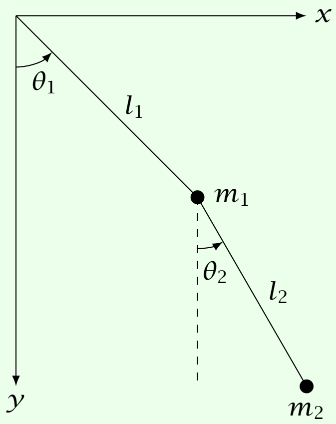


Figure 3.3. Generalized coordinates (θ_1, θ_2) for the double pendulum system.

Consider the double pendulum schematically represented in Fig. 3.3. Calling $\mathbf{r}_\alpha = (x_\alpha, y_\alpha)$ the position vector of the particle $\alpha = 1, 2$, the system's constraints are

$$\mathbf{r}_1^2 - l_1^2 = 0\,, \qquad (\mathbf{r}_2 - \mathbf{r}_1)^2 - l_2^2 = 0\,.$$

These constraints are obviously holonomic and time-independent. Moreover, the principle of virtual work clearly holds, since the constraint forces —the tension of the wire or rod connecting the first particle to the anchor point and the second particle to the first one— are respectively parallel to the vectors $\mathbf{r}_1$ and $\mathbf{r} := \mathbf{r}_2 - \mathbf{r}_1$, and thus perpendicular to the particles' infinitesimal displacements. Hence we can apply the Lagrangian formalism. We shall use as generalized coordinates the two angles θ_1 and θ_2 between the pendulums' strings and the vertical (see Fig. 3.3). Taking the y axis *downwards* (see Fig. 3.3) we then have

$$\mathbf{r}_1 = l_1(\sin\theta_1, \cos\theta_1)\,, \qquad \mathbf{r} = l_2(\sin\theta_2, \cos\theta_2)\,,$$

and therefore

$$\dot{\mathbf{r}}_1 = l_1\dot\theta_1(\cos\theta_1, -\sin\theta_1)\,, \qquad \dot{\mathbf{r}} = l_2\dot\theta_2(\cos\theta_2, -\sin\theta_2)\,.$$

Hence

$$\begin{aligned}T &= \frac{1}{2}\,m_1\dot{\mathbf{r}}_1^2 + \frac{1}{2}\,m_2(\dot{\mathbf{r}}_1 + \dot{\mathbf{r}})^2 = \frac{1}{2}\,M\dot{\mathbf{r}}_1^2 + \frac{1}{2}\,m_2\dot{\mathbf{r}}^2 + m_2\dot{\mathbf{r}}_1\cdot\dot{\mathbf{r}}\\ &= \frac{1}{2}\,Ml_1^2\dot{\theta}_1^2 + \frac{1}{2}\,m_2l_2^2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2)\\ &= \frac{1}{2}\,Ml_1^2\dot{\theta}_1^2 + \frac{1}{2}\,m_2l_2^2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1-\theta_2)\,,\end{aligned}$$

where $M := m_1 + m_2$. Likewise, the potential energy is given by

$$V = -m_1gy_1 - m_2gy_2 = -m_1gy_1 - m_2g(y + y_1) = -Mgl_1\cos\theta_1 - m_2gl_2\cos\theta_2\,,$$

where y is the vertical coordinate of the relative position vector $\mathbf{r}$, and therefore the system's Lagrangian can be taken as

$$\boxed{L = T - V = \frac{1}{2}\,Ml_1^2\dot{\theta}_1^2 + \frac{1}{2}\,m_2l_2^2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1-\theta_2) + Mgl_1\cos\theta_1 + m_2gl_2\cos\theta_2\,.}$$

The equilibria are determined by the system

$$\frac{\partial V}{\partial\theta_1} = Mgl_1\sin\theta_1 = 0\,,\qquad \frac{\partial V}{\partial\theta_1} = m_2gl_2\sin\theta_2 = 0\,,$$

and are therefore (up to integer multiples of 2π) the four points[a]

$$(0,0)\,,\quad (0,\pi),\quad (\pi,0)\,,\quad (\pi,\pi)\,.$$

It is straightforward to ascertain that the point $(0,0)$ is the unique local (in fact, global) minimum of V ($(\pi,0)$ and $(0,\pi)$ are saddle points and (π,π) is a global maximum). The exact equations of motion

$$\begin{aligned}&Ml_1^2\ddot{\theta}_1 + m_2l_1l_2\ddot{\theta}_2\cos(\theta_1-\theta_2) - m_2l_1l_2\dot{\theta}_2(\dot{\theta}_1-\dot{\theta}_2)\sin(\theta_1-\theta_2)\\ &\qquad\qquad + Mgl_1\sin\theta_1 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_1-\theta_2) = 0\,,\\ &m_2l_2^2\ddot{\theta}_2 + m_2l_1l_2\ddot{\theta}_1\cos(\theta_1-\theta_2) - m_2l_1l_2\dot{\theta}_1(\dot{\theta}_1-\dot{\theta}_2)\sin(\theta_1-\theta_2)\\ &\qquad\qquad + m_2gl_2\sin\theta_2 - m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_1-\theta_2) = 0\,,\end{aligned}$$

are a system of nonlinear coupled second-order differential equations that cannot be solved in closed form (i.e., in terms of elementary functions and their primitives). On the other hand, we can easily study the system's motion near the stable equilibrium $\theta_1 = \theta_2 = 0$, i.e., the small oscillations of the two pendulums, through the method explained above.

To begin with, taking into account that

$$\cos\theta = 1 - \frac{\theta^2}{2} + o(\theta^2)$$

we easily obtain

$$\begin{aligned}T_0 &= T(0,0,\dot{\theta}_1,\dot{\theta}_2) = \frac{1}{2}\,Ml_1^2\dot{\theta}_1^2 + \frac{1}{2}\,m_2l_2^2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\,,\\ V &= -(M+m_2)g + \frac{g}{2}\,(Ml_1\theta_1^2 + m_2l_2\theta_2^2) + o(\theta_1^2+\theta_2^2)\,,\end{aligned}$$

and thus (ignoring the irrelevant constant term in V)

$$A = \begin{pmatrix} Ml_1^2 & m_2 l_1 l_2 \\ m_2 l_1 l_2 & m_2 l_2^2 \end{pmatrix} = Ml_1^2 \begin{pmatrix} 1 & \lambda\mu \\ \lambda\mu & \lambda^2\mu \end{pmatrix}, \qquad B = g \begin{pmatrix} Ml_1 & 0 \\ 0 & m_2 l_2 \end{pmatrix} = Ml_1^2 \omega_0^2 \begin{pmatrix} 1 & 0 \\ 0 & \lambda\mu \end{pmatrix},$$

where

$$\omega_0 := \sqrt{\frac{g}{l_1}}$$

is the natural frequency of the first pendulum and we have set

$$\lambda := \frac{l_2}{l_1}, \qquad \mu := \frac{m_2}{M}.$$

Since B is diagonal, it is convenient to write the characteristic equation (3.86) in the equivalent way

$$\det\left(A - \frac{B}{\omega^2}\right) = M^2 l_1^4 \begin{vmatrix} 1 - \frac{\omega_0^2}{\omega^2} & \lambda\mu \\ \lambda\mu & \lambda\mu\left(\lambda - \frac{\omega_0^2}{\omega^2}\right) \end{vmatrix} = 0 \iff \left(1 - \frac{\omega_0^2}{\omega^2}\right)\left(\lambda - \frac{\omega_0^2}{\omega^2}\right) - \lambda\mu = 0,$$

or

$$\left(\frac{\omega_0}{\omega}\right)^4 - (\lambda + 1)\left(\frac{\omega_0}{\omega}\right)^2 + \lambda(1 - \mu) = 0.$$

The normal frequencies are thus determined by the equation

$$\frac{\omega_0^2}{\omega_\pm^2} = \frac{1}{2}\left(\lambda + 1 \pm \sqrt{(\lambda + 1)^2 - 4\lambda(1 - \mu)}\right),$$

whence

$$\omega_\pm^2 = \omega_0^2 \frac{\lambda + 1 \mp \sqrt{(\lambda + 1)^2 - 4\lambda(1 - \mu)}}{2\lambda(1 - \mu)} = \omega_0^2 \frac{\lambda + 1 \mp \sqrt{(\lambda - 1)^2 + 4\lambda\mu}}{2\lambda(1 - \mu)}.$$

In particular, when the two pendulums have the same length (i.e., for $\lambda = 1$) we simply have

$$\omega_\pm^2 = \frac{\omega_0^2}{1 \pm \sqrt{\mu}} = \frac{\omega_0^2}{1 \pm \sqrt{\frac{m_2}{m_1 + m_2}}}.$$

The two normal modes are found by solving the characteristic equation

$$\left(A - \frac{B}{\omega_\pm^2}\right)\mathbf{c}_\pm = 0,$$

i.e.,

$$\left(1 - \frac{\omega_0^2}{\omega_\pm^2}\right)c_{\pm,1} + \lambda\mu c_{\pm,2} = 0.$$

Using the previous formulas for $\omega_\pm^2$ we can rewrite the last equation as

$$\left(1 - \lambda \mp \sqrt{(\lambda - 1)^2 + 4\lambda\mu}\right)c_{\pm,1} + 2\lambda\mu c_{\pm,2} = 0,$$

thus obtaining the two (unnormalized) vectors

$$\boxed{\mathbf{c}_\pm = c\left(1, \frac{\lambda - 1 \pm \sqrt{(\lambda-1)^2 + 4\lambda\mu}}{2\lambda\mu}\right)}$$

with arbitrary $c \neq 0$. Hence the two normal mode solutions are

$$(\theta_1^{(\pm)}, \theta_2^{(\pm)}) = \mathbf{c}_\pm \cos(\omega_\pm t + \alpha_\pm)\,,$$

i.e.,

$$\boxed{\theta_1^{(\pm)} = c\cos(\omega_\pm t + \alpha_\pm)\,, \qquad \theta_2^{(\pm)} = c\,\frac{\lambda - 1 \pm \sqrt{(\lambda-1)^2 + 4\lambda\mu}}{2\lambda\mu}\cos(\omega_\pm t + \alpha_\pm)\,.}$$

Note that the quotient of the amplitudes of the oscillations of the angles θ_2 and θ_1 in these normal modes, given by

$$\boxed{\frac{\theta_2^{(\pm)}}{\theta_1^{(\pm)}} = \frac{\lambda - 1 \pm \sqrt{(\lambda-1)^2 + 4\lambda\mu}}{2\lambda\mu}\,,}$$

is positive (resp. negative) for the normal mode with the smaller frequency ω_+ (resp. the larger frequency ω_-). Thus in the normal mode with frequency ω_+ the pendulums oscillate *in phase*, whereas in the one with frequency ω_- they oscillate completely *out of phase* (i.e., θ_1 is maximum when θ_2 is minimum, and vice versa). Again, in the particular case in which $l_1 = l_2$ the quotient θ_2/θ_1 simplifies to

$$\frac{\theta_2^{(\pm)}}{\theta_1^{(\pm)}} = \pm\frac{1}{\sqrt{\mu}}\,.$$

[a] In fact, the equilibria $(\pi, 0)$ and (π, π) are not possible in practice due to the fact that the first pendulum is anchored to the ceiling, and thus $\theta_1 \in [-\pi/2, \pi/2]$. For the same reason, the equilibrium $(0, \pi)$ is only possible if $l_2 \leqslant l_1$.

Exercise. Show that for all (positive) values of $\lambda = l_2/l_1$ and $\mu = m_2/M$ we have $\omega_- < \omega_0 < \omega_+$.

Exercise. Find the normal coordinates $\theta_\pm$ for the double pendulum system.

Solution. From Eq. (3.87) we have

$$\begin{aligned}\begin{pmatrix}\theta_+\\ \theta_-\end{pmatrix} &= \begin{pmatrix}1 & 1\\ \frac{\lambda-1+\sqrt{(\lambda-1)^2+4\lambda\mu}}{2\lambda\mu} & \frac{\lambda-1-\sqrt{(\lambda-1)^2+4\lambda\mu}}{2\lambda\mu}\end{pmatrix}\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}\\ &= \frac{1}{2\sqrt{(\lambda-1)^2+4\lambda\mu}}\begin{pmatrix}1-\lambda+\sqrt{(\lambda-1)^2+4\lambda\mu} & 2\lambda\mu\\ \lambda-1+\sqrt{(\lambda-1)^2+4\lambda\mu} & -2\lambda\mu\end{pmatrix}\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}.\end{aligned}$$

In fact, since the normal coordinates Q_i are defined up to multiplication by a constant scalar, we can take as normal coordinates

$$\theta_\pm = \left[1 - \lambda \pm \sqrt{(\lambda-1)^2 + 4\lambda\mu}\right]\theta_1 + 2\lambda\mu\theta_2\,.$$

This expression simplifies considerably when the two pendulums have the same length, in which case (dropping the inessential constant factor $\pm 2\sqrt{\mu}$) we obtain

$$\theta_\pm = \theta_1 \pm \sqrt{\mu}\,\theta_2\,.$$

Example 3.13. *Longitudinal vibrations of the* CO_2 *molecule.*

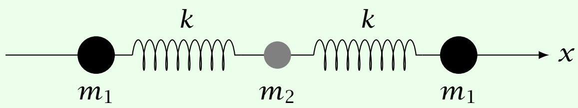


Figure 3.4. Schematic representation of the CO_2 molecule.

Consider a triatomic molecule made up of two identical atoms of mass m_1 and a single atom of mass m_2. We shall also assume that, as is the case with the CO_2 molecule, the molecule's equilibrium configuration is collinear, with the atom of mass m_2 lying between the other two atoms and separated from each of them by the same distance a. Let us choose the x axis along the line joining the equilibrium positions of the three atoms, and place the origin at the equilibrium position of the atom of mass m_2. We shall next study the longitudinal vibrations of the molecule, i.e., the motions of its atoms along the line of the molecule at equilibrium (the x axis). Calling x_1 and x_3 the x coordinates of the atoms of mass m_1 (from left to right), and x_2 that of the atom of mass m_2, the system's kinetic and potential energies are given by

$$\boxed{T = \frac{1}{2}\left[m_1(\dot{x}_1^2 + \dot{x}_3^2) + m_2\dot{x}_2^2\right], \qquad V = U(x_2 - x_1) + U(x_3 - x_2),}$$

where U is the interaction potential between the atoms of mass m_2 and each of the atoms of mass m_1. We have assumed that the interaction between the two atoms of mass m_1 is negligible compared to their interaction with the atom of mass m_2, since the strength of atomic interactions usually falls off very quickly with the distance. Although the potential U is not known in detail, we are only interested in small vibrations of the atoms about their equilibrium position $x_1 = -a$, $x_2 = 0$, $x_3 = a$. Imposing that the partial derivatives of V vanish at equilibrium we easily deduce that $U'(a) = 0$. If we now Taylor expand $U(x)$ about $x = a$ and keep only the lowest order nontrivial term we obtain

$$U(x) \simeq U(a) + \frac{k}{2}(x - a)^2,$$

where $k = U''(a)$, and thus (dropping the inessential constant $U(a)$)

$$V \simeq \frac{k}{2}\left[(x_2 - x_1 - a)^2 + (x_3 - x_2 - a)^2\right].$$

Thus in this approximation (i.e., when $x_2 - x_1$ and $x_3 - x_2$ are both close to a) the molecule behaves as a system of three collinear particles of masses m_1, m_2 and m_1 connected by springs of natural length a and constant k (cf. Fig. 3.4). The system's Lagrangian $L = T - V$ is approximately given by

$$\boxed{L_0 := \frac{1}{2}\left[m_1(\dot{x}_1^2 + \dot{x}_3^2) + m_2\dot{x}_2^2\right] - \frac{k}{2}\left[(x_2 - x_1 - a)^2 + (x_3 - x_2 - a)^2\right].}$$

Since L_0 is invariant under the translation $x_i \mapsto x_i + \varepsilon$ for arbitrary ε, the x component of the linear momentum $P = m_1(\dot{x}_1 + \dot{x}_3) + m_2\dot{x}_2$ is conserved. This suggests, as in the two-body problem, separating the center of mass motion from the particles' relative motion, i.e., to use as generalized coordinates

$$\boxed{X := \frac{1}{M}\left[m_1(x_1 + x_3) + m_2 x_2\right], \qquad q_1 := x_2 - x_1 - a, \qquad q_2 := x_3 - x_2 - a\,,}$$

where $M = 2m_1 + m_2$ is the molecule's total mass. Inverting the latter equations we readily find the following expressions for the atoms' physical coordinates in terms of the generalized ones:

$$\boxed{\begin{aligned} &x_1 = X - \frac{m_1+m_2}{M}\,q_1 - \frac{m_1}{M}\,q_2 - a\,, \qquad x_2 = X + \frac{m_1}{M}\,(q_1 - q_2\,,\\ &x_3 = X + \frac{m_1}{M}\,q_1 + \frac{m_1+m_2}{M}\,q_2 + a)\,, \end{aligned}} \tag{3.88}$$

and hence

$$\dot{x}_1 = \dot{X} - \frac{m_1+m_2}{M}\,\dot{q}_1 - \frac{m_1}{M}\,\dot{q}_2\,, \qquad \dot{x}_2 = \dot{X} + \frac{m_1}{M}\,(\dot{q}_1 - \dot{q}_2)\,, \qquad \dot{x}_3 = \dot{X} + \frac{m_1}{M}\,\dot{q}_1 + \frac{m_1+m_2}{M}\,\dot{q}_2\,.$$

Substituting these expressions into the Lagrangian L_0 and operating we obtain

$$L_0 \simeq \frac{1}{2}M\dot{X}^2 + L_1(q_1, q_2, \dot{q}_1, \dot{q}_2)\,,$$

where

$$\boxed{L_1 = \frac{m_1(m_1+m_2)}{2M}\,(\dot{q}_1^2 + \dot{q}_2^2) + \frac{m_1^2}{M}\,q_1 q_2 - \frac{k}{2}\,(q_1^2 + q_2^2)\,.}$$

Thus in the approximation of small vibrations (i.e., when $|q_1|$ and $|q_2|$ are small) the equation of motion of the center of mass coordinate is $\ddot{X} = 0$, as expected (since there are no external forces), and the motion of the coordinates (q_1, q_2) is governed by the Lagrangian L_1. From the expression of the latter Lagrangian we readily obtain the following formulas for the matrices A and B:

$$\boxed{A = \frac{m_1}{M}\begin{pmatrix} m_1+m_2 & m_1 \\ m_1 & m_1+m_2 \end{pmatrix}, \qquad B = k\mathbb{1}.}$$

As in the previous example, we can write the characteristic equation as

$$\det\left(A - \frac{B}{\omega^2}\right) = \frac{m_1^2}{M^2}\begin{vmatrix} m_1 + m_2 - \frac{kM}{m_1\omega^2} & m_1 \\ m_1 & m_1 + m_2 - \frac{kM}{m_1\omega^2} \end{vmatrix} = 0$$

$$\iff m_1 + m_2 - \frac{kM}{m_1\omega^2} = \pm m_1\,,$$

whence we easily obtain

$$\boxed{\omega_\pm = \sqrt{\frac{kM}{m_1(m_1+m_2 \mp m_1)}} = \begin{cases} \sqrt{\frac{kM}{m_1 m_2}} \\ \sqrt{\frac{k}{m_1}}\,. \end{cases}}$$

The normal mode vectors $\mathbf{c}_\pm$ are determined by the eigenvalue equation

$$\left(m_1 + m_2 - \frac{kM}{m_1\omega_\pm^2}\right)c_{\pm,1} + m_1 c_{\pm,2} = m_1(\pm c_1 + c_2) = 0\,,$$

so that

$$\mathbf{c}_{\pm} = c(1, \mp 1)$$

with $c \neq 0$ constant. Hence the two normal mode solutions are given by

$$\boxed{q_1^{(\pm)} = c\cos(\omega_{\pm}t + \alpha_{\pm})\,, \qquad q_2^{(\pm)} = \mp c\cos(\omega_{\pm}t + \alpha_{\pm})\,,}$$

and the normal mode coordinates are simply

$$\boxed{Q_{\pm} = q_1 \mp q_2}$$

(why?). The motion of the atoms' physical coordinates x_i in each of these normal modes can be easily obtained from Eqs. (3.88). Note that in the normal mode with the smaller frequency ω_- we have $q_1 = q_2$, or equivalently $x_2 - x_1 = x_3 - x_2$. Hence in this mode the distances between the atom of mass m_2 and each of the atoms of mass m_1 increase or decrease *in step*, oscillating with the same frequency ω_-. Moreover, from Eqs. (3.88) it follows that $x_2 = X$, i.e., the atom of mass m_2 is fixed at the molecule's center of mass (in particular, it is stationary in the CM frame). On the other hand, in the normal mode with the larger frequency ω_+ we have $q_1 = -q_2$, so that $x_2 - x_1 - a$ and $x_3 - x_2 - a$ have opposite signs and oscillate completely *out of phase* with the same frequency ω_+. Thus when the right half of the molecule stretches the left one contracts, and vice versa. Moreover, in this case the x_2 coordinate is given by

$$x_2 = X + \frac{2m_1}{M}\,q_1 = X + \frac{2cm_1}{M}\cos(\omega_{\pm}t + \alpha_{\pm})\,,$$

so that the position of the atom of mass m_2 oscillates with frequency ω_+ in the CM frame.

3.6 Introduction to Hamiltonian mechanics

3.6.1 Hamilton's canonical equations

Lagrange's equations of motion of a mechanical system:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\mathbf{q}}} - \frac{\partial L}{\partial \mathbf{q}} = 0\,, \tag{3.89}$$

although more versatile than Newton's, suffer from two main drawbacks. First of all, Lagrange's equations *are not in normal form*, i.e., the second derivatives $\ddot{q}_i$ are not expressed in terms of $(t, \mathbf{q}, \dot{\mathbf{q}})$. Secondly, they are *second-order* equations, so that the graph of two solutions $\mathbf{q}_1(t)$ and $\mathbf{q}_2(t)$ —i.e., two system trajectories— can intersect in the extended configuration space $\mathbb{R} \times \mathbb{R}^n$ of the variables $(t, \mathbf{q})$ without violating the existence and uniqueness theorem for systems of ordinary differential equations. Both problems can be solved if we are able to express Eqs. (3.89) as a *normal* system of *first-order* differential equations. Since Lagrange's equations are first-order in the canonical momenta p_i, the most natural way to achieve this aim is to use as dependent variables $\mathbf{q} = (q_1, \dots, q_n)$ and $\mathbf{p} := (p_1, \dots, p_n)$, in terms of which Eqs. (3.89) can be rewritten as

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \dot{\mathbf{q}}\,, \qquad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \frac{\partial L}{\partial \mathbf{q}}(t, \mathbf{q}, \dot{\mathbf{q}})\,. \tag{3.90}$$

The problem is that $\dot{\mathbf{q}}$, which appears in the RHS of these equations, *must be expressed as a function of* $(t, \mathbf{q}, \mathbf{p})$ *using the relation*

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}(t, \mathbf{q}, \dot{\mathbf{q}})\,. \tag{3.91}$$

Note that, by the inverse function theorem, for this to be possible (at least locally) we must have

$$\det\left(\frac{\partial p_i}{\partial \dot{q}_j}\right)_{1\leqslant i,j\leqslant n} = \det\left(\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j}\right)_{1\leqslant i,j\leqslant n} \neq 0\,. \tag{3.92}$$

For instance, it can be shown that this condition automatically holds in a natural mechanical system. Indeed, in such a system we have

$$\frac{\partial^2 L}{\partial \dot{q}_i \partial \dot{q}_j} = a_{ij}(t, \mathbf{q})$$

(cf. Section 3.4). Since $T > 0$ for $\dot{\mathbf{q}} \neq 0$ the matrix $\big(a_{ij}(t, \mathbf{q})\big)_{1 \leqslant i,j \leqslant n}$ is positive definite, and therefore invertible.

As we have just remarked, (3.90) should be more precisely written as

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \dot{\mathbf{q}}(t, \mathbf{q}, \mathbf{p}), \qquad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \frac{\partial L}{\partial \mathbf{q}}\big(t, \mathbf{q}, \dot{\mathbf{q}}(t, \mathbf{q}, \mathbf{p})\big),$$

where $\dot{\mathbf{q}}(t, \mathbf{q}, \mathbf{p})$ is the (vector-valued) function obtained by solving Eq. (3.91) for $\dot{\mathbf{q}}$ in terms of $(t, \mathbf{q}, \mathbf{p})$. In order to recast the latter system in a more symmetric form, it is essential to study how the Lagrangian L depends on the variables $(t, \mathbf{q}, \dot{\mathbf{q}})$. The differential of L, considered as a function of the latter variables, is given by

$$\mathrm{d}L = \frac{\partial L}{\partial t}\,\mathrm{d}t + \frac{\partial L}{\partial \mathbf{q}}\,\mathrm{d}\mathbf{q} + \frac{\partial L}{\partial \dot{\mathbf{q}}}\,\mathrm{d}\dot{\mathbf{q}} = \frac{\partial L}{\partial t}\,\mathrm{d}t + \frac{\partial L}{\partial \mathbf{q}}\,\mathrm{d}\mathbf{q} + \mathbf{p}\,\mathrm{d}\dot{\mathbf{q}}. \tag{3.93}$$

Taking into account that

$$\mathbf{p}\,\mathrm{d}\dot{\mathbf{q}} = \mathrm{d}(\mathbf{p}\,\dot{\mathbf{q}}) - \dot{\mathbf{q}}\,\mathrm{d}\mathbf{p},$$

from Eq. (3.93) we obtain

$$\mathrm{d}(\mathbf{p}\,\dot{\mathbf{q}} - L) = \mathrm{d}h = -\frac{\partial L}{\partial t}\,\mathrm{d}t - \frac{\partial L}{\partial \mathbf{q}}\,\mathrm{d}\mathbf{q} + \dot{\mathbf{q}}\,\mathrm{d}\mathbf{p}. \tag{3.94}$$

If in the previous formula we consider $\dot{\mathbf{q}}$ as a function of the variables $(t, \mathbf{q}, \mathbf{p})$ the energy integral h becomes a function

$$\boxed{H(t, \mathbf{q}, \mathbf{p}) := h\big(t, \mathbf{q}, \dot{\mathbf{q}}(t, \mathbf{q}, \mathbf{p})\big)} \tag{3.95}$$

of these variables called the system's **Hamiltonian**. Since $dH = dh$ is given by the RHS of Eq. (3.94), the partial derivatives of $H(t, \mathbf{q}, \mathbf{p})$ with respect to the independent variables $(t, \mathbf{q}, \mathbf{p})$ are simply the coefficients of $\mathrm{d}t$, $\mathrm{d}\mathbf{q}$ and $\mathrm{d}\mathbf{p}$ in the latter equation, i.e.,

$$\boxed{\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}, \qquad \frac{\partial H}{\partial \mathbf{q}} = -\frac{\partial L}{\partial \mathbf{q}}, \qquad \frac{\partial H}{\partial \mathbf{p}} = \dot{\mathbf{q}}}\,. \tag{3.96}$$

It is understood that in the RHS of these equations $\dot{\mathbf{q}}$ *must be expressed in terms of* $(t, \mathbf{q}, \mathbf{p})$ *inverting Eq.* (3.91). From Eqs. (3.96) it then follows that Lagrange's equations of motion (3.90) are equivalent to the following system of first-order ordinary differential equations in the *independent variables* $(\mathbf{q}, \mathbf{p})$:

$$\boxed{\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{p}}(t, \mathbf{q}, \mathbf{p}), \qquad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\frac{\partial H}{\partial \mathbf{q}}(t, \mathbf{q}, \mathbf{p}).} \tag{3.97}$$

Equations (3.97) are known as **Hamilton's canonical equations.**

Remark. In mathematics, the passage from the generalized coordinates and velocities $(\mathbf{q}, \dot{\mathbf{q}})$ to the canonical variables $(\mathbf{q}, \mathbf{p})$, where $\dot{\mathbf{q}}$ and $\mathbf{p}$ are related through (3.91), is called a **Legendre transformation.** This type of transformation is widely used, among other areas of physics, in thermodynamics. ■

- In order to write Hamilton's canonical equations of a mechanical system with Lagrangian $L(t, \mathbf{q}, \dot{\mathbf{q}})$ we can proceed as follows:

1. Find the canonical momenta

$$p_i = \frac{\partial L}{\partial \dot{q}_i}(t, \mathbf{q}, \dot{\mathbf{q}}), \qquad i = 1, \ldots, n.$$

2. Use the above equations to solve for the generalized velocities $\dot{q}_i$ in terms of the canonical momenta p_j:

$$\dot{q}_i = \dot{q}_i(t, \mathbf{q}, \mathbf{p}), \qquad i = 1, \ldots, n. \tag{3.98}$$

3. Compute the system's Hamiltonian

$$H(t, \mathbf{q}, \mathbf{p}) = \mathbf{p} \cdot \dot{\mathbf{q}} - L$$

using Eqs. (3.98) to express $\dot{\mathbf{q}}$ as a function of the variables $(t, \mathbf{q}, \mathbf{p})$.

4. Hamilton's canonical equations (3.97) can then be written down by computing the partial derivatives of H with respect to the canonical variables $\mathbf{q}$ and $\mathbf{p}$. In fact, the first n Hamilton's equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad i = 1, \ldots, n,$$

are actually equations (3.98), so that in practice it is only necessary to find the n remaining equations

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \qquad i = 1, \ldots, n.$$

- Recall that in a *natural* mechanical system

$$\boxed{h = \dot{\mathbf{q}} \frac{\partial L}{\partial \dot{\mathbf{q}}} - L = T + V,}$$

and therefore:

The Hamiltonian of a *natural* mechanical system is the energy $T + V$ *expressed in terms of the variables* $(t, \mathbf{q}, \mathbf{p})$.

3.6.2 Basic conservation laws

First of all, from Hamilton's equations it follows that if the Hamiltonian H is independent of a coordinate q_i the corresponding momentum p_i is conserved:

$$\boxed{\frac{\partial H}{\partial q_i} = 0 \quad \Longrightarrow \quad p_i = \text{const.}}$$

Likewise, if H es independent of the momentum p_i its corresponding coordinate q_i is conserved:

$$\boxed{\frac{\partial H}{\partial p_i} = 0 \quad \Longrightarrow \quad q_i = \text{const.}}$$

This example illustrates the great *symmetry* between the generalized coordinates q_i and their associated momenta p_i, which is in fact one of the distinctive advantages of the Hamiltonian formulation of mechanics.

From Hamilton's equations we also deduce that

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t} + \frac{\partial H}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial H}{\partial \mathbf{p}}\dot{\mathbf{p}} = \frac{\partial H}{\partial t} + \frac{\partial H}{\partial \mathbf{q}}\frac{\partial H}{\partial \mathbf{p}} - \frac{\partial H}{\partial \mathbf{p}}\frac{\partial H}{\partial \mathbf{q}} = \frac{\partial H}{\partial t}.$$

Hence the Hamiltonian is conserved if it does not depend explicitly on t:

$$\boxed{\frac{\partial H}{\partial t} = 0 \quad \Longrightarrow \quad H = \text{const.}}$$

Note that from the first Eq. (3.96), namely

$$\boxed{\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}\,,}$$

it follows that H is conserved if and only if L is independent of t. Since $H(t, q, \mathbf{p}) = h(t, \mathbf{q}, \dot{\mathbf{q}})$, this is the conservation of the energy integral deduced in the Lagrangian formalism (cf. Section 3.4).

- Another advantage of the Hamiltonian formulation of mechanics over the Lagrangian one consists in the following fact: *if the coordinate q_i is* cyclic*, it is possible to* eliminate *from Hamilton's equations the degree of freedom corresponding to this coordinate and its associated momentum p_i, reducing these equations to a system of* $2(n-1)$ *canonical equations.*

  Indeed, suppose that

$$\frac{\partial H}{\partial q_i} = 0\,,$$

so that $p_i(t) = c$ for all t. It is then immediate to check that *the equations of motion of the remaining coordinates and momenta are Hamilton's canonical equations for the Hamiltonian*

$$H\big|_{p_i=c} = H(t, q_1, \ldots, q_{i-1}, q_{i+1}, \ldots, q_n, p_1, \ldots, p_{i-1}, c, p_{i+1}, \ldots, p_n)\,,$$

which depends only on the $2(n-1)$ canonical variables (q_j, p_j) with $j \neq i$. Indeed, if $j \neq i$ we have

$$\dot{q}_j = \frac{\partial H}{\partial p_j}\bigg|_{p_i=c} = \frac{\partial}{\partial p_j}\left(H\big|_{p_i=c}\right)\,, \qquad \dot{p}_j = -\frac{\partial H}{\partial q_j}\bigg|_{p_i=c} = -\frac{\partial}{\partial q_j}\left(H\big|_{p_i=c}\right)\,.$$

Once these equations are solved, the motion of the cyclic coordinate q_i is determined simply by integrating its corresponding Hamilton equation

$$\dot{q}_i = \frac{\partial H}{\partial p_i}\bigg|_{p_i=c}\,,$$

i.e.,

$$q_i(t) = \int \frac{\partial H}{\partial p_i}(t, q_1(t), \ldots, q_{i-1}(t), q_{i+1}(t), \ldots, q_n(t), p_1(t), \ldots, p_{i-1}(t), c, p_{i+1}(t), \ldots, p_n(t))\,\mathrm{d}t\,.$$

Example 3.14. *Hamiltonian of a particle in Cartesian coordinates.*

As we saw in Section 3.2.1, in this case the Lagrangian is given by

$$L = \frac{1}{2}m\dot{\mathbf{r}}^2 - V(t, \mathbf{r})\,,$$

and the canonical momentum coincides with the linear one:

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} \quad \Longleftrightarrow \quad \dot{\mathbf{r}} = \frac{\mathbf{p}}{m}\,.$$

Since the Lagrangian is natural, the Hamiltonian is the total energy $T + V$ expressed in terms of $(t, \mathbf{r}, \mathbf{p})$:

$$H(t, \mathbf{r}, \mathbf{p}) = \frac{1}{2}m\dot{\mathbf{r}}^2 + V(t, \mathbf{r}) = \boxed{\frac{\mathbf{p}^2}{2m} + V(t, \mathbf{r})}\,.$$

If H does not depend on the coordinate x_i (i.e, if V is independent of x_i) the corresponding momentum $p_i = m\dot{x}_i$ is conserved, whereas if H is time-independent (equivalently, if V does not depend on t) H itself is conserved. These are nothing but the laws of conservation of the i-th component of the linear momentum and the total energy that we already knew.

Example 3.15. *Hamiltonian of a particle in spherical coordinates.*
As we saw in Example 3.7, the Lagrangian of a particle of mass m in spherical coordinates is given by Eq. (3.62), namely

$$L = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\varphi}^2\right) - V(t,r,\theta,\varphi)\,.$$

This Lagrangian is clearly *natural*, so that its corresponding Hamiltonian is simply the total energy

$$T + V = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\varphi}^2\right) + V(t,r,\theta,\varphi)\,,$$

expressed in terms of the canonical momenta (3.63):

$$p_r = m\dot{r}\,,\qquad p_\theta = mr^2\dot{\theta}\,,\qquad p_\varphi = mr^2\sin^2\theta\,\dot{\varphi}\,.$$

From these equations we obtain

$$\boxed{\dot{r} = \frac{p_r}{m}\,,\qquad \dot{\theta} = \frac{p_\theta}{mr^2}\,,\qquad \dot{\varphi} = \frac{p_\varphi}{mr^2\sin^2\theta}\,,} \tag{3.99}$$

which yields the following expression for the Hamiltonian:

$$\boxed{H(t,r,\theta,\varphi,p_r,p_\theta,p_\varphi) = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2\sin^2\theta}\right) + V(t,r,\theta,\varphi)\,.} \tag{3.100}$$

Hamilton's canonical equations are in this case the three equations (3.99), along with the remaining three equations for the derivatives of the momenta:

$$\boxed{\begin{aligned}
\dot{p}_r &= -\frac{\partial H}{\partial r} = -\frac{\partial V}{\partial r} + \frac{1}{mr^3}\left(p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}\right),\\
\dot{p}_\theta &= -\frac{\partial H}{\partial \theta} = -\frac{\partial V}{\partial \theta} + \frac{p_\varphi^2}{mr^2}\frac{\cos\theta}{\sin^3\theta}\,,\\
\dot{p}_\varphi &= -\frac{\partial H}{\partial \varphi} = -\frac{\partial V}{\partial \varphi}\,.
\end{aligned}}$$

As we already knew, from the last of these equations it follows that p_φ (which is equal to the z component of the angular momentum) is conserved if the potential does not depend on φ. Similarly, since

$$\frac{\partial H}{\partial t} = \frac{\partial V}{\partial t}\,,$$

if V is independent of t the Hamiltonian H, which coincides with the system's total energy, is conserved.

Example 3.16. *Hamiltonian of a charged particle in an electromagnetic field.*
As we saw in Example 3.10, using Cartesian coordinates $\mathbf{r} = (x_1, x_2, x_3)$ the Lagrangian can be taken as

$$L = \frac{1}{2}m\dot{\mathbf{r}}^2 - e\Phi(t,\mathbf{r}) + e\dot{\mathbf{r}}\cdot\mathbf{A}(t,\mathbf{r})\,.$$

Hence the canonical momenta are given by

$$p_i = m\dot{x}_i + eA_i(t,\mathbf{r}) \,, \qquad i = 1,2,3 \,, \tag{3.101}$$

and the Hamiltonian is given by

$$H = \frac{1}{2}m\dot{\mathbf{r}}^2 + e\Phi(t,\mathbf{r}) \,,$$

where it is understood that the velocities must be expressed in terms of the canonical momenta. Since

$$\boxed{\dot{x}_i = \frac{1}{m}(p_i - eA_i(t,\mathbf{r})) \,, \qquad i = 1,2,3 \,,} \tag{3.102}$$

substituting in the formula for H we obtain the expression

$$\boxed{H(t,\mathbf{r},\mathbf{p}) = \frac{1}{2m}\left(\mathbf{p} - e\mathbf{A}(t,\mathbf{r})\right)^2 + e\Phi(t,\mathbf{r}) \,.}$$

Note that in this formula $\mathbf{p}$ does *not* denote the *linear* momentum of the particle, but rather the vector whose three components are the *canonical* momenta p_i given by Eq. (3.101). Hamilton's equations are the three equations (3.102), along with

$$\boxed{\dot{p}_i = -\frac{\partial H}{\partial x_i} = -e\frac{\partial \Phi}{\partial x_i}(t,\mathbf{r}) + \frac{e}{m}\left(\mathbf{p} - e\mathbf{A}(t,\mathbf{r})\right) \cdot \frac{\partial \mathbf{A}}{\partial x_i}(t,\mathbf{r}) \,, \qquad i = 1,2,3 \,.}$$

The previous Hamiltonian can also be easily calculated in *spherical coordinates*. In fact, we know that the Lagrangian is *covariant* under coordinate changes, so that in order to obtain the Lagrangian of a charged particle in spherical coordinates it suffices to express (3.37) in these coordinates. Since

$$\dot{\mathbf{r}} \cdot \mathbf{A} = (\dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta + r\sin\theta\dot{\varphi}\mathbf{e}_\varphi) \cdot (A_r\mathbf{e}_r + A_\theta\mathbf{e}_\theta + A_\varphi\mathbf{e}_\varphi) = \dot{r}A_r + r\dot{\theta}A_\theta + r\sin\theta\dot{\varphi}A_\varphi \,,$$

substituting into Eq. (3.37) we obtain

$$\boxed{L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2) - e\Phi + e(\dot{r}A_r + r\dot{\theta}A_\theta + r\sin\theta\dot{\varphi}A_\varphi) \,.}$$

The canonical momenta are now

$$p_r = m\dot{r} + eA_r \,, \qquad p_\theta = mr^2\dot{\theta} + erA_\theta \,, \qquad p_\varphi = mr^2\sin^2\theta\dot{\varphi} + er\sin\theta A_\varphi \,,$$

so that

$$\dot{r} = \frac{1}{m}(p_r - eA_r) \,, \qquad r\dot{\theta} = \frac{1}{mr}(p_\theta - erA_\theta) \,, \qquad r\sin\theta\dot{\varphi} = \frac{1}{mr\sin\theta}(p_\varphi - er\sin\theta A_\varphi) \,.$$

Substituting in the definition of H we finally obtain

$$\begin{aligned} H &= \dot{r}p_r + \dot{\theta}p_\theta + \dot{\varphi}p_\varphi - \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2) + e\Phi - e(\dot{r}A_r + r\dot{\theta}A_\theta + r\sin\theta\dot{\varphi}A_\varphi) \\ &= \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\varphi}^2) + e\Phi \\ &= \boxed{\frac{1}{2m}\left[(p_r - eA_r)^2 + \frac{(p_\theta - erA_\theta)^2}{r^2} + \frac{(p_\varphi - er\sin\theta A_\varphi)^2}{r^2\sin^2\theta}\right] + e\Phi \,,} \end{aligned}$$

which can also be expressed as

$$H = \frac{1}{2m}\left[(p_r - eA_r)^2 + \left(\frac{p_\theta}{r} - eA_\theta\right)^2 + \left(\frac{p_\varphi}{r\sin\theta} - eA_\varphi\right)^2\right] + e\Phi\,. \tag{3.103}$$

Exercise. Write down Hamilton's canonical equations for the Hamiltonian (3.103).

3.6.3 Poisson brackets

As we have just seen, in the Hamiltonian formalism the equations of motion are *first-order* in the variables $\mathbf{q}$ (the generalized coordinates of the Lagrangian formalism) and $\mathbf{p}$ (their associated canonical momenta). Thus the motion of the system can be represented by the trajectory of a single point in the space $\mathbb{R}^{2n}$ where the canonical variables $(\mathbf{q},\mathbf{p})$ take values, usually referred to as the system's **phase space**. Note that, by the existence and uniqueness theorem for systems of first-order ordinary differential equations, there is a *unique* trajectory $(\mathbf{q}(t),\mathbf{p}(t))$ passing through any point $(\mathbf{q}_0,\mathbf{p}_0)$ in phase space at a certain initial time t_0, i.e., verifying the initial conditions $\mathbf{q}(t_0) = \mathbf{q}_0$, $\mathbf{p}(t_0) = \mathbf{p}_0$. (We are assuming, as we shall implicitly do in what follows, that the Hamiltonian $H(t,\mathbf{q},\mathbf{p})$ is of class C^2 in the variables $(\mathbf{q},\mathbf{p})$ for all t.) In other words, *the system's trajectories in phase space do not intersect.*

The rate of change of a smooth function $f(t,\mathbf{q},\mathbf{p})$ (usually called a **dynamical variable**) as the canonical variables $\mathbf{q}$ and $\mathbf{p}$ evolve with time through Hamilton's canonical equations (3.97) for a given Hamiltonian $H(t,\mathbf{q},\mathbf{p})$ is given by

$$\dot f = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{q}}\dot{\mathbf{q}} + \frac{\partial f}{\partial \mathbf{p}}\dot{\mathbf{p}} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{q}}\frac{\partial H}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}}\frac{\partial H}{\partial \mathbf{q}}\,.$$

This suggests defining the **Poisson bracket** of two functions $f(t,\mathbf{q},\mathbf{p})$ and $g(t,\mathbf{q},\mathbf{p})$ as the expression

$$\{f,g\} := \frac{\partial f}{\partial \mathbf{q}}\frac{\partial g}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}}\frac{\partial g}{\partial \mathbf{q}} \equiv \sum_{i=1}^{n}\left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right). \tag{3.104}$$

Using this definition, the previous formula for $\dot f$ can be concisely written as

$$\dot f = \frac{\partial f}{\partial t} + \{f,H\}\,; \tag{3.105}$$

in particular, if f does not depend explicitly on time t we obtain the simpler expression

$$\dot f = \{f,H\}\,.$$

Applying the previous formula to the coordinates $(\mathbf{q},\mathbf{p})$ in phase space we obtain the following formulation of Hamilton's canonical equations in terms of the Poisson bracket:

$$\dot q_i = \{q_i,H\}\,, \qquad \dot p_i = \{p_i,H\}\,, \qquad i = 1,\dots,n\,.$$

The Poisson brackets of the canonical coordinates and momenta among themselves are particularly simple:

$$\{q_i,q_j\} = \{p_i,p_j\} = 0\,, \qquad \{q_i,p_j\} = \delta_{ij}\,, \qquad i,j = 1,\dots,n\,, \tag{3.106}$$

where δ_{ij} is Kronecker's delta.

The following properties of the Poisson bracket follow immediately from its definition:

1. *Antisymmetry*: $\{f,g\} = -\{g,f\}$. In particular, $\{f,f\} = 0$.
2. *Bilinearity*: $\{\lambda f + \mu g, h\} = \lambda\{f,h\} + \mu\{g,h\}$, where λ, μ are constant (or, more generally, functions only of t). (By antisymmetry, the analogous property holds for the Poisson bracket $\{f, \lambda g + \mu h\}$.)
3. *Leibniz's rule*: $\{fg,h\} = f\{g,h\} + \{f,h\}g$ (and similarly for $\{f,gh\}$).

On the other hand, a long but straightforward calculation shows that the Poisson bracket satisfies the so called **Jacobi identity**

$$\{\{f,g\},h\} + \{\{g,h\},f\} + \{\{h,f\},g\} = 0\,. \tag{3.107}$$

From the elementary properties of partial derivatives it easily follows that

$$\frac{\partial}{\partial t}\{f,g\} = \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\}.$$

Using this relation and the Jacobi identity we can derive an important generalization of the latter result, known as the *Jacobi–Poisson identity*:

$$\frac{\mathrm{d}}{\mathrm{d}t}\{f,g\} = \{\dot f, g\} + \{f, \dot g\}\,. \tag{3.108}$$

Indeed,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\{f,g\} &= \frac{\partial}{\partial t}\{f,g\} + \{\{f,g\},H\} = \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\} - \{\{g,H\},f\} - \{\{H,f\},g\} \\
&= \left\{\frac{\partial f}{\partial t}, g\right\} + \{\{f,H\},g\} + \left\{f, \frac{\partial g}{\partial t}\right\} + \{f,\{g,H\}\} = \left\{\frac{\partial f}{\partial t} + \{f,H\}, g\right\} + \left\{f, \frac{\partial g}{\partial t} + \{g,H\}\right\} \\
&= \{\dot f, g\} + \{f, \dot g\}\,,
\end{aligned}$$

where we have used the Jacobi identity in the second equality. An important corollary of the Jacobi–Poisson identity is the so called **Jacobi–Poisson theorem**, of fundamental importance for obtaining first integrals of Hamiltonian systems:

If $f(t,\mathbf{q},\mathbf{p})$ and $g(t,\mathbf{q},\mathbf{p})$ are two first integrals of Hamilton's canonical equations (3.97), so is their Poisson bracket $\{f,g\}$.

Example 3.17. As we saw in Example 3.14, the Hamiltonian of a particle of mass m in Cartesian coordinates is given by

$$H = \frac{\mathbf{p}^2}{2m} + V(t,\mathbf{r})\,,$$

where $\mathbf{r} = (x_1, x_2, x_3)$ plays the role of $\mathbf{q}$ and $\mathbf{p} = m\dot{\mathbf{r}}$ is the linear momentum. We can compute the Poisson bracket of any two components of the angular momentum

$$\mathbf{J} = \mathbf{r} \times \mathbf{p} = (x_2p_3 - x_3p_2, x_3p_1 - x_1p_3, x_1p_2 - x_2p_1)$$

by applying the properties of the Poisson bracket reviewed above and the fundamental brackets (3.106). For instance,

$$\begin{aligned}
\{J_1, J_2\} &= \{x_2p_3 - x_3p_2, x_3p_1 - x_1p_3\} = \{x_2p_3, x_3p_1\} - \{x_2p_3, x_1p_3\} - \{x_3p_2, x_3p_1\} + \{x_3p_2, x_1p_3\} \\
&= -x_2p_1 + p_2x_1 = J_3\,.
\end{aligned}$$

Proceeding in this way we obtain the important relations

$$\boxed{\{J_i, J_j\} = J_k\,, \qquad (i,j,k) = \textit{cyclic}\text{ permutation of } (1,2,3)\,.}$$

Suppose now that any two components of the angular momentum are conserved, for instance J_1 and J_2. By the Jacobi–Poisson theorem the remaining component $J_3 = \{J_1, J_2\}$ will also be conserved. In other words, *if any two components of the angular moment* **J** *are conserved then* **J** *is conserved.* Likewise, suppose that the projection of the linear momentum **p** (which in this case coincides with the canonical one) along a certain direction **n** and the angular momentum are conserved. Choosing the coordinates appropriately, we can assume that p_1 and **J** are conserved. The relation

$$\{p_1, J_2\} = \{p_1, x_3p_1 - x_1p_3\} = -\{p_1, x_1p_3\} = p_3$$

then implies, by the Jacobi–Poisson theorem, that p_3 is also conserved. Similarly, from the Poisson bracket

$$\{p_1, J_3\} = \{p_1, x_1p_2 - x_2p_1\} = \{p_1, x_1p_2\} = -p_2$$

we deduce that p_2 is conserved. Hence in this case the linear momentum **p** is conserved.

We have already remarked in the previous subsection that in the Hamiltonian formulation the canonical variables **q** and **p** have identical status. It is therefore reasonable to try to simplify Hamilton's equations (3.97) using general changes of variables of the form

$$\tilde{\mathbf{q}} = \tilde{\mathbf{q}}(t, \mathbf{q}, \mathbf{p})\,, \qquad \tilde{\mathbf{p}} = \tilde{\mathbf{p}}(t, \mathbf{q}, \mathbf{p}) \tag{3.109}$$

involving both coordinates and momenta. The problem is that, in general, such a transformation maps the system (3.97) into a first-order system that need *not* be in general of Hamiltonian type, i.e., of the form

$$\dot{\tilde{\mathbf{q}}} = \frac{\partial \tilde{H}}{\partial \tilde{\mathbf{p}}}\,, \qquad \dot{\tilde{\mathbf{p}}} = -\frac{\partial \tilde{H}}{\partial \tilde{\mathbf{q}}} \tag{3.110}$$

for a certain function $\tilde{H}(t, \tilde{\mathbf{q}}, \tilde{\mathbf{p}})$. The transformation (3.109) is said to be **canonical** provided that it maps Hamilton's equations of *any* Hamiltonian $H(t, \mathbf{q}, \mathbf{p})$ into the canonical equations of another Hamiltonian $\tilde{H}(t, \tilde{\mathbf{q}}, \tilde{\mathbf{p}})$.

Example 3.18. The transformation

$$\tilde{\mathbf{q}} = \mathbf{p}\,, \qquad \tilde{\mathbf{p}} = \mathbf{q}$$

is canonical, since it transforms Hamilton's equations of any Hamiltonian $H(t, \mathbf{q}, \mathbf{p})$ into those of the Hamiltonian $\tilde{H}(t, \tilde{\mathbf{q}}, \tilde{\mathbf{p}}) = -H(t, \mathbf{q}, \mathbf{p})$. Indeed,

$$\dot{\tilde{\mathbf{q}}} = \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} = \frac{\partial \tilde{H}}{\partial \tilde{\mathbf{p}}}\,, \qquad \dot{\tilde{\mathbf{p}}} = \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} = -\frac{\partial \tilde{H}}{\partial \tilde{\mathbf{q}}}\,.$$

The transformation

$$\tilde{\mathbf{q}} = \mathbf{p}\,, \qquad \tilde{\mathbf{p}} = -\mathbf{q}\,,$$

is also canonical, with $\tilde{H}(t, \tilde{\mathbf{q}}, \tilde{\mathbf{p}}) = H(t, \mathbf{q}, \mathbf{p})$.

Exercise. Show that the transformation

$$\tilde{q} = p^2\,, \qquad \tilde{p} = q$$

is *not* canonical.

• An important result in Hamiltonian mechanics states that *the transformation* (3.109) *is canonical if and only if the Poisson brackets of the transformed canonical variables* $\tilde{\mathbf{q}}$ *and* $\tilde{\mathbf{p}}$ *verify*

$$\{\tilde{q}_i, \tilde{q}_j\} = \{\tilde{p}_i, \tilde{p}_j\} = 0\,, \qquad \{\tilde{q}_i, \tilde{p}_j\} = \lambda\delta_{ij}\,, \qquad i, j = 1, \ldots, n\,,$$

with $\lambda \neq 0$ *constant*[10]. In particular, if $\lambda = 1$ the variables $\tilde{\mathbf{q}}$ and $\tilde{\mathbf{p}}$ are said to be **canonically conjugate**.

• It can also be shown that it is always possible to find a canonical transformation (3.109) mapping Hamilton's equations (3.97) of *any* Hamiltonian $H(t, \mathbf{q}, \mathbf{p})$ into the canonical equations of the Hamiltonian $\tilde{H} = 0$, that is to say, into the trivial system

$$\dot{\tilde{\mathbf{q}}} = 0\,, \qquad \dot{\tilde{\mathbf{p}}} = 0\,.$$

The general solution of the latter system is obviously

$$\tilde{\mathbf{q}} = \tilde{\mathbf{q}}_0\,, \qquad \tilde{\mathbf{p}} = \tilde{\mathbf{p}}_0\,,$$

with $\tilde{\mathbf{q}}_0, \tilde{\mathbf{p}}_0$ arbitrary constant vectors. The general solution of the canonical equations of the original Hamiltonian $H(t, \mathbf{q}, \mathbf{p})$ is then obtained inverting the relations

$$\tilde{\mathbf{q}}(t, \mathbf{q}, \mathbf{p}) = \tilde{\mathbf{q}}_0\,, \qquad \tilde{\mathbf{p}}(t, \mathbf{q}, \mathbf{p}) = \tilde{\mathbf{p}}_0$$

to express $\mathbf{q}$ and $\mathbf{p}$ in terms of t and the $2n$ constants $(\tilde{\mathbf{q}}_0, \tilde{\mathbf{p}}_0)$. In fact, this is one of the most effective methods for solving Hamilton's equations, using the so-called *Hamilton–Jacobi equation* for finding a canonical transformation mapping the original Hamiltonian H into $\tilde{H} = 0$.

• The fundamental Poisson brackets (3.106) make it possible to establish a formal analogy between classical and quantum mechanics. Indeed, in quantum mechanics the dynamical variables dynamic (q_j, p_j) are replaced (in the so called Schrödinger picture) by the *self-adjoint operators*

$$\boxed{Q_j = q_j\,, \qquad P_j = -\mathrm{i}\hbar\frac{\partial}{\partial q_j}}\,,$$

whose action on a complex valued *wave function* (probability amplitude) $\psi(\mathbf{q})$ is given by

$$(Q_j\psi)(\mathbf{q}) = q_j\psi(\mathbf{q})\,, \qquad (P_j\psi)(\mathbf{q}) = -\mathrm{i}\hbar\,\frac{\partial\psi}{\partial q_j}(\mathbf{q})\,.$$

The fundamental operators (Q_i, P_j) satisfy *commutation relations* totally analogous to Eq. (3.106):

$$\boxed{[Q_i, Q_j] = [P_i, P_j] = 0\,, \qquad [Q_i, P_j] = \mathrm{i}\hbar\left[\frac{\partial}{\partial q_j}, q_i\right] = \mathrm{i}\hbar\,\delta_{ij}\,,} \tag{3.111}$$

where the *commutator* of two operators A, B is defined by

$$[A, B] = AB - BA\,.$$

Any other function $f(\mathbf{q}, \mathbf{p})$ is represented in quantum mechanics by a self-adjoint operator $F(\mathbf{Q}, \mathbf{P})$ such that

$$F(\mathbf{q}, \mathbf{p}) = f(\mathbf{q}, \mathbf{p})\,.$$

[10]Usually only canonical transformations with $\lambda = 1$ (called *proper*) are considered. This does not entail any real restriction, since if (3.109) is a canonical transformation with $\lambda \neq 1$ the transformation $(\mathbf{q}, \mathbf{p}) \mapsto (\tilde{\mathbf{q}}, \tilde{\mathbf{p}}/\lambda)$ is another canonical transformation with $\lambda = 1$

This fact is known as *Bohr's correspondence principle*. It is important to realize in this respect that, since the product of operators is not commutative in general, the operator F determines the classical function f but not vice versa. For instance,

$$F_1(Q,P) = PQ^2P \neq F_2(Q,P) = \frac{1}{2}(Q^2P^2 + P^2Q^2)$$

(in fact, $F_1 - F_2 = \hbar^2$), but nevertheless $f_1(q,p) = f_2(q,p) = q^2p^2$. This fact is not surprising, since classical mechanics is the limit as $\hbar \to 0$ of quantum mechanics, so the former theory must be determined by the latter. The converse, however, is not necessarily true, as there may exist different theories with the same limit as $\hbar \to 0$.

The commutator $[A,B]$ of two operators A and B has algebraic properties formally analogous to those of the Poisson bracket. Indeed, it is obviously antisymmetric and linear in each of its arguments. In addition, if A, B, and C are three operators then it is immediate to verify that

$$[AB,C] = A[B,C] + [A,C]B\,. \tag{3.112}$$

This identity is similar to Leibniz's rule satisfied by the Poisson bracket, with the only difference that the *order* in which operators appear in Eq. (3.112) is essential for its validity. Finally, it is straightforward to show that the commutator also verifies the *Jacobi identity*

$$[[A,B],C] + [[B,C],A] + [[C,A],B] = 0\,,$$

where the order is again essential. If $F(\mathbf{Q},\mathbf{P})$ and $G(\mathbf{Q},\mathbf{P})$ are two self-adjoint operators depending *polynomially* on $(\mathbf{Q},\mathbf{P})$ (and not explicitly dependent on $\hbar$), by repeatedly applying Eq. (3.112) we can always express the commutator $[F,G]$ in terms of the canonical commutators in Eq. (3.111). From Leibniz's rule satisfied by the Poisson bracket it then follows that the Poisson bracket $\{f,g\}$ of the corresponding classical functions $f(\mathbf{q},\mathbf{p}) = F(\mathbf{q},\mathbf{p})$, $g(\mathbf{q},\mathbf{p}) = G(\mathbf{q},\mathbf{p})$ will satisfy the *same* expression replacing Q_i by q_i, P_i by p_i and the canonical commutators (3.111) by the canonical Poisson brackets (3.106) (see next example). It follows that if

$$[F,G] = \mathrm{i}\hbar K\,,$$

where $K(\mathbf{Q},\mathbf{P})$ is a polynomial independent of $\hbar$, the classical Poisson bracket $\{f,g\}$ will be given by

$$\{f,g\} = k$$

with $k(\mathbf{q},\mathbf{p}) = K(\mathbf{q},\mathbf{p})$. In other words, *the commutator in quantum mechanics determines the Poisson bracket in classical mechanics through the relation*

$$\boxed{\frac{1}{\mathrm{i}\hbar}[F,G] \to \{f,g\}\,.}$$

The opposite route (from the Poisson bracket in classical mechanics to the commutator in quantum mechanics) *is not well defined in general*, since as we have remarked different self-adjoint operators $F(\mathbf{Q},\mathbf{P})$ can yield the same function $f(\mathbf{q},\mathbf{p})$.

Example 3.19. Consider, for example, the commutator $[Q^2,P^2]$. Using repeatedly Eq. (3.112) we obtain

$$\begin{aligned}[Q^2,P^2] &= [Q\cdot Q,P^2] = Q[Q,P^2] + [Q,P^2]Q = QP[Q,P] + Q[Q,P]P + P[Q,P]Q + [Q,P]PQ\\ &= 2\mathrm{i}\hbar(QP+PQ)\,.\end{aligned}$$

At the classical level, applying repeatedly Leibniz's rule to the Poisson bracket $\{q^2,p^2\}$ we arrive at

$$\{q^2, p^2\} = q\{q, p^2\} + \{q, p^2\}q = qp\{q, p\} + q\{q, p\}p + p\{q, p\}q + \{q, p\}pq = 2(qp + pq)$$
$$= 4qp\,,$$

which is indeed obtained from $[Q^2, P^2]/(\mathrm{i}\hbar)$ replacing q by Q and P by p.

4 Motion relative to a non-inertial frame

4.1 Angular velocity of a reference frame with respect to another

We shall study in this chapter the description of the motion of a particle in a non-inertial reference frame. Consider, to begin with, two reference frames S and S' with the same origin, and denote by $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}'_i, \mathbf{e}'_2, \mathbf{e}'_3\}$ the orthonormal positively oriented frames determining the axes of S and S'. We shall always assume in this chapter that the frame S' is *inertial*, and denote by $O(t)$ the linear application relating the vectors $\mathbf{e}'_i$ (**fixed axes**) with the vectors $\mathbf{e}_i$ (**moving axes**):

$$\boxed{\mathbf{e}_i(t) = O(t)\,\mathbf{e}'_i\,, \qquad i = 1, 2, 3\,.} \tag{4.1}$$

We shall often identify in what follows the operator $O(t)$ with its *matrix* in the basis $\{\mathbf{e}'_i\}_{i=1}^{3}$, whose *columns* are the coordinates of the vectors $\mathbf{e}_i(t)$ with respect to the latter basis. Since $O(t)$ transforms a positively oriented orthonormal frame into another such frame, this operator is an element of the **special orthogonal group** SO(3) of all linear operators $M : \mathbb{R}^3 \to \mathbb{R}^3$ (or, equivalently, 3×3 real matrices M) satisfying the the conditions

$$\boxed{M^{\mathsf{T}} M = M M^{\mathsf{T}} = \mathbb{1}\,, \qquad \det M = 1\,.}$$

A theorem first proved by Euler states that *every element M of* SO(3) *is a rotation around a certain axis* $\mathbf{n}$. The proof of this theorem is as follows. First of all, taking the determinant of both members of the equality

$$M^{\mathsf{T}}(M - \mathbb{1}) = \mathbb{1} - M^{\mathsf{T}}$$

and using the elementary identities

$$\det M = \det M^{\mathsf{T}} = 1\,, \qquad \det(\mathbb{1} - M^{\mathsf{T}}) = \det\big((\mathbb{1} - M)^{\mathsf{T}}\big) = \det(\mathbb{1} - M)$$

we obtain

$$\det(M - \mathbb{1}) = \det(\mathbb{1} - M) = -\det(M - \mathbb{1}) \quad \Longrightarrow \quad \det(M - \mathbb{1}) = 0\,.$$

Hence $\lambda = 1$ is an *eigenvalue* of M. In other words, there exists a nonzero vector $\mathbf{n} \in \mathbb{R}^3$ (which we can take w.l.o.g. of unit length) such that $M\mathbf{n} = \mathbf{n}$. Let us next show that M is a rotation around the axis $\mathbf{n}$. Indeed, taking $\mathbf{e}'_3 = \mathbf{n}$ the matrix M is of the form

$$M = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where $A \equiv (a_{ij})_{1 \leqslant i,j \leqslant 2}$ is an orthogonal 2×2 matrix with unit determinant (recall that the columns of an orthogonal matrix are unit vectors perpendicular with one another). Since $a_{11}^2 + a_{21}^2 = 1$, we can take

$$a_{11} = \cos\theta, \quad a_{21} = \sin\theta\,, \qquad \text{with } \theta \in [0, 2\pi)\,.$$

Similarly,

$$a_{12} = \cos\psi, \quad a_{22} = \sin\psi\,, \qquad \text{with } \psi \in [0, 2\pi)\,.$$

Imposing the orthogonality of the columns of A we obtain

$$\cos\theta\cos\psi + \sin\theta\sin\psi = \cos(\theta-\psi) = 0\,.$$

Hence $\psi = \theta \pm \frac{\pi}{2}$ (up to an integer multiple of 2π), and

$$M = \begin{pmatrix} \cos\theta & \mp\sin\theta & 0 \\ \sin\theta & \pm\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Actually, the solution $\psi = \theta - \frac{\pi}{2}$ is unacceptable, since it implies that $\det M = -1$. Thus $\psi = \theta + \pi/2$, and

$$M = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} =: R_3(\theta)$$

is indeed a (counterclockwise) rotation of angle θ around the axis $\mathbf{e}'_3 = \mathbf{n}$.

Exercise. Show that the rotation angle θ of a matrix $M \in SO(3)$ is determined by the equation $1 + 2\cos\theta = \operatorname{tr} M$, where $\operatorname{tr} M := \sum_{i=1}^{3} M_{ii}$ denotes the **trace** of the matrix M.

Solution. We have just seen that if $M \in SO(3)$ and $\mathbf{n}$ (with $|\mathbf{n}| = 1$) is an eigenvector of M of eigenvalue 1 then M is a rotation around the axis $\mathbf{n}$. To determine the angle of rotation θ, note that $M = UR_3(\theta)U^{-1}$, where U is the change of basis matrix from the original basis $\{\mathbf{e}'_i\}_{i=1}^{3}$ to the basis with $\mathbf{n} = \mathbf{e}'_3$. Taking the trace of this equality and remembering that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ we obtain

$$\operatorname{tr} M = \operatorname{tr}\left(UR_3(\theta)U^{-1}\right) = \operatorname{tr}\left(U^{-1}UR_3(\theta)\right) = \operatorname{tr} R_3(\theta) = 1 + 2\cos\theta\,.$$

Consider again the rotation matrix $R_3(\theta)$ around the axis $\mathbf{e}'_3$. A direct calculation shows that

$$\frac{\mathrm{d}R_3}{\mathrm{d}\theta}(0) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

hence, if $\mathbf{c} = \sum_{i=1}^{3} c'_i \mathbf{e}'_i \in \mathbb{R}^3$ is an arbitrary vector we have

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\bigg|_{\theta=0} R_3(\theta)\mathbf{c} = \frac{\mathrm{d}R_3}{\mathrm{d}\theta}(0)\,\mathbf{c} = -c_2\mathbf{e}'_1 + c_1\mathbf{e}'_2 = \mathbf{e}'_3 \times \mathbf{c}\,.$$

In general, if $R_{\mathbf{n}}(\theta)$ denotes the matrix implementing a rotation around the axis $\mathbf{n}$ by an angle θ we must accordingly have

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\bigg|_{\theta=0} R_{\mathbf{n}}(\theta)\mathbf{c} = \frac{\mathrm{d}R_{\mathbf{n}}}{\mathrm{d}\theta}(0)\,\mathbf{c} = \mathbf{n} \times \mathbf{c}\,.$$

We can symbolically write

$$\frac{\mathrm{d}R_{\mathbf{n}}}{\mathrm{d}\theta}(0) = \mathbf{n}\times\,, \tag{4.2}$$

with the understanding that both sides are equal when applied to an arbitrary vector $\mathbf{c} \in \mathbb{R}^3$. Another consequence of the previous result is that, since $R_{\mathbf{n}}(0) = \mathbb{1}$, for θ small we have

$$R_{\mathbf{n}}(\theta)\mathbf{c} = \mathbf{c} + \theta\,\mathbf{n} \times \mathbf{c} + O(\theta^2)\,. \tag{4.3}$$

For this reason the transformation

$$\mathbf{c} \mapsto \mathbf{c} + \theta\,\mathbf{n} \times \mathbf{c}$$

is called an *infinitesimal rotation* of angle θ around $\mathbf{n}$.

Exercise. Show that $R_{\mathbf{n}}(\theta)\mathbf{r} = \cos\theta\,\mathbf{r} + (1-\cos\theta)(\mathbf{n}\cdot\mathbf{r})\mathbf{n} + \sin\theta\,\mathbf{n}\times\mathbf{r}$.

Solution. If $\mathbf{r}$ is parallel to $\mathbf{n}$ the formula is clearly true. On the other hand, if $\mathbf{r}$ is not parallel to $\mathbf{n}$ the vectors $\mathbf{n}$, $\mathbf{n}\times\mathbf{r}$ and $\mathbf{r}-(\mathbf{n}\cdot\mathbf{r})\mathbf{n}$ are mutually orthogonal and nonzero. Moreover, the last two vectors have the same length $l > 0$, since

$$|\mathbf{r}-(\mathbf{n}\cdot\mathbf{r})\mathbf{n}|^2 = \mathbf{r}^2 - (\mathbf{n}\cdot\mathbf{r})^2 .$$

It follows that the vectors

$$\mathbf{e}_1 = \frac{1}{l}\big(\mathbf{r}-(\mathbf{n}\cdot\mathbf{r})\mathbf{n}\big), \quad \mathbf{e}_2 = \frac{1}{l}\,\mathbf{n}\times\mathbf{r}, \quad \mathbf{e}_3 = \mathbf{n}$$

make up a positively oriented orthonormal basis. Using this basis and the equation of the rotation $R_3(\theta)$ we obtain

$$\begin{aligned} R_{\mathbf{n}}(\theta)\mathbf{r} &= R_{\mathbf{n}}(\theta)\big(l\mathbf{e}_1 + (\mathbf{n}\cdot\mathbf{r})\mathbf{n}\big) = lR_{\mathbf{n}}(\theta)\mathbf{e}_1 + (\mathbf{n}\cdot\mathbf{r})\mathbf{n} = l\cos\theta\,\mathbf{e}_1 + l\sin\theta\,\mathbf{e}_2 + (\mathbf{n}\cdot\mathbf{r})\mathbf{n} \\ &= \cos\theta\big(\mathbf{r}-(\mathbf{n}\cdot\mathbf{r})\mathbf{n}\big) + \sin\theta\,\mathbf{n}\times\mathbf{r} + (\mathbf{n}\cdot\mathbf{r})\mathbf{n}, \end{aligned}$$

which yields the sought-for formula.

Let now $O(t) \in \mathrm{SO}(3)$ for all t, and suppose that O is of class C^1 (i.e., that the matrix elements of O are continuously differentiable functions of t). We shall next compute the derivative $\dot O(t)$ at an arbitrary time t. To this end, we differentiate with respect to t the identity

$$O(t)O(t)^{\mathsf{T}} = \mathbb{1},$$

obtaining

$$0 = \dot O(t)O(t)^{\mathsf{T}} + O(t)\dot O(t)^{\mathsf{T}} = \dot O(t)O(t)^{\mathsf{T}} + \big[\dot O(t)O(t)^{\mathsf{T}}\big]^{\mathsf{T}}.$$

Thus

$$\Omega(t) := \dot O(t)O(t)^{\mathsf{T}}$$

is an *antisymmetric* 3×3 matrix. Since $O(t)^{\mathsf{T}} = O(t)^{-1}$, from the previous relation we obtain

$$\boxed{\dot O(t) = \Omega(t)O(t)\,.} \tag{4.4}$$

The antisymmetric matrix $\Omega(t)$ can be written as

$$\Omega(t) = \begin{pmatrix} 0 & -\omega_3(t) & \omega_2(t) \\ \omega_3(t) & 0 & -\omega_1(t) \\ -\omega_2(t) & \omega_1(t) & 0 \end{pmatrix}$$

for appropriate real numbers $\omega_i(t)$. It is then straightforward to check that if $\mathbf{c} \in \mathbb{R}^3$ is any vector we have

$$\Omega(t)\mathbf{c} = \boldsymbol{\omega}(t)\times\mathbf{c},$$

where $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ is the vector with components $\omega_i(t)$. From Eq. (4.4) and the previous identity (with $O(t)\mathbf{c}$ instead of $\mathbf{c}$) we finally deduce that

$$\boxed{\dot O(t)\mathbf{c} = \boldsymbol{\omega}(t)\times O(t)\mathbf{c}, \qquad \forall t \in \mathbb{R}.} \tag{4.5}$$

The vector $\boldsymbol{\omega}(t) \in \mathbb{R}^3$, which is in general time-dependent, is determined by the relation

$$\boldsymbol{\omega}(t)\times = \Omega(t) = \dot O(t)O(t)^T = \dot O(t)O(t)^{-1},$$

which can also be written as

$$\boldsymbol{\omega}(t_0)\times = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=t_0} O(t)O(t_0)^{-1}. \tag{4.6}$$

Exercise. Show that the matrix elements Ω_{ij} of the antisymmetric matrix Ω and the components ω_k of the vector $\boldsymbol{\omega}$ are related by

$$\Omega_{ij} = -\sum_{k=1}^{3} \varepsilon_{ijk}\omega_k\,, \qquad \omega_k = -\frac{1}{2}\sum_{i,j=1}^{3} \varepsilon_{ijk}\Omega_{ij}\,,$$

where ε_{ijk} is Levi-Civita's completely antisymmetric tensor defined in Eq. (1.13).

Example 4.1. If $O(t)$ is a rotation around a fixed axis $\mathbf{n}$ by a time-dependent angle $\alpha(t)$ we have

$$\boldsymbol{\omega}(t) = \dot{\alpha}(t)\mathbf{n}\,.$$

Indeed, from Eq. (4.6) we obtain

$$O(t) = R_{\mathbf{n}}(\alpha(t)) \quad \Longrightarrow \quad O(t)O(t_0)^{-1} = R_{\mathbf{n}}(\alpha(t))R_{\mathbf{n}}(\alpha(t_0))^{-1} = R_{\mathbf{n}}(\alpha(t)-\alpha(t_0))\,,$$

and thus, by Eq. (4.2),

$$\boldsymbol{\omega}(t_0)\times = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} R_{\mathbf{n}}(\alpha(t)-\alpha(t_0)) = \dot{\alpha}(t_0)\left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=0} R_{\mathbf{n}}(\theta) = \dot{\alpha}(t_0)\mathbf{n}\times \quad \Longrightarrow \quad \boldsymbol{\omega}(t_0) = \dot{\alpha}(t_0)\mathbf{n}\,.$$

Applying Eq. (4.5) to Eq. (4.1), which relates the moving axes unit vectors $\mathbf{e}_i(t) = O(t)\mathbf{e}'_i$ with the fixed ones $\mathbf{e}'_i$, we obtain the important formula

$$\dot{\mathbf{e}}_i(t) = \boldsymbol{\omega}(t)\times O(t)\mathbf{e}'_i = \boldsymbol{\omega}(t)\times\mathbf{e}_i(t)\,, \tag{4.7}$$

where $\dot{\mathbf{e}}_i(t)$ denotes the time derivative of the vector $\mathbf{e}_i(t)$ with respect to the inertial (fixed) frame S'. We thus have

$$\mathbf{e}_i(t+\Delta t) = \mathbf{e}_i(t) + \boldsymbol{\omega}(t)\Delta t\times\mathbf{e}_i(t) + O(\Delta t^2)\,.$$

Comparing with Eq. (4.3) we deduce that to first order in Δt each vector $\mathbf{e}_i(t+\Delta t)$ is obtained from $\mathbf{e}_i(t)$ applying an *infinitesimal rotation of angle $\Delta\theta$ and axis* $\mathbf{n}(t)$ such that

$$\Delta\theta\,\mathbf{n}(t) = \Delta t\,\boldsymbol{\omega}(t)\,.$$

Letting Δt tend to zero in the previous equation we obtain the relation

$$\boldsymbol{\omega}(t) = \dot{\theta}(t)\,\mathbf{n}(t)\,.$$

In other words:

The direction and the magnitude of the vector $\boldsymbol{\omega}(t)$ are respectively equal to the *instantaneous axis of rotation* and the magnitude of the *instantaneous angular velocity* of the moving axes $\{\mathbf{e}_i\}_{i=1}^3$ with respect to the fixed ones $\{\mathbf{e}'_i\}_{i=1}^3$.

For this reason, the vector $\boldsymbol{\omega}(t)$ is called the **instantaneous angular velocity vector** (at time t) of the moving axes $\{\mathbf{e}_i\}_{i=1}^3$ with respect to the fixed ones $\{\mathbf{e}'_i\}_{i=1}^3$.

4.2 Time derivative in the fixed and moving frames

To study the relation between the inertial frame S' (**fixed frame**) and its non-inertial counterpart S (**moving frame**), we shall analyze in this section how the time derivative of a vector $\mathbf{A}(t)$ is expressed in each of these reference frames. To this end, let us start by expanding $\mathbf{A}(t)$ in the *moving* frame $\{\mathbf{e}_i\}_{i=1}^3$:

$$\mathbf{A}(t) = \sum_{i=1}^{3} A_i(t)\,\mathbf{e}_i\,.$$

From now on, to avoid confusion we shall respectively denote by $\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)_{\mathrm{f}}$ and $\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)_{\mathrm{m}}$ the time derivatives with respect to the fixed and moving reference frames. Differentiating the previous equation *in the fixed frame* and using the latter notation we obtain the identity

$$\left(\frac{\mathrm{d}\mathbf{A}(t)}{\mathrm{d}t}\right)_{\mathrm{f}} = \sum_{i=1}^{3} \dot{A}_i(t)\,\mathbf{e}_i + \sum_{i=1}^{3} A_i(t)\left(\frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}t}\right)_{\mathrm{f}}\,, \tag{4.8}$$

where we have taken into account that the functions $A_i(t)$ are *scalars*, and therefore their time derivative is the same in any frame. Using Eq. (4.7), which in the notation just introduced is written

$$\left(\frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}t}\right)_{\mathrm{f}} = \boldsymbol{\omega}(t)\times\mathbf{e}_i\,, \tag{4.9}$$

Eq. (4.8) becomes

$$\left(\frac{\mathrm{d}\mathbf{A}(t)}{\mathrm{d}t}\right)_{\mathrm{f}} = \sum_{i=1}^{3} \dot{A}_i(t)\,\mathbf{e}_i + \boldsymbol{\omega}(t)\times\mathbf{A}(t)\,. \tag{4.10}$$

On the other hand, in the moving frame the vectors $\mathbf{e}_i$ are *constant*, so that $\left(\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t}\right)_{\mathrm{m}}$ is simply given by

$$\left(\frac{\mathrm{d}\mathbf{A}(t)}{\mathrm{d}t}\right)_{\mathrm{m}} = \sum_{i=1}^{3} \dot{A}_i(t)\,\mathbf{e}_i\,.$$

Comparing the last two equations we obtain the important relation

$$\left(\frac{\mathrm{d}\mathbf{A}(t)}{\mathrm{d}t}\right)_{\mathrm{f}} = \left(\frac{\mathrm{d}\mathbf{A}(t)}{\mathrm{d}t}\right)_{\mathrm{m}} + \boldsymbol{\omega}(t)\times\mathbf{A}(t)\,. \tag{4.11}$$

The above expression is valid for *any* time-dependent vector $\mathbf{A}(t)$; in particular, if we apply it to the instantaneous angular velocity $\boldsymbol{\omega}(t)$ we obtain

$$\left(\frac{\mathrm{d}\boldsymbol{\omega}(t)}{\mathrm{d}t}\right)_{\mathrm{f}} = \left(\frac{\mathrm{d}\boldsymbol{\omega}(t)}{\mathrm{d}t}\right)_{\mathrm{m}} =: \dot{\boldsymbol{\omega}}(t)\,. \tag{4.12}$$

4.3 Dynamics in a non-inertial reference frame

Consider next the most general situation in which the origin of S is displaced from that of S' by a time-dependent vector $\mathbf{R}(t)$. If $\mathbf{r}$ is the position vector of a particle with respect to the non-inertial frame S, its position vector in the inertial frame S' will be given by

$$\mathbf{r}' = \mathbf{r} + \mathbf{R}\,.$$

Differentiating this equality with respect to the fixed (inertial) reference frame we obtain

$$\left(\frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}t}\right)_{\mathrm{f}} = \left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right)_{\mathrm{m}} + \boldsymbol{\omega}\times\mathbf{r} + \mathbf{V}\,, \tag{4.13}$$

where

$$\boxed{\mathbf{V} := \left(\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t}\right)_{\mathrm{f}}} \tag{4.14}$$

is the velocity of the origin of S measured in the inertial frame S'. Denoting by

$$\mathbf{v}_{\mathrm{f}} = \left(\frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}t}\right)_{\mathrm{f}}\,, \qquad \mathbf{v}_{\mathrm{m}} = \left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right)_{\mathrm{m}} \tag{4.15}$$

the velocity of the particle with respect to the fixed and moving frames, we can rewrite Eq. (4.13) in the more compact form

$$\boxed{\mathbf{v}_{\mathrm{f}} = \mathbf{v}_{\mathrm{m}} + \boldsymbol{\omega}\times\mathbf{r} + \mathbf{V}\,.} \tag{4.16}$$

Differentiating again this relation in the inertial frame we obtain

$$\begin{aligned}\left(\frac{\mathrm{d}\mathbf{v}_{\mathrm{f}}}{\mathrm{d}t}\right)_{\mathrm{f}} &= \left(\frac{\mathrm{d}\mathbf{v}_{\mathrm{m}}}{\mathrm{d}t}\right)_{\mathrm{m}} + \boldsymbol{\omega}\times\mathbf{v}_{\mathrm{m}} + \dot{\boldsymbol{\omega}}\times\mathbf{r} + \boldsymbol{\omega}\times(\mathbf{v}_{\mathrm{m}} + \boldsymbol{\omega}\times\mathbf{r}) + \left(\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}\right)_{\mathrm{f}} \\ &= \left(\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}\right)_{\mathrm{f}} + \left(\frac{\mathrm{d}\mathbf{v}_{\mathrm{m}}}{\mathrm{d}t}\right)_{\mathrm{m}} + 2\boldsymbol{\omega}\times\mathbf{v}_{\mathrm{m}} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) + \dot{\boldsymbol{\omega}}\times\mathbf{r}\,.\end{aligned}$$

Taking into account that

$$\left(\frac{\mathrm{d}\mathbf{v}_{\mathrm{f}}}{\mathrm{d}t}\right)_{\mathrm{f}} = \left(\frac{\mathrm{d}^2\mathbf{r}'}{\mathrm{d}t^2}\right)_{\mathrm{f}} = \mathbf{a}_{\mathrm{f}}\,, \qquad \left(\frac{\mathrm{d}\mathbf{v}_{\mathrm{m}}}{\mathrm{d}t}\right)_{\mathrm{m}} = \left(\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}\right)_{\mathrm{m}} = \mathbf{a}_{\mathrm{m}}$$

is the particle's *acceleration* in the fixed and moving frames, and denoting by

$$\mathbf{A} := \left(\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}\right)_{\mathrm{f}} = \left(\frac{\mathrm{d}^2\mathbf{R}}{\mathrm{d}t^2}\right)_{\mathrm{f}}$$

the acceleration of the origin of the moving frame S with respect to the fixed one S', we finally obtain the important relation

$$\boxed{\mathbf{a}_{\mathrm{f}} = \mathbf{a}_{\mathrm{m}} + \mathbf{A} + 2\boldsymbol{\omega}\times\mathbf{v}_{\mathrm{m}} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) + \dot{\boldsymbol{\omega}}\times\mathbf{r}\,.} \tag{4.17}$$

From the previous equation it follows that if the particle is acted upon by a force **F**, as measured in the *inertial* frame S', its equation of motion in the *moving* frame S will be

$$\boxed{m\mathbf{a}_{\mathrm{m}} = \mathbf{F} - m\mathbf{A} - 2m\boldsymbol{\omega}\times\mathbf{v}_{\mathrm{m}} - m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) - m\dot{\boldsymbol{\omega}}\times\mathbf{r} =: \mathbf{F} + \mathbf{F}_{\mathrm{in}}\,.} \tag{4.18}$$

Hence in the moving frame Newton's second law must be modified by adding to the *real force* **F** the *fictitious force*

$$\boxed{\mathbf{F}_{\mathrm{in}} = -m\mathbf{A} - 2m\boldsymbol{\omega}\times\mathbf{v}_{\mathrm{m}} - m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) - m\dot{\boldsymbol{\omega}}\times\mathbf{r}\,.} \tag{4.19}$$

It is important to note that this fictitious force is an *inertial* force, since it is proportional to the particle's *mass* m.

The first term in the fictitious force $\mathbf{F}_{\mathrm{in}}$ is simply due to the *acceleration* of the origin of the non-inertial frame S with respect to the inertial one S', and thus vanishes if the latter point moves with *constant velocity* relative to S'. The remaining terms in $\mathbf{F}_{\mathrm{i}}$ are due to the *rotation of*

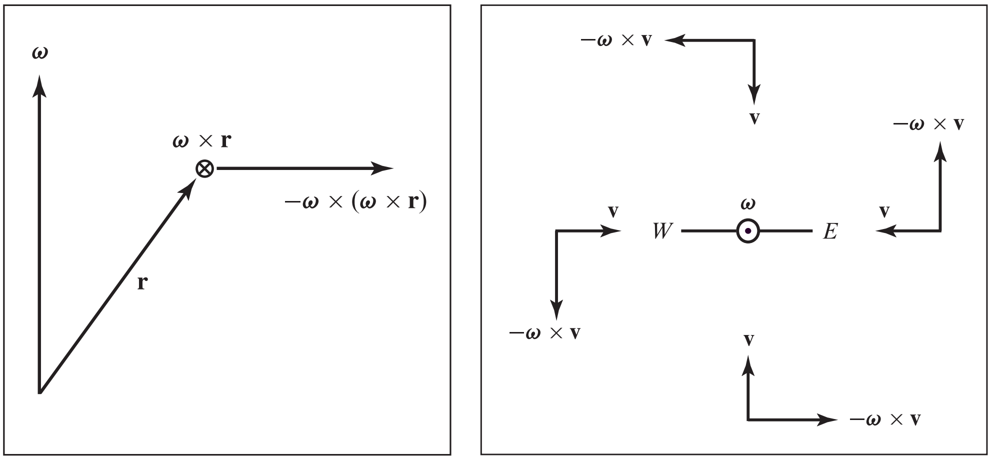


Figure 4.1. Centrifugal (left) and Coriolis (right) forces.

the axes of the moving system. While the last term vanishes if the angular velocity $\boldsymbol{\omega}$ is *constant*, the second and the third terms are in general nonzero even when $\boldsymbol{\omega}$ is constant. The term $-m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})$ is the so called **centrifugal force**, since it is a vector in the plane determined by $\boldsymbol{\omega}$ and $\mathbf{r}$, perpendicular to $\boldsymbol{\omega}$ and pointing *away* from the axis determined by the latter vector (cf. Fig. 4.1). The term $-2m\boldsymbol{\omega}\times\mathbf{v}_{\mathrm{m}}$, which depends on the particle's velocity, is known as the **Coriolis force** (see again Fig. 4.1). Note that the centrifugal force is of *second order* in $\boldsymbol{\omega}$, while the Coriolis force is of *first order*. It is therefore to be expected that the former force should be negligible compared to the latter for small angular velocities $|\boldsymbol{\omega}|$.

From the previous remarks it is also clear that the fictitious force $\mathbf{F}_{\mathrm{i}}$ vanishes identically if and only if

$$\left(\frac{\mathrm{d}^2\mathbf{R}}{\mathrm{d}t^2}\right)_{\mathrm{f}} = \boldsymbol{\omega} = 0$$

for all t. By Eq. (4.5), the vanishing of $\boldsymbol{\omega}(t)$ for all t is equivalent to the condition that the rotation matrix $O(t)$ be *constant*. From the discussion of Section 1.3.4 on Galileo's relativity principle, this is the same as saying that S is also an *inertial* frame. In other words, *the only reference frames in which inertial forces are absent are the inertial ones.*

4.4 Motion of a particle relative to the rotating Earth

We shall apply in this section the equation of motion (4.18) obtained above to study the dynamics of a particle moving near Earth's surface. We shall neglect Earth's motion around the Sun, and assume that Earth rotates around its north-south axis in the west-east direction with constant angular velocity of magnitude[1]

$$\omega = \frac{2\pi \text{ rad}}{1 \text{ sidereal day}} \simeq \frac{2\pi \text{ rad}}{86164.1 \text{ s}} \simeq 7.29212 \cdot 10^{-5} \,\text{rad}\,\text{s}^{-1}\,.$$

Let us choose a (moving) frame of **terrestrial axes** in the manner indicated in Fig. 4.2. More precisely, the origin O of the terrestrial frame S is a point on Earth's surface with *latitude* λ and *longitude* φ, the vector $\mathbf{e}_3$ (z axis) is directed along the vector $\mathbf{R}$ joining Earth's center with the point O, the vector $\mathbf{e}_1$ (x axis) is tangent to the *meridian* passing through O (in a southerly

[1]By definition, a *sidereal day* is the time taken by Earth to perform a complete rotation around its axis, while a *solar day* (equal to 24 hours) is the interval between two consecutive transits of the Sun across the meridian of any point on Earth's surface. Due to Earth's rotation around the Sun, the sidereal day is about 4 minutes shorter than the solar day.

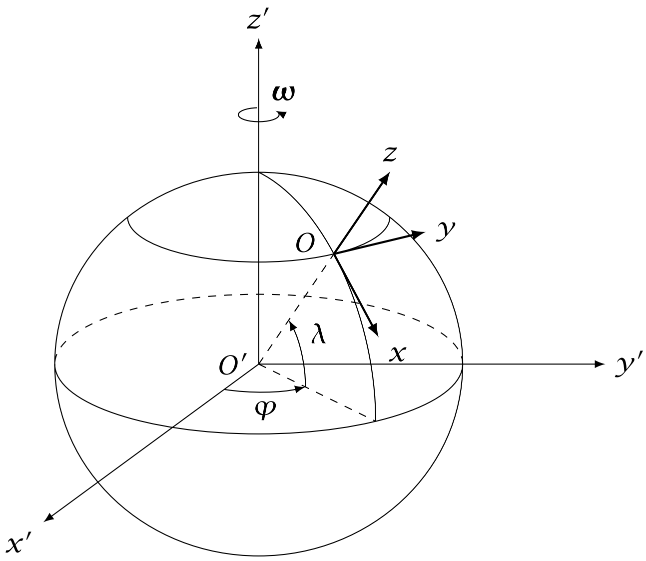


Figure 4.2. Terrestrial axes at a point O on Earth's surface.

direction), and the vector $\mathbf{e}_2$ (y axis) is then tangent to the *parallel* passing through O (in an easterly direction). In other words:

The x axis is directed towards the *south*, the y axis towards the *east*, and the z axis along the *vertical*.

As fixed axes we shall take a frame with origin O' located at Earth's center such that the vector $\mathbf{e}'_3$ is directed along the South Pole-North Pole axis. Hence Earth's angular velocity is given by

$$\boldsymbol{\omega} = \omega \mathbf{e}'_3 \,.$$

Note that the vectors $\mathbf{e}_i$ of the moving (terrestrial) frame are respectively the unit vectors $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$, and $\mathbf{e}_r$ of the spherical coordinate system at the point $\mathbf{r}$, with $\theta = \frac{\pi}{2} - \lambda$. Using Eqs. (1.4) we thus obtain

$$\begin{aligned}
\mathbf{e}_1 &= \sin\lambda\cos\varphi\,\mathbf{e}'_1 + \sin\lambda\sin\varphi\,\mathbf{e}'_2 - \cos\lambda\,\mathbf{e}'_3\,,\\
\mathbf{e}_2 &= -\sin\varphi\,\mathbf{e}'_1 + \cos\varphi\,\mathbf{e}'_2\,,\\
\mathbf{e}_3 &= \cos\lambda\cos\varphi\,\mathbf{e}'_1 + \cos\lambda\sin\varphi\,\mathbf{e}'_2 + \sin\lambda\,\mathbf{e}'_3\,.
\end{aligned}$$

From the previous equations (or simply from Fig. 4.2) it follows that in the terrestrial frame Earth's angular velocity is given by

$$\boldsymbol{\omega} = \omega(-\cos\lambda\,\mathbf{e}_1 + \sin\lambda\,\mathbf{e}_3)\,. \tag{4.20}$$

Let us next write down the equations of motion (4.18) relative to the terrestrial frame for a particle of mass m moving in the vicinity of the point O. We shall assume, for the time being, that the only force acting on the particle is Earth's gravitational attraction $m\mathbf{g}_0$, where

$$\mathbf{g}_0 = -\frac{GM}{r'^3}\,\mathbf{r}'\,,$$

M is Earth's mass and $\mathbf{r}' = \mathbf{R} + \mathbf{r}$ is the particle's position vector relative to the fixed frame. If the particle remains close enough to the point O on Earth's surface we can replace the vector $\mathbf{r}'$ by $\mathbf{R}$, and thus take

$$\mathbf{g}_0 = -\frac{GM}{R^3}\,\mathbf{R} = -\frac{GM}{R^2}\,\mathbf{e}_3 = -g_0\mathbf{e}_3$$

with

$$g_0 := \frac{GM}{R^2} \simeq 9.80665\ \mathrm{m\,s^{-2}}$$

is the *acceleration due to gravity* (also called *standard gravity*) on Earth's surface. In the rest of this chapter we shall write, for simplicity,

$$\mathbf{v}_\mathrm{m} = \dot{\mathbf{r}}\,, \qquad \mathbf{a}_\mathrm{m} = \ddot{\mathbf{r}}\,.$$

The particle's equations of motion in the terrestrial frame are therefore

$$\ddot{\mathbf{r}} = \mathbf{g}_0 - \mathbf{A} - 2\boldsymbol{\omega}\times\dot{\mathbf{r}} - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})\,.$$

This expression can be simplified taking into account that in this case

$$\mathbf{V} = \left(\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t}\right)_\mathrm{f} = \left(\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t}\right)_\mathrm{m} + \boldsymbol{\omega}\times\mathbf{R} = \boldsymbol{\omega}\times\mathbf{R}\,,$$

since $\mathbf{R} = R\mathbf{e}_3$ is constant in the terrestrial frame. Differentiating with respect to t (and taking into account that $\boldsymbol{\omega}$ is constant) we obtain

$$\mathbf{A} = \left(\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t}\right)_\mathrm{f} = \boldsymbol{\omega}\times\mathbf{V} = \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{R})\,.$$

Thus the particle's equations of motion reduce to

$$\ddot{\mathbf{r}} = \mathbf{g} - 2\boldsymbol{\omega}\times\dot{\mathbf{r}} - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})\,, \tag{4.21a}$$

where the vector

$$\mathbf{g} := \mathbf{g}_0 - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{R})\,, \tag{4.21b}$$

which is *constant* in the terrestrial frame, is known as the **effective gravity** at the point O (i.e., the acceleration relative to the terrestrial frame experienced by a particle instantaneously at rest at the point O on Earth's surface). Obviously, if apart from gravity an additional force $\mathbf{F}$ is exerted on the particle its equation of motion is

$$\ddot{\mathbf{r}} = \frac{\mathbf{F}}{m} + \mathbf{g} - 2\boldsymbol{\omega}\times\dot{\mathbf{r}} - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})\,. \tag{4.22}$$

Exercise. Show that at a point of latitude λ the plumb line deviates from the vertical by an angle $\delta(\lambda)$ given by

$$\tan\delta(\lambda) = \frac{\omega^2 R\sin\lambda\cos\lambda}{g_0 - \omega^2 R\cos^2\lambda}\,. \tag{4.23}$$

Find the latitude λ for which $\delta(\lambda)$ is maximum and the maximum value of $\delta(\lambda)$.

Solution. By definition, the plumb line is the direction determined by a string from which a mass hangs at rest, i.e., the direction opposite to the string's tension $\mathbf{T}$ at equilibrium. To find this direction, it is enough to note that the equation of motion of the mass is obtained substituting $\mathbf{F} = \mathbf{T}$ in Eq. (4.22), that is

$$\ddot{\mathbf{r}} = \mathbf{g} - 2\boldsymbol{\omega}\times\dot{\mathbf{r}} - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) + \frac{\mathbf{T}}{m}\,.$$

Since the mass is at rest $\dot{\mathbf{r}} = \ddot{\mathbf{r}} = 0$, and thus

$$\frac{\mathbf{T}}{m} = -\mathbf{g} + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) \simeq -\mathbf{g}\,, \tag{4.24}$$

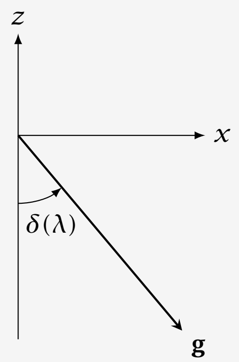


Figure 4.3. Effective gravity $\mathbf{g}$ (in the Northern Hemisphere).

where we have neglected the term $\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$ taking into account that

$$|\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})| \leqslant \omega^2 r \ll |\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{R})| = \omega^2 R \cos\lambda$$

except near the Poles ($\lambda = \pm\pi/2$). From Eq. (4.24) it follows that the direction of the plumb line is approximately that of the effective gravity $\mathbf{g}$. Taking into account that $\mathbf{g}_0 = -g_0\mathbf{e}_3$ and

$$\begin{aligned}\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{R}) &= (\boldsymbol{\omega} \cdot \mathbf{R})\boldsymbol{\omega} - \omega^2\mathbf{R} = \omega^2 R \sin\lambda(-\cos\lambda\mathbf{e}_1 + \sin\lambda\mathbf{e}_3) - \omega^2 R\mathbf{e}_3 \\ &= -\omega^2 R\cos\lambda(\sin\lambda\mathbf{e}_1 + \cos\lambda\mathbf{e}_3)\end{aligned}$$

we obtain

$$\mathbf{g} = g_0\left(\gamma\sin\lambda\cos\lambda\mathbf{e}_1 - (1-\gamma\cos^2\lambda)\mathbf{e}_3\right), \qquad \text{with} \quad \gamma := \frac{\omega^2 R}{g_0} \simeq 3.458 \cdot 10^{-3}.$$

The vector $\mathbf{g}$ has a component

$$g_1 = \gamma g_0 \sin\lambda\cos\lambda = \frac{\omega^2 R}{2}\sin 2\lambda$$

in the direction of the basis vector $\mathbf{e}_1$ (cf. Fig. 4.3). Therefore $\mathbf{g}$ deviates from the vertical (i.e., the direction of the z axis) to the south in the Northern Hemisphere ($\lambda > 0$) and to the north in the Southern one ($\lambda < 0$). The tangent of the angle $\delta(\lambda)$ between the vector $\mathbf{g}$ is given by

$$\tan\delta(\lambda) = \frac{g_1}{|g_3|} = \frac{\gamma\sin\lambda\cos\lambda}{1-\gamma\cos^2\lambda} = \boxed{\frac{\gamma\sin 2\lambda}{2-\gamma-\gamma\cos 2\lambda}},$$

which coincides with Eq. (4.23) by the definition of γ. Differentiating the previous expression we obtain

$$\frac{1}{2\gamma}\frac{\mathrm{d}}{\mathrm{d}\lambda}\tan\delta(\lambda) = \frac{(2-\gamma-\gamma\cos 2\lambda)\cos 2\lambda - \gamma\sin^2 2\lambda}{(2-\gamma-\gamma\cos 2\lambda)^2} = \frac{(2-\gamma)\cos 2\lambda - \gamma}{(2-\gamma-\gamma\cos 2\lambda)^2}.$$

Thus the angle $\delta(\lambda)$ will be maximum (in absolute value) when

$$\boxed{\cos 2\lambda = \frac{\gamma}{2-\gamma},} \tag{4.25}$$

and its maximum value $\delta_{\max}$ verifies

$$\tan\delta_{\max} = \operatorname{sgn}\lambda\,\frac{\gamma\sqrt{1-\frac{\gamma^2}{(2-\gamma)^2}}}{2-\gamma-\frac{\gamma^2}{2-\gamma}} = \frac{\gamma\operatorname{sgn}\lambda}{\sqrt{(2-\gamma)^2-\gamma^2}} = \boxed{\frac{\gamma\operatorname{sgn}\lambda}{2\sqrt{1-\gamma}}} \simeq 1.732\cdot 10^{-3}\operatorname{sgn}\lambda\,,$$

or equivalently

$$\delta_{\max} \simeq 5.955'\ \operatorname{sgn}\lambda\,.$$

Since γ is of the order of 10^{-3}, it follows from Eq. (4.25) that the latitude $\lambda_{\max}$ for which $\delta(\lambda)$ is maximum can be expressed as $\lambda_{\max} = \pm(\frac{\pi}{4}-\varepsilon)$, with $\varepsilon > 0$ small. The value of ε can be approximately computed expanding $\cos 2\lambda_{\max}$ to first order in ε:

$$\cos 2\lambda_{\max} = \cos\Big(\frac{\pi}{2}-2\varepsilon\Big) = \sin 2\varepsilon \simeq 2\varepsilon = \frac{\gamma}{2-\gamma} \simeq \frac{\gamma}{2}$$
$$\implies \quad \varepsilon \simeq \frac{\gamma}{4} \simeq 8.646\cdot 10^{-4}\ \text{rad} = 2.972'\,.$$

The equation of motion (4.21) is *exact*. In fact, the latter equation is a system of (inhomogeneous) *linear* second-order ordinary differential equations with *constant coefficients* in the components of the vector $\mathbf{r}$. As shown in the course of *Mathematical Methods I*, this type of systems can in principle be exactly solved, for instance, by transforming them into a first-order system in $(\mathbf{r},\dot{\mathbf{r}})$ and using the matrix exponential. In practice, it is preferable to first simplify Eq. (4.21) taking into account the different orders of magnitude of its terms. More precisely, the second term of (4.21b) is at most of order $\gamma \sim 10^{-3}$ with respect to the first one, while the last term of (4.21a) is at most of order $\gamma r/R$ with respect to the first. Thus, if $r \ll R$ the equation of motion can be approximated by

$$\boxed{\ddot{\mathbf{r}} = \mathbf{g}_0 - 2\boldsymbol{\omega}\times\dot{\mathbf{r}}\,.} \tag{4.26}$$

Integrating once with respect to t we obtain

$$\dot{\mathbf{r}} = \mathbf{g}_0 t - 2\boldsymbol{\omega}\times\mathbf{r} + \mathbf{c}\,,$$

where $\mathbf{c}$ is a constant vector in the terrestrial frame. Although this system can again be *exactly* solved (it is an inhomogeneous linear system of first-order ordinary differential equations with constant coefficients), it is more convenient in practice to take advantage of the fact that for small speeds $|\dot{\mathbf{r}}|$ the first term of the RHS of Eq. (4.26) is much larger than the second one, since

$$\frac{g_0}{\omega} \simeq 1.34483\cdot 10^5\ \text{m}\,\text{s}^{-1}\,.$$

This fact makes it possible to obtain an approximate solution of Eq. (4.26), considered as the first-order equation in the velocity

$$\boxed{\dot{\mathbf{v}} = \mathbf{g}_0 - 2\boldsymbol{\omega}\times\mathbf{v}}\,, \tag{4.27}$$

by expanding $\mathbf{v}$ in powers of ω:

$$\boxed{\mathbf{v}(t) = \mathbf{v}_1(t) + \omega\mathbf{v}_2(t) + O(\omega^2)\,,}$$

with $\mathbf{v}_{1,2}$ independent of ω and

$$\mathbf{v}(0) := \mathbf{v}_0 \quad \implies \quad \mathbf{v}_1(0) = \mathbf{v}_0\,, \quad \mathbf{v}_2(0) = 0\,.$$

Substituting into Eq. (4.27) we have

$$\dot{\mathbf{v}}_1 + \omega\dot{\mathbf{v}}_2 = \mathbf{g}_0 - 2\boldsymbol{\omega}\times\mathbf{v}_1 + O(\omega^2)\,,$$

whence, equating to zero the terms $O(1)$ and $O(\omega)$ in both sides of the latter expression we obtain

$$\dot{\mathbf{v}}_1 = \mathbf{g}_0 \,, \qquad \omega\dot{\mathbf{v}}_2 = -2\boldsymbol{\omega}\times\mathbf{v}_1 \,.$$

Solving for $\mathbf{v}_1$ in the first of these equations and substituting the result into the second one we have

$$\mathbf{v}_1 = \mathbf{g}_0 t + \mathbf{v}_0 \,, \qquad \omega\dot{\mathbf{v}}_2 = -2\boldsymbol{\omega}\times\mathbf{v}_0 - 2t\boldsymbol{\omega}\times\mathbf{g}_0 \overset{\mathbf{v}_2(0)=0}{\Longrightarrow} \omega\mathbf{v}_2 = -2t\boldsymbol{\omega}\times\mathbf{v}_0 - t^2\boldsymbol{\omega}\times\mathbf{g}_0 \,.$$

Hence

$$\mathbf{v} \simeq \mathbf{v}_1 + \omega\mathbf{v}_2 = \mathbf{v}_0 + \mathbf{g}_0 t - 2t\boldsymbol{\omega}\times\mathbf{v}_0 - t^2\boldsymbol{\omega}\times\mathbf{g}_0 \,,$$

and integrating with respect of t we finally obtain

$$\mathbf{r} \simeq \mathbf{r}_0 + \mathbf{v}_0 t + \mathbf{g}_0 \frac{t^2}{2} - t^2\boldsymbol{\omega}\times\mathbf{v}_0 - \frac{t^3}{3}\,\boldsymbol{\omega}\times\mathbf{g}_0 \,. \tag{4.28}$$

Exercise. A particle is thrown vertically from a point on Earth's surface with latitude λ until it reaches a height h. Show that the particle lands at a point $(4/3)\sqrt{8h^3/g_0}\,\omega\cos\lambda$ west of the starting point. (Consider only small heights h and neglect air resistance.)

Solution. We have

$$\mathbf{r}_0 = 0\,, \quad \mathbf{v}_0 = v_0\mathbf{e}_3 \quad \Longrightarrow \quad -\boldsymbol{\omega}\times\mathbf{v}_0 = \omega v_0(\cos\lambda\,\mathbf{e}_1 - \sin\lambda\,\mathbf{e}_3)\times\mathbf{e}_3 = -\omega v_0\cos\lambda\,\mathbf{e}_2$$

and

$$-\boldsymbol{\omega}\times\mathbf{g}_0 = \omega g_0\cos\lambda\,\mathbf{e}_2 \,.$$

From Eq. (4.28) with the previous values of $\mathbf{r}_0$ and $\mathbf{v}_0$ we (approximately) obtain

$$\mathbf{r} = v_0 t\,\mathbf{e}_3 - \frac{g_0}{2}\,t^2\,\mathbf{e}_3 + \Big(-\omega v_0\cos\lambda\,t^2 + \frac{\omega g_0}{3}\,t^3\cos\lambda\Big)\,\mathbf{e}_2 \,.$$

Hence the law of motion is

$$x = 0\,, \qquad y = \omega v_0\cos\lambda\,t^2\left(\frac{g_0 t}{3v_0} - 1\right), \qquad z = v_0 t - \frac{g_0}{2}\,t^2 \,.$$

The particle lands at the time $t_0 > 0$ for which $z = 0$, namely

$$t_0 = \frac{2v_0}{g_0} \,.$$

The value of the y coordinate at this time is thus

$$y(t_0) = -\frac{4}{3}\,\frac{\omega v_0^3}{g_0^2}\,\cos\lambda \leqslant 0 \,.$$

We thus see that the particle deviates to the *west*, both in the northern and in the southern hemisphere, with maximum deviation at the equator ($\lambda = 0$). In order to express the deviation $y(t_0)$ in terms of the maximum height h, it suffices to note that this height is reached when $\dot{z} = 0$:

$$\dot{z} = v_0 - g_0 t = 0 \quad \Longrightarrow \quad t = \frac{v_0}{g_0} \quad \Longrightarrow \quad z = h = \frac{v_0^2}{2g_0} \,.$$

Hence

$$y(t_0) = -\frac{4\omega}{3g_0^2}(2g_0 h)^{3/2}\cos\lambda = \boxed{-\frac{4}{3}\,\omega\cos\lambda\sqrt{\frac{8h^3}{g_0}}}\,.$$

For instance,

$$h = 100\,\mathrm{m}\,, \quad \lambda = 40^\circ\,25' \;\text{(Madrid's latitude)} \quad \Longrightarrow \quad y(t_0) = -6.686\ \mathrm{cm}\,.$$

Exercise. Redo the previous problem assuming that the particle is dropped from a height h over the vertical.

Solution. In this case

$$\mathbf{r}_0 = h\mathbf{e}_3\,, \qquad \mathbf{v}_0 = 0\,,$$

and substituting into Eq. (4.28) we obtain

$$x = 0\,, \qquad y = \frac{\omega g_0}{3}\,t^3\cos\lambda\,, \qquad z = h - \frac{g_0}{2}\,t^2\,.$$

The particle lands when

$$t_0 = \sqrt{\frac{2h}{g_0}}\,,$$

and its deviation in the y direction is thus

$$\boxed{y(t_0) = \frac{\omega}{3}\cos\lambda\sqrt{\frac{8h^3}{g_0}}\,.}$$

Since $y(t_0) \geqslant 0$, the particles deviates *eastwards* in both hemispheres.

4.5 Foucault's pendulum

In 1851 J.B.L. Foucault experimentally demonstrated Earth's rotation using the pendulum that nowadays bears his name, schematically represented in Fig. 4.4.

Since we are only interested in studying the small oscillations, we shall assume that the pendulum's length l is very large compared to the coordinates x, y, z of its bob. The total force acting on the pendulum's bob is thus

$$\mathbf{F} = m\mathbf{g}_0 + \mathbf{T}\,,$$

where the tension $\mathbf{T}$ of the pendulum's string is given by

$$\mathbf{T} = T\,\frac{l\mathbf{e}_3 - \mathbf{r}}{|l\mathbf{e}_3 - \mathbf{r}|} \simeq T\left(\mathbf{e}_3 - \frac{\mathbf{r}}{l}\right).$$

With this approximation the equation of motion reads

$$\boxed{\ddot{\mathbf{r}} = \mathbf{g}_0 + \frac{T}{m}\left(\mathbf{e}_3 - \frac{\mathbf{r}}{l}\right) - 2\boldsymbol{\omega}\times\dot{\mathbf{r}}\,,}$$

where

$$\boldsymbol{\omega}\times\dot{\mathbf{r}} = \omega\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ -\cos\lambda & 0 & \sin\lambda \\ \dot{x} & \dot{y} & \dot{z} \end{vmatrix} = \omega\big[-\dot{y}\sin\lambda\,\mathbf{e}_1 + (\dot{x}\sin\lambda + \dot{z}\cos\lambda)\mathbf{e}_2 - \dot{y}\cos\lambda\,\mathbf{e}_3\big],$$

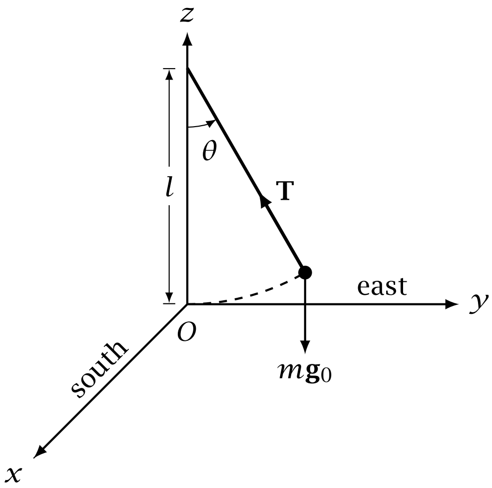


Figure 4.4. Foucault's pendulum.

and thus

$$\boxed{\begin{aligned} \ddot{x} &= -\frac{Tx}{lm} + 2\omega\dot{y}\sin\lambda \\ \ddot{y} &= -\frac{Ty}{lm} - 2\omega(\dot{x}\sin\lambda + \dot{z}\cos\lambda) \\ \ddot{z} &= -g_0 + \frac{T}{m}\Big(1 - \frac{z}{l}\Big) + 2\omega\dot{y}\cos\lambda\,. \end{aligned}} \tag{4.29}$$

Note that $z \ll \sqrt{x^2 + y^2}$, since calling θ the angle between the pendulum and the vertical we have

$$\sqrt{x^2 + y^2} = l\sin\theta \simeq l\theta\,, \qquad z = l(1 - \cos\theta) \simeq \frac{l\theta^2}{2}\,.$$

We can thus neglect in Eqs. (4.29) the quantities z, $\dot{z}$ and $\ddot{z}$ compared to x, y and their derivatives. In particular, from the last equation we obtain

$$\frac{T}{m} \simeq g_0 - 2\omega\dot{y}\cos\lambda \simeq g_0\,,$$

since g_0/ω is of the order of $10^5\ \mathrm{m\,s^{-1}}$. Substituting into the first two equations (4.29) and dropping the term proportional to $\dot{z}$ we finally obtain the following system for the coordinates (x, y):

$$\boxed{\begin{aligned} \ddot{x} + \alpha^2 x &= 2\omega\dot{y}\sin\lambda \\ \ddot{y} + \alpha^2 y &= -2\omega\dot{x}\sin\lambda\,, \end{aligned}} \tag{4.30}$$

where

$$\boxed{\alpha := \sqrt{\frac{g_0}{l}}}$$

is the pendulum's natural frequency.

The latter equations are easily solved introducing the complex variable

$$u = x + \mathrm{i}y\,,$$

in terms of which they adopt the simple form

$$\boxed{\ddot{u} + 2\mathrm{i}\Omega\,\dot{u} + \alpha^2 u = 0\,, \qquad \text{with } \Omega := \omega\sin\lambda\,.} \tag{4.31}$$

This is a linear homogeneous second-order ordinary differential equation with constant coefficients, whose characteristic polynomial

$$p(s) = s^2 + 2\mathrm{i}\Omega\, s + \alpha^2$$

possesses the two pure imaginary roots

$$s_\pm = -\mathrm{i}\Omega \pm \mathrm{i}\sqrt{\Omega^2 + \alpha^2}\,.$$

For all practical purposes, we can neglect the term Ω^2 in the radical compared to α^2, since[2]

$$\frac{\Omega^2}{\alpha^2} \leqslant \frac{\omega^2 l}{g_0} = \frac{l}{1.84422 \cdot 10^9\,\mathrm{m}}\,.$$

We thus have

$$s_\pm \simeq -\mathrm{i}\Omega \pm \mathrm{i}\,\alpha\,,$$

and the general solution of Eq. (4.31) is therefore given by

$$\boxed{u = \mathrm{e}^{-\mathrm{i}\Omega t}\left(c_1 \mathrm{e}^{\mathrm{i}\alpha t} + c_2 \mathrm{e}^{-\mathrm{i}\alpha t}\right)\,,} \tag{4.32}$$

where the constants c_1, c_2 are in general *complex*.

Let us find, for instance, the solution of Eqs. (4.30) with the initial conditions

$$x(0) = x_0 > 0\,, \quad y(0) = 0\,, \quad \dot{x}(0) = \dot{y}(0) = 0\,, \tag{4.33}$$

i.e., when the pendulum's bob is initially at rest in the Oxz plane at a distance x_0 from the vertical. If Earth did not rotate around its north-south axis, i.e., if $\omega = 0$, the solution of the equations of motion (4.30) with the initial conditions (4.33) would be

$$x = x_0 \cos(\alpha t)\,, \qquad y = 0\,.$$

In other words, the pendulum would oscillate with frequency α and amplitude x_0 around the vertical in the Oxz plane. On the other hand, when $\omega > 0$ the solution (4.32) verifying the initial conditions (4.33) is easily found taking into account that

$$u(0) = x(0) + \mathrm{i}y(0) = x_0\,, \qquad \dot{u}(0) = \dot{x}(0) + \mathrm{i}\dot{y}(0) = 0\,. \tag{4.34}$$

We thus have

$$\begin{cases} c_1 + c_2 = x_0\,, \\ \mathrm{i}\alpha(c_1 - c_2) - \mathrm{i}\Omega(c_1 + c_2) = \mathrm{i}\alpha(c_1 - c_2) - \mathrm{i}\Omega x_0 = 0 \quad \Longrightarrow \quad c_1 - c_2 = \dfrac{\Omega}{\alpha}\, x_0 \simeq 0\,, \end{cases}$$

whose approximate solution is

$$c_1 = c_2 = \frac{1}{2}\, x_0\,.$$

Thus the sought-for solution of Eq. (4.31) is approximately

$$\boxed{u = x_0 \mathrm{e}^{-\mathrm{i}\Omega t} \cos(\alpha t)}\,.$$

The complex number $\mathrm{e}^{-\mathrm{i}\Omega t}$ is the point on the unit circle making an angle $-\Omega t$ with the real (x)

[2]More precisely, we have

$$s_\pm = -\mathrm{i}\Omega \pm \mathrm{i}\alpha\sqrt{1 + \frac{\Omega^2}{\alpha^2}} = \mathrm{i}\alpha\left(\pm 1 - \frac{\Omega}{\alpha} + O\left(\frac{\Omega^2}{\alpha^2}\right)\right),$$

so that the term discarded is of order $O(\Omega/\alpha)$ compared to the smallest term retained ($-\mathrm{i}\Omega$).

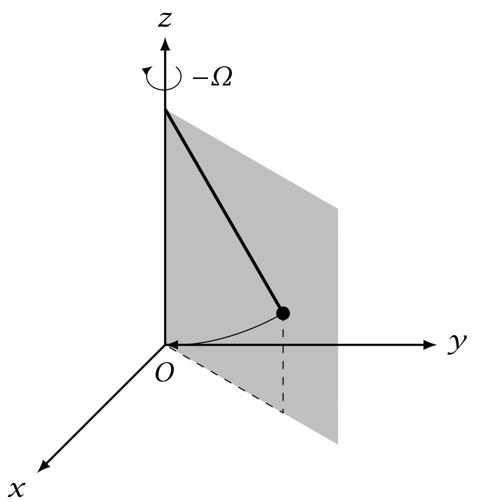


Figure 4.5. Rotation of the plane of Foucault's pendulum (shaded gray).

axis, obtained rotating the unit coordinate vector $\mathbf{e}_1$ by an angle $-\Omega t$ around the z axis. Since $x_0 \cos(\alpha t)$ is real, the previous equation can be rewritten in real terms as

$$(x, y) = x_0 \cos(\alpha t)\, \mathbf{n}(t)\,, \qquad \text{with} \quad \boxed{\mathbf{n}(t) := R_3(-\Omega t)\mathbf{e}_1\,.}$$

From this equation it follows that at each instant t *the pendulum's plane,* determined by the vectors $\mathbf{e}_3$ and $\mathbf{n}(t)$, *makes an angle $-\Omega t$ with the Oxz plane.* The pendulum's motion can thus be viewed as the composition of two periodic motions, namely a "fast" oscillation with period $2\pi/\alpha$ in the plane determined by the vectors $\mathbf{e}_3$ and $\mathbf{n}(t)$ and a "slow" rotation of the latter plane around the z axis with period $2\pi/\Omega \gg 2\pi/\alpha$ (cf. Fig. 4.5). In particular:

In the Northern Hemisphere the pendulum's plane rotates *clockwise*, i.e., in the *east-south* direction (since $\dot{\varphi} = -\Omega = -\omega \sin\lambda < 0$), *with angular velocity $\Omega = \omega \sin\lambda$.* In the Southern Hemisphere the rotation of the pendulum's plane is *counterclockwise* (since $\sin\lambda < 0$), and in the equator ($\lambda = 0$) no such rotation occurs.

Note that in each period $2\pi/\alpha$ of the pendulum (time in which $\cos(\alpha t)$ performs a complete oscillation) the angle between the pendulum's plane and the Oxz plane increases by $-2\pi\Omega/\alpha$, which as noted before is a very small quantity. The *period* of the rotation of the pendulum's plane is given by

$$\tau = \frac{2\pi}{\Omega} = \frac{2\pi}{\omega} \csc\lambda = \csc\lambda \ \text{sidereal days}\,.$$

For instance, at a latitude of 30° the period is 2 sidereal days, while in Madrid ($\lambda = 40°\,25'$) it is 1.5424 sidereal days. Note, finally, that (with the approximations made) the motion of the pendulum's bob is not exactly periodic unless the ratio α/Ω is a rational number (cf. Fig. 4.6).

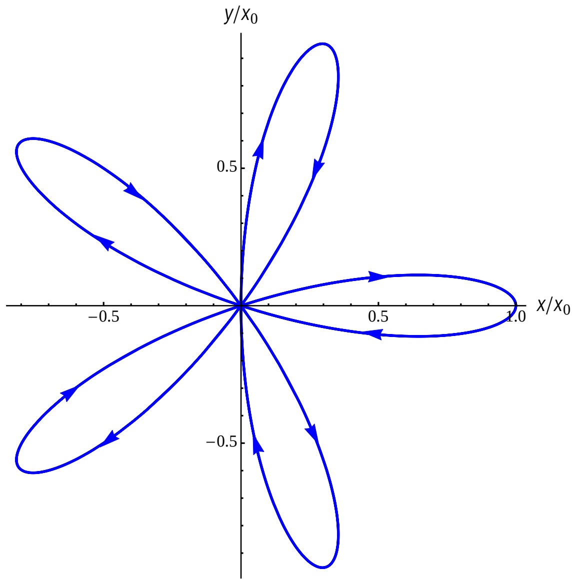


Figure 4.6. Projection onto the Oxy plane of the trajectory of the pendulum's bob for $\alpha/\Omega = 5$.

5 Rigid body motion

5.1 Degrees of freedom

A **rigid body** is a system of particles of mass m_α ($\alpha = 1, \dots, N$) in which the *distance* $|\mathbf{r}_\alpha - \mathbf{r}_\beta|$ between any two particles is *constant*. In other words, a rigid body is a mechanical system of N particles subject to the $N(N-1)/2$ time-independent holonomic constraints (not all of them independent!)

$$\boxed{(\mathbf{r}_\alpha - \mathbf{r}_\beta)^2 = l_{\alpha\beta}^2 = \text{const.}, \qquad 1 \leqslant \alpha < \beta \leqslant N.} \tag{5.1}$$

- We shall assume in what follows that Newton's third law holds in its *strongest version*, i.e., that the constraint force $\mathbf{F}_{\alpha\beta}$ exerted by particle β on particle α satisfies

$$\boxed{\mathbf{F}_{\alpha\beta} = -\mathbf{F}_{\beta\alpha} \parallel \mathbf{r}_\alpha - \mathbf{r}_\beta\,.}$$

It is easy to see that if this is the case the constraints (5.1) are *ideal*, i.e., that the *principle of virtual work* holds. Indeed, the work done by the constraint forces in an infinitesimal displacement $\mathrm{d}\mathbf{r}_\alpha$ ($\alpha = 1, \dots, N$) of the system's particles is given by[1]

$$\sum_{\alpha\neq\beta} \mathbf{F}_{\alpha\beta} \cdot \mathrm{d}\mathbf{r}_\alpha = \frac{1}{2} \sum_{\alpha\neq\beta} (\mathbf{F}_{\alpha\beta} \cdot \mathrm{d}\mathbf{r}_\alpha + \mathbf{F}_{\beta\alpha} \cdot \mathrm{d}\mathbf{r}_\beta) = \frac{1}{2} \sum_{\alpha\neq\beta} \mathbf{F}_{\alpha\beta} \cdot (\mathrm{d}\mathbf{r}_\alpha - \mathrm{d}\mathbf{r}_\beta)\,. \tag{5.2}$$

On the other hand, differentiating the constraint equation we obtain

$$(\mathbf{r}_\alpha - \mathbf{r}_\beta) \cdot (\mathrm{d}\mathbf{r}_\alpha - \mathrm{d}\mathbf{r}_\beta) = 0\,, \qquad 1 \leqslant \alpha < \beta \leqslant N\,,$$

whence it follows (since $\mathbf{F}_{\alpha\beta}$ is parallel to the vector $\mathbf{r}_\alpha - \mathbf{r}_\beta$) that

$$\mathbf{F}_{\alpha\beta} \cdot (\mathrm{d}\mathbf{r}_\alpha - \mathrm{d}\mathbf{r}_\beta) = 0\,, \qquad 1 \leqslant \alpha < \beta \leqslant N\,.$$

We thus see that all the terms in the last sum in Eq. (5.2) vanish identically, and as a consequence the total work done by the constraint forces is indeed zero.

- We shall say that a rigid body is *generic* if it contains three non-collinear particles[2].

> In a generic rigid body it is always possible to construct a set of moving axes, the so-called **body axes**, with respect to which all of the body's particles are *fixed*, i.e., *at rest*.

In other words, the position vectors $\mathbf{r}_\alpha$ ($1 \leqslant \alpha \leqslant N$) of *all* the particles making up the body are *constant* in the frame of body axes.

Proof. Let P, Q, R be three non-collinear points in the rigid body. A set of body axes is obtained, for example, taking as origin the point P, the x axis in the direction of the vector $\overrightarrow{PQ}$, the y axis in the direction of the line in the plane PQR perpendicular to $\overrightarrow{PQ}$, oriented so that the y coordinate of the point R is (say) positive, and the z axis in the direction of $\overrightarrow{PQ} \times \overrightarrow{PR}$. Indeed, in

[1] In what follow, sums over *Greek* indices α, β, γ, $\dots$ will implicitly run from 1 to N, while *Latin* ones i, j, k, $\dots$ will take the values $1, 2, 3$.

[2] It is easy to show that if three points of a rigid body are collinear at some instant they must remain collinear at any other time.

this frame the points P, Q, and R are fixed by construction (exercise). We shall next show that the coordinates (x, y, z) of any other point S in the body relative to this frame are constant. To this end, let $(a, 0, 0)$ and $(b, c, 0)$ denote the coordinates of Q and R in the frame just constructed (where $a, c > 0$ by construction). If r_1, r_2 and r_3 are the (fixed) distances of S to the points P, Q, R we have

$$x^2 + y^2 + z^2 = r_1^2, \quad (x-a)^2 + y^2 + z^2 = r_2^2, \quad (x-b)^2 + (y-c)^2 + z^2 = r_3^2 . \tag{5.3}$$

Subtracting the second equation from the first we obtain

$$2ax = r_1^2 - r_2^2 + a^2 \quad \Longrightarrow \quad x = \frac{a}{2} + \frac{r_1^2 - r_2^2}{2a} ,$$

so that the x coordinate is constant. Likewise, subtracting the third equation from the second one we deduce that

$$y = \frac{c}{2} + \frac{1}{2c}\left[2(a-b)x + r_2^2 - r_3^2 + b^2 - a^2\right]$$

is also constant. Finally, from the above and any of the three equations (5.3) it follows that z^2 is constant, which implies (by continuity) that z is also constant. ∎

• Obviously, there is an infinite number of body axes, obtained from the frame we have just constructed by translating the origin to any point fixed in the body and applying a *constant* rotation to the axes.

More precisely, we shall say that a point is **fixed in the body** if its coordinates are constant (i.e., time-independent) in a frame of body axes. Such a point is, for instance, the body's *center of mass*. Indeed, let O and O' respectively denote the origin of the frame of body axes and of the inertial frame, and let $\mathbf{a} = \overrightarrow{OO'}$. The position vector of the body's CM in the inertial frame is by definition the vector

$$\overrightarrow{O'C} = \frac{1}{M}\sum_\alpha m_\alpha \mathbf{r}'_\alpha ,$$

where $M = \sum_\alpha m_\alpha$ is the body's total mass and $\mathbf{r}'_\alpha$ is the position vector of the α-th particle relative to the inertial frame. Moreover, if $\mathbf{r}$ and $\mathbf{r}'$ respectively denote the position vectors of a point in space in the frame of body axes and in the inertial frame we have

$$\mathbf{r} = \mathbf{a} + \mathbf{r}' .$$

In particular, the position vector of the body's CM in the frame of body axes is given by

$$\overrightarrow{OC} = \mathbf{a} + \frac{1}{M}\sum_\alpha m_\alpha \mathbf{r}'_\alpha = \mathbf{a} + \frac{1}{M}\sum_\alpha m_\alpha(\mathbf{r}_\alpha - \mathbf{a}) = \mathbf{a} + \frac{1}{M}\sum_\alpha m_\alpha \mathbf{r}_\alpha - \frac{\mathbf{a}}{M}\sum_\alpha m_\alpha = \frac{1}{M}\sum_\alpha m_\alpha \mathbf{r}_\alpha ,$$

where $\mathbf{r}_\alpha$ denotes the position vector of the α-th particle in the latter frame. Since the components of the position vectors $\mathbf{r}_\alpha$ are by construction constant in the frame of body axes, so are the coordinates of the CM in this frame. Note, finally, that from the previous argument it also follows that the position vector of the CM in the frame of body axes is still given by the usual formula $(1/M)\sum_\alpha m_\alpha \mathbf{r}_\alpha$.

A *generic* rigid body has 6 degrees of freedom.

Indeed, in order to determine the coordinates of any particle in the rigid body at an arbitrary time t, given its coordinates at the initial instant t_0, it suffices to know the position of the origin P of a set of body axes $\{\mathbf{e}_i\}_{i=1}^3$ together with the rotation matrix $O(t)$ relating these axes to those of an inertial (fixed) reference frame $\{\mathbf{e}'_i\}_{i=1}^3$ (cf. Eq. (4.1)). Indeed, all the vectors $\mathbf{r}_\alpha$

are known at $t = t_0$, and since they are *constant* the coordinates $\mathbf{r}'_\alpha(t)$ of the position vector of particle α in the fixed frame at an arbitrary time t can be computed through the formula[3]

$$\mathbf{r}'_\alpha(t) = \overrightarrow{O'P} + O(t)\mathbf{r}_\alpha\,.$$

The vector $\overrightarrow{O'P}$ (i.e., the position of the point P) is determined by three parameters (for instance, its Cartesian coordinates), while the matrix $O(t) \in \mathrm{SO}(3)$ can be specified by another three independent parameters (for example, the two spherical coordinates of the rotation axis $\mathbf{n}$ and the rotation angle $\theta \in [0, \pi]$). (In practice, the most widespread way of determining the rotation matrix $O(t)$ is through three angles, the so-called *Euler angles*).

> *Exercise.* How many degrees of freedom has a *rotor* (rigid body all of whose particles are collinear)?

The above considerations also apply to the *continuous version* of a rigid body, which consists of a continuous mass distribution of density $\rho(\mathbf{r})$ over a volume $\Omega \subset \mathbb{R}^3$ whose *shape* does not change with time. In other words, the location and shape of the body at any instant t is obtained applying a *rigid motion* (overall translation followed by a rotation, or vice versa) to the set Ω. More precisely, at any instant t the mass distribution is concentrated on the set $\Omega(t) \subset \mathbb{R}^3$ given by

$$\Omega(t) = O(t)\Omega + \mathbf{X}(t),$$

for some vector $\mathbf{X}(t) \in \mathbb{R}^3$ and rotation matrix $O(t) \in \mathrm{SO}(3)$ about a fixed point in the body (for instance, its center of mass). The state of the system is thus completely determined by the three components of the vector $\mathbf{X}(t)$ together with the three parameters needed to specify the rotation matrix $O(t)$. As in the discrete case, this implies that *a continuous rigid body has* 6 *degrees of freedom.* Of course, we can also have continuous bodies whose mass density is concentrated on a surface, or even a curve, in $\mathbb{R}^3$. Note, finally, that the *center of mass* of a continuous rigid body is naturally defined by

$$\boxed{\mathbf{R} = \frac{1}{M}\int_\Omega \rho(\mathbf{r})\mathbf{r}\,\mathrm{d}^3\mathbf{r}\,,}$$

where

$$\boxed{M = \int_\Omega \rho(\mathbf{r})\,\mathrm{d}^3\mathbf{r}}$$

is the body's total mass. In particular, if the mass density ρ is constant then

$$\boxed{\mathbf{R} = \frac{1}{V}\int_\Omega \mathbf{r}\,\mathrm{d}^3\mathbf{r}\,,}$$

where V is the body's volume. Similar considerations apply to a continuous rigid body whose mass is distributed over a surface or on a curve in $\mathbb{R}^3$, replacing the volume element by the surface or line element and the volume density by the surface or line density.

5.2 Angular momentum and kinetic energy

We shall next compute the angular momentum of the rigid body with respect to an *inertial frame*, which we shall often call for short (as in the previous chapter) the *fixed* (or *space*) *frame*.

[3]In the formula that follows $\mathbf{r}_\alpha$ denotes the vector whose components are the coordinates of the position of particle α in the body frame. As explained in Chapter 1.3.3, to obtain the coordinates of the same vector in the fixed frame we have to multiply $\mathbf{r}_\alpha$ by the rotation matrix $O(t)$.

From now on, unless otherwise stated we shall take the rigid body's *center of mass* as the origin of the set of body axes.

Let us denote, as usual, by $\mathbf{R}(t)$ the vector joining the origin O' of the fixed frame with the center of mass O (i.e., the origin of the set of body axes). If $\mathbf{r}_\alpha$ and $\mathbf{r}'_\alpha$ are respectively the position vectors of the α-th particle with respect to the body axes (which plays the role of the *moving frame* in the last chapter) and the fixed ones we have

$$\boxed{\mathbf{r}'_\alpha = \mathbf{R} + \mathbf{r}_\alpha\,.}$$

In this case $\dot{\mathbf{r}}_\alpha = 0$, since the particles which make up the rigid body are at rest with respect to the set of body axes. Hence Eq. (4.16) reduces to

$$\boxed{\mathbf{v}'_\alpha = \mathbf{V} + \boldsymbol{\omega} \times \mathbf{r}_\alpha\,,} \tag{5.4}$$

where $\boldsymbol{\omega}$ denotes the instantaneous angular velocity of the set of body axes with respect to the fixed ones and

$$\boxed{\mathbf{v}'_\alpha := \left(\frac{\mathrm{d}\mathbf{r}'_\alpha}{\mathrm{d}t}\right)_\mathrm{f}\,, \qquad \mathbf{V} := \left(\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t}\right)_\mathrm{f}\,.}$$

- By Eq. (5.4), the infinitesimal change of the position vector of particle α from a time t to a time $t + \mathrm{d}t$ is given by

$$\mathrm{d}\mathbf{r}'_\alpha = \mathbf{v}'_\alpha\,\mathrm{d}t = \mathbf{V}\,\mathrm{d}t + \boldsymbol{\omega}\,\mathrm{d}t \times \mathbf{r}_\alpha = \mathrm{d}\mathbf{R} + \boldsymbol{\omega}\,\mathrm{d}t \times \mathbf{r}_\alpha\,.$$

Hence:

The instantaneous motion of the body, as seen from the fixed reference frame, can be viewed as an infinitesimal translation followed by an infinitesimal rotation around the axis parallel to $\boldsymbol{\omega}$ passing through the CM by an angle $\omega(t)\,\mathrm{d}t$.

The latter assertion can also be proved directly taking into account that the rate of change of the position vector with respect to the CM of any particle α, as measured in the fixed frame, is given by

$$\left(\frac{\mathrm{d}\mathbf{r}_\alpha}{\mathrm{d}t}\right)_\mathrm{f} = \dot{\mathbf{r}}_\alpha + \boldsymbol{\omega}(t) \times \mathbf{r}_\alpha = \boldsymbol{\omega}(t) \times \mathbf{r}_\alpha\,. \tag{5.5}$$

Hence from the point of view of the fixed frame the position vectors of all particles in the rigid body relative to the CM rotate *instantaneously* with the *same* angular velocity $\omega(t)$ around an axis parallel to the vector $\boldsymbol{\omega}(t)/\omega(t)$ passing through the CM.

The rigid body's linear momentum with respect to the inertial frame is given by

$$\boxed{\mathbf{P} = \sum_\alpha m_\alpha \mathbf{v}'_\alpha = \sum_\alpha m_\alpha \mathbf{V} + \boldsymbol{\omega} \times \sum_\alpha m_\alpha \mathbf{r}_\alpha = M\mathbf{V}\,,} \tag{5.6}$$

where

$$M = \sum_\alpha m_\alpha$$

is the body's total mass and we have used the identity

$$\sum_\alpha m_\alpha \mathbf{r}_\alpha = 0 \tag{5.7}$$

(since the LHS is proportional to the position vector of the CM with respect to the CM itself). As expected, Eq. (5.6) coincides with Eq. (1.79) in Chapter 1.

Let us next find the rigid body's angular momentum with respect to the origin O' of the set of fixed axes (as measured in the latter frame), defined by

$$\boxed{\mathbf{L} = \sum_\alpha m_\alpha \mathbf{r}'_\alpha \times \mathbf{v}'_\alpha \,.}$$

Using Eq. (5.4) for $\mathbf{v}'_\alpha$ and the identity (5.7) we easily obtain

$$\begin{aligned}\mathbf{L} &= \sum_\alpha m_\alpha(\mathbf{R}+\mathbf{r}_\alpha)\times(\mathbf{V}+\boldsymbol{\omega}\times\mathbf{r}_\alpha) = M\mathbf{R}\times\mathbf{V} + \mathbf{R}\times\Big(\boldsymbol{\omega}\times\sum_\alpha m_\alpha\mathbf{r}_\alpha\Big) + \Big(\sum_\alpha m_\alpha\mathbf{r}_\alpha\Big)\times\mathbf{V}\\ &\quad + \sum_\alpha m_\alpha\mathbf{r}_\alpha\times(\boldsymbol{\omega}\times\mathbf{r}_\alpha) = \boxed{M\mathbf{R}\times\mathbf{V} + \sum_\alpha m_\alpha\mathbf{r}_\alpha\times(\boldsymbol{\omega}\times\mathbf{r}_\alpha)}\,. \end{aligned} \tag{5.8}$$

Note that this expression is nothing but Eq. (1.84) from Chapter 1, on account of Eq. (5.5). The first term in Eq. (5.8) is simply the angular momentum of a particle located at the CM with mass equal to the body's total mass. To interpret the second term, note first that the angular momentum of the rigid body with respect to any point P, as measured in the fixed frame S', is by definition

$$\mathbf{L}_P := \sum_\alpha m_\alpha(\mathbf{r}'_\alpha - \overrightarrow{O'P})\times\mathbf{v}'_\alpha\,,$$

where $\mathbf{r}'_\alpha - \overrightarrow{O'P}$ is the position vector of the α-th particle with respect to the point P. In particular, taking P as the CM we have

$$\boxed{\mathbf{L}_{\mathrm{CM}} = \sum_\alpha m_\alpha\mathbf{r}_\alpha\times\mathbf{v}'_\alpha = \sum_\alpha m_\alpha\mathbf{r}_\alpha\times(\boldsymbol{\omega}\times\mathbf{r}_\alpha)\,,} \tag{5.9}$$

where we have used again Eq. (5.4) for $\mathbf{v}'_\alpha$ and the identity (5.7). By Eq. (5.8) we then have

$$\boxed{\mathbf{L} = M\mathbf{R}\times\mathbf{V} + \mathbf{L}_{\mathrm{CM}}\,.} \tag{5.10}$$

It is important to note that, although $\mathbf{L}_{\mathrm{CM}}$ is the rigid body's angular momentum with respect to the CM, *it is computed in the space frame* S', since the particle's velocities $\mathbf{v}'_\alpha$ in Eq. (5.9) are measured in the latter frame.

Proceeding in the same way we can compute the body's kinetic energy (with respect to the inertial frame)

$$\boxed{T = \frac{1}{2}\sum_\alpha m_\alpha \mathbf{v}'^2_\alpha\,.}$$

Indeed, using again Eq. (5.4) for $\mathbf{v}'_\alpha$ and the identity (5.7) we obtain the expression

$$T = \frac{1}{2}\sum_\alpha m_\alpha(\mathbf{V}+\boldsymbol{\omega}\times\mathbf{r}_\alpha)^2 = \frac{1}{2}M\mathbf{V}^2 + \frac{1}{2}\sum_\alpha m_\alpha(\boldsymbol{\omega}\times\mathbf{r}_\alpha)^2\,, \tag{5.11}$$

which again coincides with Eq. (1.88) of Chapter 1 on account of Eq. (5.5). We thus have

$$\boxed{T = \frac{1}{2}M\mathbf{V}^2 + T_{\mathrm{rot}}\,,\qquad T_{\mathrm{rot}} = \frac{1}{2}\sum_\alpha m_\alpha(\boldsymbol{\omega}\times\mathbf{r}_\alpha)^2\,,} \tag{5.12}$$

where the first term in T is the CM's translational energy while the second one is the body's **rotational energy** around its CM, since

$$\frac{1}{2}\sum_\alpha m_\alpha(\boldsymbol{\omega}\times\mathbf{r}_\alpha)^2 = \frac{1}{2}\sum_\alpha m_\alpha\left(\frac{\mathrm{d}\mathbf{r}_\alpha}{\mathrm{d}t}\right)_{\mathrm{f}}^2 .$$

Using the identities

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}\,, \qquad (\mathbf{a}\times\mathbf{b})^2 = \mathbf{a}^2\mathbf{b}^2 - (\mathbf{a}\cdot\mathbf{b})^2$$

Eqs. (5.10) and (5.12) can be recast in the alternative form

$$\mathbf{L}_{\mathrm{CM}} = \sum_\alpha m_\alpha\big[r_\alpha^2\,\boldsymbol{\omega} - (\boldsymbol{\omega}\cdot\mathbf{r}_\alpha)\mathbf{r}_\alpha\big]\,, \qquad T_{\mathrm{rot}} = \frac{1}{2}\sum_\alpha m_\alpha\big[\omega^2 r_\alpha^2 - (\boldsymbol{\omega}\cdot\mathbf{r}_\alpha)^2\big]\,, \tag{5.13}$$

whence it follows the important identity

$$T_{\mathrm{rot}} = \frac{1}{2}\,\boldsymbol{\omega}\cdot\mathbf{L}_{\mathrm{CM}}\,. \tag{5.14}$$

Note that in all of the previous formulas the vectors $\mathbf{L}_{\mathrm{CM}}$ and $\boldsymbol{\omega}$, and hence the rotational energy T_{rot}, are in general functions of time.

5.3 Inertia tensor

5.3.1 Definition and elementary properties

The expressions obtained in the previous section for the angular momentum with respect to the CM and the rotational energy of a rigid body can be greatly simplified with the help of the so-called *inertia tensor*. Since the rotational energy is expressed in terms of $\mathbf{L}_{\mathrm{CM}}$ through Eq. (5.14), we can restrict ourselves to the angular momentum. The key observation is that Eq. (5.13) clearly indicates that, although in general $\mathbf{L}_{\mathrm{CM}}$ is *not* parallel to the angular velocity $\boldsymbol{\omega}$, it is a *linear function* thereof. In other words, we can write

$$\mathbf{L}_{\mathrm{CM}} = I\boldsymbol{\omega}\,, \tag{5.15}$$

where $I:\mathbb{R}^3\to\mathbb{R}^3$ is a linear map, which can be represented by a 3×3 matrix whose entries we shall now compute. To this end it suffices to note that, if $x_{\alpha i}$ ($i=1,2,3$) denotes the i-th component of the vector $\mathbf{r}_\alpha$, the i-th component of $\mathbf{L}_{\mathrm{CM}}$ (in the same basis) is given by

$$\begin{aligned} L_{\mathrm{CM},i} &= \omega_i\sum_\alpha m_\alpha r_\alpha^2 - \sum_\alpha m_\alpha x_{\alpha i}\sum_j \omega_j x_{\alpha j} = \omega_i\sum_\alpha m_\alpha r_\alpha^2 - \sum_j\omega_j\sum_\alpha m_\alpha x_{\alpha i}x_{\alpha j} \\ &\sum_j\omega_j\delta_{ij}\sum_\alpha m_\alpha r_\alpha^2 - \sum_j\omega_j\sum_\alpha m_\alpha x_{\alpha i}x_{\alpha j} = \sum_j\omega_j\sum_\alpha m_\alpha(\delta_{ij}r_\alpha^2 - x_{\alpha i}x_{\alpha j})\,. \end{aligned}$$

We thus have

$$L_{\mathrm{CM},i} = \sum_j I_{ij}\omega_j\,, \tag{5.16a}$$

where the matrix element I_{ij} is given by

$$I_{ij} = \sum_\alpha m_\alpha(\delta_{ij}r_\alpha^2 - x_{\alpha i}x_{\alpha j})\,. \tag{5.16b}$$

The linear map I with matrix elements given by Eq. (5.16b) is known as the rigid body's **inertia tensor**[4]. It is important to note that, although both $\mathbf{L}_{\rm CM}$ and $\boldsymbol{\omega}$ in general depend on t, the matrix elements (5.16b) of the inertia tensor are *constant*, since the Cartesian coordinates $x_{\alpha i}$ $(i = 1, 2, 3)$ of the body's particles in a frame of body axes do not depend on time. In other words:

The inertia tensor is a *constant matrix* characteristic of the rigid body, depending only on the initial choice of the body axes.

- From Eq. (5.16b) it immediately follows that the inertia tensor is *symmetric*:

$$\boxed{I_{ij} = I_{ji}\,, \qquad i, j = 1, 2, 3\,.}$$

The diagonal matrix elements of the inertia tensor are given by

$$I_{ii} = \sum_\alpha m_\alpha (x_{\alpha j}^2 + x_{\alpha k}^2)\,, \qquad i = 1, 2, 3\,,$$

with (i, j, k) different from each other. In other words,

$$\boxed{I_{ii} = \sum_\alpha m_\alpha d_{\alpha i}^2\,,} \tag{5.17}$$

where $d_{\alpha i}$ is the distance of the α-th particular to the i-th axis. Hence the matrix element I_{ii} is the so-called **moment of inertia** of the body with respect to the axis $\mathbf{e}_i$. Likewise, the off-diagonal matrix elements of I

$$\boxed{I_{ij} = -\sum_\alpha m_\alpha x_{\alpha i} x_{\alpha j}\,, \qquad 1 \leqslant i \neq j \leqslant 3\,,}$$

are the negatives of the body's **products of inertia**. For a continuous rigid body Ω with mass density $\rho(\mathbf{r})$, the latter expressions must be replaced by their obvious continuous analogues

$$\boxed{I_{ij} = \int_\Omega \rho(\mathbf{r})\,(\delta_{ij} r^2 - x_i x_j)\,\mathrm{d}^3\mathbf{r}\,, \qquad i, j = 1, 2, 3\,,} \tag{5.18}$$

or, in more detail,

$$I_{ii} = \int_\Omega \rho(\mathbf{r})\,(x_j^2 + x_k^2)\,\mathrm{d}^3\mathbf{r}\,, \qquad i = 1, 2, 3\,,$$

(with (i, j, k) different from each other) and

$$I_{ij} = -\int_\Omega \rho(\mathbf{r}) x_i x_j\,\mathrm{d}^3\mathbf{r}\,, \qquad 1 \leqslant i \neq j \leqslant 3\,.$$

Analogous expressions are obtained for continuous body whose mass is distributed on a surface or along a curve replacing the volume element $\mathrm{d}^3\mathbf{r}$ with the surface element $\mathrm{d}S$ or the line element $\mathrm{d}s$.

- From the identity (5.14) it follows that the body's rotational energy can be expressed in terms of its angular velocity $\boldsymbol{\omega}$ and the inertia tensor I through the formula

$$\boxed{T_{\rm rot} = \frac{1}{2}\,\boldsymbol{\omega}\cdot(I\boldsymbol{\omega})\,.} \tag{5.19}$$

[4]The name "tensor" is due to the fact that in general a linear map is a tensor with one covariant and one contravariant indices. Note, however, that in *orthogonal* Cartesian coordinates there is no distinction between covariant and contravariant indices.

Note also that the previous expression can be written using the usual matrix notation as

$$T_{\rm rot} = \frac{1}{2}\,\omega^{\mathsf T}\mathsf{I}\,\omega\,, \tag{5.20}$$

if we interpret ω as the *column* vector $\begin{pmatrix}\omega_1\\ \omega_2\\ \omega_3\end{pmatrix}$ and I (sans serif I) denotes the real 3×3 matrix with elements I_{ij}. In other words, $T_{\rm rot}$ is a *quadratic form* in the components of $\boldsymbol{\omega}$, whose matrix elements are the matrix elements (5.16b) of the inertia tensor. Since $T_{\rm rot} \geqslant 0$ for all $\boldsymbol{\omega}$, this quadratic form —or, equivalently, the inertia tensor I— is *positive semi-definite*. In fact:

The inertia tensor is *positive definite* if and only if the body is generic.

Indeed, if I were not positive definite, by Eq. (5.12) there would exist a nonzero vector $\boldsymbol{\omega}$ such that

$$2\mathsf{T}_{\rm rot} = \sum_\alpha m_\alpha(\boldsymbol{\omega}\times\mathbf{r}_\alpha)^2 = 0\,.$$

Since all the terms in the sum are nonnegative, the latter equality is only possible if $\boldsymbol{\omega}\times\mathbf{r}_\alpha$ vanishes for all $\alpha = 1,\dots,N$, i.e., if all the particles lie on the line parallel to $\boldsymbol{\omega}$ passing through the CM.

- From Eq. (5.19) it follows that a rigid body's rotational energy can also be expressed as

$$T_{\rm rot} = \frac{1}{2}\,\omega^2\,\mathbf{n}\cdot I\mathbf{n}\,,$$

where $\mathbf{n} = \boldsymbol{\omega}/\omega$ is the direction of instantaneous axis of rotation of the body axes. Since $I_{ii} = \mathbf{e}_i\cdot I\mathbf{e}_i$, by Eq. (5.17) we can also write

$$\mathbf{n}\cdot I\mathbf{n} = \sum_\alpha m_\alpha d_\alpha(\mathbf{n})^2 =: I_{\mathbf{n}},$$

where $d_\alpha(\mathbf{n})$ and $I_{\mathbf{n}}$ respectively denote the distance of particle α to the line through the CM parallel to the vector $\mathbf{n}$ and the body's moment of inertia with respect to the latter axis. It follows that the rotational energy $T_{\rm rot}$ can be expressed in terms of $I_{\mathbf{n}}$ as

$$T_{\rm rot} = \frac{1}{2}\,I_{\mathbf{n}}\,\omega^2\,.$$

5.3.2 Steiner's theorem

We shall next see how the inertia tensor changes when we compute it with respect to a point P *fixed in the body* that does not necessarily coincide with the center of mass C. If we denote by $\tilde{\mathbf{r}}_\alpha$ the position vector of the α-th particle with respect to the point P, the inertia tensor I_P with respect to the point P is by definition

$$(I_P)_{ij} = \sum_\alpha m_\alpha\big(\delta_{ij}\tilde{\mathbf{r}}_\alpha^2 - \tilde{x}_{\alpha i}\tilde{x}_{\alpha j}\big)\,. \tag{5.21}$$

Taking into account that

$$\tilde{\mathbf{r}}_\alpha = \mathbf{r}_\alpha - \mathbf{a}\,, \qquad \mathbf{a} := \overrightarrow{CP}\,,$$

we obtain

$$(I_P)_{ij} = \sum_\alpha m_\alpha[\delta_{ij}(\mathbf{r}_\alpha - \mathbf{a})^2 - (x_{\alpha i} - a_i)(x_{\alpha j} - a_j)] = I_{ij} + M(\mathbf{a}^2\delta_{ij} - a_i a_j) \\ - 2\delta_{ij}\mathbf{a}\cdot\sum_\alpha m_\alpha \mathbf{r}_\alpha + a_i\sum_\alpha m_\alpha x_{\alpha j} + a_j\sum_\alpha m_\alpha x_{\alpha i}\,.$$

The last three terms vanish on account of the identity (5.7), so that we finally have

$$(I_P)_{ij} = I_{ij} + M(\mathbf{a}^2\delta_{ij} - a_i a_j)\,. \tag{5.22}$$

The latter formula is known as *Steiner's theorem.*

It often happens that there is a point P *fixed in the body* which is *also fixed in an inertial frame*; for example, if the body is rotating around a *fixed* axis we can choose as P any point on the axis of rotation. When this is the case it is possible —and, in fact, usually advantageous— to *take* P *as the origin* O' *of the inertial frame.* It then follows that the vector $\mathbf{R} = \overrightarrow{O'O} = \overrightarrow{PC}$ is *constant in the frame of body axes*, since its endpoints are both fixed in the body. Thus in this case $\dot{\mathbf{R}} = 0$, and the velocity of the CM in the inertial frame can be simply expressed as

$$\mathbf{V} = \boldsymbol{\omega}\times\mathbf{R}\,,$$

so that

$$\mathbf{v}'_\alpha = \boldsymbol{\omega}\times\mathbf{R} + \boldsymbol{\omega}\times\mathbf{r}_\alpha = \boldsymbol{\omega}\times(\mathbf{r}_\alpha + \mathbf{R}) = \boldsymbol{\omega}\times\mathbf{r}'_\alpha\,.$$

The angular momentum $\mathbf{L}$ with respect to $O' = P$ is given by

$$\mathbf{L} = \sum_\alpha m_\alpha \mathbf{r}'_\alpha \times \mathbf{v}'_\alpha = \sum_\alpha m_\alpha \mathbf{r}'_\alpha \times (\boldsymbol{\omega}\times\mathbf{r}'_\alpha)\,,$$

i.e., is obtained replacing $\mathbf{r}_\alpha$ by $\mathbf{r}'_\alpha$ in Eq. (5.10) for $\mathbf{L}_{\text{CM}}$. In other words, in this case the body's *total* angular momentum is given by

$$\mathbf{L} = I_P\boldsymbol{\omega}\,. \tag{5.23}$$

Note that I_P is still a *constant* matrix characteristic of the rigid body considered, since in this case $\mathbf{r}'_\alpha = \mathbf{R} + \mathbf{r}_\alpha$ is a constant vector in the body frame. Likewise

$$T = \frac{1}{2}\sum_\alpha m_\alpha \mathbf{v}'^2_\alpha = \frac{1}{2}\sum_\alpha m_\alpha(\boldsymbol{\omega}\times\mathbf{r}'_\alpha)^2 = \frac{1}{2}\,\boldsymbol{\omega}\cdot I_P\boldsymbol{\omega}\,. \tag{5.24}$$

The inertia tensor I_P can be computed from I applying Steiner's theorem (5.22), taking into account that in this case $\mathbf{a} = \overrightarrow{CP} = -\overrightarrow{PC} = -\overrightarrow{O'O} = -\mathbf{R}$:

$$(I_P)_{ij} = I_{ij} + M(\mathbf{R}^2\delta_{ij} - X_i X_j)\,,$$

where X_i ($i = 1, 2, 3$) are the components of the vector $\mathbf{R}$. Note that the last term in Eq. (5.22) is nothing but the inertia tensor with respect to O' of a particle of mass M located at the CM.

Note: from now on we shall usually *omit the subindex* when the point with respect to which the inertia tensor is computed is clear from the context.

5.3.3 Principal axes of inertia

Let us next see how the components of the inertia tensor (5.16b) (with respect to the CM or, more generally, to any point P *fixed in the body*) change when we perform a *constant* rotation of the frame of body axes. More precisely, let

$$\boxed{\tilde{\mathbf{e}}_i = \sum_j a_{ji}\mathbf{e}_j\,, \qquad i = 1,2,3\,,} \tag{5.25}$$

be a second positively oriented frame fixed in the body. Then the *change of basis matrix*

$$\boxed{A := (a_{ij})_{1\leqslant i,j\leqslant 3}}$$

is a *constant* proper orthogonal matrix (i.e., $A \in \mathrm{SO}(3)$ is time independent). As is well known, the coordinates (or, in general, the components of any vector) in both frames are related by the dual equation

$$\boxed{x_i = \sum_j a_{ij}\tilde{x}_j\,;}$$

indeed,

$$\sum_j \tilde{x}_j\tilde{\mathbf{e}}_j = \sum_j \tilde{x}_j \sum_i a_{ij}\mathbf{e}_i = \sum_{i,j} a_{ij}\tilde{x}_j\mathbf{e}_i = \sum_i x_i\mathbf{e}_i \quad \Longrightarrow \quad x_i = \sum_j a_{ij}\tilde{x}_j\,.$$

Denoting by x (sans serif x) the column vector whose components are the coordinates x_i, and similarly $\tilde{\mathsf{x}} = (x_1\, x_2\, x_3)^{\mathsf{T}}$, we can rewrite the previous relation in matrix form as

$$\boxed{\mathsf{x} = A\,\tilde{\mathsf{x}}\,.}$$

Using this notation, and denoting by I the matrix of the inertia tensor with respect to the set of axes $\mathbf{e}_i$, we obtain

$$\mathsf{L}_{\mathrm{CM}} = \mathsf{I}\,\omega = \mathsf{I}\,A\,\tilde{\omega} = A\,\tilde{\mathsf{L}}_{\mathrm{CM}} \quad \Longrightarrow \quad \tilde{\mathsf{L}}_{\mathrm{CM}} = A^{-1}\mathsf{I}\,A\,\tilde{\omega} =: \tilde{\mathsf{I}}\,\tilde{\omega}\,.$$

Hence the matrix of the inertia tensor in the new set of axis is given by

$$\boxed{\tilde{\mathsf{I}} = A^{-1}\mathsf{I}\,A = A^{\mathsf{T}}\mathsf{I}\,A\,,}$$

since A is orthogonal. Note that in the new set of body axes the body's rotational energy can be expressed as

$$T_{\mathrm{rot}} = \frac{1}{2}\,\omega^{\mathsf{T}}\mathsf{I}\,\omega = \frac{1}{2}\,\tilde{\omega}^{\mathsf{T}}A^{\mathsf{T}}\mathsf{I}\,A\tilde{\omega} = \frac{1}{2}\,\tilde{\omega}^{\mathsf{T}}\tilde{\mathsf{I}}\,\tilde{\omega}\,,$$

which agrees with the expression we just derived for the matrix $\tilde{\mathsf{I}}$.

It is well known that *a real symmetric matrix can be diagonalized by means of a proper orthogonal transformation*[5]. In other words, it is always possible to find a matrix $A \in \mathrm{SO}(3)$ such that in the new set of body axes (5.25) we have

$$\tilde{I}_{ij} = \delta_{ij}I_i\,, \qquad 1 \leqslant i,j \leqslant 3\,,$$

where I_1, I_2 and I_3 are the three *eigenvalues* of the inertia tensor I. Note that A is a *constant* (i.e., time-independent) matrix, since the matrix elements of the inertia tensor are also constant. If the

[5]This is essentially due to the following facts: i) every real symmetric matrix is diagonalizable; ii) its eigenvalues are all real, and iii) two eigenvectors of a real symmetric matrix corresponding to different eigenvalues are orthogonal. From these three facts it easily follows that there exists a (positively oriented) orthonormal basis of eigenvectors of any real orthogonal matrix.

vectors $\tilde{\mathbf{e}}_i$ are defined by (5.25), where $A \in \mathrm{SO}(3)$ is the proper orthogonal matrix diagonalizing I, then

$$I\tilde{\mathbf{e}}_i = I_i\tilde{\mathbf{e}}_i\,, \qquad i = 1,2,3\,. \tag{5.26}$$

In other words, the vector $\tilde{\mathbf{e}}_i$, whose components with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are the i-th column of the change of basis matrix A, is an *eigenvector* of the inertia tensor I with eigenvalue I_i. Since A is a proper orthogonal matrix, the vectors $\tilde{\mathbf{e}}_i$ ($i = 1,2,3$) are a positively oriented orthonormal basis of $\mathbb{R}^3$. These vectors are *fixed in the body*, since their components with respect to the original set of body axes $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are the matrix elements of the *constant* matrix A. Hence:

It is always possible to find a set of body axes whose unit vectors $\tilde{\mathbf{e}}_i$ are all *eigenvectors* of the linear map I. In this set of body axes the inertia tensor is represented by the diagonal matrix

$$I = \begin{pmatrix} I_1 & & \\ & I_2 & \\ & & I_3 \end{pmatrix},$$

where I_i is the eigenvalue corresponding to the eigenvector $\tilde{\mathbf{e}}_i$.

The vectors $\tilde{\mathbf{e}}_i$ satisfying the relations (5.26) are known as the body's **principal axes of inertia**, and their corresponding eigenvalues I_i are called its **principal moments of inertia**. As is well known, the eigenvalues of the matrix (I_{ij}), i.e., the principal moments of inertia, are the roots of the **secular equation**

$$\det(I_{ij} - \lambda\delta_{ij}) = 0\,.$$

Note that the principal axes of inertia, i.e., the directions of the eigenvectors of the matrix (I_{ij}), are not uniquely determined (up to a sign) unless all the eigenvalues of the inertia tensor are distinct (i.e., they are *simple* roots of the secular equation).

- If the set of body axes $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is a set of principal axes of inertia Eqs. (5.15) and (5.19) reduce to

$$\mathbf{L}_{\mathrm{CM}} = \sum_i I_i\omega_i\mathbf{e}_i\,, \qquad T_{\mathrm{rot}} = \frac{1}{2}\sum_i I_i\omega_i^2\,. \tag{5.27}$$

In particular, if the body rotates around its i-th principal axis of inertia we have

$$\mathbf{L}_{\mathrm{CM}} = I_i\boldsymbol{\omega}\,, \qquad T_{\mathrm{rot}} = \frac{1}{2}I_i\omega^2\,.$$

If the origin of the fixed frame is also a point fixed in the body, expressions analogous to the previous ones are valid for $\mathbf{L}$ and T replacing I by the inertia tensor $I_{O'}$ with respect to the point O'.

- Rigid bodies can be classified into the following three categories, depending on the multiplicity of the eigenvalues of their inertia tensor:

  1. *Asymmetric tops:* $I_i \neq I_j$ for all $i \neq j$
  2. *Axially symmetric tops:* $I_i = I_j \neq I_k$ (with (i,j,k) distinct)
  3. *Spherically symmetric tops*: $I_1 = I_2 = I_3$.

5.3.4 Symmetries

We shall next examine how the *symmetries* of a rigid body Ω of mass density ρ result in simplifications of its inertia tensor.

1) *If Ω and ρ are invariant under the reflection $x_i \mapsto -x_i$, then*

$$\boxed{I_{ij} = 0\,, \qquad \forall\, j \neq i\,.}$$

Indeed, suppose that (for instance) Ω is invariant under reflection of the x_1 coordinate and $\rho(-x_1, x_2, x_3) = \rho(x_1, x_2, x_3)$. Performing the change of variables

$$x_1 = -x_1'\,, \quad x_2 = x_2'\,, \quad x_3 = x_3'$$

in the integral for I_{1j} (with $j \neq 1$), which by hypothesis maps Ω to itself, we obtain

$$-I_{1j} = \int_\Omega \rho(\mathbf{r}) x_1 x_j \,\mathrm{d}^3\mathbf{r} = -\int_{\Omega'} \rho(-x_1', x_2', x_3') x_1' x_j' \,\mathrm{d}^3\mathbf{r}' = -\int_\Omega \rho(x_1', x_2', x_3') x_1' x_j' \,\mathrm{d}^3\mathbf{r}' = I_{1j}$$
$$\implies \quad I_{1j} = 0\,, \qquad j \neq 1\,.$$

2) *If Ω and ρ are invariant under the exchange $x_i \mapsto x_j$, then*

$$\boxed{I_{ii} = I_{jj}\,, \qquad I_{ik} = I_{jk} \quad (k \neq i, j)\,.}$$

Indeed, if (for instance) Ω is invariant under $x_1 \mapsto x_2$ and $\rho(x_2, x_1, x_3) = \rho(x_1, x_2, x_3)$, performing the change of variable

$$x_1 = x_2'\,, \quad x_2 = x_1'\,, \quad x_3 = x_3'$$

in the integral for I_{11} we obtain

$$I_{11} = \int_\Omega \rho(\mathbf{r})(x_2^2 + x_3^2)\,\mathrm{d}^3\mathbf{r} = \int_{\Omega'} \rho(x_2', x_1', x_3')(x_1'^2 + x_3'^2)\,\mathrm{d}^3\mathbf{r}' = \int_\Omega \rho(x_1', x_2', x_3')(x_1'^2 + x_3'^2)\,\mathrm{d}^3\mathbf{r}' = I_{22}\,.$$

Likewise,

$$-I_{13} = \int_\Omega \rho(\mathbf{r}) x_1 x_3 \,\mathrm{d}^3\mathbf{r} = \int_{\Omega'} \rho(x_2', x_1', x_3') x_2' x_3' \,\mathrm{d}^3\mathbf{r}' = \int_\Omega \rho(x_1', x_2', x_3') x_2' x_3' \,\mathrm{d}^3\mathbf{r}' = -I_{23}\,.$$

Analogous results hold for the coordinates of the body's *center of mass*. For instance, if Ω and ρ are invariant under the reflection $x_i \mapsto -x_i$ then the i-th coordinate of the CM vanishes, since (taking, for definiteness, $i = 1$)

$$MX_1 = \int_\Omega \rho(\mathbf{r}) x_1 \,\mathrm{d}^3\mathbf{r} = -\int_{\Omega'} \rho(-x_1', x_2', x_3') x_1' \,\mathrm{d}^3\mathbf{r}' = -\int_\Omega \rho(x_1', x_2', x_3') x_1' \,\mathrm{d}^3\mathbf{r}' = -MX_1$$
$$\implies \quad X_1 = 0\,.$$

Similarly, if Ω and ρ are invariant under the permutation $x_i \mapsto x_j$ then $X_i = X_j$.

Example 5.1. Consider a *homogeneous* (i.e., $\rho = $ const.) rigid body Ω in the shape of a *solid of revolution* around a certain axis. Taking the z axis of the set of body axes in the direction of the body's axis of revolution, in cylindrical coordinates[a] (r, φ, z) the body is described by an equation of the form

$$0 \leqslant r \leqslant f(z)\,, \quad z_1 \leqslant z \leqslant z_2\,, \quad 0 \leqslant \varphi \leqslant 2\pi\,.$$

The symmetry under rotations around the z axis implies the invariance of the body under the transformations

$$x_1 \mapsto -x_1\,, \qquad x_2 \mapsto -x_2\,, \qquad x_1 \mapsto x_2\,.$$

Hence the body's center of mass is a point on the z axis, which we shall take as the origin of coordinates since we are interested in computing the inertia tensor with respect to the CM. By the above symmetries, the components of the inertia tensor satisfy

$$I_{11} = I_{22}\,, \qquad I_{ij} = 0 \quad (i \neq j)\,.$$

Hence in this case the inertia tensor is diagonal, with principal moments of inertia $I_i = I_{ii}$ given by

$$\begin{aligned} I_1 = I_2 &= \rho \int_{z_1}^{z_2} \mathrm{d}z \int_0^{f(z)} \mathrm{d}r \int_0^{2\pi} r\,\mathrm{d}\varphi \cdot (z^2 + r^2 \sin^2\varphi) \\ &= \boxed{\pi\rho \int_{z_1}^{z_2} z^2 f^2(z)\,\mathrm{d}z + \frac{\pi\rho}{4} \int_{z_1}^{z_2} f^4(z)\,\mathrm{d}z}\,, \\ I_3 &= \rho \int_{z_1}^{z_2} \mathrm{d}z \int_0^{f(z)} \mathrm{d}r \int_0^{2\pi} r\,\mathrm{d}\varphi \cdot r^2 = \boxed{\frac{\pi\rho}{2} \int_{z_1}^{z_2} f^4(z)\,\mathrm{d}z}\,. \end{aligned}$$

The mass density ρ can be expressed in terms of the body's total mass M through the formula

$$\rho = \frac{M}{V} = \frac{M}{\pi \displaystyle\int_{z_1}^{z_2} f^2(z)\,\mathrm{d}z}\,.$$

Note that, in general, $I_1 = I_2 \neq I_3$; more precisely, we have

$$I_1 = I_2 = \frac{1}{2} I_3 + \pi\rho \int_{z_1}^{z_2} z^2 f^2(z)\,\mathrm{d}z\,.$$

Thus a solid of revolution is in general an axially symmetric top. Note also that in this case the axis of revolution is a principal axis of inertia, as is any axis perpendicular to it.

For instance, in the case of a *cylinder* of radius a and height h we can take $f(z) = a$, $z_1 = -h/2$ and $z_2 = h/2$, since by symmetry the CM of a cylinder is equidistant from its bases. Hence

$$\boxed{\begin{aligned} I_3 &= \frac{\pi}{2}\rho a^4 h = \frac{1}{2} M a^2\,, \\ I_1 = I_2 &= \frac{1}{2} I_3 + \pi\rho a^2 \int_{-h/2}^{h/2} z^2\,\mathrm{d}z = \frac{1}{4} M a^2 + 2\pi\rho a^2 \int_0^{h/2} z^2\,\mathrm{d}z = \frac{1}{4} M a^2 + \frac{1}{12}\pi\rho a^2 h^3 \\ &= \frac{1}{4} M\left(a^2 + \frac{h^2}{3}\right). \end{aligned}}$$

In particular, a cylinder is a spherically symmetric top if and only if $h = \sqrt{3}\,a$.

[a]We are temporarily denoting the distance to the z axis as r instead of the usual notation ρ to avoid confusion with the mass density.

5.4 Equations of motion of a rigid body

5.4.1 Equations of motion of a rigid body in an inertial frame

Since a rigid body has (in general) 6 degrees of freedom, it should be expected that its motion is determined by 6 differential equations. The first three of these equations are obviously the equations of motion of the body's CM, which, as we saw in Chapter 1, read

$$\boxed{M\left(\frac{\mathrm{d}^2\mathbf{R}}{\mathrm{d}t^2}\right)_{\mathrm{f}} = \mathbf{F}\,.} \tag{5.28}$$

In the RHS

$$\mathbf{F} = \sum_\alpha \mathbf{F}_\alpha$$

denotes the sum of the *external* forces acting on the particles making up the body (remember that, by Newton's third law, the sum of the internal forces vanishes). The remaining three differential equations can be taken as the equations of motion of the body's angular momentum with respect to the origin O' of the fixed frame, namely

$$\boxed{\left(\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}\right)_{\mathrm{f}} = \mathbf{N}\,.} \tag{5.29}$$

Here

$$\mathbf{N} = \sum_\alpha \mathbf{r}'_\alpha \times \mathbf{F}_\alpha\,,$$

denotes the total torque (with respect to O') of the *external* forces acting on the body. Indeed, as we saw in Chapter 1, if we assume that Newton's third law holds in its stronger sense the torque of the internal forces vanishes.

In equation (5.29) both the angular momentum and the total torque of the external forces are computed with respect to the origin O' of the fixed frame. In fact, *Eq.* (5.29) *still holds if we replace* $\mathbf{L}$ *by* $\mathbf{L}_{\mathrm{CM}}$ *and* $\mathbf{N}$ *by the total torque of the external forces with respect to the CM*, given by

$$\boxed{\mathbf{N}_{\mathrm{CM}} = \sum_\alpha \mathbf{r}_\alpha \times \mathbf{F}_\alpha\,.} \tag{5.30}$$

Indeed, from the relation $\mathbf{L} = M\mathbf{R}\times\mathbf{V} + \mathbf{L}_{\mathrm{CM}}$ and the CM's equation of motion it follows that

$$\left(\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}\right)_{\mathrm{f}} = \left(\frac{\mathrm{d}\mathbf{L}_{\mathrm{CM}}}{\mathrm{d}t}\right)_{\mathrm{f}} + \mathbf{R}\times\mathbf{F} = \mathbf{N} = \sum_\alpha (\mathbf{R}+\mathbf{r}_\alpha)\times\mathbf{F}_\alpha = \mathbf{R}\times\mathbf{F} + \mathbf{N}_{\mathrm{CM}}\,,$$

and thus

$$\boxed{\left(\frac{\mathrm{d}\mathbf{L}_{\mathrm{CM}}}{\mathrm{d}t}\right)_{\mathrm{f}} = \mathbf{N}_{\mathrm{CM}}\,.} \tag{5.31}$$

In general, if the total force acting on the body *vanishes* the torque $\mathbf{N}$ is independent of the point with respect to which it is computed.

Indeed, if $\mathbf{F} := \sum_\alpha \mathbf{F}_\alpha = 0$ and $\mathbf{a}$ is a fixed vector we have

$$\sum_\alpha (\mathbf{r}'_\alpha + \mathbf{a})\times\mathbf{F}_\alpha = \sum_\alpha \mathbf{r}'_\alpha\times\mathbf{F}_\alpha + \mathbf{a}\times\mathbf{F} = \sum_\alpha \mathbf{r}'_\alpha\times\mathbf{F}_\alpha\,.$$

Exercise. Show that the necessary and sufficient condition for a generic rigid body to be at equilibrium in an inertial frame is that $\mathbf{F} = \mathbf{N} = 0$.

Solution. In general, the body is at equilibrium in an inertial frame —i.e., $\mathbf{v}'_\alpha = 0$ for all t and for all α— if and only of $\mathbf{V} = \boldsymbol{\omega} = 0$, since[a]

$$\mathbf{v}'_\alpha = \mathbf{V} + \boldsymbol{\omega}\times\mathbf{r}_\alpha\,.$$

Suppose, to begin with, that the body is at equilibrium. Then $\mathbf{V} = \boldsymbol{\omega} = 0$, and therefore $\mathbf{L} = M\mathbf{R}\times\mathbf{V} + I\boldsymbol{\omega} = 0$. Substituting into the equations of motion (5.28)-(5.29) we immediately obtain $\mathbf{F} = \mathbf{N} = 0$.

Conversely, suppose that $\mathbf{F} = \mathbf{N} = 0$ *and that initially the body is at rest in some inertial frame.* We shall then show that the body is at equilibrium in that frame, i.e., that $\mathbf{V}(t) = \boldsymbol{\omega}(t) = 0$ for all t. Indeed, by hypothesis the body's particles are all instantaneously at rest in an inertial frame, i.e., $\mathbf{V}(0) = \boldsymbol{\omega}(0) = 0$. From $\mathbf{F} = 0$ and the equation of motion for $\mathbf{R}$ we deduce that $\mathbf{V}$ is constant *in the fixed frame*. Since initially $\mathbf{V}(0) = 0$, we must therefore have $\mathbf{V}(t) = 0$ for all t. Moreover, $\mathbf{F} = 0$ implies that $\mathbf{N}_{\mathrm{CM}} = \mathbf{N} = 0$, and therefore

$$\left(\frac{\mathrm{d}\mathbf{L}_{\mathrm{CM}}}{\mathrm{d}t}\right)_{\mathrm{f}} = \mathbf{N}_{\mathrm{CM}} = 0 \quad \Longrightarrow \quad \mathbf{L}_{\mathrm{CM}} = \mathbf{L}_{\mathrm{CM}}(0) = I\boldsymbol{\omega}(0) = 0 \;\; \forall t\,.$$

If the body is generic its inertia tensor I is invertible, and hence

$$I\boldsymbol{\omega} = 0 \quad \Longrightarrow \quad \boldsymbol{\omega} = 0 \quad \forall t\,.$$

We have thus shown that $\mathbf{V} = \boldsymbol{\omega} = 0$, so that by Eq. (5.4) the body is indeed at rest in the inertial frame considered.

[a]Indeed,

$$\mathbf{V} + \boldsymbol{\omega} \times \mathbf{r}_\alpha = 0\,, \;\; \forall \alpha \quad \Longrightarrow \quad \boldsymbol{\omega} \times (\mathbf{r}_\alpha - \mathbf{r}_\beta) = 0\,, \;\; \forall \alpha \neq \beta\,.$$

It follows that $\boldsymbol{\omega} = 0$, which implies that $\mathbf{V} = 0$. Indeed, if $\boldsymbol{\omega} \neq 0$ then $\mathbf{r}_\alpha - \mathbf{r}_\beta$ would be parallel to $\boldsymbol{\omega}$ for fixed β and all $\alpha \neq \beta$, and hence the body would lie on the straight line parallel to $\boldsymbol{\omega}$ passing through $\mathbf{r}_\beta$.

Remark. The condition $\mathbf{V}(0) = \boldsymbol{\omega}(0) = 0$ is *essential* to guarantee that the body is at equilibrium when $\mathbf{F} = \mathbf{N} = 0$. Indeed, we shall see in Section 5.5 that when $\mathbf{F}$ and $\mathbf{N}$ vanish the body can still rotate with constant angular velocity around a principal axis of inertia passing through the CM. ■

Exercise. Show that the condition for equilibrium of a rigid body found above is equivalent to $\mathbf{F} = \mathbf{N}_{\mathrm{CM}} = 0$.

Solution. If $\mathbf{F} = 0$ the torque of the external forces does not depend on the point with respect to which it is taken, so that in this case $\mathbf{N} = \mathbf{N}_{\mathrm{CM}}$.

A particular case which often occurs in practice arises when the external forces $\mathbf{F}_\alpha$ acting on the rigid body are due to a *constant external field* $\mathbf{f}$, to which the particles couple through a "charge" λ. In this case

$$\mathbf{F}_\alpha = \lambda_\alpha \mathbf{f}\,, \qquad \alpha = 1, \dots, N\,, \tag{5.32}$$

where $\mathbf{f}$ is independent of α and λ_α is the charge of particle α. For instance, *Earth's gravitational force* is of this form if the body's extension is small compared to its distance to Earth's center, so that Earth's gravitational field is approximately uniform inside the body (in this case $\lambda_\alpha = m_\alpha$, $\mathbf{f} = \mathbf{g}$). The same is true for the electric force due to a *uniform electric field* (in this case $\lambda_\alpha = e_\alpha$ is the electric charge, $\mathbf{f} = \mathbf{E}$). If the external forces are of the form (5.32) we have

$$\mathbf{F} = \sum_\alpha \mathbf{F}_\alpha = \mathbf{f} \sum_\alpha \lambda_\alpha = \Lambda \mathbf{f}\,, \qquad \mathbf{N} = \Big(\sum_\alpha \lambda_\alpha \mathbf{r}'_\alpha\Big) \times \mathbf{f}\,,$$

$\Lambda := \sum_\alpha \lambda_\alpha$ being the body's total charge. If (as is the case with the gravitational force) $\Lambda \neq 0$, we define the body's *center of charge* by the equation

$$\mathbf{X} = \frac{1}{\Lambda} \sum_\alpha \lambda_\alpha \mathbf{r}'_\alpha = \frac{1}{\Lambda} \sum_\alpha \lambda_\alpha (\mathbf{R} + \mathbf{r}_\alpha) = \mathbf{R} + \frac{1}{\Lambda} \sum_\alpha \lambda_\alpha \mathbf{r}_\alpha\,. \tag{5.33}$$

Note that the point $\mathbf{X}$ is *fixed* in the body, since its position vector with respect to the CM (also fixed in the body)

$$\mathbf{X} - \mathbf{R} = \frac{1}{\Lambda} \sum_\alpha \lambda_\alpha \mathbf{r}_\alpha =: \mathbf{X}_{\mathrm{CM}}$$

is a constant vector in the frame of body axes. In particular, for the gravitational field $\mathbf{X} = \mathbf{R}$ and $\mathbf{X}_{\mathrm{CM}} = 0$. In terms of the vector $\mathbf{X}$, the torque $\mathbf{N}$ can be concisely expressed as

$$\boxed{\mathbf{N} = \mathbf{X} \times \mathbf{F}\,.} \tag{5.34}$$

In other words:

If the total charge doesn't vanish, the total torque of the external forces coincides with the torque of the total external force applied at the *center of charge*. In particular, when computing the torque of the gravitational forces acting on a rigid body we can always assume that they are applied at its *center of mass*.

Likewise, if $\Lambda \neq 0$ the torque of the external forces with respect to the CM is given by

$$\mathbf{N}_{\text{CM}} = \Big(\sum_\alpha \lambda_\alpha \mathbf{r}_\alpha\Big) \times \mathbf{f} = \mathbf{X}_{\text{CM}} \times \mathbf{F}\,.$$

In the case of the gravitational force $\mathbf{X}_{\text{CM}} = 0$, and hence $\mathbf{N}_{\text{CM}} = 0$. In other words:

The torque with respect to the CM of the gravitational forces $\mathbf{F}_\alpha = m_\alpha \mathbf{g}$ *vanishes* (assuming again that the body's size is negligible compared to its distance to Earth's center).

Note, finally, that the forces (5.32) are clearly *conservative*, with potential

$$\boxed{U = -\mathbf{f} \cdot \sum_\alpha \lambda_\alpha \mathbf{r}'_\alpha = -\Lambda \mathbf{f} \cdot \mathbf{X} = -\mathbf{F} \cdot \mathbf{X}\,,}$$

where the last two equalities are valid only when $\Lambda \neq 0$. Thus (assuming again that $\Lambda \neq 0$) *when computing the potential energy we can assume in this case that the total constant external force* $\mathbf{F}$ *is applied at the point* $\mathbf{X}$. In particular, the potential energy of a rigid body due to Earth's gravitational field is simply

$$\boxed{U = -M\mathbf{g} \cdot \mathbf{R}\,.}$$

Note. If the total charge Λ vanishes the center of charge cannot be defined by Eq. (5.33). In this case the total torque of the external forces is clearly independent of the point with respect to which it is taken, since the total external force $\mathbf{F} = \Lambda \mathbf{f}$ vanishes.

5.4.2 Euler's equations

Since the relation between angular momentum and angular velocity is particularly simple in a frame of principal axes of inertia fixed in the body, it is convenient to formulate the equation of motion of the angular momentum in such a frame. To this end, we shall assume that the point P with respect to which $\mathbf{L}$, I and $\mathbf{N}$ are computed is either the CM or (when it exists) a point simultaneously *fixed* in the body and in an inertial frame, that we shall take as the origin O' of the latter frame. To cover both situations we shall use the more descriptive notation $\mathbf{L}_P$, $\mathbf{N}_P$, and I_P to respectively denote the angular momentum, torque, and inertia tensor taking as origin the point P (in our old notation, $\mathbf{L}_C \equiv \mathbf{L}_{\text{CM}}$, $\mathbf{L}_{O'} \equiv \mathbf{L}$, $\mathbf{N}_C \equiv \mathbf{N}_{\text{CM}}$, $\mathbf{N}_{O'} \equiv \mathbf{N}$, and $I_C \equiv I$). With this notation we can write

$$\boxed{\left(\frac{\mathrm{d}\mathbf{L}_P}{\mathrm{d}t}\right)_{\mathrm{f}} = \mathbf{N}_P\,, \qquad \mathbf{L}_P = I_P\boldsymbol{\omega}\,, \qquad E := \frac{1}{2}\,\boldsymbol{\omega} \cdot \mathbf{L}_P = \begin{cases} T_{\text{rot}}\,, & P = C \\ T\,, & P = O'\,, \end{cases}}$$

where I_P is *constant* in the frame of body axes (since P is by hypothesis fixed in the body). *We shall usually drop the subindex, and simply write*

$$\boxed{\left(\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}\right)_{\mathrm{f}} = \mathbf{N}\,, \qquad \mathbf{L} = I\boldsymbol{\omega}, \qquad \frac{1}{2}\,\boldsymbol{\omega} \cdot \mathbf{L} = E} \tag{5.35}$$

to deal with both cases at the same time. From Eq. (4.11) we then obtain the relation

$$\dot{\mathbf{L}} + \boldsymbol{\omega} \times \mathbf{L} = \mathbf{N}\,,$$

where as usual the dot denotes time derivative with respect to the frame of body axes. Using the relation between $\mathbf{L}$ and $\boldsymbol{\omega}$, and taking into account that I_{ij} is constant in a frame of body axes, we immediately obtain

$$I\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times (I\boldsymbol{\omega}) = \mathbf{N}\,. \tag{5.36}$$

If the body axes are principal axes of inertia the i-th component of this vector equation is simply

$$I_i\dot{\omega}_i + \omega_j(I_k\omega_k) - \omega_k(I_j\omega_j) = N_i\,, \qquad i = 1,2,3\,,$$

or equivalently

$$I_i\dot{\omega}_i - (I_j - I_k)\omega_j\omega_k = N_i\,, \qquad i = 1,2,3\,, \tag{5.37a}$$

where

$$(i,j,k) = \textit{cyclic}\text{ permutation of }(1,2,3)\,. \tag{5.37b}$$

Equations (5.37), i.e., the system

$$\begin{aligned} I_1\dot{\omega}_1 - (I_2 - I_3)\omega_2\omega_3 &= N_1\,,\\ I_2\dot{\omega}_2 - (I_3 - I_1)\omega_1\omega_3 &= N_2\,,\\ I_3\dot{\omega}_3 - (I_1 - I_2)\omega_1\omega_2 &= N_3\,, \end{aligned} \tag{5.38}$$

are known as **Euler's equations**. We emphasize that these equations are valid if both $\mathbf{N}$ and I are computed either with respect to the CM or (when possible) to a point simultaneously fixed in the body and in the inertial frame. Moreover, the quantities ω_i and N_i appearing in Euler's equations are the components of the vectors $\boldsymbol{\omega}$ and $\mathbf{N}$ *in a frame of principal axes of inertia* (in general *not* inertial).

If the total torque of the external forces vanishes and the origin O' of the inertial frame is a point fixed in the body, $\mathbf{L}_{O'} \equiv \mathbf{L}$ and T are conserved. Similarly, if $\mathbf{N}_{\mathrm{CM}}$ vanishes then $\mathbf{L}_{\mathrm{CM}}$ and the rotational energy T_{rot} are conserved.

Proof. Using the notation in Eq. (5.35), we only need to show that

$$\mathbf{N} = 0 \quad \Longrightarrow \quad \left(\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}\right)_{\mathrm{f}} = 0, \quad \dot{E} = 0\,.$$

(Note that E is a *scalar*, so that it is not necessary to specify the frame with respect to which the time derivative is taken.) The conservation of $\mathbf{L}$ (in the fixed frame) follows immediately from its equation of motion (5.35), while the conservation of E is obtained differentiating the third relation (5.35) *in the frame of body axes* (which is correct, since E is a *scalar*). Indeed,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{\omega}\cdot I\boldsymbol{\omega}) = \frac{\mathrm{d}}{\mathrm{d}t}\sum_i I_i\omega_i^2 = 2\sum_i I_i\omega_i\dot{\omega}_i = 2\boldsymbol{\omega}\cdot I\dot{\boldsymbol{\omega}}$$

From Euler's equations in their vector form (5.36) with $\mathbf{N} = 0$ it then follows that

$$\dot{E} = \boldsymbol{\omega}\cdot I\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega}\cdot(\boldsymbol{\omega}\times(I\boldsymbol{\omega})) = 0$$

(as $\boldsymbol{\omega}\times(I\boldsymbol{\omega})$ is perpendicular to $\boldsymbol{\omega}$). ■

Remark. Since $L = |\mathbf{L}|$ and E are both *scalars*, when $\mathbf{N} = 0$ they are also constant (i.e., time-independent) *in the body frame.* ■

5.5 Inertial motion of a symmetric top

We shall study in this section the rotational motion of an *axially symmetric* top when the total torque of the external forces with respect to either the CM, or a point simultaneously fixed in the body and in an inertial frame (when such a point exists), vanishes. This will obviously happen (in both cases) if the body is *free*, that is, in the absence of external forces. More generally, as we saw at the end of Section 5.4.1, the torque $\mathbf{N}_{\mathrm{CM}}$ will vanish provided that the only external force acting on the body is Earth's gravity (assuming the body's size to be negligible compared to its distance to Earth's center). Before starting our analysis, it is convenient to prove the following fact regarding the angular velocity vector $\boldsymbol{\omega}$:

> The instantaneous angular velocity $\boldsymbol{\omega}$ of a set of axes with respect to another is *additive*.

In other words, let S_0, S_1, and S_2 be three sets of axes, and suppose that at a certain time t the axes of S_1 have angular velocity $\boldsymbol{\omega}_1$ relative to those of S_0, and that the axes of S_2 have in turn angular velocity $\boldsymbol{\omega}_2$ with respect to those of S_1. Then the angular velocity of the axes of S_2 relative to those of S_0 is

$$\boxed{\boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2 \,.} \tag{5.39}$$

Indeed, let $\{\mathbf{e}'_i\}_{1\leqslant i\leqslant 3}$, $\{\mathbf{e}_i\}_{1\leqslant i\leqslant 3}$, and $\{\mathbf{e}''_i\}_{1\leqslant i\leqslant 3}$ respectively denote the axes of the frames S_0, S_1, and S_2. By definition of angular velocity,

$$\left(\frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}t}\right)_0 = \boldsymbol{\omega}_1 \times \mathbf{e}_i \,, \qquad \left(\frac{\mathrm{d}\mathbf{e}''_i}{\mathrm{d}t}\right)_1 = \boldsymbol{\omega}_2 \times \mathbf{e}''_i \,, \qquad \left(\frac{\mathrm{d}\mathbf{e}''_i}{\mathrm{d}t}\right)_0 = \boldsymbol{\omega} \times \mathbf{e}''_i \,.$$

But then

$$\left(\frac{\mathrm{d}\mathbf{e}''_i}{\mathrm{d}t}\right)_0 = \left(\frac{\mathrm{d}\mathbf{e}''_i}{\mathrm{d}t}\right)_1 + \boldsymbol{\omega}_1 \times \mathbf{e}''_i = \boldsymbol{\omega}_2 \times \mathbf{e}''_i + \boldsymbol{\omega}_1 \times \mathbf{e}''_i = (\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) \times \mathbf{e}''_i \,,$$

whence Eq. (5.39) follows.

Note. Recall that we are denoting by $\mathbf{L}$ and I the angular momentum and the inertia tensor with respect to either O' or the CM, depending on whether $\mathbf{N} = 0$ or $\mathbf{N}_{\mathrm{CM}} = 0$ (in the first case, it is assumed that O' is simultaneously fixed in the body and the inertial system). ■

By definition, in an axially symmetric top two principal moments of inertia, which we shall take as I_1 and I_2, coincide, while the third one (i.e., I_3) differs from the other two. In other words, we have $I_1 = I_2 \neq I_3$. In this case the $\mathbf{e}_3$ axis is a principal axis of inertia (with moment of inertia I_3), as is any axis perpendicular to it (with moment of inertia I_1). In particular, from the discussion in Example 5.1 it follows that a solid of revolution around the x_3 axis is an axially symmetric top with symmetry axis along the vector $\mathbf{e}_3$. This is not, however, the most general example; for instance, a homogeneous rectangular parallelepiped with sides $l_1 = l_2 \neq l_3$ is also an axially symmetric top with $I_1 = I_2 \neq I_3$ (exercise).

Substituting $\mathbf{N} = 0$ and $I_1 = I_2$ in Euler's equations (5.38) we obtain the simpler system

$$\boxed{\begin{aligned} I_1\dot{\omega}_1 - (I_1 - I_3)\omega_2\omega_3 &= 0\,, \\ I_1\dot{\omega}_2 - (I_3 - I_1)\omega_1\omega_3 &= 0\,, \\ I_3\dot{\omega}_3 &= 0\,, \end{aligned}} \tag{5.40}$$

whence it immediately follows (assuming that $I_3 \neq 0$, i.e., that the body is not collinear) that

$$\boxed{\omega_3 = \text{const.}}$$

Calling

$$\Omega := \frac{I_3 - I_1}{I_1}\,\omega_3\,,$$

the first two equations read

$$\dot\omega_1 = -\Omega\,\omega_2\,, \qquad \dot\omega_2 = \Omega\,\omega_1$$

or, in complex notation,

$$\dot\omega_1 + \mathrm{i}\dot\omega_2 = \mathrm{i}\Omega(\omega_1 + \mathrm{i}\omega_2)\,.$$

The solution of this linear first-order equation is

$$\omega_1 + \mathrm{i}\omega_2 = (\omega_1(0) + \mathrm{i}\omega_2(0))\mathrm{e}^{\mathrm{i}\Omega t}\,. \tag{5.41}$$

From (5.41) it follows that

$$\omega_1^2 + \omega_2^2 = |\omega_1 + \mathrm{i}\omega_2|^2 = \omega_1(0)^2 + \omega_2(0)^2 =: \omega_0^2$$

is constant, and so are $\omega = \sqrt{\omega_0^2 + \omega_3^2}$ and the angle $\alpha = \arctan(\omega_0/\omega_3)$ between the vectors $\boldsymbol{\omega}$ and $\mathbf{e}_3$. We have thus shown the following:

The magnitude of the projection of $\boldsymbol{\omega}$ onto the plane perpendicular to $\mathbf{e}_3$, ω_3, ω, and the angle α between the vectors $\boldsymbol{\omega}$ and $\mathbf{e}_3$ are all constant.

In real terms, Eq. (5.41) can be written as

$$\omega_1\mathbf{e}_1 + \omega_2\mathbf{e}_2 = R_3(\Omega t)\cdot(\omega_1\mathbf{e}_1(0) + \omega_2\mathbf{e}_2(0))\,,$$

where $R_3(\varphi)$ is a rotation around the $\mathbf{e}_3$ axis by an angle φ. On the other hand,

$$\omega_3\mathbf{e}_3 = R_3(\Omega t)\cdot(\omega_3\mathbf{e}_3) = R_3(\Omega t)\cdot(\omega_3(0)\mathbf{e}_3)\,,$$

since ω_3 is constant. Adding both equations we finally obtain

$$\boldsymbol{\omega} = R_3(\Omega t)\cdot\boldsymbol{\omega}(0)\,.$$

In other words:

In the *frame of body axes* the vector $\boldsymbol{\omega}$ rotates around the $\mathbf{e}_3$ axis with constant angular velocity Ω (cf. Fig. 5.1).

The latter result could have been deduced directly from Euler's equations, since

$$\dot{\boldsymbol{\omega}} = \dot\omega_1\mathbf{e}_1 + \dot\omega_2\mathbf{e}_2 = \Omega(-\omega_2\mathbf{e}_1 + \omega_1\mathbf{e}_2) = \Omega\mathbf{e}_3\times\boldsymbol{\omega}\,.$$

Note also that the angular velocity Ω is positive for $I_3 > I_1$ ("flat" body), whereas it is negative for $I_3 < I_1$ ("tall" body); indeed, $I_3 - I_1 = \int_\Omega \rho(\mathbf{r})(x_1^2 - x_3^2)\,\mathrm{d}^3\mathbf{r}$.

Remark. A particular solution of the Euler equations consists in a *rotation around a principal axis of inertia with constant angular velocity.* Indeed, in this case $\dot{\boldsymbol{\omega}} = 0$ and $I\boldsymbol{\omega} \parallel \boldsymbol{\omega}$, so that $I\dot{\boldsymbol{\omega}} + \boldsymbol{\omega}\times(I\boldsymbol{\omega}) = 0$. (Note that in this case the axis of rotation $\boldsymbol{\omega}/\omega$ is also fixed in space, since $\dot{\boldsymbol{\omega}}$ holds both in the body and in the fixed frame). *We shall in what follows disregard these* (trivial) *solutions.* For this reason, we can assume that ω_3 and α do not vanish. Indeed, if $\alpha = 0$ then $\boldsymbol{\omega} = \omega_3\mathbf{e}_3$ is constant and hence the body rotates with constant angular velocity around its symmetry axis, which is a principal axis of inertia. Likewise, if $\omega_3 = 0$ then $\Omega = 0$,

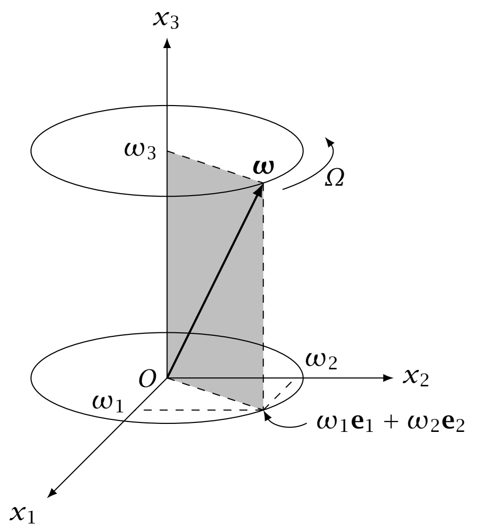


Figure 5.1. Procession of the vector $\boldsymbol{\omega}$ about the $\mathbf{e}_3$ body axis.

$\boldsymbol{\omega} = \omega_1\mathbf{e}_1 + \omega_2\mathbf{e}_2$ is constant, and thus the body rotates again around a principal axis of inertia perpendicular to the axis of symmetry with constant angular velocity. Furthermore, since we are assuming that $\omega_3 \neq 0$ we can choose the direction of the $\mathbf{e}_3$ axis so that $\omega_3 > 0$ and hence $\alpha \in (0, \pi/2)$. In terms of the angle α we can write

$$\boxed{\omega_3 = \omega\cos\alpha\,, \qquad \omega_0 = \omega\sin\alpha\,, \qquad \omega_1 + \mathrm{i}\omega_2 = \omega\sin\alpha\,\mathrm{e}^{\mathrm{i}(\Omega t+\beta)}\,,} \tag{5.42}$$

where β is the angle between the vectors $\omega_1(0)\mathbf{e}_1 + \omega_2(0)\mathbf{e}_2$ and $\mathbf{e}_1$ (which could be taken as zero choosing appropriately the initial time). ■

The previous results can be expressed in a more geometric language as follows:

Relative to the *frame of body axes,* the vector $\boldsymbol{\omega}$ moves tracing out a cone with axis $\mathbf{e}_3$ and half-angle α, with constant angular velocity Ω. This cone is called the **body cone** or, more correctly, *cone fixed in the body.*

The motion of the angular momentum $\mathbf{L}$ *relative to the body axes* is easily determined from the relation (5.27), which can be written using complex notation as

$$\boxed{L_3 = I_3\omega_3 = I_3\omega\cos\alpha = \text{const.}\,, \qquad L_1 + \mathrm{i}L_2 = I_1(\omega_1 + \mathrm{i}\omega_2) = I_1\omega\sin\alpha\,\mathrm{e}^{\mathrm{i}(\Omega t+\beta)}\,.} \tag{5.43}$$

In other words, $\mathbf{L}$ lies on the plane determined by the vectors $\mathbf{e}_3$ and $\boldsymbol{\omega}$, with L_3, $L_1^2 + L_2^2$, L and the angle θ between $\mathbf{L}$ and $\mathbf{e}_3$ all constant with respect to the frame of body axes. Note that the fact that L is constant in the latter frame follows also from the fact that it is constant in the fixed frame (since $\mathbf{N} = 0$). Similarly, that the vectors $\mathbf{e}_3$, $\boldsymbol{\omega}$, and $\mathbf{L}$ are coplanar can be proved directly remarking that

$$\mathbf{e}_3 \cdot (\boldsymbol{\omega} \times \mathbf{L}) = (I_2 - I_1)\omega_1\omega_2 = 0\,.$$

From the previous discussion it follows that:

Relative to the *frame of body axes,* the angular momentum $\mathbf{L}$ rotates with constant angular velocity Ω around the $\mathbf{e}_3$ axis.

The angle θ between $\mathbf{L}$ and $\mathbf{e}_3$ is easily computed noting that

$$\boxed{\tan\theta = \frac{\sqrt{L_1^2 + L_2^2}}{L_3} = \frac{I_1\omega_0}{I_3\omega_3} = \frac{I_1}{I_3}\tan\alpha\,;}$$

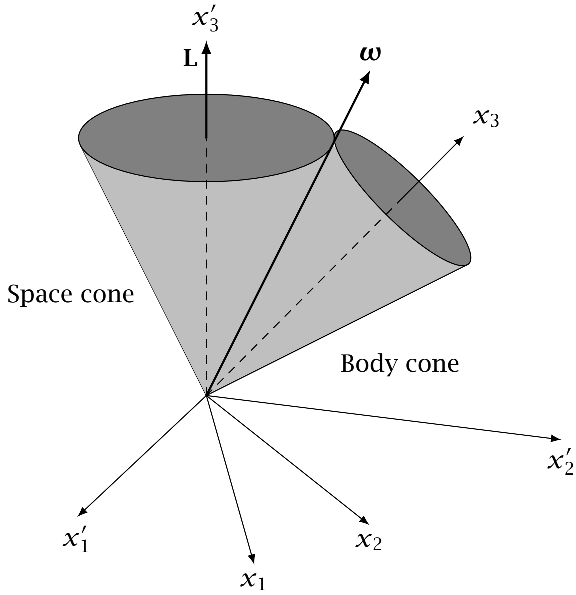


Figure 5.2. Space and body cones (in the case $I_1 > I_3$).

in particular, $\theta > \alpha$ for a "tall" body. Note also that the angle between $\mathbf{L}$ and ω is $|\theta - \alpha|$.

Relative to the *fixed frame* the vector $\mathbf{L}$ is *constant*, since by hypothesis the torque of the external forces vanishes. The direction of the latter vector, which is therefore constant with respect to the fixed frame, is known as the **invariant direction** and is usually taken as the $\mathbf{e}_3'$ axis:

$$\boxed{\mathbf{e}_3' = \frac{\mathbf{L}}{L}\,.}$$

The vectors $\boldsymbol{\omega}$ and $\mathbf{e}_3$ both rotate around $\mathbf{L}$ since, as we have just seen, the angle between the latter vectors and the angular momentum, as well as their magnitude, are constant. Moreover, $\boldsymbol{\omega}$ and $\mathbf{e}_3$ rotate with the *same angular velocity* Ω_p, since they are coplanar with $\mathbf{L}$. In other words:

Relative to the *fixed frame*, the vector $\boldsymbol{\omega}$ moves tracing out a cone of axis $\mathbf{L}$ and half-angle $|\theta - \alpha|$ with angular velocity Ω_p. This cone is known as the **space cone** (more precisely, *cone fixed in space*).

Note that the body and space cones are tangent at all times along their common generatrix parallel to the vector $\boldsymbol{\omega}$ (cf. Fig. 5.2).

To compute the angular velocity Ω_p, consider a third system of axes $\mathbf{e}_i''$ $(i = 1, 2, 3)$ with

$$\mathbf{e}_3'' = \mathbf{e}_3\,, \qquad \mathbf{e}_1'' = \frac{L_1\mathbf{e}_1 + L_2\mathbf{e}_2}{\sqrt{L_1^2 + L_2^2}}\,, \qquad \mathbf{e}_2'' = \mathbf{e}_3'' \times \mathbf{e}_1'' = \frac{\mathbf{e}_3 \times \mathbf{L}}{|\mathbf{e}_3 \times \mathbf{L}|} = \frac{\mathbf{e}_3 \times \mathbf{L}}{\sqrt{L_1^2 + L_2^2}}\,.$$

Note that by construction the vectors $\mathbf{e}_1''$ and $\mathbf{e}_3'' = \mathbf{e}_3$ span the same plane as $\mathbf{L}$ and $\mathbf{e}_3$, i.e.,

$$\operatorname{lin}\{\mathbf{e}_1'', \mathbf{e}_3''\} = \operatorname{lin}\{\mathbf{L}, \mathbf{e}_3\}\,.$$

Thus the angular velocity $\boldsymbol{\omega}'$ of the axes $\{\mathbf{e}_i''\}_{1\leqslant i\leqslant 3}$ with respect to the fixed frame $\{\mathbf{e}_i'\}_{1\leqslant i\leqslant 3}$ is equal to the angular velocity with which the plane spanned by $\mathbf{e}_3$ and $\mathbf{L}$ rotates around $\mathbf{e}_3' = \mathbf{L}/L$, which coincides with the angular velocity $\Omega_p\mathbf{e}_3' = \Omega_p\mathbf{L}/L$ of the rotation of the vector $\mathbf{e}_3$ around the invariant direction $\mathbf{e}_3'$. On the other hand, the angular velocity $\boldsymbol{\omega}''$ of the body axes $\{\mathbf{e}_i\}_{1\leqslant i\leqslant 3}$ relative to the frame $\{\mathbf{e}_i''\}_{1\leqslant i\leqslant 3}$ is equal to $-\Omega\mathbf{e}_3$, since the plane $\operatorname{lin}\{\mathbf{L}, \mathbf{e}_3\} = \operatorname{lin}\{\mathbf{e}_1'', \mathbf{e}_3''\}$ rotates

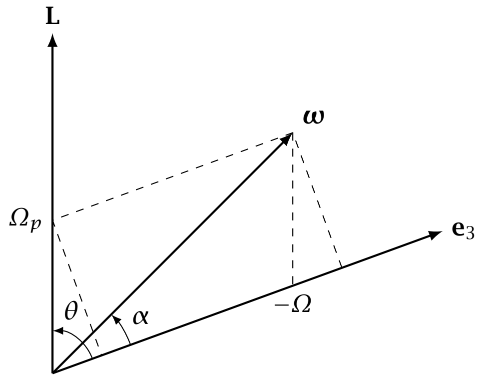


Figure 5.3. Vectors $\mathbf{L}$, $\boldsymbol{\omega}$, and $\mathbf{e}_3$ in the inertial motion of a rigid body symmetric about the $\mathbf{e}_3$ axis (in the case $I_1 > I_3$).

with angular velocity Ω around $\mathbf{e}_3$ relative to the body axes $\{\mathbf{e}_i\}_{1\leqslant i\leqslant 3}$. By the additivity of angular velocities, the angular velocity $\boldsymbol{\omega}$ of the body axes $\{\mathbf{e}_i\}_{1\leqslant i\leqslant 3}$ with respect to the fixed frame $\{\mathbf{e}'_i\}_{1\leqslant i\leqslant 3}$, is given by

$$\boxed{\boldsymbol{\omega} = \boldsymbol{\omega}' + \boldsymbol{\omega}'' = \Omega_p \frac{\mathbf{L}}{L} - \Omega \mathbf{e}_3 \,.} \tag{5.44}$$

The previous equation, together with the argument leading to its proof, shows that the body's motion can be described as the composition of a *precession* of its symmetry axis $\mathbf{e}_3$ about the invariant direction (i.e., $\mathbf{L}/|\mathbf{L}|$) with angular velocity Ω_p and a *rotation* around its symmetry axis with angular velocity $-\Omega$.

From Fig. 5.3 it follows that

$$\omega_0 = \omega \sin\alpha = \Omega_p \sin\theta \,,$$

and hence

$$\Omega_p = \omega \frac{\sin\alpha}{\sin\theta} = \omega \sin\alpha \frac{L}{\sqrt{L_1^2 + L_2^2}} = \frac{L}{I_1} \,. \tag{5.45}$$

From Eqs. (5.43) we finally obtain

$$\Omega_p = \frac{\omega}{I_1}\sqrt{I_1^2 \sin^2\alpha + I_3^2\cos^2\alpha} = \omega\sqrt{1 + \frac{I_3^2 - I_1^2}{I_1^2}\cos^2\alpha} \,.$$

In particular, $\Omega_p < \omega$ for a "tall" body, whereas $\Omega_p > \omega$ for a "flat" one.

Exercise. Deduce Eq. (5.45) directly from Eq. (5.44).

Solution. Taking the scalar product of Eq. (5.44) with $\mathbf{e}_3$ we obtain

$$\Omega_p \frac{L_3}{L} = \Omega_p \frac{\omega_3 I_3}{L} = \omega_3 + \Omega = \omega_3 + \frac{I_3 - I_1}{I_1}\omega_3 = \frac{I_3\omega_3}{I_1} \quad\Longrightarrow\quad \Omega_p = \frac{L}{I_1} \,,$$

as before.

Exercise. Study the stability of the rotation of an asymmetric top (i.e., $I_i \neq I_j$ for $i \neq j$) around one of its principal axes of inertia in the case of inertial motion.

Solution. Suppose, for instance, that the body is rotating around its principal axis of inertia $\mathbf{e}_3$ with angular velocity $\boldsymbol{\omega} = \omega_3 \mathbf{e}_3$. Note, first of all, that ω_3 must be *constant*, as is easily deduced from Euler's equations with $\mathbf{N} = 0$ and $\omega_1 = \omega_2 = 0$. Let us next see what happens if we slightly modify the initial conditions

$$\omega_1(0) = \omega_2(0) = 0\,, \qquad \omega_3(0) = \omega_3$$

leading to the previous solution. To first order in ω_1 and ω_2 the product $\omega_1\omega_2$ can be taken as zero, so that the third Euler equation implies that ω_3 remains approximately constant. Differentiating with respect to time the first two Euler equations (with ω_3 constant) we easily obtain

$$\ddot{\omega}_i + \frac{\omega_3^2}{I_1 I_2}\,(I_1 - I_3)(I_2 - I_3)\omega_i = 0\,, \qquad i = 1, 2\,.$$

The solution $\omega_1 = \omega_2 = 0$, $\omega_3 = \text{const.}$ will be stable provided that the solutions of the latter equations are *oscillatory*, and unstable otherwise. Thus the stability condition is that the product $(I_1 - I_3)(I_2 - I_3)$ be *positive*, namely that either $I_3 < I_{1,2}$ or $I_3 > I_{1,2}$. In other words:

The inertial rotation around a principal axis of inertia of an asymmetric top is *stable* if and only if the corresponding principal moment of inertia is either *maximum* or *minimum*.

Exercise. Study the inertial motion of an *asymmetric* rigid body with $I_1 > I_2 > I_3$ in the case $L^2 = 2I_2E$.

Solution. In a frame of principal axes of inertia the conservation of the magnitude of the angular momentum and the (rotational) kinetic energy read

$$\sum_i I_i^2\omega_i^2 = L^2\,, \qquad \sum_i I_i\omega_i^2 = 2E\,.$$

Combining these equations we obtain

$$I_2(I_2 - I_1)\omega_2^2 + I_3(I_3 - I_1)\omega_3^2 = L^2 - 2I_1E\,, \quad I_1(I_1 - I_3)\omega_1^2 + I_2(I_2 - I_3)\omega_2^2 = L^2 - 2I_3E\,,$$

whence it follows that

$$\omega_1^2 = \frac{L^2 - 2I_3E - I_2(I_2 - I_3)\omega_2^2}{I_1(I_1 - I_3)} = \frac{I_2 - I_3}{I_1(I_1 - I_3)}\,(2E - I_2\omega_2^2)\,,$$

$$\omega_3^2 = \frac{L^2 - 2I_1E - I_2(I_2 - I_1)\omega_2^2}{I_3(I_3 - I_1)} = \frac{I_1 - I_2}{I_3(I_1 - I_3)}\,(2E - I_2\omega_2^2) = \frac{I_1(I_1 - I_2)}{I_3(I_2 - I_3)}\,\omega_1^2\,.$$

The Euler equation for ω_2 is thus

$$\dot{\omega}_2 = \pm\sqrt{\frac{(I_2 - I_3)(I_1 - I_2)}{I_1 I_3}}\,\left(\frac{L^2}{I_2^2} - \omega_2^2\right),$$

whose general solution is

$$\pm\nu(t - t_0) = \text{arctanhh}(I_2\omega_2/L) \iff \omega_2 = \pm\frac{L}{I_2}\tanh\bigl(\nu(t - t_0)\bigr)$$

with

$$\nu := \frac{L}{I_2}\sqrt{\frac{(I_2 - I_3)(I_1 - I_2)}{I_1 I_3}}\,.$$

Choosing appropriately the origin of t and the direction of the x_2 principal axis we can simply write

$$\boxed{\omega_2 = \frac{L}{I_2}\tanh(\nu t)\,.}$$

From the previous equations for $\omega_{1,3}$ we then obtain (taking into account that, by the Euler equation for ω_2, ω_1 and ω_3 must have opposite signs, and changing the direction of the x_1 principal axis if necessary)

$$\boxed{\omega_1 = L\sqrt{\frac{I_2 - I_3}{I_1 I_2 (I_1 - I_3)}}\ \mathrm{sech}(\nu t)\,, \qquad \omega_3 = -L\sqrt{\frac{I_1 - I_2}{I_2 I_3 (I_1 - I_3)}}\ \mathrm{sech}(\nu t)\,.}$$

Thus when $t \to \infty$ we have

$$\omega_2 \to \frac{L}{I_2}\,, \qquad \omega_{1,3} \to 0\,;$$

in other words, in the limit $t \to \infty$ the body rotates around the x_2 principal axis of inertia with constant angular velocity.

Remark. In the generic case $I_i \neq I_j$ for $i \neq j$, *the solution of Euler's equations with* $\mathbf{N} = 0$ *can be found by quadratures* using the conservation of E and L^2. Indeed, let us suppose as in the previous exercise that $I_1 > I_2 > I_3$. Proceeding as before we can solve for $\omega_{1,3}$ in terms of L and E to find

$$\omega_1^2 = \frac{L^2 - 2I_3E - I_2(I_2 - I_3)\omega_2^2}{I_1(I_1 - I_3)}\,, \qquad \omega_3^2 = \frac{2I_1E - L^2 - I_2(I_1 - I_2)\omega_2^2}{I_3(I_1 - I_3)}\,. \tag{5.46}$$

From these equations it follows that the motion is only possible if $2EI_3 \leqslant L^2 \leqslant 2EI_1$. Note that $L^2 = 2EI_3$ implies that

$$\omega_1 = \omega_2 = 0, \qquad \omega_3^2 = \frac{2I_1E - L^2}{I_3(I_1 - I_3)}\,,$$

so that the body is rotating around its third principal axis of inertia with constant angular velocity, and similarly if $L^2 = 2EI_1$. To exclude these trivial solutions, we shall suppose in what follows that $2EI_3 < L^2 < 2EI_1$. From the Euler equation for ω_2 we then obtain

$$\int \frac{\mathrm{d}\omega_2}{\sqrt{[L^2 - 2I_3E - I_2(I_2 - I_3)\omega_2^2][2I_1E - L^2 - I_2(I_1 - I_2)\omega_2^2]}} = \pm\frac{t - t_0}{I_2\sqrt{I_1 I_3}}\,. \tag{5.47}$$

Once $\omega_2(t)$ is found from this equation (which can be explicitly done in terms of Jacobian elliptic functions), $\omega_1(t)$ and $\omega_3(t)$ can be immediately found using Eq. (5.46). ∎

Exercise. Show that ω_2 in Eq. (5.47) behaves as the x coordinate of a particle moving in a certain effective one-dimensional potential. Find the particle's effective energy and potential. What can you say about the qualitative features of the motion of ω_2?

Solution. From Eq. (5.47) it immediately follows that

$$\frac{1}{2}\,\dot{\omega}_2^2 + U(\omega_2) = 0\,,$$

with

$$U(\omega_2) = -(2I_1I_2^2I_3)^{-1}\Big(L^2 - 2I_3E - I_2(I_2 - I_3)\omega_2^2\Big)\Big(2I_1E - L^2 - I_2(I_1 - I_2)\omega_2^2\Big)\,.$$

Thus ω_2 behaves as the x coordinate of a particle of unit mass and zero energy subject to the one-dimensional potential $U(\omega_2)$. Since the coefficients

$$L^2 - 2I_3E =: a_1^2, \quad I_2(I_2 - I_3) =: b_1^2, \quad 2I_1E - L^2 := a_2^2, \quad I_2(I_1 - I_2) =: b_2^2,$$

are all positive, the effective potential behaves as shown in Fig. 5.4 with

$$r_1 = \min\left\{\frac{a_1}{b_1}, \frac{a_2}{b_2}\right\}, \qquad r_2 = \max\left\{\frac{a_1}{b_1}, \frac{a_2}{b_2}\right\}.$$

Since, by Eq. (5.46)

$$a_1^2 - b_1^2\omega_2^2 = I_1(I_1 - I_3)\omega_1^2 \geqslant 0\,, \qquad a_2^2 - b_2^2\omega_2^2 = I_3(I_1 - I_3)\omega_1^2 \geqslant 0,$$

we conclude that $|\omega_2(t)| \leqslant \min\left\{\frac{a_1}{b_1}, \frac{a_2}{b_2}\right\} = r_1$. Hence the motion of ω_2 (and, as a consequence, ω_1 and ω_3) is *always bounded*. Moreover, an elementary calculation shows that the absolute maximum $U_{\max}$ of $U(\omega_2)$ is reached for

$$\omega_2 = \pm\left(\frac{r_1^2 + r_2^2}{2}\right)^{1/2},$$

and that

$$U_{\max} = \frac{b_1^2 b_2^2}{8I_1 I_2^2 I_3}\,(r_2^2 - r_1^2)^2 \geqslant 0\,.$$

Since the effective energy is zero, it follows from Fig. (5.4) that $\omega_2(t)$ —and, hence, $\omega_1(t)$ and $\omega_3(t)$— *is a periodic function of t if* $U_{\max} > 0$, whereas *when* $U_{\max} = 0$ *the motion of* $\omega_2(t)$ *is bounded but not periodic* (and similarly for $\omega_1(t)$ and $\omega_3(t)$). Note, finally, that

$$\begin{aligned} U_{\max} = 0 \quad &\Longleftrightarrow \quad r_1^2 = r_2^2 \quad \Longleftrightarrow \quad a_1^2 b_2^2 = a_2^2 b_1^2 \quad \Longleftrightarrow \quad (L^2 - 2I_3E)(I_1 - I_2) = (2I_1E - L^2)(I_2 - I_3) \\ &\Longleftrightarrow \quad L^2 = 2EI_2 \end{aligned}$$

(cf. the previous exercise).

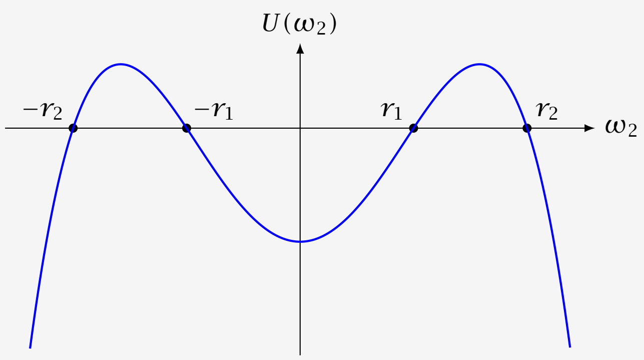


Figure 5.4. Effective potential $U(\omega_2)$.

6 Introduction to relativistic mechanics

6.1 The principles of special relativity

As we saw in Chapter 1, *the laws of mechanics have the same form in* all *inertial frames.* More precisely, consider (for instance) the Galilean boost

$$\boxed{t' = t\,, \qquad x_1' = x_1 - vt\,, \qquad x_2' = x_2\,, \qquad x_3' = x_3\,,} \tag{6.1}$$

relating the space-time coordinates (t, x_1, x_2, x_3) of an event in an inertial reference frame S with their counterparts (t', x_1', x_2', x_3') in another inertial frame S', whose origin O' moves with *constant* velocity $v\mathbf{e}_1$ with respect to S and whose axes are parallel to those of S. Newton's second law in the S frame

$$m\mathbf{a} = \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$$

then becomes in S'

$$m\mathbf{a}' = \mathbf{F}'(t', \mathbf{r}', \dot{\mathbf{r}}')\,, \qquad \text{with} \quad \mathbf{F}'(t', \mathbf{r}', \dot{\mathbf{r}}') = \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}}) \quad (\mathbf{r}' = \mathbf{r} - vt\mathbf{e}_1,\ \dot{\mathbf{r}}' = \dot{\mathbf{r}} - v\mathbf{e}_1).$$

In other words, in both frames the particle's acceleration is the quotient between the force acting on it and its mass, the force being the *same* in both frames but expressed in terms of the particle's coordinates and velocities in each of them. An equivalent way of stating this principle, known as **Galileo's relativity principle**, is the following:

No *mechanical* experiment can discriminate between two inertial frames.

Indeed, mechanical experiments are ultimately based on Newton's second law, which determines the *acceleration* of particles, and this acceleration is the same in S as in S':

$$\boxed{\mathbf{a}' = \ddot{\mathbf{r}}' = \frac{\mathrm{d}^2}{\mathrm{d}t^2}(\mathbf{r} - vt\mathbf{e}_1) = \ddot{\mathbf{r}} = \mathbf{a}\,.}$$

In other words, the following (Galilean) *relativity principle* holds:

All inertial frames are *equivalent* from the point of view of Newtonian mechanics.

At the end of the 19th century, the question arose whether Galileo's relativity principle also applied to Maxwell's equations, which govern electromagnetic phenomena —in particular, the propagation of *electromagnetic waves*, including *light.* Stated differently: is it possible to distinguish between two inertial frames by some kind of electromagnetic (in particular, optical) phenomenon? To answer this question, recall that in empty space the electromagnetic potentials $A_0 := \Phi/c$ and $\mathbf{A} = (A_1, A_2, A_3)$ obey the wave equation[1]

$$\frac{1}{c^2}\frac{\partial^2 A_\mu}{\partial t^2} - \sum_{i=1}^{3}\frac{\partial^2 A_\mu}{\partial x_i^2} = 0\,, \qquad \mu = 0, \dots, 3\,, \tag{6.2}$$

where

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$$

[1]We shall suppose in this chapter that the electromagnetic potentials verify the Lorenz gauge (1.51).

is a *universal constant* depending on the *constant* parameters ε_0, μ_0 appearing in Maxwell's equations. How does Eq. (6.2) transform under the Galilean boost (6.1)? To answer this question, note first of all that

$$A'_\mu(t',\mathbf{r}') = A_\mu(t,\mathbf{r}),$$

since $A_0 = \Phi/c$ is a scalar and, although $\mathbf{A}$ is a vector, the axes of S and S' are parallel by hypothesis. Hence the transformed potentials $A'_\mu(t',\mathbf{r}')$ also satisfy Eq. (6.2), that is[2]

$$\frac{1}{c^2}\frac{\partial^2 A'_\mu}{\partial t^2} - \sum_{i=1}^{3}\frac{\partial^2 A'_\mu}{\partial x_i^2} = 0, \qquad \mu = 0,\dots,3.$$

Taking into account that

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - v\frac{\partial}{\partial x'_1}, \qquad \frac{\partial}{\partial x_i} = \frac{\partial}{\partial x'_i}, \qquad i = 1,2,3,$$

we immediately obtain

$$\frac{1}{c^2}\frac{\partial^2 A'_\mu}{\partial t'^2} - \left(1 - \frac{v^2}{c^2}\right)\frac{\partial^2 A'_\mu}{\partial x_1'^2} - \sum_{i=2}^{3}\frac{\partial^2 A'_\mu}{\partial x_i'^2} - \frac{2v}{c^2}\frac{\partial^2 A'_\mu}{\partial t'\partial x'_1} = 0, \qquad \mu = 0,\dots,3,$$

which is *not* a wave equation in the space-time coordinates (t', x'_1, x'_2, x'_3) for any value of[3] $v \neq 0$.

The fact that the wave equation (6.2) (or, equivalently, Maxwell's equations) is not invariant under Galilean transformations raises the following three possibilities, which can only be decided by experiment:

1. There exists a privileged reference frame in which Eqs. (6.2) (or, equivalently, Maxwell's equations) are valid, and electromagnetic waves propagate with speed $c = 1/\sqrt{\varepsilon_0\mu_0}$. Consequently, the relativity principle —namely, *the equivalence of all inertial frames*— holds only for mechanics but not for electromagnetism.

2. The relativity principle holds both for mechanics and electromagnetism, but Maxwell's equations are incorrect.

3. The relativity principle holds both for mechanics and electromagnetism, but Eq. (6.1) —which follows from a fundamental tenet of Newtonian mechanics, namely the absolute character of time— is not the correct formula relating the space-time coordinates of the same event in two inertial frames.

At the end of the 19th century it was generally believed that the correct hypothesis was the first one. The theoretical basis for this opinion was the belief that electromagnetic waves propagated in a material medium filling all space called the *ether*, and therefore that Eqs. (6.2) —or, equivalently, Maxwell's equations— only hold in an inertial frame at rest with respect to the ether. It was also thought that this privileged inertial frame coincided with that of distant stars, usually identified with Newton's "absolute space". If this hypothesis were true, it would be possible in principle to experimentally detect the motion of an inertial frame relative to the ether ("absolute motion") by measuring the velocity of electromagnetic waves in it.

In 1887, Michelson and Morley conducted a very sensitive interferometric experiment to detect Earth's motion with respect to the ether. The experiment was based on studying the trajectory

[2] In fact, since the wave equation (6.2) is *linear* this result would still hold if the potentials transformed linearly among themselves, i.e., if

$$A'_\mu(t',\mathbf{r}') = \sum_{\nu=0}^{3}\Lambda_\mu^{\ \nu}(v)A_\nu(t,\mathbf{r}), \qquad \mu = 0,\dots,3.$$

[3] Note, however, that for $v \ll c$ the wave equation is approximately invariant under a Galilean boost.

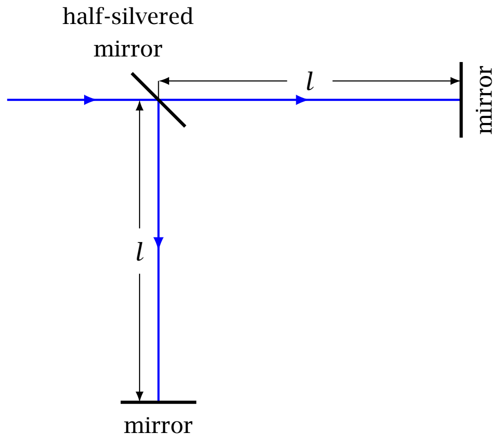


Figure 6.1. The Michelson–Morley experiment.

of a light beam that is divided by a half-silvered mirror into two perpendicular rays (cf. Fig. 6.1), so that the time taken by each of these rays to return to the mirror is different if the device is moving with respect to the ether. Even if this effect is very small (of the order of $v^2/c^2 \sim 10^{-8}$, where v is Earth's speed relative to the ether, believed to be approximately equal to its speed with respect to the Sun), it is possible to observe it by studying the interference fringes produced when both rays recombine. Although the experiment was repeated numerous times, *a negative result was always obtained*, i.e., no relative speed of Earth with respect to the ether was detected. This result was totally unexpected and certainly surprising, since, even admitting that at some point in its orbit Earth's speed relative to the ether vanishes, Earth's velocity varies along its orbit as well as throughout the day (due to Earth's rotation around its axis).

During almost two decades Michelson–Morley's experiment remained without an explanation consistent with other known phenomena (like the aberration of light or the speed of light in moving material media) discarding the theory of ether drag. Finally, in 1905 Einstein observed that the negative result of this experiment (as well as the all of the above mentioned phenomena) can be explained on the basis of the following two fundamental assumptions:

1. The laws of *physics* are the same in *all* inertial frames (**relativity principle**).
2. The speed of electromagnetic waves *in vacuo* is the *universal constant* $c = 1/\sqrt{\varepsilon_0\mu_0}$.

These two postulates are the foundations of the **special theory of relativity**[4] (SR). The first postulate is evidently an extension of Galileo's relativity principle to *all* laws of physics (including electromagnetism), not just mechanics. Combining this postulate with the second one we immediately reach the following conclusion:

The speed of electromagnetic waves *in vacuo* is equal to c in *all* inertial frames.

Of course, this principle satisfactorily explains the negative result of Michelson–Morley's experiment, since it implies that the two light rays in the latter experiment travel with the same speed c. It is, however, profoundly anti-intuitive from the point of view of Newtonian mechanics, since it violates the familiar *law of addition of velocities*

$$\dot{\mathbf{r}} = \dot{\mathbf{r}}' + v\mathbf{e}_1$$

which follows immediately differentiating Eq. (6.1). As a consequence, *the Galilean transformation* (6.1) cannot *be correct*. The same conclusion is reached by noting that the wave equation

[4]The general theory extends the relativity principle to non-inertial frames, thereby developing a theory of gravitation (based on space-time geometry) compatible with the postulates of special relativity.

(6.2) for the electromagnetic potentials —which is equivalent to Maxwell's equations— is not invariant under the Galilean boost (6.1), in contradiction with the two postulates of special relativity.

6.2 Lorentz transformations

6.2.1 Deduction of the equations of the transformation

As we have just remarked, the Galilean boost (6.1) is *not* compatible with the postulates of the special theory of relativity. We shall apply in this section the latter postulates, together with the *homogeneity* and *isotropy* of space-time[5], to deduce the correct equations of the transformation relating the space-time coordinates $(t, \mathbf{r})$ and $(t', \mathbf{r}')$ of the same event in two different inertial frames S and S'. We shall assume, as in the previous section, that the axes of both frames are parallel, their origins coincide at some instant and the velocity of the origin O' of S' relative to S is[6] $\mathbf{v} = v\mathbf{e}_1$. Choosing suitably the origin of time in S and S', we can always arrange for O and O' to coincide at $t = t' = 0$, so that

$$x_\mu = 0\,, \quad \mu = 0, \dots, 3 \quad \Longrightarrow \quad x'_\mu = 0\,, \quad \forall \mu = 0, \dots, 3\,, \tag{6.3}$$

where we have introduced the notation

$$\boxed{x_0 := ct\,, \qquad x'_0 := ct'\,.}$$

From now on we shall tacitly assume that condition (6.3) is satisfied, unless otherwise stated.

i) To begin with, using the *homogeneity* of space-time it can be shown that the transformation relating the coordinates x'_μ and x_μ is *linear*, i.e., that

$$x'_\mu = \sum_{\nu=0}^{3} \Lambda_{\mu\nu}(v)\, x_\nu\,, \qquad \mu = 0, \dots, 3\,,$$

where the coefficients $\Lambda_{\mu\nu}(v)$ depend only on the relative velocity between both frames.

Indeed, consider a clock moving with *constant* velocity with respect to S, and hence (by the first postulate of SR) to S'. If $x_i(t)$ and $x'_i(t')$ $(i = 1, 2, 3)$ are the clock's spatial coordinates in the frames S and S' we then have

$$\frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} = \frac{\mathrm{d}^2 x'_i}{\mathrm{d}t'^2} = 0\,, \qquad i = 1, 2, 3\,.$$

On the other hand, from the homogeneity of space-time it follows that the time τ measured by a moving clock must satisfy

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = \text{const.}\,, \qquad \frac{\mathrm{d}\tau}{\mathrm{d}t'} = \text{const.}$$

From this equation it easily follows that

$$\frac{\mathrm{d}^2 x_i}{\mathrm{d}\tau^2} = \frac{\mathrm{d}^2 x'_i}{\mathrm{d}\tau^2} = 0\,, \qquad i = 1, 2, 3\,,$$

and hence, calling $x_0 = ct$, $x'_0 = ct'$,

$$\frac{\mathrm{d}^2 x_\mu}{\mathrm{d}\tau^2} = \frac{\mathrm{d}^2 x'_\mu}{\mathrm{d}\tau^2} = 0\,, \qquad i = 0, \dots, 3\,.$$

[5] Space-time must be *homogeneous*, i.e., all its points must be equivalent. Likewise, space should be *isotropic*, by which it is meant that all spatial directions must be equivalent.

[6] Throughout this chapter, v shall usually denote the x_1 component of the velocity of the origin of the frame S' relative to S, which can thus be positive or negative. To avoid confusion, the magnitude of the velocity shall be denoted by $|\mathbf{v}| = |v|$.

But then

$$\frac{dx'_\mu}{d\tau} = \sum_{\nu=0}^{3} \frac{\partial x'_\mu}{\partial x_\nu} \frac{dx_\nu}{d\tau}, \quad \frac{d^2 x'_\mu}{d\tau^2} = \sum_{\nu,\sigma=0}^{3} \frac{\partial^2 x'_\mu}{\partial x_\nu \partial x_\sigma} \frac{dx_\nu}{d\tau} \frac{dx_\sigma}{d\tau} = 0 \quad \Longrightarrow \quad \frac{\partial^2 x'_\mu}{\partial x_\nu \partial x_\sigma} = 0, \qquad \forall \nu, \sigma = 0, \ldots, 3,$$

since $dx_\mu / d\tau$ ($\mu = 0, \ldots, 3$) is arbitrary.

ii) Secondly, it is easy to check that the spatial coordinates transversal to the velocity $\mathbf{v}$ must be equal in both frames, i.e., that

$$\boxed{x'_2 = x_2\,, \qquad x'_3 = x_3\,.}$$

Indeed (for instance), since the transformation $x_\mu \mapsto x'_\mu$ is linear we must have

$$x'_2 = \sum_\mu a_\mu(v) x_\mu\,,$$

where the summation index ranges from 0 to 3 (in general, from now on *Greek* indices will always range from 0 to 3 and *Latin* ones from 1 to 3). Since $x_2 = 0$ implies $x'_2 = 0$ (recall that the axes of S and S' are parallel), all the coefficients a_μ vanish except for a_2, and hence

$$x'_2 = a_2(v) x_2\,.$$

By the *isotropy of space*, the coefficient a_2 can only depend on $|v|$, i.e.,

$$a_2(v) = a_2(-v)\,.$$

By the relativity principle, the velocity of O relative to O' must be $-v\mathbf{e}_1$, and hence

$$x_2 = a_2(-v)x'_2 = a_2(-v)a_2(v)x_2 = a_2^2(v)x_2 \quad \Longrightarrow \quad a_2(v) = \pm 1\,.$$

By continuity (since $a_2(0) = 1$) we must have $a_2(v) = 1$, which implies the equality of x_2 and x'_2. Obviously, a similar argument applies to x_3 and x'_3.

iii) Since the origin of S' moves with velocity $v\mathbf{e}_1$ relative to S, the coordinate x'_1 must vanish when $x_1 - vt = 0$, and thus (since the relation between x'_μ and x_ν is *linear*)

$$x'_1 = \gamma(v)(x_1 - vt)\,, \tag{6.4}$$

where γ is an *even* function of v by the isotropy of space. The relativity principle implies the analogous relation

$$x_1 = \gamma(v)(x'_1 + vt')\,. \tag{6.5}$$

Solving for t' in the latter equation and using the value of x'_1 from Eq. (6.4) we obtain

$$x_1 = \gamma^2(v)(x_1 - vt) + \gamma(v)vt' \quad \Longrightarrow \quad t' = \gamma(v)\Big[t + \big(\gamma(v)^{-2} - 1\big)\,\frac{x_1}{v}\Big]. \tag{6.6}$$

Equations (6.4)-(6.6) determine the transformation $(t, x_1) \mapsto (t', x'_1)$ in terms of the unknown coefficient $\gamma(v)$.

So far we have only applied the relativity principle (the first postulate of SR) and the homogeneity and isotropy of space-time. If we assumed at this point that $t' = t$ (i.e., that time is *absolute*), from Eq. (6.6) we would immediately conclude that $\gamma(v) = 1$, which yields the Galilean boost (6.1). We know, however, that this transformation is incorrect, so that necessarily $t' \neq t$. This contradicts one of the fundamental assumptions of Newtonian mechanics, namely the absolute character of time. In fact, in order to find the correct relation between t and t' we must apply Einstein's second postulate, which so far had played no role in our argument. More precisely, according to this postulate the equation $x_1 - ct = 0$ (which describes the propagation of a plane electromagnetic wave in the x_1 direction emitted at $t = t' = 0$ from the origin of both frames) must imply $x_1' - ct' = 0$. Substituting these relations into Eqs. (6.4) and (6.5) we obtain

$$ct' = \gamma(v)(c - v)t\,, \qquad ct = \gamma(v)(c + v)t'\,.$$

Multiplying both equations and canceling the common factor tt' we easily arrive at the relation

$$c^2 = \gamma^2(v)(c^2 - v^2) \quad \Longrightarrow \quad \gamma(v) = \pm\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\,.$$

We must again take, by continuity, the "+" sign (since $v = 0$ we must have $t = t'$, and hence $\gamma(0) = 1$), so that

$$\gamma(v) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\,. \tag{6.7}$$

Substituting into Eqs. (6.5)-(6.6) we finally arrive at the sought-for equations relating the coordinates x_μ and x_μ' in both inertial frames:

$$t' = \gamma(v)\Big(t - \frac{vx_1}{c^2}\Big)\,, \qquad x_1' = \gamma(v)(x_1 - vt)\,, \qquad x_k' = x_k \qquad (k = 2, 3)\,. \tag{6.8}$$

The transformation (6.7)-(6.8) between the coordinates x_μ and x_μ', which replaces the Galilean boost (6.1), is known as a **Lorentz boost** (in the x_1 direction). Note that in terms of the coordinate $x_0 = ct$ (which has dimensions of length), and using the dimensionless parameter $\beta := v/c$, the previous equation adopt the more symmetric form

$$x_0' = \gamma(v)(x_0 - \beta x_1)\,, \qquad x_1' = \gamma(v)(x_1 - \beta x_0)\,, \qquad x_k' = x_k \qquad (k = 2, 3)\,. \tag{6.9}$$

The following facts are a direct consequence of Eq. (6.8) for a Lorentz boost:

i) From Eq. (6.7) for the function $\gamma(v)$ it follows that *the relative speed between two inertial frames must be* strictly less *than the speed c of electromagnetic waves* in vacuo.

ii) In particular, *the speed of all material particles* (i.e., with non-vanishing mass) *is necessarily less than* c, since a set of such particles can be used to construct a reference frame.

iii) In fact, it is easy to show that *the propagation speed of* any *physical signal cannot exceed* c, where by physical "signal" is meant the exchange of *information* between two observers.

To prove the last assertion, suppose that a signal is sent from a point P to a second point Q with speed $u > c$ measured in an inertial frame S. Let us choose the axes of S in such a way that P and Q both lie on the x_1 axis with a spatial separation $\Delta x_1 > 0$, and let $\Delta t > 0$ be the time taken by the signal to reach Q according to S (cf. Fig. 6.2). By Eqs. (6.7)-(6.8), the corresponding time measured in the second inertial frame S' is

$$\Delta t' = \gamma(v)\Big(\Delta t - \frac{v\Delta x_1}{c^2}\Big) = \gamma(v)\Delta t\Big(1 - \frac{uv}{c^2}\Big)\,.$$

If the speed v of the origin of S' relative to S satisfies

$$\frac{c^2}{u} < v < c\,,$$

which is possible since $u > c$ by hypothesis, we shall have $\Delta t' < 0$. In other words, according to S' the signal is received by Q *before* it was emitted by P, which violates the *causality principle* (a cause must always precede its effect).

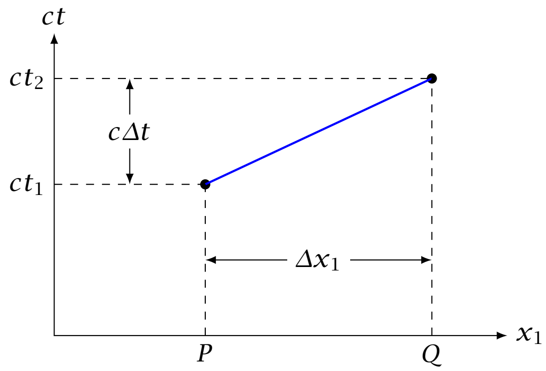


Figure 6.2. Transmission of a signal from P to Q with sped $u > c$ relative to an inertial frame S.

Exercise. Show that in general the Lorentz transformation between to inertial frames S and S' with parallel axes is given by

$$t' = \gamma(v)\left(t - \frac{\mathbf{v}\cdot\mathbf{x}}{c^2}\right), \qquad \mathbf{x}' = \mathbf{x} - \gamma(v)\mathbf{v}t + \big(\gamma(v) - 1\big)\frac{\mathbf{v}\cdot\mathbf{x}}{v^2}\,\mathbf{v}\,, \tag{6.10}$$

where $\mathbf{v}$ is the velocity of O' relative to S.

Solution. Setting $\mathbf{n} = \mathbf{e}_1 = \mathbf{v}/v$ we have

$$x_1 = \mathbf{n}\cdot\mathbf{x} = \frac{\mathbf{x}\cdot\mathbf{v}}{v}\,, \qquad x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = \mathbf{x} - x_1\mathbf{n} = \mathbf{x} - \frac{\mathbf{x}\cdot\mathbf{v}}{v^2}\,\mathbf{v}$$

and hence

$$t' = \gamma(v)\left(t - \frac{\mathbf{v}\cdot\mathbf{x}}{c^2}\right), \qquad \mathbf{x}' = \gamma(v)\left(\frac{\mathbf{x}\cdot\mathbf{v}}{v} - vt\right)\frac{\mathbf{v}}{v} + \mathbf{x} - \frac{\mathbf{x}\cdot\mathbf{v}}{v^2}\,\mathbf{v}\,,$$

which upon simplification yields Eq. (6.10).

6.2.2 Relativistic addition of velocities

Although we have just shown that the two postulates of SR lead to the equations (6.7)-(6.8) of a Lorentz boost, we must still check that this transformation is actually consistent with the latter postulates. As to the first postulate, using the the Lorentz transformation equations and setting

$$u_i := \frac{\mathrm{d}x_i}{\mathrm{d}t}\,, \quad u_i' := \frac{\mathrm{d}x_i'}{\mathrm{d}t'} = \frac{\mathrm{d}x_i'}{\mathrm{d}t}\Big/\frac{\mathrm{d}t'}{\mathrm{d}t}\,, \qquad i = 1, 2, 3,$$

we immediately obtain

$$u_1' = \frac{\mathrm{d}x_1'}{\mathrm{d}t'} = \frac{u_1 - v}{1 - \dfrac{u_1 v}{c^2}}\,, \qquad u_k' = \frac{\mathrm{d}x_k'}{\mathrm{d}t'} = \frac{u_k}{\gamma(v)\left(1 - \dfrac{u_1 v}{c^2}\right)} \qquad (k = 2, 3)\,.$$

Thus if a particle moves with constant velocity $\mathbf{u}$ with respect to S it also moves with constant velocity $\mathbf{u}'$ relative to S'. This is consistent with *Newton's first law* (i.e., if the *law of inertia* applies in the inertial frame S it will also apply in S').

The expression of $\mathbf{u}$ as a function of $\mathbf{u}'$ can be obtained solving for u_i in terms of u_i' from the previous equations, or simply (by the relativity principle) replacing v by $-v$ and u_i' by u_i in the latter equations:

$$u_1 = \frac{u_1' + v}{1 + \dfrac{u_1' v}{c^2}}\,, \qquad u_k = \frac{u_k'}{\gamma(v)\left(1 + \dfrac{u_1' v}{c^2}\right)} \qquad (k = 2, 3)\,. \tag{6.11}$$

This is the **relativistic law for the addition of velocities** which replaces its Galilean analogue $\mathbf{u} = \mathbf{u}' + v\mathbf{e}_1$. From Eq. (6.11) it easily follows that

$$\left(1 + \frac{u_1' v}{c^2}\right)^2 (\mathbf{u}^2 - c^2) = u_1'^2 + \frac{1}{\gamma^2(v)}(u_2'^2 + u_3'^2) + v^2 - c^2 - \frac{v^2 u_1'^2}{c^2} = \frac{\mathbf{u}'^2 - c^2}{\gamma^2(v)}\,.$$

Since $|u_1'| \leqslant c$ and $|v| < c$ the term in parentheses in the LHS is always positive, and therefore $\mathbf{u}^2 - c^2$ and $\mathbf{u}'^2 - c^2$ have the same sign. In particular, if $|\mathbf{u}'| = c$ then $|\mathbf{u}| = c$, which is consistent with the second postulate of special relativity. Moreover, from the previous equation it also follows that if $|\mathbf{u}'| < c$ then $|\mathbf{u}| < c$. In other words, the addition of two speeds smaller than the speed of light produces a speed which is also smaller than c.

Exercise. Show that in general (i.e, when the relative velocity $\mathbf{v}$ between the inertial frames S and S' with parallel axes is not necessarily directed along the $\mathbf{e}_1 = \mathbf{e}_1'$ direction) the relation between the velocities $\mathbf{u}$ and $\mathbf{u}'$ is

$$\mathbf{u} = \frac{\mathbf{u}'}{\gamma(v)\left(1 + \dfrac{\mathbf{u}\cdot\mathbf{v}}{v^2}\right)} + \mathbf{v}\,.$$

Solution. When $\mathbf{v} = v\mathbf{e}_1$ equation (6.11) yields

$$\begin{aligned}\mathbf{u} &= \frac{\mathbf{u}'}{\gamma(v)\left(1 + \dfrac{u_1' v}{c^2}\right)} + \frac{v\mathbf{e}_1}{1 + \dfrac{u_1' v}{c^2}} + \left(1 - \gamma(v)^{-2}\right)\frac{u_1'\mathbf{e}_1}{1 + \dfrac{u_1' v}{c^2}} \\ &= \frac{\mathbf{u}'}{\gamma(v)\left(1 + \dfrac{u_1' v}{c^2}\right)} + \frac{v + \dfrac{v^2 u_1'}{c^2}}{1 + \dfrac{u_1' v}{c^2}}\mathbf{e}_1 = \frac{\mathbf{u}'}{\gamma(v)\left(1 + \dfrac{u_1' v}{c^2}\right)} + v\mathbf{e}_1\,.\end{aligned}$$

Setting $v\mathbf{e}_1 = \mathbf{v}$ and $u_1' = \mathbf{u}'\cdot\mathbf{v}/v$ we obtain the sought-for relation.

6.2.3 Interval

Consider the propagation of a light signal (in general, an electromagnetic pulse) emitted from the origin of S at the time $t = 0$, governed by the equation

$$c^2 t^2 - \mathbf{x}^2 = 0\,, \qquad \mathbf{x} := (x_1, x_2, x_3)$$

in an inertial frame S. By the second postulate, the equation of the wave front in the frame S' whose origin O' coincides with O at $t = t' = 0$ should be

$$c^2 t'^2 - \mathbf{x}'^2 = 0\,.$$

Hence $c^2t^2 - \mathbf{x}^2 = 0$ must imply $c^2t'^2 - \mathbf{x}'^2 = 0$. In fact, using the Lorentz transformation equations (6.7)-(6.8) in the expression $c^2t'^2 - \mathbf{x}'^2$ we readily obtain

$$\begin{aligned}c^2t'^2 - \mathbf{x}'^2 &= \gamma^2(v)\left(ct - \frac{vx_1}{c}\right)^2 - \gamma^2(v)(x_1 - vt)^2 - x_2^2 - x_3^2 \\ &= \gamma^2(v)(c^2 - v^2)t^2 - \gamma^2(v)\left(1 - \frac{v^2}{c^2}\right)x_1^2 - x_2^2 - x_3^2 = c^2t^2 - \mathbf{x}^2\,.\end{aligned}$$

We have thus shown the following fundamental property of Lorentz transformations:

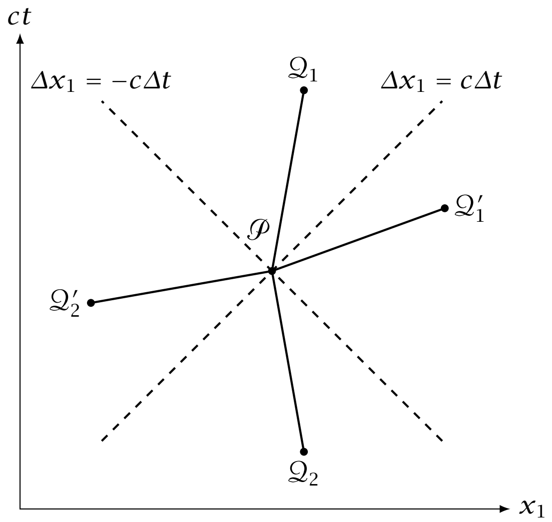


Figure 6.3. In this figure we have represented 5 events $\mathcal{P}$, $\mathcal{Q}_i$, $\mathcal{Q}'_i$ $(i = 1, 2)$, where for the sake of simplicity we have taken $x_2 = x_3 = 0$. The intervals $\mathcal{Q}_i - \mathcal{P}$ are time-like, whereas the remaining intervals $\mathcal{Q}'_i - \mathcal{P}$ are space-like. The event $\mathcal{Q}_1$ is in the future of $\mathcal{P}$ $(t(\mathcal{Q}_1) - t(\mathcal{P}) > 0)$, while $\mathcal{Q}_2$ is in its past $(t(\mathcal{Q}_2) - t(\mathcal{P}) < 0)$. The event $\mathcal{Q}'_2$ cannot have influenced $\mathcal{P}$, nor $\mathcal{Q}'_1$ have been influenced by $\mathcal{P}$.

The quadratic form $c^2t^2 - \mathbf{x}^2 \equiv x_0^2 - \mathbf{x}^2$ is *invariant* under the Lorentz transformation (6.7)-(6.8).

In general, the **interval** between two events with space-time coordinates x_μ and $x_\mu + \Delta x_\mu$ (with $x_0 = ct$) is defined as

$$\Delta s^2 := c^2 \Delta t^2 - \sum_{i=1}^{3} \Delta x_i^2 = \Delta x_0^2 - \sum_{i=1}^{3} \Delta x_i^2 = \Delta x_0^2 - \Delta \mathbf{x}^2 \,. \tag{6.12}$$

Note that, in spite of what the notation might suggest, the interval Δs^2 *may be negative.* Since the Lorentz transformation (6.7)-(6.8) is *linear*, the differences Δx_μ transform in the same way as the coordinates x_μ, and hence:

The interval between two events is invariant under the Lorentz transformation (6.7)-(6.8):

$$\Delta s^2 = \Delta x_0^2 - \Delta \mathbf{x}^2 = \Delta x_0'^2 - \Delta \mathbf{x}'^2 \,. \tag{6.13}$$

Thus the interval between two events is an *intrinsic* property of their mutual relation, independent of the reference frame used to describe them.

By definition, the interval between two events is **time-like** if $\Delta s^2 > 0$, **light-like** if $\Delta s^2 = 0$, and **space-like** if $\Delta s^2 < 0$ (cf. Fig. (6.3)).

Note that

$$\begin{aligned}
\Delta s^2 > 0 \quad &\Longleftrightarrow \quad \Delta x_0 \neq 0\,, \quad \left|\frac{\Delta \mathbf{x}}{\Delta x_0}\right| < 1\,,\\
\Delta s^2 < 0 \quad &\Longleftrightarrow \quad |\Delta \mathbf{x}| \neq 0\,, \quad \frac{|\Delta x_0|}{|\Delta \mathbf{x}|} < 1\,,\\
&\Longleftrightarrow \quad \Delta x_0 = 0, \quad |\Delta \mathbf{x}| > 0 \quad \text{or} \quad \Delta x_0 \neq 0\,, \quad \left|\frac{\Delta \mathbf{x}}{\Delta x_0}\right| > 1\,,\\
\Delta s^2 = 0 \quad &\Longleftrightarrow \quad \Delta x_0 = |\Delta \mathbf{x}| = 0 \quad \text{or} \quad \Delta x_0 \neq 0\,, \quad \left|\frac{\Delta \mathbf{x}}{\Delta x_0}\right| = 1\,.
\end{aligned}$$

It follows that *two events separated by a time-like* (*or light-like*) *interval can influence each other* (in particular, one can be the cause of the other), since it is possible to transmit a signal from one to the other at a speed $|\Delta \mathbf{x}|/|\Delta t|$ not exceeding the speed of light. On the contrary, *two events separated by a space-like interval cannot influence each other,* since a hypothetical signal transmitted from one to the other would travel at a speed greater than c.

If the interval between two events is *time-like*, there exists an inertial reference frame relative to which they occur at the *same point* in space.

Indeed, let us choose the axes of the original inertial frame S so that

$$\Delta x_2 = \Delta x_3 = 0\,,$$

and consider a second inertial frame S' moving with velocity $v\mathbf{e}_1$ relative to S. Since

$$\Delta x_2' = \Delta x_3' = 0\,, \qquad \Delta x_1' = \gamma(\Delta x_1 - v\Delta t)\,,$$

in order to guarantee that $\Delta \mathbf{x}' = 0$ it suffices to take

$$v = \frac{\Delta x_1}{\Delta t}\,.$$

This is certainly possible, since $\Delta s^2 > 0$ implies that

$$\frac{|v|}{c} = \left|\frac{\Delta x_1}{\Delta x_0}\right| < 1\,.$$

The time lapse $\Delta t'$ between both events, measured in the frame S' relative to which they take place at the same point in space, is known as the **proper time lapse** and is usually denoted by $\Delta\tau$. It follows from the invariance of the interval that in any other inertial frame S we have

$$\Delta s^2 = c^2\Delta t^2 - \Delta \mathbf{x}^2 = c^2\Delta t'^2 = c^2\Delta\tau^2 \quad \Longrightarrow \quad \boxed{\Delta\tau = \Delta t\sqrt{1 - \frac{\Delta \mathbf{x}^2}{\Delta x_0^2}}}\,,$$

where we have taken into account that $\Delta\tau = \Delta t'$ and Δt have the same sign[7] (cf. Eq. (6.14) below). Thus *the coordinate time lapse* Δt *is always* greater than or equal to *the proper time lapse*, and in fact only coincides with the latter in the inertial frame with respect to which both events take place at the same point in space.

[7]Since $\mathscr{P}$ and $\mathscr{Q}$ are separated by a time-like interval, it is possible to transmit a signal from $\mathscr{P}$ to $\mathscr{Q}$ or vice versa. If Δt and $\Delta t'$ had opposite signs the effect would precede the cause in either S or S', which would of course violate the causality principle.

If the interval between two events is *space-like*, it is possible to find an inertial frame relative to which they appear to be *simultaneous*.

Indeed (supposing, again, that $\Delta x_2 = \Delta x_3 = 0$), since

$$\Delta t' = \gamma\left(\Delta t - \frac{v\Delta x_1}{c^2}\right)$$

$\Delta t'$ will vanish provided that

$$v = \frac{c^2 \Delta t}{\Delta x_1} = \frac{c\Delta x_0}{\Delta x_1}\,.$$

This is again possible, since $\Delta s^2 < 0$ implies that

$$\frac{|v|}{c} = \left|\frac{\Delta x_0}{\Delta x_1}\right| < 1\,.$$

Note also that in this case

$$\sqrt{-\Delta s^2} = |\Delta x_1'|$$

coincides with the distance between both events in the reference frame relative to which they are simultaneous, known as the events' **proper distance**. Since

$$|\Delta x_1'| = \sqrt{-\Delta s^2} = \sqrt{\Delta \mathbf{x}^2 - \Delta x_0^2} \leqslant |\Delta \mathbf{x}|\,,$$

the proper distance is always *less than or equal to* the spatial distance $|\Delta \mathbf{x}|$ in any other inertial frame, and only coincides with the latter in a frame in which both events are simultaneous. Note, however, that the concept of *proper distance* is different (and less useful) than that of *proper length* that we shall define below.

Finally, if two events are separated by a light-like interval

$$\Delta s^2 = c^2 \Delta t^2 - \Delta \mathbf{x}^2 = 0\,,$$

and hence both events lie along the path of a light ray.

• Consider two events separated by a space-like interval, like $\mathscr{P}$ and $\mathscr{Q}_1'$ in Fig. 6.3. As we have just seen, although in the inertial frame S the event $\mathscr{P}$ precedes $\mathscr{Q}_1'$ there is an inertial frame relative to which $\mathscr{P}$ and $\mathscr{Q}_1'$ are simultaneous. In fact, it is easy to show that there are inertial frames S'' in which $\mathscr{Q}_1'$ precedes $\mathscr{P}$ (exercise). In other words:

The concept of *simultaneity* is not absolute, but depends on the inertial reference frame used.

This fact is called the **relativity of simultaneity**, and is one of the most radical differences between the special theory of relativity and the Newtonian concept of time.

• It is important to realize that *the relativity of simultaneity does not violate the causality principle*, since it applies to events separated by a *space-like* interval, between which there can be no transfer of information (indeed, a hypothetical signal connecting both events would travel with a speed $|\Delta \mathbf{x}|/|\Delta t|$ greater than c). Thus two events separated by a space-like interval cannot influence one another.

• On the other hand, if the interval between two events $\mathscr{P} \neq \mathscr{Q}$ is *time-like* or *light-like*, and $\mathscr{P}$ precedes $\mathscr{Q}$ in an inertial frame S, the same must be true in *any* other inertial frame S' related to S by a Lorentz transformation (6.7)-(6.8), since otherwise the causality principle would be violated. Indeed, the coordinate time lapses Δt and $\Delta t'$ between both events satisfy[8]

$$\Delta t' = \gamma(v)\Delta t\left(1 - \frac{v}{c}\,\frac{\Delta x_1}{\Delta x_0}\right), \tag{6.14}$$

where the term in parentheses is always positive if $\Delta s^2 > 0$ (since $|v|/c$ and $|\Delta x_1|/|\Delta x_0|$ are both less than 1).

[8] Recall that if two different events are separated by a time-like or light-like interval the coordinate time lapse Δt cannot vanish in any inertial frame.

6.2.4 Minkowski product

If $x = (x_0, \mathbf{x})$ and $y = (y_0, \mathbf{y})$ denote the coordinates of two events in a certain inertial frame S, from the invariance of the interval and of the quadratic form $x_0^2 - \mathbf{x}^2$ it follows that

$$\begin{aligned}(y_0 - x_0)^2 - (\mathbf{y} - \mathbf{x})^2 &= y_0^2 - \mathbf{y}^2 + x_0^2 - \mathbf{x}^2 - 2(x_0 y_0 - \mathbf{x}\mathbf{y}) \\ &= (y_0' - x_0')^2 - (\mathbf{y}' - \mathbf{x}')^2 = y_0'^2 - \mathbf{y}'^2 + x_0'^2 - \mathbf{x}'^2 - 2(x_0' y_0' - \mathbf{x}'\mathbf{y}') \\ &= y_0^2 - \mathbf{y}^2 + x_0^2 - \mathbf{x}^2 - 2(x_0' y_0' - \mathbf{x}'\mathbf{y}'),\end{aligned}$$

and thus

$$x_0 y_0 - \mathbf{x}\mathbf{y} = x_0' y_0' - \mathbf{x}'\mathbf{y}' . \tag{6.15}$$

In other words, *the bilinear form*

$$x \cdot y := x_0 y_0 - \mathbf{x}\mathbf{y}, \tag{6.16}$$

known as the **Minkowski product**, *is also invariant under Lorentz transformations* (6.7)-(6.8). Note that, according to this definition,

$$x^2 := x \cdot x = x_0^2 - \mathbf{x}^2, \qquad \Delta s^2 = (\Delta x)^2 . \tag{6.17}$$

The vector space $\mathbb{R}^4$ whose elements are the space-time coordinates of events (or **space-time**, for short), endowed with the Minkowski product (6.16), is usually known as **Minkowski space.** Note that, since the quadratic form (6.17) associated with the Minkowski product (essentially, the interval) is *not* positive definite, Eq. (6.16) does *not* define a true scalar product in Minkowski space. We can, however, use the Minkowski product to define a geometric structure in Minkowski space which is of great help in uncovering the properties of space-time in SR.

6.2.5 Lorentz group

Let S and S' be two inertial frames whose origins coincide for $t = t' = 0$, and denote by $\mathbf{v}$ the velocity of the origin of S' relative to S. In order to find the relation between the space-time coordinates x and x' of a certain event respectively in S and S', we can proceed as follows. First of all, consider an inertial frame S'' at rest relative to S, whose x_1'' axis is in the direction of the relative velocity $\mathbf{v}$. We then have

$$x'' = R_1 x ,$$

where R_1 is a rotation of the spatial coordinates:

$$x_0'' = x_0, \quad \mathbf{x}'' = \mathcal{R}\mathbf{x},$$

with $\mathcal{R} \in \mathrm{SO}(3)$. Secondly, let S''' be a new inertial frame moving with velocity $\mathbf{v} = v\mathbf{e}_1''$ (with $v = |\mathbf{v}| > 0$) relative to S'', with axes parallel to those of S'' and whose origin coincides with that of S'' for $t'' = t''' = 0$. Hence the coordinates in S'' and S''' are related by

$$x''' = L(v)x'' ,$$

where $L(v)$ is the Lorentz transformation (6.7)-(6.8) with x replaced by x'' and x' by x'''. Finally, since S' and S''' move with the same velocity $\mathbf{v}$ with respect to S, and their origins initially coincide, the space-time coordinates x' and x''' are related simply by a spatial rotation R_2, i.e.,

$$x' = R_2 x''' .$$

Combining these equations we finally obtain

$$\boxed{x' = R_2 L(v) R_1 x =: \Lambda x\,.} \tag{6.18}$$

The transformation Λ, which is known as a **general Lorentz transformation**, is the most general transformation relating the coordinates of the same event in two inertial frames whose space-time origins coincide (i.e., $t = x_i = 0 \iff t' = x_i' = 0$). Obviously, if we do not make the latter assumption then we obtain the **Poincaré transformation**

$$\boxed{x' = \Lambda x + a\,,}$$

with $a \in \mathbb{R}^4$ constant.

The Lorentz transformation (6.7)-(6.8) can be written in matrix form as

$$x' = L(v)x\,, \tag{6.19}$$

where $L(v)$ is the 4×4 matrix

$$L(v) = \begin{pmatrix} \gamma(v) & -\beta(v)\gamma(v) & 0 & 0 \\ -\beta(v)\gamma(v) & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \beta(v) := \frac{v}{c}\,. \tag{6.20}$$

Using matrix notation, the Minkowski product of two **four-vectors** $x, y \in \mathbb{R}^4$ can be expressed as

$$\boxed{x \cdot y = x^{\mathsf{T}} G y\,,}$$

where x, y in the RHS are regarded as column vectors and G is the diagonal matrix

$$G = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{6.21}$$

The invariance of the Minkowski product under the transformation (6.8) can be written in matrix form as

$$x' \cdot y' = (L(v)x)^{\mathsf{T}} G (L(v)y) = x^{\mathsf{T}} (L(v)^{\mathsf{T}} G L(v)) y = x \cdot y = x^{\mathsf{T}} G y\,, \qquad \forall x, y \in \mathbb{R}^4\,,$$

or equivalently

$$L(v)^{\mathsf{T}} G L(v) = G\,. \tag{6.22}$$

On the other hand, the Minkowski product is also invariant under *rotations*, since they do not affect time and leave invariant the scalar product of the spatial components of two four-vectors:

$$R^{\mathsf{T}} G R = G \tag{6.23}$$

for any rotation R. If $\Lambda = R_2 L(v) R_1$ is a general Lorentz transformation, from Eqs. (6.22)-(6.23) it immediately follows that

$$\boxed{\Lambda^{\mathsf{T}} G \Lambda = G\,.} \tag{6.24}$$

In other words:

The Minkowski product, and hence the interval, are *invariant* under general Lorentz transformations.

From the mathematical point of view, the set of matrices satisfying Eqs. (6.24) make up a group usually denoted by $\mathrm{O}(1,3)$ and known as the **Lorentz group**, of fundamental importance in physics. It can be shown that general Lorentz transformations (6.18) are a *subgroup* of the Lorentz group, denoted by $\mathrm{SO}_+(1,3)$ and known as *proper orthochronous*, defined by Eq. (6.24) and the additional conditions $\det \Lambda = 1$ and $\Lambda_{00} > 0$.

- Consider, again, the Lorentz boost in the x_1 direction (6.7)-(6.8). Since $\beta(v) \in (-1,1)$, there is a unique $\phi \in \mathbb{R}$ such that

$$\beta(v) = \tanh\phi \,.$$

In terms of this parameter, usually called *rapidity*, $\gamma(v)$ is given by

$$\gamma(v) = \frac{1}{\sqrt{1-\beta(v)^2}} = \frac{1}{\sqrt{1-\tanh^2\phi}} = \cosh\phi \,,$$

and thus the matrix $L(v)$ adopts the following simple expression

$$L(v) = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{6.25}$$

Suppose that we successively perform two Lorentz boosts with velocities $v_1 = c\tanh\phi_1$ and $v_2 = c\tanh\phi_2$. Using the addition formulas satisfied by cosh and sinh it is easy to show that the resulting transformation is another Lorentz boost, with rapidity $\phi_1 + \phi_2$. The velocity of this boost is thus

$$v = c\tanh(\phi_1+\phi_2) = c\,\frac{\tanh\phi_1 + \tanh\phi_2}{1+\tanh\phi_1\tanh\phi_2} = \frac{v_1+v_2}{1+\dfrac{v_1 v_2}{c^2}} \,.$$

We obtain in this way the relativistic law for the addition of two parallels velocities $v_1\mathbf{e}_1$ and $v_2\mathbf{e}_2$, which is a particular case of Eq. (6.11)).

6.3 Physical consequences of Lorentz transformations

Equations (6.7)-(6.8) have important physical consequences that we shall briefly review in this section.

6.3.1 Time dilation

Let, again, S and S' be two inertial frames with parallel axes[9] moving with relative velocity $v\mathbf{e}_1$, and consider a clock attached at the origin of S'. According to S', $\mathbf{x}' = 0$ for all t' along the clock's trajectory. Hence when the clock records a time t' the corresponding time t measured in S is given by

$$t = \gamma(v)\left(t' + \frac{v x_1'}{c^2}\right) = \gamma(v)t' = \frac{t'}{\sqrt{1-\frac{v^2}{c^2}}} > t' \,. \tag{6.26}$$

[9] From now on, we shall tacitly assume that the origins of S and S' coincide at $t = t' = 0$ unless otherwise stated.

In other words, the clock attached to the origin of S' appears to run *slow* relative to the clocks in S. For small velocities v compared to the speed of light c the difference $t - t'$ is very small, since

$$t = t'\left(1 + \frac{v^2}{2c^2} + O(v^4/c^4)\right).$$

However, for velocities comparable to c this difference can be arbitrarily large, since it tends to infinity as $v \to c$. For instance, if $v = 3c/5$ we have $t = 5t'/4$. It is important to bear in mind the following considerations:

- The effect just described, known as **time dilation**, is *symmetric among both inertial frames*, in accordance with Einstein's first postulate. In other words, if we attach a clock to the origin of S the relation between the time t recorded by this clock and the corresponding time t' measured by the clocks in S' is

$$t' = \gamma(v)t\,, \tag{6.27}$$

  since now $\mathbf{x} = 0$ for all t.

- The apparent discrepancy between Eqs. (6.26) and (6.27) is explained taking into account that in these equations both t and t' denote *different* times. The point is that in both cases there is a clear *asymmetry* between the **proper time** measured by a *single* clock at rest in a certain inertial frame and the **coordinate time** recorded by the *clocks* in another frame relative to which the clock is moving —necessarily *more than one*, since the "ticks" of a clock at rest in an inertial frame occur in *different* positions as seen from another inertial frame. It would be incorrect to say that time flows more slowly in S than in S', or vice versa, since *all inertial frames are equivalent*, and there is no *absolute* motion (or rest). It is however true that *the proper time of a clock runs more slowly than the* coordinate *time measured by the clocks in any inertial frame in motion with respect to the clock.*

- Time dilation is constantly being verified in experiments measuring the *half-life* of elementary particles. By definition, the half-life $t_{1/2}$ of a certain particle is the lapse of time after the particle is produced for which the probability that the particle decays reaches $1/2$. In other words, given a large sample of such particles produced at $t = 0$ about half of the sample will have decayed at $t = t_{1/2}$. If the half-life of a particle is Δt_0 in an inertial frame relative to which the particle is at rest —i.e., in the particle's *rest frame*—, its half-life in the laboratory frame will be

$$\Delta t = \gamma(v)\Delta t_0\,, \tag{6.28}$$

  v being the particle's velocity with respect to the latter frame. The half-life Δt_0 can often be computed using quantum field theory techniques, which makes it possible to check the validity of Eq. (6.28) by measuring v and Δt. All the (extremely numerous) experiments performed to date have confirmed the validity of Eq. (6.28). For instance, muons present in cosmic rays can reach a speed

$$v = 0.999\,c$$

  when they enter Earth's atmosphere. For this value of v, the muon's half-life measured in Earth's frame (approximately inertial) is given by

$$\Delta t = \frac{\Delta t_0}{\sqrt{1 - (1 - 10^{-3})^2}} \simeq 22.3663\,\Delta t_0\,.$$

  In the case of muons,

$$\Delta t_0 \simeq 1.5 \cdot 10^{-6}\,\text{s} \quad \Longrightarrow \quad \Delta t \simeq 3.35 \cdot 10^{-5}\,\text{s}\,.$$

Note that the distance traveled by the muon in the time Δt is

$$v\Delta t \sim 10\,\mathrm{Km}\,,$$

while the distance traveled at that speed during the time Δt_0 is merely

$$v\Delta t_0 \sim 450\,\mathrm{m}\,.$$

Thus if it weren't for time dilation muons in cosmic rays would decay long before reaching Earth's surface.

Example 6.1. *Twins paradox.* Suppose that a traveler departs from the origin O of an inertial frame S with velocity $v\mathbf{e}_1$ and after a certain time $\Delta t/2$ (measured in S) reverts its velocity, arriving back to O at a time Δt (cf. Fig. 6.4). What is the time elapsed according to the traveler? In the first part of the trip (until reaching the event denoted by $\mathscr{P}$ in Fig. 6.4 left), the traveler's reference frame is an inertial frame S' moving with constant velocity $v\mathbf{e}_1$ with respect to S. Thus the time assigned to the event $\mathscr{P}$ by the traveler is

$$\Delta t_1' = \frac{\Delta t}{2\gamma(v)}$$

(cf. Eq. (6.26)). In the second part of the trip (from $\mathscr{P}$ on), the traveler's reference frame is *another* inertial frame S'' whose velocity with relative to S is $-v\mathbf{e}_1$. The travel time according to S for this part of the trip is again (by symmetry) $\Delta t/2$, while for the traveler the corresponding time lapse will be

$$\Delta t_2' = \frac{\Delta t}{2\gamma(-v)} = \frac{\Delta t}{2\gamma(v)}\,.$$

Thus the trip's total duration according to the traveler is

$$\Delta t' = \Delta t_1' + \Delta t_2' = \frac{\Delta t}{\gamma(v)} = \sqrt{1 - \frac{v^2}{c^2}}\,\Delta t\,,$$

which can be considerably less than Δt if v/c is close to 1. This result may seem paradoxical, since one might think that from the point of view of the traveler it is the observer at O who has moved with speed $\mp v\mathbf{e}_1$, and hence the duration of the trip measured by the traveler should be $\gamma(v)\Delta t > \Delta t$. The fallacy consists in assuming that the relation between the observer at O and the traveler is *symmetric*, which is very far from being true. Indeed, while O is at rest in an *inertial* reference frame at *all* times, the traveling twin is at rest with respect to *no* inertial frame during the *whole* trip, due to the change in the direction of his velocity at $\mathscr{P}$. In other words, while the observer has not been subject to any acceleration, the traveler has felt an (infinite) acceleration when changing course. It is clear that this will happen *regardless of the trajectory* described by the traveler. Indeed, since this trajectory begins and ends at the origin of the inertial frame S, the traveler must necessarily feel an acceleration at some point (otherwise he or she would move away from the observer with constant speed).

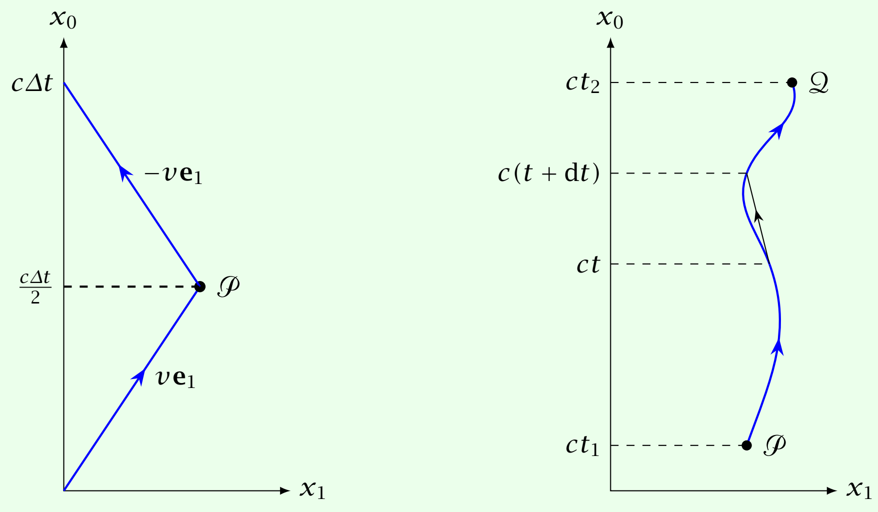


Figure 6.4. World line of the traveler in the twins paradox (left) and of the material particle used in the definition of proper time (right).

More generally, suppose that a material particle follows a trajectory C with equation

$$\mathbf{x} = \mathbf{x}(t)\,, \qquad t_1 \leqslant t \leqslant t_2\,,$$

relative to an inertial frame S. We shall define the particle's **proper time** lapse as the time elapsed between the two events $\mathscr{P} = (ct_1, \mathbf{x}(t_1))$ and $\mathscr{Q} = (ct_2, \mathbf{x}(t_2))$ according to a clock (i.e., an observer) traveling with the particle, i.e., for which the particle is at rest at all times. Since such an observer does not define an inertial frame unless its velocity $\dot{\mathbf{x}}(t)$ is constant, in order to compute the proper time we subdivide the particle's trajectory in Minkowski space, known as its **world line**, in small, approximately straight arcs. In each of these arcs the coordinate time of S varies between t and $t + \mathrm{d}t$, and the particle's velocity is approximately constant and equal to $\dot{\mathbf{x}}(t)$ (cf. Fig. 6.4 right). Hence the proper time $\mathrm{d}\tau$ taken by the particle to trace out this infinitesimal arc is equal to the proper time lapse measured by an inertial frame S' moving with speed $\dot{\mathbf{x}}(t)$ relative to S, namely

$$\boxed{\mathrm{d}\tau = \sqrt{1 - \frac{\dot{\mathbf{x}}^2(t)}{c^2}}\;\mathrm{d}t\,.} \tag{6.29}$$

"Adding up" all these infinitesimal proper times $\mathrm{d}\tau$, i.e., integrating with respect to t, we obtain the following expression for the total lapse of proper time $\Delta\tau(C)$ as the particle travels from $\mathscr{P}$ to $\mathscr{Q}$ along C:

$$\boxed{\Delta\tau(C) = \int_{t_1}^{t_2} \sqrt{1 - \frac{\dot{\mathbf{x}}^2(t)}{c^2}}\;\mathrm{d}t\,.} \tag{6.30}$$

Note that $\Delta\tau(C)$ is invariant under Lorentz transformations by its own definition. This can also be checked directly, since by Eq. (6.29) we have

$$\mathrm{d}\tau^2 = \frac{1}{c^2}(c^2\,\mathrm{d}t^2 - \mathrm{d}\mathbf{x}^2) = \frac{\mathrm{d}s^2}{c^2}\,.$$

Obviously $\Delta\tau(C)$ is always *less than or equal to* the coordinate time lapse $\Delta t = t_2 - t_1$, and $\Delta\tau(C) = \Delta t$ if and only if $\dot{\mathbf{x}}(t) = 0$ for all $t \in [t_1, t_2]$, i.e., if the particle is *at rest* relative to S. It is also important to realize that *the proper time $\Delta\tau$ depends in general on the trajectory followed by the particle,* and not just on the initial and final events $(t_i, \mathbf{x}(t_i))$, $i = 1, 2$, thereof (cf. Fig. 6.5).

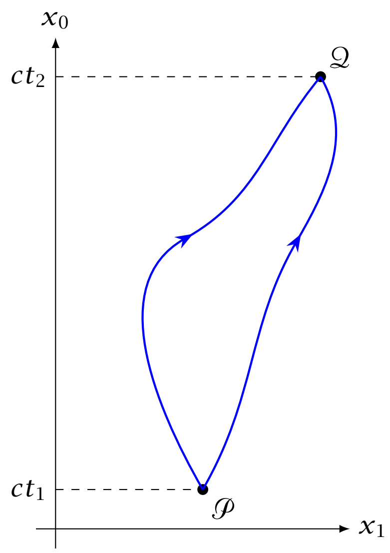


Figure 6.5. World lines connecting two events $\mathscr{P}$ and $\mathscr{Q}$.

Exercise. Let $\mathscr{P}$ and $\mathscr{Q}$ be two events separated by a *time-like* interval. Prove that the world line with endpoints $\mathscr{P}$ and $\mathscr{Q}$ along which the elapsed proper time is *maximum* is a straight line (corresponding to rectilinear motion with constant speed). What happens if the two points are separated instead by a *space-like* or *light-like* interval?

Solution. Suppose, first of all, that the events $\mathscr{P}$ and $\mathscr{Q}$ are separated by a time-like interval, and let C denote the straight world line from $\mathscr{P}$ to $\mathscr{Q}$. Since $\Delta\tau$ is Lorentz invariant, we can compute the proper time $\Delta\tau(C)$ along the world line C in any inertial frame. In particular, choosing the frame S in which $\mathscr{P}$ and $\mathscr{Q}$ take place at the same point in space (i.e., the *proper frame* for the latter events) we have

$$\Delta\tau(C) = t_2 - t_1 \geqslant \int_{t_1}^{t_2} \sqrt{1 - \frac{\dot{\mathbf{x}}^2(t)}{c^2}}\, \mathrm{d}t\,,$$

where the last expression is the proper time elapsed along an arbitrary path $\mathbf{x} = \mathbf{x}(t)$. This shows that the proper time $\Delta\tau$ is maximum along C, as claimed. On the other hand, if $\mathscr{P}$ and $\mathscr{Q}$ are joined by a space- or light-like interval, *no* curve joining $\mathscr{P}$ and $\mathscr{Q}$ can be the world line of a material particle (i.e., no material particle can travel from $\mathscr{P}$ to $\mathscr{Q}$). Indeed, along the world line of a material particle we must have

$$|\Delta\mathbf{x}| = \left|\int_{t_1}^{t_2} \dot{\mathbf{x}}(t)\, \mathrm{d}t\right| \leqslant \int_{t_1}^{t_2} |\dot{\mathbf{x}}(t)|\, \mathrm{d}t < c\Delta t \quad \Longrightarrow \quad \Delta s^2 > 0\,.$$

(Note also that if the interval separating $\mathscr{P}$ and $\mathscr{Q}$ is space-like then both events are *simultaneous* in an appropriate inertial frame, which also implies that no material or even massless particle can travel from $\mathscr{P}$ to $\mathscr{Q}$.)

Note. If $\mathscr{P}$ and $\mathscr{Q}$ are separated by a *light-like* interval, and we only assume that $v \leqslant c$ along the path from $\mathscr{P}$ to $\mathscr{Q}$, it can be shown that the only possible world line joining both events is that of a light ray. Indeed, from the above argument it follows that in this case $|\dot{\mathbf{x}}(t)| = c$ for all t (exercise). If $l(C)$ denotes the (spatial) length of the trajectory C we then have

$$\Delta t = \frac{l(C)}{c} = \frac{|\Delta\mathbf{x}|}{c} \quad \Longrightarrow \quad l(C) = |\Delta\mathbf{x}|\,,$$

so that the path is indeed a straight line traced out with constant speed c.

6.3.2 Lorentz–Fitzgerald contraction

Let, again, S and S' be two reference frames with parallel axes moving with relative speed $v\mathbf{e}_1$. Consider a ruler *at rest* in S', which we can assume to be determined by two marks at the points x'_1 and $x'_1 + l_0$ on the x'_1 axis (with $l_0 > 0$). The distance l_0, i.e., the ruler's length in its **proper frame** S', is known as the ruler's **rest length.** To determine the ruler's length l in the frame S, it is necessary to measure the coordinates x_1 and $x_1 + l$ of its endpoints *at the same time* t. Using Eqs. (6.7)-(6.8) of the Lorentz transformation relating S to S' we obtain

$$\Delta x'_1 = l_0 = \gamma(v)(\Delta x_1 - v\Delta t) = \gamma(v)\Delta x_1 = \gamma(v)l \quad \Longrightarrow \quad l = \frac{l_0}{\gamma(v)} = l_0\sqrt{1 - \frac{v^2}{c^2}} < l_0 \,.$$

Thus in the frame S the ruler appears to be *contracted* by a factor $1/\gamma(v) = \left(1 - \frac{v^2}{c^2}\right)^{1/2}$, a phenomenon known as the **Lorentz–Fitzgerald contraction.**

• Note that *this contraction only occurs in the direction of the velocity of the inertial frame* S' (relative to which the ruler is *at rest*) *with respect to* S, since in the transversal directions $x_k = x'_k$ ($k = 2, 3$).

• Again, it should be stressed that this phenomenon is *symmetric with respect to both reference frames*. In other words, rulers at rest in S also appear to be contracted in S' (along the x'_1 direction) by the same factor $1/\gamma(v)$.

• The *asymmetry* is again between the inertial reference frame in which the ruler is *at rest* and any other inertial frame. Indeed, in the ruler's proper frame its length can be determined *directly* (comparing it, for example, with a calibrated ruler), without the need of measuring *simultaneously* the spatial coordinates of its two endpoints.

More precisely, in the ruler's proper frame the world lines of its endpoints are the vertical lines

$$(t', 0, 0, 0)\,, \qquad (t', l_0, 0, 0)\,,$$

where for simplicity's sake we have taken $x'_1 = 0$. In another inertial frame S these world lines become the lines

$$\Big(\gamma(v)\,t', \gamma(v)vt', 0, 0\Big)\,, \qquad \Big(\gamma(v)(t' + \tfrac{vl_0}{c^2}), \gamma(v)(l_0 + vt'), 0, 0\Big)\,.$$

According to S, when the observer in the ruler's proper frame measures the distance between its endpoints he/she is doing it at two *different* instants $t = \gamma(v)t'$ and $t + \Delta t$, separated by a time difference

$$\Delta t = \frac{\gamma(v)vl_0}{c^2}\,.$$

In this time Δt the right endpoint has moved, according to the observer in S, by

$$v\Delta t = \frac{v^2}{c^2}\,\gamma(v)l_0\,.$$

Thus from the latter observer's point of view at the time $t = \gamma(v)t'$ the endpoints of the ruler are located at the points

$$x_1 = \gamma(v)vt'\,, \qquad x_1 + \Delta x_1 = \gamma(v)(l_0 + vt') - \frac{v^2}{c^2}\,\gamma(v)l_0\,,$$

and the ruler's length measured in S is therefore

$$l = \Delta x_1 = \gamma(v) l_0 - \frac{v^2}{c^2}\,\gamma(v) l_0 = \gamma(v) l_0 \left(1 - \frac{v^2}{c^2}\right) = \frac{l_0}{\gamma(v)}\,.$$

We see, in particular, that the Lorentz–Fitzgerald contraction is closely related to the *relativity of simultaneity*.

6.4 Four-velocity and four-momentum. Relativistic kinetic energy

In Newtonian mechanics, the velocity and momentum of a particle of mass m are related by

$$\mathbf{p} = m\mathbf{v}\,, \tag{6.31}$$

and the particle's equation of motion is Newton's second law

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{F}\,. \tag{6.32}$$

The previous relations are *incompatible with the postulates of special relativity*. For instance, if m and $\mathbf{F}$ are constant the previous equation implies that

$$\mathbf{v}(t) = \mathbf{v}(0) + \frac{\mathbf{F}}{m}t\,,$$

so the particle's speed will become greater than c for $|t|$ large enough. It is clear, therefore, that Eqs. (6.31)-(6.32) cannot be valid (at least for speeds comparable to c), and thus the question arises of what are the correct equations that should replace them. A fundamental guiding principle in this endeavor is the *principle of relativity*, according to which the correct equations should have the *same form* in all inertial reference frames. In other words, they must be *Lorentz covariant*, i.e., they should maintain their form when we apply to them *any* Lorentz (or more generally, Poincaré) transformation. In general, the easiest way of obtaining Lorentz covariant equations is writing down a relation between two scalars (such as the Minkowski product $x \cdot y$, the interval $x^2 = x \cdot x$, etc.), vectors (such as space-time coordinates x) or, in general, *tensors*, under Lorentz transformations. The problem here is that $\mathbf{v}$, $\mathbf{p}$ and $\mathbf{F}$ are vectors in $\mathbb{R}^3$, covariant only under *rotations*. An even more serious issue is that, while in Newtonian mechanics the time t is a *scalar* (essentially invariant under Galilean transformations), according to the theory of special relativity t actually *depends on the reference frame*.

The simplest generalization of the Newtonian definition of velocity

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,, \qquad \mathbf{x} = (x_1, x_2, x_3)\,,$$

which is manifestly covariant under Lorentz transformations is the **four-velocity**

$$\boxed{u = \frac{\mathrm{d}x}{\mathrm{d}\tau}\,.} \tag{6.33}$$

In the latter equation τ is the particle's proper time, which as we know is related to the coordinate time t in *any* inertial reference frame by

$$\mathrm{d}\tau = \sqrt{1 - \frac{\mathbf{v}^2}{c^2}}\,\mathrm{d}t = \frac{\mathrm{d}t}{\gamma(v)}\,. \tag{6.34}$$

To show that u transforms as a vector under a general Lorentz transformation $x' = \Lambda x$, it suffices to note that

$$\mathrm{d}x' = \Lambda\,\mathrm{d}x\,,$$

whereas $d\tau$ is a Lorentz *scalar* ($d\tau = d\tau'$), and therefore

$$u' = \frac{dx'}{d\tau'} = \frac{dx'}{d\tau} = \Lambda \frac{dx}{d\tau},$$

i.e.,

$$\boxed{u' = \Lambda u.}$$

This shows that $u \in \mathbb{R}^4$ is indeed a *vector under Lorentz transformations*, since it transforms in the same way as the coordinates x of a space-time event. In fact, u is actually a vector under *Poincaré transformations* $x' = \Lambda x + a$, since differentiating the latter equation it still follows that $dx' = \Lambda\, dx$. Let us write

$$\boxed{u =: (u_0, \mathbf{u}), \qquad \text{with } \mathbf{u} = (u_1, u_2, u_3) \in \mathbb{R}^3.}$$

The spatial coordinates of the four-velocity in an arbitrary inertial frame are then given by

$$\mathbf{u} = \frac{d\mathbf{x}}{d\tau} = \frac{d\mathbf{x}}{dt}\frac{dt}{d\tau} = \gamma(v)\mathbf{v}. \tag{6.35}$$

In particular, if the particle's velocity is much smaller than c then $\gamma(v) \simeq 1$ and $\mathbf{u} \simeq \mathbf{v}$. As to the time-like coordinate u_0,

$$u_0 = \frac{dx_0}{d\tau} = c\,\frac{dt}{d\tau} = c\gamma(v), \tag{6.36}$$

and hence

$$\boxed{u = \gamma(v)(c, \mathbf{v}).} \tag{6.37}$$

From the previous equation it immediately follows the important relation

$$\boxed{u^2 = c^2.} \tag{6.38}$$

This identity can also be deduced directly from the definition of u, since

$$dx^2 = c^2\, dt^2 - d\mathbf{x}^2 = c^2\, d\tau^2.$$

Note. The vector $\mathbf{u}$ is *not* the particle's velocity in any inertial frame. For instance, since

$$\mathbf{u}^2 = \gamma^2(v)v^2 = \frac{v^2}{1 - \frac{v^2}{c^2}},$$

$|\mathbf{u}| > c$ if $v > c/\sqrt{2}$, and in fact $|\mathbf{u}| \to \infty$ for $v \to c$. ■

In view of the definition of the four-velocity, it is natural to define the **four-momentum** p by

$$\boxed{p = mu,} \tag{6.39}$$

where $m > 0$ is the particle's mass. By Eqs. (6.37)-(6.38), the components of the four-momentum are

$$\boxed{p = m\gamma(v)(c, \mathbf{v}),} \tag{6.40}$$

and its square is given by

$$\boxed{p^2 = m^2c^2.} \tag{6.41}$$

In particular,

$$p_i = m\gamma(v)v_i, \qquad i = 1, 2, 3, \tag{6.42}$$

so that for small velocities compared to c we have

$$p_i \simeq mv_i \qquad (v \ll c).$$

From now on, *we shall denote by* $\mathbf{p}$ *the vector*

$$\mathbf{p} = (p_1, p_2, p_3) = m\gamma(v)\mathbf{v}\,, \tag{6.43}$$

which coincides with the non-relativistic momentum $m\mathbf{v}$ only in the limit $v \to 0$. We shall refer to $\mathbf{p}$ as the *relativistic three-momentum.*

On the other hand, the time-like component p_0 of p is given by

$$\boxed{p_0 = mc\gamma(v) \geqslant mc > 0\,.}$$

Using the identity (6.41), written as

$$p_0^2 = \mathbf{p}^2 + m^2c^2\,, \tag{6.44}$$

and taking into account that $p_0 > 0$, we obtain

$$\boxed{p_0 = \sqrt{\mathbf{p}^2 + m^2c^2}\,.} \tag{6.45}$$

Since u and p are proportional we have

$$\frac{\mathbf{u}}{u_0} = \frac{\mathbf{v}}{c} = \frac{\mathbf{p}}{p_0} \quad \Longrightarrow \quad \boxed{\mathbf{v} = \frac{c\mathbf{p}}{p_0}\,,} \tag{6.46}$$

and hence, by Eq. (6.45),

$$\boxed{\mathbf{v} = \frac{c\mathbf{p}}{\sqrt{\mathbf{p}^2 + m^2c^2}} = \frac{\mathbf{p}/m}{\sqrt{1 + \frac{\mathbf{p}^2}{m^2c^2}}}\,.} \tag{6.47}$$

Note that the previous equation implies that the velocity of a material particle (with non-vanishing mass) must be less than c, in accordance with the principles of special relativity. We can also use Eq. (6.45) to solve for $\gamma(v)$ in terms of $\mathbf{p}$:

$$\boxed{\gamma(v) = \frac{p_0}{mc} = \frac{1}{mc}\sqrt{\mathbf{p}^2 + m^2c^2}\,.} \tag{6.48}$$

If $v \ll c$, expanding cp_0 in powers of v/c and keeping only the first non-constant term we obtain

$$cp_0 = mc^2\left(1 - \frac{v^2}{c^2}\right)^{-1/2} = mc^2\left(1 + \frac{v^2}{2c^2} + O(v^4/c^4)\right) = mc^2 + \frac{1}{2}mv^2 + O(v^4/c^2)\,, \tag{6.49}$$

which, apart form the constant mc^2, coincides to first order in v^2/c^2 with the non-relativistic kinetic energy. The previous equation suggests defining the **relativistic kinetic energy** T by

$$\boxed{T = cp_0 - mc^2 = mc^2(\gamma(v) - 1)\,,} \tag{6.50}$$

and hence

$$\boxed{p_0 = \frac{1}{c}(mc^2 + T)\,.} \tag{6.51}$$

6.5 Four-momentum conservation. Relativistic energy

Newton's first law establishes the conservation of the (non-relativistic) momentum $\mathbf{p} = m\mathbf{v}$ of a particle subject to no external forces. The most natural *Lorentz covariant* generalization of this principle is the *conservation of four-momentum* for a relativistic particle moving in the absence of external forces, namely

$$\boxed{p = \text{const.}\,,}$$

or equivalently

$$cp_0 = mc^2 + T = \text{const.}\,, \qquad p_i = m\gamma(v)v_i = \text{const.}\,.$$

These equations reduce to the conservation of non-relativistic kinetic energy and momentum in the limit $v \ll c$. As in the Newtonian case, by Eq. (6.46) both conservation laws are equivalent to the constancy of the components v_i of the ordinary velocity (in any inertial frame).

Consider next the collision of N particles of mass m_n ($n = 1, \ldots, N$) on which no external forces act. The **total four-momentum** P is then defined by

$$\boxed{P = \sum_{n=1}^{N} p_n =: (P_0, \mathbf{P})\,,} \tag{6.52}$$

where p_n is the four-momentum of the n-th particle. Hence

$$\boxed{P_0 = \sum_{n=1}^{N} p_{n,0} = c\sum_{n=1}^{N} m_n\gamma(v_n)\,, \qquad \mathbf{P} = \sum_{n=1}^{N} \mathbf{p}_n = \sum_{n=1}^{N} m_n\gamma(v_n)\mathbf{v}_n\,.} \tag{6.53}$$

According to Newtonian mechanics, even if the collision is not elastic the system's total linear momentum should be conserved. Moreover, this non-relativistic momentum tends to $\mathbf{P}$ in the limit in which the speeds v_n of all the particles are small compared to c. This fact makes it plausible to postulate the conservation of $\mathbf{P}$ also in relativistic mechanics, i.e.,

$$\mathbf{P}_i = \mathbf{P}_f\,, \tag{6.54}$$

where P_i and P_f denote the total four-momentum respectively before and after the collision. This equation is not *Lorentz covariant*, since only involves the *spatial* components of a four-vector. However, if Eq. (6.54) holds in *all* inertial frames then the *full* four-momentum P is also necessarily conserved, i.e., we must have

$$\boxed{P_i = P_f\,.} \tag{6.55}$$

Indeed, suppose that $\mathbf{P}_i = \mathbf{P}_f$ holds in some inertial frame S, and consider a second inertial frame S' moving with velocity $w\mathbf{e}_1$ relative to S. Since by hypothesis $\mathbf{P}'_i = \mathbf{P}'_f$ should also hold in S', it follows that

$$P'_{i,1} = \gamma(w)\left(P_{i,1} - \frac{w}{c}P_{i,0}\right) = P'_{f,1} = \gamma(w)\left(P_{f,1} - \frac{w}{c}P_{f,0}\right).$$

From $P_{i,1} = P_{f,1}$ and the latter equation it follows that $P_{i,0} = P_{i,0}$, and hence $P_i = P_f$. Actually, the relativistic **law of four-momentum conservation** (6.55) has been (and is being) experimentally verified in multiple situations for speeds arbitrarily close to c, especially in the analysis of collisions taking place in particle accelerators.

The conservation of the time-like component of the four-momentum can be expressed as

$$\sum_n (m_n c^2 + T_n)_i = \sum_n (m_n c^2 + T_n)_f\,,$$

or equivalently

$$(Mc^2 + T)_i = (Mc^2 + T)_f\,,$$

where

$$M = \sum_n m_n\,, \qquad T = \sum_n T_n$$

respectively denote the system's *total mass* and *total kinetic energy*. It is important to note at this point that in relativistic mechanics *the number of particles before and after a collision need not be the same*, since, as we shall see below, particles can be created or destroyed under the appropriate conditions. For this reason, from now on it shall be understood that the sums over n appearing in expressions like the previous ones are tacitly extended to *all* particles in the system before or after the collision, without explicitly specifying their number N_i (before the collision) or N_f (after the collision).

In Newtonian mechanics the total mass M is conserved[10], and therefore the conservation of P_0 is equivalent to that of the system's kinetic energy

$$T_i = T_f\,.$$

According to what we have just seen, however, in relativistic mechanics only $cP_0 = Mc^2 + T$ need be conserved, not M and T separately. In particular:

There may be processes in which the system's total mass decreases (resp. increases), provided that this decrease (resp. increase) is compensated by a corresponding increase (resp. decrease) of the kinetic energy.

More precisely, denoting by $\Delta M = M_f - M_i$ and $\Delta T = T_f - T_i$, the conservation of P_0 can be expressed as

$$\Delta T = -\Delta(Mc^2)\,. \tag{6.56}$$

In other words:

Kinetic energy can be transformed into mass, and vice versa, the conversion factor energy/mass being equal to the square of the velocity of EM waves *in vacuo*.

This is one of the most important predictions of the theory of special relativity, which so far has been experimentally corroborated without exception.

By the previous discussion, we are practically forced to interpret the quantity

$$cP_0 = \sum_n cp_{n,0} = \sum_n (m_n c^2 + T_n) = Mc^2 + T$$

as the system's **total relativistic energy** E (in the absence of external forces). We thus have

$$cP_0 = Mc^2 + T = E\,, \tag{6.57}$$

and the system's total momentum can be expressed as

$$P = (E/c, \mathbf{P})$$

[10] The conservation of the total mass in classical mechanics is a *consequence* of the conservation of total momentum and *Galilean invariance*. Indeed, applying a Galilean boost with velocity $\mathbf{w}$ to the equality $\mathbf{P}_i = \mathbf{P}_f$ we obtain:

$$\mathbf{P}'_i = \sum_n m_n \mathbf{v}'_{n,i} = \sum_n m_n(\mathbf{v}_{n,i} - \mathbf{w}) = \mathbf{P}_i - M_i\mathbf{w} = \mathbf{P}'_f = \mathbf{P}_f - M_f\mathbf{w} \quad \Longrightarrow \quad M_i = M_f\,.$$

For a single particle

$$p = (p_0, \mathbf{p}) = (E/c, m\gamma(v)\mathbf{v}),$$

and from Eq. (6.48) it follows that relativistic energy can be expressed in terms of the particle's velocity by the formula

$$\boxed{E = cp_0 = mc^2\gamma(v).} \tag{6.58}$$

Note that the total relativistic energy E is necessarily positive. In particular, when $v = 0$ the particle possesses a **rest energy**

$$\boxed{E_0 = mc^2.}$$

Note also that from Eq. (6.44) and (6.57) it follows the important relation

$$\boxed{E = c\sqrt{\mathbf{p}^2 + m^2c^2}} \tag{6.59}$$

between relativistic energy and momentum. Writing this relation as

$$E = mc^2\sqrt{1 + \frac{\mathbf{p}^2}{m^2c^2}}$$

and expanding in powers of $\mathbf{p}^2$ we obtain

$$E = mc^2 + \frac{\mathbf{p}^2}{2m} + O(|\mathbf{p}|^4/(m^3c^2)).$$

Note finally that from Eqs. (6.46) and (6.59) we obtain the following relations follow the particle's velocity, energy and momentum:

$$\boxed{\mathbf{v} = \frac{c^2\mathbf{p}}{E}.} \tag{6.60}$$

Note. An alternative formulation of the previous results consists in defining a *velocity dependent mass*

$$m(v) := m\gamma(v) = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}},$$

in terms of which the relativistic momentum and energy are simply given by

$$\mathbf{p} = m(v)\mathbf{v}, \qquad E = m(v)c^2.$$

Note, however, that the previous formula for the kinetic energy

$$T = (m(v) - m)c^2,$$

does not reduce to the classical expression replacing m by $m(v)$. At any rate, we shall not use the concept of variable mass in these notes.

6.6 Massless particles

As we have just seen, the four-momentum p of a particle of mass $m > 0$ has components

$$p = (E/c, \mathbf{p}), \qquad \text{with} \quad E = c\sqrt{\mathbf{p}^2 + m^2c^2}.$$

These relations also make sense if the particle's mass vanishes. Indeed, if $m = 0$ the last equation reduces to

$$\boxed{E = c|\mathbf{p}|,} \tag{6.61}$$

and hence

$$\boxed{p = (|\mathbf{p}|, \mathbf{p})\,.} \tag{6.62}$$

Moreover, for a massive particle velocity and relativistic three-momentum are related by Eq. (6.60). Taking the limit as $m \to 0$ of this equation, and using Eq. (6.61), we immediately obtain

$$\boxed{\mathbf{v} = c\,\frac{\mathbf{p}}{|\mathbf{p}|}\,.} \tag{6.63}$$

Thus *the speed of a massless particle is equal to c*.

The only known massless particle[11] is the **photon**, which is the *quantum of energy* of the electromagnetic field (i.e., the particle carrying the EM field's quanta of energy-momentum). According to quantum mechanics, the relation between the energy of a photon and the frequency ω of its associated electromagnetic wave is given by *Planck's equation*

$$\boxed{E = \hbar\omega = h\nu = \frac{hc}{\lambda}\,,} \tag{6.64}$$

where λ is the *wavelength* and

$$h = 2\pi\hbar = 6.62606957 \cdot 10^{-34}\,\mathrm{J\,s}$$

is Planck's constant. From the relations[12]

$$\boxed{\omega = c\,|\mathbf{k}|\,, \qquad \mathbf{v} = c\,\frac{\mathbf{k}}{|\mathbf{k}|}} \tag{6.65}$$

and Eqs. (6.61), (6.63), and (6.64), it follows that the wave vector $\mathbf{k}$ of the EM wave associated to the photon is given by

$$\mathbf{k} = \frac{\omega}{c}\frac{\mathbf{v}}{c} = \frac{\omega}{c}\frac{\mathbf{p}}{|\mathbf{p}|} = \frac{\omega\mathbf{p}}{E} = \frac{\mathbf{p}}{\hbar} \quad \Longrightarrow \quad \mathbf{p} = \hbar\mathbf{k}\,.$$

This suggests defining a **wave four-vector** $k = (k_0, \mathbf{k})$ by

$$\boxed{k = p/\hbar\,,}$$

with time-like component

$$\boxed{k_0 = \frac{p_0}{\hbar} = \frac{E}{\hbar c} = \frac{\omega}{c} = \frac{2\pi}{\lambda} = |\mathbf{k}|\,.}$$

[11]The existence of a massless particle mediating strong interactions, called *gluon*, has been experimentally confirmed, although gluons are not directly observable because they are confined inside hadrons. For theoretical reasons, it is believed that a similar massless particle known as *graviton* should also exist for the gravitational field.

[12]Recall that in a plane wave propagating with speed c the angular frequency ω, the period τ, the wave vector $\mathbf{k}$, the wavelength λ and the propagation velocity $\mathbf{v}$ are related by

$$\omega = \frac{2\pi}{\tau}\,, \qquad |\mathbf{k}| = \frac{2\pi}{\lambda}\,, \qquad \mathbf{v} = \frac{c\mathbf{k}}{|\mathbf{k}|}\,, \qquad c = \frac{\lambda}{\tau} = \frac{\lambda\omega}{2\pi} = \frac{\omega}{|\mathbf{k}|}\,.$$

All of these relations easily follow from the fact that in a plane wave the wave fronts are moving planes with equation $\omega t - \mathbf{k}\cdot\mathbf{x} = \text{const}$.

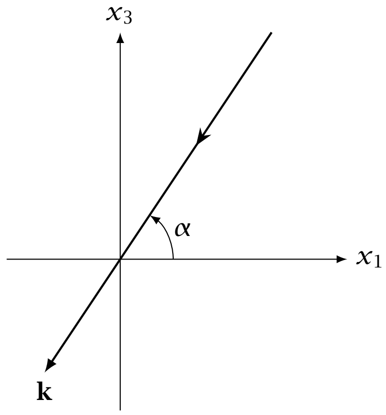


Figure 6.6. Geometry of the relativistic Doppler effect.

It is important to note that k is a *vector under Lorentz transformations*, being proportional to the four-momentum p of the wave's photons. In other words, if S' is another inertial system and $x' = \Lambda x$ we have

$$k' = \Lambda k\,. \tag{6.66}$$

More generally, if two inertial frames S and S' are related by a *Poincaré transformation*

$$x' = \Lambda x + a\,, \tag{6.67}$$

where Λ is a general Lorentz transformation, we know that $u' = \Lambda u$. For $m > 0$ momentum and velocity are proportional, and thus

$$p' = \Lambda p\,.$$

Since this relation is independent of the particle's mass, taking the limit $m \to 0$ we conclude that it must also hold for massless particles. Finally, since k is proportional to p we conclude that the transformation law of the wave four-vector k under the Poincaré transformation (6.67) is still given by Eq. (6.66). In other words, *the wave four-vector is a vector under Poincaré transformations.*

6.6.1 Relativistic Doppler effect

The Lorentz covariance of the wave four-vector k makes it easy to deduce the equations of the **relativistic Doppler effect**. Indeed, let S' be an inertial frame traveling with speed $v\mathbf{e}_1$ relative to the laboratory inertial frame S, with axes parallel to those of S and origin not necessarily coinciding with that of S at $t = 0$. Suppose that an electromagnetic wave with frequency ω_0 and wavelength $\lambda_0 = 2\pi c/\omega_0$ is emitted from S' (ω_0 and λ_0 are respectively called the wave's *proper frequency* and *proper wavelength*). Let us choose the axes so that the wave's propagation direction lies in the plane $x_2 = 0$ and makes an angle $\pi + \alpha$ with the x_1 axis according to the observer at S (cf. Fig 6.6).

$$\mathbf{k} = -|\mathbf{k}|(\cos\alpha, 0, \sin\alpha)\,.$$

(Note that we can assume without loss of generality that $0 \leqslant \alpha \leqslant \pi/2$, changing the orientation of the x_1 axis if necessary.) By the remark at the end of the previous section, we can find the wave four-vector k' in the frame S' in which the wave was emitted by applying to the wave four-vector k a Lorentz transformation $L(v)$ with velocity $v\mathbf{e}_1$, namely

$$k' = L(v)k\,.$$

Since $k'_2 = k_2 = 0$, the spatial components of k' are also of the form

$$\mathbf{k}' = -|\mathbf{k}'|(\cos\alpha', 0, \sin\alpha').$$

On the other hand, the time-like component k'_0 is given by

$$k'_0 = \frac{\omega_0}{c} = \gamma\Big(k_0 - \beta k_1\Big) = \gamma\Big(\frac{\omega}{c} + \beta|\mathbf{k}|\cos\alpha\Big) = \frac{\gamma\omega}{c}(1 + \beta\cos\alpha),$$

and hence

$$\boxed{\omega = \frac{\omega_0}{\gamma(1+\beta\cos\alpha)} \quad \Longrightarrow \quad \lambda = \gamma(1+\beta\cos\alpha)\,\lambda_0\,.} \tag{6.68}$$

The relation between the angles α and α' is also easily computed from the equations

$$k'_1 = \gamma(k_1 - \beta k_0) = \gamma(k_1 - \beta|\mathbf{k}|) = -\gamma|\mathbf{k}|(\cos\alpha + \beta), \qquad k'_3 = k_3 = -|\mathbf{k}|\sin\alpha,$$

whence

$$\boxed{\tan\alpha' = \frac{k'_3}{k'_1} = \frac{\sin\alpha}{\gamma(\beta + \cos\alpha)}\,.} \tag{6.69}$$

A particularly important case is the so-called *longitudinal Doppler effect*, in which $\alpha' = 0$, i.e, the electromagnetic wave propagates in the direction of the relative motion between the observer S and the source S'. From the above formulas it follows that $\alpha = 0$ and hence

$$\boxed{\lambda = \gamma(1+\beta)\,\lambda_0 = \sqrt{\frac{1+\beta}{1-\beta}}\,\lambda_0\,.} \tag{6.70}$$

Thus if the source S' moves *away* from the observer S (i.e, if $\beta > 0$) then $\lambda > \lambda_0$, so that the observer perceives a *shift towards the red* in the wavelength of the electromagnetic wave emitted by S'. On the contrary, if the source moves *towards* the observer then $\beta < 0$, and hence $\lambda < \lambda_0$. Thus in this case the the wavelength of the electromagnetic wave emitted by S' appears *shifted towards the blue* to the observer in S.

On the other hand, if $\alpha = \pi/2$, i.e, when according to the observer in S the wavefront is *perpendicular* to the direction of the emitter's velocity, from Eqs. (6.68)-(6.69) we obtain

$$\boxed{\tan\alpha' = \frac{1}{\gamma\beta}, \qquad \lambda = \gamma\lambda_0 > \lambda_0\,.}$$

Hence in this case the observer perceives a *shift towards the red* regardless of the sign of v. This is the so-called *transversal Doppler effect*, which does not have a classical analogue.

6.6.2 Compton effect

We shall consider next the so-called **Compton effect**, which occurs when a photon is scattered by an electron. In the inertial frame in which the electron is at rest (which usually coincides with the laboratory frame) the initial momenta of the photon and the electron are respectively

$$p_\gamma = \Big(\frac{E}{c}, |\mathbf{p}|, 0, 0\Big) = \frac{E}{c}(1, 1, 0, 0), \qquad p_e = (mc, 0, 0, 0),$$

m being the electron's mass. Let us choose the axes of the frame S so that the collision takes place in the $x_3 = 0$ plane, and denote by θ the angle between the three-momentum of the scattered photon and the x_1 axis. The photon's momentum after the collision is then given by

$$p'_\gamma = \frac{E'}{c}(1, \cos\theta, \sin\theta, 0).$$

By the law of momentum conservation we must then have

$$p_\gamma + p_e = p'_\gamma + p'_e \,,$$

or equivalently

$$p_e + (p_\gamma - p'_\gamma) = p'_e \,.$$

Squaring and taking into account that

$$p_\gamma^2 = p_\gamma'^2 = 0\,, \qquad p_e^2 = p_e'^2 = m^2c^2$$

we arrive at the relation

$$p_e(p_\gamma - p'_\gamma) = p_\gamma p'_\gamma \,,$$

in which we have eliminated the momentum p'_e of the scattered electron. Substituting the previous expressions for p_γ, p'_γ, and p_e we obtain

$$m(E - E') = \frac{EE'}{c^2}(1 - \cos\theta) \quad \Longrightarrow \quad mc^2\left(\frac{1}{E'} - \frac{1}{E}\right) = 1 - \cos\theta \,,$$

and taking into account Eq. (6.64) we finally arrive at the relation

$$\lambda' - \lambda = \frac{h}{mc}(1 - \cos\theta) \tag{6.71}$$

known as **Compton's equation.** We thus see that the wavelength of the scattered photon is always *greater than or equal* to that of the incoming one.

Exercise. Show that the angle $-\theta_e$ between the scattered electron and the x_1 axis and its kinetic energy T_e are determined by the equations

$$\cot\theta_e = \left(1 + \frac{E}{mc^2}\right)\tan(\theta/2)\,, \qquad T_e = \frac{E}{1 + \frac{mc^2}{2E}\csc^2(\theta/2)}\,.$$

6.7 Relativistic collisions

The conservation of the (four-)momentum of a system of particles on which no external forces act is of fundamental importance in the study of *collisions* in the framework of the special theory relativity. Indeed, as we saw in the previous sections, in the absence of external forces the system's total momentum P is conserved, so in particular the momentum P_i immediately before a collision must coincide with the momentum P_f after it (cf. Eq. (6.55)). This conservation law is equivalent to the *conservation of relativistic energy*

$$P_0 = \sum_n p_{n,0} = \sum_n \gamma(v_n) m_n c \tag{6.72}$$

along with the *conservation of three-momentum*

$$\mathbf{P} = \sum_n \mathbf{p}_n = \sum_n \gamma(v_n) m_n \mathbf{v}_n \,. \tag{6.73}$$

6.7.1 Center of momentum frame

The relation (6.55) is valid in any inertial reference frame. In the analysis of the collisions of a system of ultra-relativistic particles (moving at speeds comparable to c) there is, however, a particularly useful inertial frame known as the **center of momentum** (CM) **frame**. This is a frame, analogous to the center of mass system in Newtonian mechanics, in which the spatial components of the system's total momentum vanish, i.e., in which the equality

$$\mathbf{P} = 0$$

holds. In order to establish the existence of such a frame, it suffices to show that the total momentum P of a system of particles is a *time-like* four-vector, i.e., that $P^2 > 0$ (cf. the discussion on p. 196). In fact, this fact is a consequence of the following general result:

> The sum $P = \sum_n p_n$ of any number of *future time-like* four-vectors p_n (that is, $p_n^2 > 0$ and $p_{n,0} > 0$ for all n) is also a future time-like four-vector.

Proof. Indeed, since p_n is a future time-like vector we have

$$p_n^2 = p_{n,0}^2 - \mathbf{p}_n^2 > 0 \quad \Longrightarrow \quad |p_{n,0}| = p_{n,0} > |\mathbf{p}_n| \,.$$

Thus, if p_m is another such vector then

$$\mathbf{p}_n \cdot \mathbf{p}_m \leqslant |\mathbf{p}_n||\mathbf{p}_m| < p_{n,0}p_{m,0}$$

and therefore

$$p_n \cdot p_m = p_{n,0}p_{m,0} - \mathbf{p}_n \cdot \mathbf{p}_m > 0 \,.$$

Hence

$$P^2 = \Big(\sum_n p_n\Big)^2 = \sum_{n,m} p_n \cdot p_m = \sum_n p_n^2 + \sum_{n \neq m} p_n \cdot p_m > 0 \,,$$

and of course (since $p_{n,0} > 0$ for all n)

$$P_0 = \sum_n p_{n,0} > 0 \,. \qquad \blacksquare$$

• It is easy to see that the previous result extends to the case in which some of the four-vectors (but *not all*) are light-like, i.e., it is valid as long as $p_n^2 \geqslant 0$ for all n and $p_k^2 > 0$ for some k (with, as before, $p_{n,0} > 0$ for all n).

6.7.2 Threshold energy

Consider a process like

$$a + b \to a + b + c \,,$$

in which two particles a and b collide producing a third particle c as a result of the collision. In the laboratory frame one of the particles (for instance, b) is the target (i.e., $\mathbf{p}_b = 0$), while the other one (the projectile) has a three-momentum $\mathbf{p}_a \neq 0$. What is the **threshold energy** of particle a, that is, the minimum energy that this particle must have so that the creation of the c particle is possible? Obviously, the conservation of relativistic energy requires that

$$\frac{E_a}{c^2} + m_b = m_a\gamma(v_a') + m_b\gamma(v_b') + m_c\gamma(v_c') \,,$$

where the primes indicate the speeds after the collision in the laboratory frame. Since $\gamma(v_i') \geqslant 1$, from this relation it follows that

$$E_a \geqslant (m_a + m_c)c^2 \,.$$

However, in order to achieve equality in the previous inequality it is necessary that $\gamma(v'_a) = \gamma(v'_b) = \gamma(v'_c) = 1$, i.e., $v'_a = v'_b = v'_c = 0$. This is, however, *impossible*, since by momentum conservation $\mathbf{p}'_a + \mathbf{p}'_b + \mathbf{p}'_c = \mathbf{p}_a \neq 0$, so the speeds of all three particles cannot vanish after the collision. Hence the threshold energy for the process is *strictly greater* than $(m_a + m_c)c^2$.

Let us next compute the threshold energy $E_{\rm min}$ in the more general case

$$a + b \to c_1 + \cdots + c_N\,, \tag{6.74}$$

in which the production of an arbitrary number of additional particles c_i of mass $m_i > 0$ is allowed. To this end, we analyze the collision in the center of momentum (CM) frame, in which the total momentum (before or after the collision) is given by

$$P_{\rm CM} = \frac{E_{\rm CM}}{c}\,(1,0,0,0)\,.$$

Computing the CM energy $E_{\rm CM}$ after the collision we obtain

$$E_{\rm CM} = \sum_i m_i \gamma(v_i) c^2 \geqslant \sum_i m_i c^2 = Mc^2\,.$$

Note that in this case equality can be achieved if all the particles are at rest in the CM system —i.e., if all of them move with the same speed $\mathbf{v}$ in the laboratory frame—, which is of course possible since none of them has zero mass. Therefore the minimum value of the energy in the CM frame is simply Mc^2:

$$E_{\rm CM} \geqslant Mc^2\,.$$

In order to find the threshold energy of particle a in the laboratory frame it suffices to apply the law of momentum conservation and the invariance of the Minkowski product, which yield the relation

$$P_{\rm CM}^2 = \frac{E_{\rm CM}^2}{c^2} = P_L^2 = (p_a + p_b)^2 = c^2(m_a^2 + m_b^2) + 2p_a \cdot p_b\,. \tag{6.75}$$

Here P_L is the initial momentum in the laboratory frame and p_a, p_b the momenta of particles a and b before the collision *in the laboratory frame:*

$$p_a = \left(\frac{E_a}{c}, \mathbf{p}_a\right), \qquad p_b = m_b c(1,0,0,0)\,.$$

Substituting into Eq. (6.75) and operating we obtain

$$\frac{E_{\rm CM}^2}{c^2} = c^2(m_a^2 + m_b^2) + 2E_a m_b\,.$$

Thus the energy of particle a in the laboratory frame is given by

$$E_a = \frac{c^2}{2m_b}\left(\frac{E_{\rm CM}^2}{c^4} - m_a^2 - m_b^2\right).$$

In particular, replacing $E_{\rm CM}$ by its minimum value Mc^2 we obtain the formula

$$E_{\rm min} = \frac{c^2}{2m_b}(M^2 - m_a^2 - m_b^2)\,. \tag{6.76}$$

Note that the previous result is also valid if the a particle (the projectile) is massless.

Exercise. A proton collides with another proton at rest in the laboratory frame, producing a proton-antiproton pair as a result of the collision ($p+p \to p+p+p+\overline{p}$). What is the minimum kinetic energy of the incident proton for this process to be possible?

Solution. Since the mass of a particle is the same as that of its antiparticle, we can apply the latter equation with

$$m_a = m_b =: m\,, \qquad M = 4m\,,$$

where $m \simeq 938.272046\,\mathrm{MeV}/c^2$ is the proton's mass. We thus obtain

$$E_{\min} = \frac{c^2}{2m}(16m^2 - 2m^2) = 7mc^2\,.$$

Hence the minimum kinetic energy of the incident proton is

$$T_{\min} = E_{\min} - mc^2 = 6mc^2 \simeq 5.63\,\mathrm{GeV}\,.$$

Exercise. Show that an *isolated* photon cannot decay into an electron-positron pair ($\gamma \not\to e^- + e^+$). Prove that, however, the process $\gamma + N \to N + e^- + e^+$ (where N is a heavy nucleus) is possible, and that the photon's threshold energy is in this case approximately equal to $2m_e c^2$.

Solution. Let us check, to begin with, that the process $\gamma \to e^- + e^+$ is impossible regardless of the photon's energy. Indeed, in the center of momentum frame of the e^--e^+ pair the final three-momentum $\mathbf{P}$ vanishes, and hence the photon's three-momentum should also vanish in this frame. But this is impossible, since for a massless particle $\mathbf{p} = 0$ implies that $E = c|\mathbf{p}| = 0$, i.e., the particle would have zero energy or momentum. (According to the special theory of relativity the energy of any particle must be strictly positive, even when for zero mass.) Let us next consider the process

$$\gamma + N \to N + e^- + e^+$$

mediated by a heavy nucleus N. Using Eq. (6.76) with

$$m_a = 0\,, \qquad m_b = m_N\,, \qquad M = 2m_e + m_N$$

we obtain

$$E_{\min} = \frac{c^2}{2m_N}\left[(2m_e + m_N)^2 - m_N^2\right] = 2m_e c^2\left(1 + \frac{m_e}{m_N}\right) \gtrsim 2m_e c^2\,,$$

since $m_e \ll m_N$.

6.8 Relativistic dynamics

6.8.1 Four-force and relativistic force

In Newtonian mechanics, the motion of a material particle is governed by Newton's law

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{F}\,, \tag{6.77}$$

which holds in any inertial frame. From the point of view of the special theory of relativity, the most natural generalization of the previous equation is

$$\frac{\mathrm{d}p}{\mathrm{d}\tau} = f\,, \tag{6.78}$$

where

$$f := (f_0, \mathbf{f}) \in \mathbb{R}^4 \tag{6.79}$$

is a four-vector known as **four-force**, depending in general on the particle's space-time coordinates and velocity. Indeed, this equation is *Lorentz covariant*, since p is a vector under Lorentz transformations and the proper time τ is a scalar. In addition, we shall next see that Eq. (6.78) essentially reduces to Newton's second law for small speeds compared to c.

By analogy with Newtonian mechanics, we shall *define* the **relativistic force F** so that Newton's second law (6.77) holds if we interpret $\mathbf{p}$ as the *relativistic* three-momentum. Since

$$\frac{d\mathbf{p}}{dt} = \frac{d\mathbf{p}}{d\tau}\frac{d\tau}{dt} = \frac{1}{\gamma(v)}\frac{d\mathbf{p}}{d\tau} = \frac{\mathbf{f}}{\gamma(v)},$$

where v is the particle's velocity, the four-force and the relativistic force are related by

$$\boxed{\mathbf{F} = \frac{\mathbf{f}}{\gamma(v)}\,.} \tag{6.80}$$

Note that Eq. (6.77) can be written as

$$\frac{d}{dt}(\gamma(v)m\mathbf{v}) = \frac{d}{dt}\left(\frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}}\right) = \mathbf{F}\,. \tag{6.81}$$

Obviously, for a given force $\mathbf{F}$ (for instance, for constant $\mathbf{F}$) the previous equation tends to its classical analogue for particle speeds much smaller than c.

Remark. The fact that the relativistic force $\mathbf{F}$ is related to the space components $\mathbf{f}$ of a four-vector f by Eq. (6.80) ensures that if Eq. (6.77) is valid in an inertial frame it is valid in *all* of them. Of course, Eq. (6.80) imposes very stringent conditions on relativistic forces; in particular, note that although $\mathbf{F}$ is a vector under rotations it does *not* transform as the spatial components of a four-vector under Lorentz transformations. ■

Let us next show that the time-like component of the four-force is determined by the spatial ones. To this end, it suffices to differentiate with respect to τ the identity

$$p^2 = p \cdot p = m^2c^2,$$

which yields

$$\boxed{p \cdot f = 0\,.} \tag{6.82}$$

In other words, *the four-force and the four-momentum are orthogonal* (with respect to the Minkowski product) *at all times.* From the definition of Minkowski product we thus obtain the relation

$$f_0 = \frac{\mathbf{f}\cdot\mathbf{p}}{p_0} = \frac{\mathbf{f}\cdot\mathbf{v}}{c} = \frac{\gamma(v)}{c}\mathbf{F}\cdot\mathbf{v}, \tag{6.83}$$

where we have taken into account Eq. (6.46). Hence the four-force f can be expressed in terms of the relativistic force $\mathbf{F}$ by the equation

$$\boxed{f = \gamma(v)\left(\frac{\mathbf{F}\cdot\mathbf{v}}{c}, \mathbf{F}\right).} \tag{6.84}$$

In Newtonian mechanics

$$\mathbf{F}\cdot\mathbf{v} = \frac{dT}{dt}, \tag{6.85}$$

where

$$T = \frac{1}{2}m\mathbf{v}^2$$

is the particle's kinetic energy. The relativistic analogue of this equation is obtained from the time-like component of the equation of motion (6.78), namely

$$\frac{\mathrm{d}p_0}{\mathrm{d}\tau} = f_0\,.$$

Indeed, by Eq. (6.83) we have

$$\frac{\mathrm{d}p_0}{\mathrm{d}\tau} = \frac{\mathrm{d}p_0}{\mathrm{d}t}\,\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma(v)\,\frac{\mathrm{d}p_0}{\mathrm{d}t} = f_0 = \frac{\gamma(v)}{c}\,\mathbf{F}\cdot\mathbf{v}\,,$$

which yields the identity

$$\frac{\mathrm{d}}{\mathrm{d}t}(cp_0) = \frac{\mathrm{d}}{\mathrm{d}t}(mc^2 + T) = \boxed{\frac{\mathrm{d}T}{\mathrm{d}t} = \mathbf{F}\cdot\mathbf{v}\,.} \tag{6.86}$$

Thus Eq. (6.85) is still valid, if we interpret T as the relativistic kinetic energy and $\mathbf{F}$ as the relativistic force.

Suppose now that, *in a certain inertial frame* S, the relativistic force $\mathbf{F}$ can be obtained from a time-independent scalar potential $V(\mathbf{x})$ through the usual equation

$$\mathbf{F} = -\frac{\partial V(\mathbf{x})}{\partial \mathbf{x}}\,. \tag{6.87}$$

This is the case, for instance, for a constant time-independent force, with $V = -\mathbf{F}\cdot\mathbf{x}$ linear in the particle's spatial coordinates. If Eq. (6.87) holds we have

$$\mathbf{F}\cdot\mathbf{v} = \mathbf{F}\cdot\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = -\frac{\partial V}{\partial \mathbf{x}}\cdot\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = -\frac{\mathrm{d}V}{\mathrm{d}t}\,,$$

and Eq. (6.86) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(cp_0 + V(\mathbf{x})\big) = 0\,. \tag{6.88}$$

Thus in this case the **total relativistic energy**

$$E = cp_0 + V(\mathbf{x}) = mc^2 + T + V(\mathbf{x}) = mc^2\gamma(v) + V(\mathbf{x}) \tag{6.89}$$

is conserved.

Exercise. Find the general solution of the equation of motion of a relativistic particle moving in one dimension under a potential $V(x)$ (in a certain inertial frame).

Solution. By conservation of energy we must have

$$mc^2\gamma(\dot{x}) + V(x) = E\,,$$

where the constant E is the relativistic energy. Since $\gamma(\dot{x}) \geqslant 1$, the motion is only possible in the region $V(x) \leqslant E - mc^2$, where $E - mc^2$ is the analogue of the non-relativistic energy (indeed, for small velocities $|\dot{x}|$ we have $E - mc^2 = \frac{1}{2}m\dot{x}^2 + V(x) + O(|\dot{x}|^4/c^2)$). Squaring and solving for $\dot{x}$ we obtain

$$1 - \frac{\dot{x}^2}{c^2} = \frac{m^2c^4}{\big(E - V(x)\big)^2} \quad \Longrightarrow \quad \dot{x} = \pm c\sqrt{1 - \frac{m^2c^4}{\big(E - V(x)\big)^2}}$$

and hence

$$t = \pm\frac{1}{c}\int \frac{\mathrm{d}x}{\sqrt{1 - \frac{m^2c^4}{(E-V(x))^2}}}\,.$$

Note that the expression under the radical is non-negative, on account of the inequality $V(x) \leqslant E - mc^2$.

6.8.2 Lorentz force

The most important example of a relativistic force is the *Lorentz force*

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{6.90}$$

where $\mathbf{v}$ denotes the particle's velocity. Indeed, it is an experimental fact that the equation of motion of a particle of charge q in an electric field $\mathbf{E}$ and a magnetic one $\mathbf{B}$ is *exactly* (even at speeds arbitrarily close to c) Eq. (6.77) with the Lorentz force (6.90). We shall check in this section that if the fields $\mathbf{E}$ and $\mathbf{B}$ transform appropriately under Lorentz transformations the equation of motion (6.77)-(6.90) is indeed valid in *any* inertial frame.

Notation. As is customary in most textbooks on relativistic mechanics, in this section we shall denote by a^μ the components in a given inertial frame S of a four-vector under Lorentz transformations $a := (a^0, \mathbf{a})$; in particular, $\mathbf{a} = (a^1, a^2, a^3)$. The components a'^μ of the same vector in another inertial frame S', related to S by a Poincaré transformation $x' = \Lambda x + x_0'$, can be obtained from the formula

$$a'^\mu = \sum_\nu \Lambda_\mu{}^\nu a_\nu \tag{6.91}$$

in terms of the matrix elements $(\Lambda_\mu{}^\nu)_{0 \leqslant \mu,\nu \leqslant 3}$ of the Lorentz transformation Λ. This relation can be expressed in matrix form as

$$a' = \Lambda a,$$

where $a = (a^0\, a^1\, a^2\, a^3)^\mathsf{T}$ is a column vector (and similarly a'). Consider next a linear map $B : \mathbb{R}^4 \to \mathbb{R}$, which in the inertial frame S is defined by an equation of the form

$$B(x) = \sum_\mu b_\mu x^\mu$$

for certain coefficients $b_\mu \in \mathbb{R}$. In a different inertial frame S' this equation becomes

$$B(x) = \sum_\mu b'_\mu x'^\mu .$$

To find the relation between the components (b_μ) and (b'_μ) of the linear form B in the frames S and S' it suffices to note that

$$B(x) = \sum_\mu b'_\mu x'^\mu = \sum_{\mu,\nu} b'_\mu \Lambda^\mu{}_\nu x^\nu = \sum_\nu b_\nu x^\nu \quad \implies \quad b_\nu = \sum_\mu \Lambda^\mu{}_\nu b'_\mu .$$

Note that this relation can be written in matrix form as

$$b = b' \Lambda ,$$

where $b = (b^0\, b^1\, b^2\, b^3)$ is a *row* vector (and similarly b'). We can invert this relation using the defining equation (6.24) of Lorentz transformations, which yields

$$b' = b\, G^{-1} \Lambda^\mathsf{T} G .$$

Note that, although $G^{-1} = G$ by Eq. (6.21), we have chosen to distinguish G^{-1} from G since in general relativity G^{-1} need not be equal to G. Denoting by $(g_{\mu\nu})_{0\leqslant\mu,\nu\leqslant 3}$ and $(g^{\mu\nu})_{0\leqslant\mu,\nu\leqslant 3}$ the matrix elements of G and G^{-1}, respectively, we have

$$b'_\mu = \sum_{\nu,\mu',\nu'} b_\nu g^{\nu\nu'} \Lambda^{\mu'}{}_{\nu'} g_{\mu'\mu} =: \sum_\nu \Lambda_\mu{}^\nu b_\nu \,, \qquad \Lambda_\mu{}^\nu = \sum_{\mu',\nu'} g_{\mu\mu'} g^{\nu\nu'} \Lambda^{\mu'}{}_{\nu'} \,, \tag{6.92}$$

where we have made use of the identity $g_{\mu\mu'} = g_{\mu'\mu}$.

We thus see that the transformation law (6.91) of the components (a^μ) of a four-vector a is different from the analogous law (6.92) for the components (b_μ) of a linear form B. To emphasize this distinction linear forms are often called *covectors*, and four-vectors are referred to as *contravariant vectors*. An important example of a covector is the (four-)gradient of a scalar function $\phi(x)$, whose components $\phi_{,\mu}(x)$ are defined by

$$\phi_{,\mu}(x) := \frac{\partial\phi(x)}{\partial x^\mu} \,.$$

Indeed, it suffices to note that the functions $\phi_{,\mu}(x)$ are the components of the linear form (in $\mathrm{d}x$)

$$\mathrm{d}\phi(x) = \sum_\mu \frac{\partial\phi}{\partial x^\mu}\,\mathrm{d}x^\mu \,.$$

Given a contravariant vector $a := (a^\mu)$, it is straightforward to check that the quantities

$$a_\mu = \sum_\nu g_{\mu\nu} a^\nu$$

transform under Lorentz transformations as the components of a covector. Indeed, since

$$\sum_\mu a_\mu x^\mu = \sum_{\mu,\nu} g_{\mu\nu} a^\nu x^\mu = a \cdot x \,,$$

the numbers (a_μ) are the components of the linear form $x \mapsto a \cdot x$. It is thus customary to refer to the numbers (a_μ) as the *covariant components* of the vector a, and in the same vein call (a^μ) its *contravariant components*. Note that from the definition (6.21) of G it follows that

$$a_0 = a^0 \,, \qquad a_i = -a^i \,.$$

In general, an *r-contravariant and s-covariant tensor* T in Minkowski space can be defined as a set of 4^{r+s} quantities $(T^{\mu_1\dots\mu_r}_{\nu_1\dots\nu_s})$ —or, more precisely, an assignment of 4^{r+s} numbers $(T^{\mu_1\dots\mu_r}_{\nu_1\dots\nu_s})$ to each inertial frame S— which transform under a Lorentz transformation Λ as

$$T'^{\mu'_1\dots\mu'_r}_{\nu'_1\dots\nu'_s} = \sum_{\substack{\mu_1,\dots,\mu_r\\ \nu_1,\dots,\nu_s}} \Lambda^{\mu'_1}{}_{\mu_1} \cdots \Lambda^{\mu'_r}{}_{\mu_r} \Lambda_{\nu'_1}{}^{\nu_1} \cdots \Lambda_{\nu'_s}{}^{\nu_s} T^{\mu_1\dots\mu_r}_{\nu_1\dots\nu_s} \,.$$

In other words, the *contravariant indices* μ_i transform as the components of a *contravariant vector* (cf. Eq. (6.91)), while the *covariant indices* ν_j transform as the components of a *covector* (cf. Eq. (6.92)). (More formally, T is a mapping from $(V^*)^r \times V^s$ to $\mathbb{R}$, where $V = \mathbb{R}^4$ and V^* is the dual space of V, which is *linear* in each of its $r + s$ arguments. Note, in this respect, that a vector space V is canonically isomorphic to its *bidual* $V^{**} := (V^*)^*$.)

The condition (6.82), or equivalently $u \cdot f = 0$, is a strong constraint on the form of the covariant four-force f. To begin with, it implies that a *nonzero four-force f must necessarily depend on the particle's velocity,* since even if $\mathbf{f}$ is independent of $\mathbf{v}$ in a certain inertial frame S its time-like component $f^0 = \mathbf{f} \cdot \mathbf{v}$ is velocity-dependent. The simplest example of nontrivial covariant four-force is a linear function of the four-velocity u, i.e.,

$$f^\mu = \sum_\nu F^\mu{}_\nu(x) u^\nu \,. \tag{6.93}$$

From the condition $u \cdot f = 0$ we then obtain

$$\sum_\mu f^\mu u_\mu = \sum_{\mu,\nu} F^\mu{}_\nu(x) u^\nu u_\mu = \sum_{\mu,\nu,\sigma} g_{\mu\sigma} F^\mu{}_\nu(x) u^\nu u^\sigma = 0\,,$$

or equivalently

$$\sum_{\mu,\nu} F^{\mu\nu}(x) u_\mu u_\nu = 0\,, \qquad \text{with} \quad F^{\mu\nu}(x) := \sum_\sigma g^{\nu\sigma} F^\mu{}_\sigma(x)\,.$$

Note that from the equality

$$G^{-1}G = \mathbb{1} \iff \sum_\mu g^{\rho\mu} g_{\mu\sigma} = \delta^\rho{}_\sigma\,,$$

where $\delta^\rho{}_\sigma$ is Kronecker's delta, it follows that

$$F^\mu{}_\nu(x) = \sum_\sigma g_{\nu\sigma} F^{\mu\sigma}(x) = g_{\nu\nu} F^{\mu\nu}(x)\,. \tag{6.94}$$

Since the condition

$$\sum_{\mu,\nu} F^{\mu\nu}(x) u_\mu u_\nu = \frac{1}{2} \sum_{\mu\leqslant\nu} \Big(F^{\mu\nu}(x) + F^{\nu\mu}(x)\Big) u_\mu u_\nu = 0$$

must hold for all values of u_μ, it follows that

$$\boxed{F^{\mu\nu}(x) = -F^{\nu\mu}(x)\,.}$$

Thus the matrix $(F^{\mu\nu}(x))_{0\leqslant\mu,\nu\leqslant3}$ is *antisymmetric*, and has therefore 6 independent components. The relativistic force f associated with the linear four-force (6.93) has components

$$F^i = \frac{f^i}{\gamma(v)} = \sum_\nu F^i{}_\nu \frac{u^\nu}{\gamma(v)} = cF^i{}_0(x) + \sum_j F^i{}_j(x) v^j = cF^{i0}(x) - \sum_j F^{ij}(x) v^j\,.$$

This expression is reminiscent of the Lorentz force acting on a charged particle, since it consists of a term independent of the velocity (proportional to the electric field strength) and another one linear in the velocity (associated with the magnetic field). In fact, let us define

$$E^i(x) := cF^{i0}(x) = -cF^{0i}(x)\,. \tag{6.95}$$

Moreover, since $F^{ij}(x)$ is antisymmetric we can write

$$F^{ij}(x) = -\sum_k \varepsilon_{ijk} B^k(x)\,, \tag{6.96}$$

where ε_{ijk} is Levi-Civita's completely antisymmetric symbol; indeed, $B^k = -\frac{1}{2}\sum_{i,j} \varepsilon_{ijk} F^{ij}(x)$. We then have

$$F^i = E^i(x) + \sum_{j,k} \varepsilon_{ijk} v^j B^k(x) \iff \mathbf{F} = \mathbf{E}(x) + \mathbf{v} \times \mathbf{B}(x)\,.$$

This is the correct form of the electromagnetic force for a unit charge. For an arbitrary charge q the electromagnetic four-force is therefore

$$\boxed{f^\mu = q \sum_\mu F^\mu{}_\nu(x) u^\nu = q \sum_\mu F^{\mu\nu}(x) u_\nu\,,} \tag{6.97}$$

where the elements of the 4×4 matrix $F^{\mu\nu}(x)$ are related to the electric and magnetic fields $\mathbf{E}(x)$ and $\mathbf{B}(x)$ by Eqs. (6.95)-(6.96). The corresponding relativistic force is the familiar Lorentz force

$$\mathbf{F} = q\big(\mathbf{E}(x) + \mathbf{v} \times \mathbf{B}(x)\big) . \tag{6.98}$$

It follows that the relativistic equation of motion of a charge q in an electromagnetic field $(\mathbf{E}(x), \mathbf{B}(x))$ is simply

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = q\big(\mathbf{E}(x) + \mathbf{v} \times \mathbf{B}(x)\big) ,$$

where $\mathbf{p} = m\gamma(v)\mathbf{v}$ is the relativistic three-momentum. Note that this equation is *exact* (i.e., it holds for particle speeds arbitrarily close to c) and is valid in *every* inertial frame.

From Eq. (6.97), and the fact that f and u are contravariant vectors, it follows that the quantities $(F^{\mu\nu})$ are the components of a twice contravariant antisymmetric tensor under Lorentz transformations, which in turn implies that $(F^{\mu}{}_{\nu})$ is a once covariant and once contravariant tensor (cf. next exercise). By Eqs. (6.95)-(6.96), the components of the tensor $(F^{\mu\nu})$, which is known as the *electromagnetic field tensor*, are related to the fields $\mathbf{E}$ and $\mathbf{B}$ by

$$(F^{\mu\nu}) = \begin{pmatrix} 0 & -E^1/c & -E^2/c & -E^3/c \\ E^1/c & 0 & -B^3 & B^2 \\ E^2/c & B^3 & 0 & -B^1 \\ E^3/c & -B^2 & B^1 & 0 \end{pmatrix} . \tag{6.99}$$

Exercise. Show that the quantities $(F^{\mu\nu})$ transform under a Lorentz transformation $x' = \Lambda x$ between two inertial frames S and S' as the components of a twice contravariant tensor.

Solution. Indeed, note that, since

$$\frac{u \cdot f}{q} = \sum_{\mu,\nu} F^{\mu}{}_{\nu} u_{\mu} u^{\nu} = \sum_{\mu,\nu} F^{\mu\nu} u_{\mu} u_{\nu}$$

is a Lorentz scalar, we must have

$$\sum_{\mu',\nu'} F'^{\mu'\nu'} u'_{\mu'} u'_{\nu'} = \sum_{\mu,\nu} F^{\mu\nu} u_{\mu} u_{\nu} = \sum_{\mu,\nu,\mu',\nu'} F^{\mu\nu} \Lambda^{\mu'}{}_{\mu} \Lambda^{\nu'}{}_{\nu} u'_{\mu'} u'_{\nu'} , \qquad \forall u \in \mathbb{R}^4 ,$$

and hence

$$F'^{\mu'\nu'} = \sum_{\mu,\nu} \Lambda^{\mu'}{}_{\mu} \Lambda^{\nu'}{}_{\nu} F^{\mu\nu} , \tag{6.100}$$

as claimed. Note, finally, that from the latter equation it easily follows that $(F^{\mu}{}_{\nu})$ is a once covariant and once contravariant tensor. Indeed,

$$\begin{aligned} F'^{\mu'}{}_{\nu'} &= \sum_{\rho'} g_{\nu'\rho'} F^{\mu'\rho'} = \sum_{\rho',\mu,\nu} g_{\nu'\rho'} \Lambda^{\mu'}{}_{\mu} \Lambda^{\rho'}{}_{\nu} F^{\mu\nu} = \sum_{\rho',\mu,\nu,\sigma} g_{\nu'\rho'} \Lambda^{\mu'}{}_{\mu} \Lambda^{\rho'}{}_{\nu} g^{\nu\sigma} F^{\mu}{}_{\sigma} \\ &= \sum_{\mu,\sigma} \Lambda^{\mu'}{}_{\mu} \Lambda_{\nu'}{}^{\sigma} F^{\mu}{}_{\sigma} , \end{aligned}$$

where in the last step we have made use of Eq. (6.92). Likewise, $(F_{\mu\nu})$ transforms as a rank 2 (antisymmetric) covariant tensor, which is defined by some authors as the electromagnetic field tensor instead of $(F^{\mu\nu})$. Since

$$F_{\mu\nu} = \sum_{\rho,\sigma} g_{\mu\rho} g_{\nu\sigma} F^{\rho\sigma} = g_{\mu\mu} g_{\nu\nu} F^{\mu\nu},$$

by Eq. (6.99), the components of $(F_{\mu\nu})$ are given by

$$\boxed{(F_{\mu\nu}) = \begin{pmatrix} 0 & E^1/c & E^2/c & E^3/c \\ -E^1/c & 0 & -B^3 & B^2 \\ -E^2/c & B^3 & 0 & -B^1 \\ -E^3/c & -B^2 & B^1 & 0 \end{pmatrix}}$$

Exercise. Find the transformation law of the fields **E** and **B** under a Lorentz boost in the x^1 direction with velocity v.

Solution. In this case the only nonzero components elements of $(\Lambda^\mu{}_\nu)$

$$\Lambda^0{}_0 = \Lambda^1{}_1 = \gamma(v), \qquad \Lambda^0{}_1 = \Lambda^1{}_0 = -\beta\gamma(v), \qquad \Lambda^2{}_2 = \Lambda^3{}_3 = 1,$$

so that after an elementary calculation we obtain

$$E'^1 = E^1, \qquad E'^2 = \gamma(E^2 - vB^3), \qquad E'^3 = \gamma(E^3 + vB^2), \tag{6.101}$$

$$B'^1 = B^1, \qquad B'^2 = \gamma\left(B^2 + \frac{v}{c^2}E^3\right), \qquad B'^3 = \gamma\left(B^3 - \frac{v}{c^2}E^2\right). \tag{6.102}$$

Denoting respectively by $\mathbf{E}_\parallel$ and $\mathbf{E}_\perp$ the components of **E** parallel and perpendicular to the velocity **v**, and similarly for **B**, the above equations can be written as

$$\boxed{\mathbf{E}'_\parallel = \mathbf{E}_\parallel, \quad \mathbf{E}_\perp = \gamma(\mathbf{E}_\perp + \mathbf{v}\times\mathbf{B}_\perp); \qquad \mathbf{B}'_\parallel = \mathbf{B}_\parallel, \quad \mathbf{B}'_\perp = \gamma\left(\mathbf{B}_\perp - \frac{\mathbf{v}}{c^2}\times\mathbf{E}_\perp\right).}$$

Taking into account that

$$\mathbf{E}_\parallel = \frac{\mathbf{E}\cdot\mathbf{v}}{v^2}\mathbf{v}, \qquad \mathbf{E}_\perp = \mathbf{E} - \mathbf{E}_\parallel$$

(and similarly for $\mathbf{E}'$, **B**, $\mathbf{B}'$), and using the identity

$$1 - \gamma = \frac{1-\gamma^2}{1+\gamma} = -\frac{\beta^2\gamma^2}{1+\gamma},$$

after a straightforward calculation we obtain

$$\boxed{\mathbf{E}' = \gamma(\mathbf{E} + \mathbf{v}\times\mathbf{B}) - \frac{\gamma^2}{1+\gamma}(\mathbf{E}\cdot\boldsymbol{\beta})\boldsymbol{\beta}, \qquad \mathbf{B}' = \gamma\left(\mathbf{B} - \frac{\mathbf{v}}{c^2}\times\mathbf{E}\right) - \frac{\gamma^2}{1+\gamma}(\mathbf{B}\cdot\boldsymbol{\beta})\boldsymbol{\beta},}$$

where $\boldsymbol{\beta} := \mathbf{v}/c$. Note that, since these equations are written in vector form, they are in fact valid regardless of the direction of the relative velocity **v** between the reference frames S and S'.

6.8.3 Hyperbolic motion

The simplest example of relativistic force is that of a constant force[13]

$$\boxed{\mathbf{F} = m\mathbf{a}\,,}$$

with $\mathbf{a} \in \mathbb{R}^3$ a constant vector with dimensions of acceleration. We shall next see that in this case, just as in non-relativistic mechanics, the particle's equation of motion can be exactly solved. We shall suppose, for the sake of simplicity, that the particle is initially at rest at the origin of coordinates, i.e.,

$$\mathbf{x}(0) = \mathbf{p}(0) = 0\,.$$

Integrating the equation of motion

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = m\mathbf{a}$$

with the initial condition $\mathbf{p}(0) = 0$ we have

$$\mathbf{p} = m\mathbf{a}t\,.$$

Substituting in Eq. (6.47) we obtain

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \frac{c\mathbf{p}}{\sqrt{\mathbf{p}^2 + m^2c^2}} = \frac{mc\mathbf{a}t}{\sqrt{m^2c^2 + m^2a^2t^2}} = \boxed{\frac{\mathbf{a}t}{\sqrt{1 + \frac{a^2t^2}{c^2}}}}\,. \tag{6.103}$$

Note that, regardless of the magnitude of the force $\mathbf{F}$ (i.e., of the constant acceleration $\mathbf{a}$), from the previous equation it follows that $v < c$ for all t. Integrating the last equation with respect to t and taking into account that $\mathbf{x}(0) = 0$ we derive the law of motion:

$$\mathbf{x} = \mathbf{a}\int_0^t \frac{s\,\mathrm{d}s}{\sqrt{1 + \frac{a^2s^2}{c^2}}} = \boxed{\frac{c^2\mathbf{a}}{a^2}\left(\sqrt{1 + \frac{a^2t^2}{c^2}} - 1\right)}\,. \tag{6.104}$$

Note that for $a|t| \ll c$ Eqs. (6.103) and (6.104) approximately reduce to their analogues in Newtonian mechanics

$$\mathbf{v} = \mathbf{a}t\,, \qquad \mathbf{x} = \frac{1}{2}\mathbf{a}t^2\,. \tag{6.105}$$

On the contrary, for $t \to \pm\infty$ the velocity $\mathbf{v}$ tends to $\pm c\mathbf{a}/a$ and, therefore, the particle's speed tends to c (cf. Fig. 6.7), whereas $\mathbf{x} \sim c|t|\mathbf{a}/a$.

If we choose the axes so that $\mathbf{a} = a\mathbf{e}_1$, the law of motion (6.104) reduces to

$$x_1 = \frac{c^2}{a}\left(\sqrt{1 + \frac{a^2t^2}{c^2}} - 1\right) \quad\Longrightarrow\quad \boxed{\left(x_1 + \frac{c^2}{a}\right)^2 - x_0^2 = \frac{c^4}{a^2}\,, \quad x_1 \geqslant 0}\,.$$

This is the equation of a (branch of an) *equilateral hyperbola* (cf. Fig. 6.8) centered at the point $(0, -c^2/a)$, whose axis is the x_1 axis and having as asymptotes the straight lines

$$x_1 + \frac{c^2}{a} = \pm x_0\,.$$

Note that in Newtonian mechanics the particle's world line is the *parabola*

$$x_1 = \frac{a}{2c^2}\,x_0^2$$

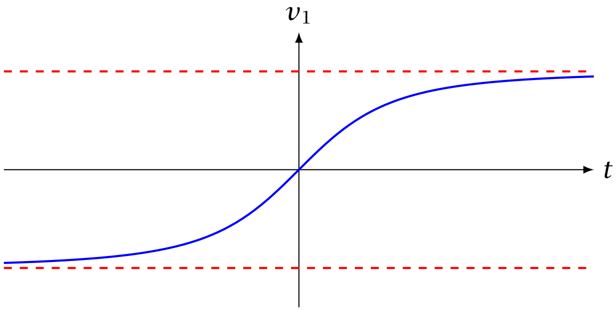


Figure 6.7. Component in the $\mathbf{e}_1$ direction of the velocity of a relativistic particle of mass m subject to a constant force $ma\mathbf{e}_1$ as a function of time (blue curve) and its two asymptotes $v_1 = \pm c$ (dashed red lines).

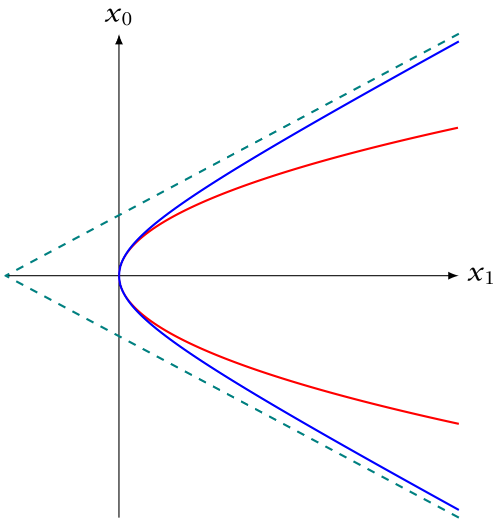


Figure 6.8. World line of a relativistic particle of mass m subject to a constant force ma in the x_1 direction (blue line), along with its analogue in Newtonian mechanics (red line). The dashed green lines represent the asymptotes $x_1 = -\frac{c^2}{a} \pm x_0$ of the particle's world line.

(cf. Eq. (6.105)).

From Eq. (6.103) it immediately follows that

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = \frac{1}{\gamma} = \left(1 + \frac{a^2t^2}{c^2}\right)^{-1/2} \quad \Longrightarrow \quad \tau = \int_0^t \frac{\mathrm{d}s}{\sqrt{1 + \frac{a^2s^2}{c^2}}}\,,$$

where for simplicity's sake we have taken $\tau(0) = 0$. Performing the change of variable $as/c = \sinh z$ in the integral we easily obtain

$$\boxed{\tau = \frac{c}{a}\operatorname{arcsinh}(at/c) = \frac{c}{a}\log\left(\frac{at}{c} + \sqrt{1 + \frac{a^2t^2}{c^2}}\right).} \tag{6.106}$$

[13]The statement that the force acting on a particle is constant is *not* Lorentz invariant, but depends on the inertial frame considered. In other words, even if **F** is constant in a given inertial frame S it need not be constant in another frame S' in motion relative to S. It can be shown, however, that if **F** is constant in an inertial frame S it will remain constant in any other frame S' whose velocity with respect to S has the same direction as **F**, and in this case $\mathbf{F}' = \mathbf{F}$ (see the exercise at the end of this section).

Thus the coordinate time t is related to the proper time τ by

$$\boxed{t = \frac{c}{a}\sinh(a\tau/c)\,.} \tag{6.107}$$

Note, in particular, that for $\tau \gg c/a$ we have

$$t \simeq \frac{c}{2a}\,\mathrm{e}^{a\tau/c} \qquad (\tau \gg c/a)\,,$$

i.e., coordinate time increases exponentially with proper time.

It is also of interest to compute $\beta(v)$ and $\gamma(v)$ as functions of the proper time τ. First of all (taking, as before, $\mathbf{a} = a\mathbf{e}_1$), the parameter $\beta(v)$ is easily obtained from Eqs. (6.103) and (6.107):

$$\beta(v) = \frac{v_1}{c} = \frac{\sinh(a\tau/c)}{\sqrt{1+\sinh^2(a\tau/c)}} = \tanh(a\tau/c)\,. \tag{6.108}$$

As for $\gamma(v)$, it can be derived from the previous expression or more directly taking into account that it is just the derivative of the coordinate time t with respect to the proper time τ:

$$\gamma(v) = \frac{\mathrm{d}t}{\mathrm{d}\tau} = \cosh(a\tau/c)\,. \tag{6.109}$$

By Eq. (6.50), the particle's kinetic energy is given by

$$T = mc^2(\gamma(v)-1) = mc^2(\cosh(a\tau/c)-1) = \boxed{2mc^2\sinh^2(a\tau/(2c))}\,. \tag{6.110}$$

This is the energy that must be supplied to the particle to maintain its constant acceleration $\mathbf{a}$ between the *proper* times 0 and τ. By the law of conservation of relativistic energy (6.89), this energy must be equal to the work $\mathbf{F}\cdot\mathbf{x} = m\mathbf{a}\cdot\mathbf{x}$ done by the constant force $\mathbf{F} = m\mathbf{a}$ during this period of time. This fact is also easily verified using Eqs. (6.104) and (6.107):

$$m\mathbf{a}\cdot\mathbf{x} = mc^2\left(\sqrt{1+\frac{a^2t^2}{c^2}}-1\right) = mc^2(\cosh(a\tau/c)-1)\,. \tag{6.111}$$

Again, for proper times $\tau \gg c/a$ this energy increases exponentially τ:

$$T \simeq \frac{1}{2}mc^2\,\mathrm{e}^{a\tau/c} \qquad (\tau \gg c/a)\,.$$

Exercise. Show that if a force $\mathbf{F}$ is constant in an inertial frame S it is also constant in any other inertial frame S' moving in the direction of $\mathbf{F}$ relative to S, and that moreover $\mathbf{F}' = \mathbf{F}$.

Solution. Let us take the x_1 axis of S in the direction of the force $\mathbf{F}$ and the axes of S' parallel to those of S, and denote by $\mathbf{w} = w\mathbf{e}_1$ the velocity of the origin O' of S' relative to S. In the original frame S the four-force f has components

$$f_1 = \gamma(v)F\,, \qquad f_0 = \frac{\gamma(v)}{c}\mathbf{F}\cdot\mathbf{v} = \gamma(v)\frac{v_1}{c}F\,, \qquad f_2 = f_3 = 0\,,$$

and therefore

$$f = F\gamma(v)\left(\frac{v_1}{c},1,0,0\right).$$

The components of the four-force in the frame S' are obtained applying a Lorentz boost of velocity w in the direction of the x_1 axis:

$$f_0' = \gamma(w)\left(f_0-\frac{w}{c}f_1\right) = F\gamma(v)\gamma(w)\frac{v_1-w}{c}\,, \quad f_1' = \gamma(w)\left(f_1-\frac{w}{c}f_0\right) = F\gamma(v)\gamma(w)\left(1-\frac{v_1w}{c^2}\right)$$

and of course $f_2 = f_3 = 0$. Taking into account that the x_1 component of the particle's velocity in the frame S' is given by the relativistic law of addition of velocities

$$v_1' = \frac{v_1 - w}{1 - \frac{v_1 w}{c^2}}$$

we obtain

$$f' = F\gamma(v)\gamma(w)\left(1 - \frac{v_1 w}{c^2}\right)\left(\frac{v_1'}{c}, 1, 0, 0\right).$$

From the identity

$$\gamma(v)\gamma(w)\left(1 - \frac{v_1 w}{c^2}\right) = \gamma(v')$$

(exercise) it then follows that

$$f' = F\gamma(v')\left(\frac{v_1'}{c}, 1, 0, 0\right),$$

and in particular

$$\mathbf{F}' = \frac{f_1'}{\gamma(v')}\,\mathbf{e}_1 = F\mathbf{e}_1 = \mathbf{F}.$$

Exercise. The *proper acceleration* of a particle is its instantaneous acceleration relative to its proper inertial frame. i) Express the proper acceleration as a function of the particle's acceleration $\mathbf{a} = \mathrm{d}\mathbf{v}/\,\mathrm{d}t$ measured in an arbitrary inertial frame. ii) If the particle's velocity relative to a certain inertial frame is always parallel to the vector $\mathbf{e}_1$, show that in that frame the proper acceleration equals $\mathrm{d}\mathbf{u}/\mathrm{d}t$.

Solution. i) To compute the proper acceleration at a certain time t, let us first find how the acceleration $\mathbf{a} = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}$ measured in a certain inertial frame S transforms under a Lorentz boost with velocity $w\mathbf{e}_1$. To this end, it suffices to differentiate the law of relativistic addition of velocities, with the result

$$a_1' = \frac{\mathrm{d}v_1'}{\mathrm{d}t}\bigg/\frac{\mathrm{d}t'}{\mathrm{d}t} = \frac{Da_1 + (v_1 - w)\frac{a_1 w}{c^2}}{\gamma(w)D^3} = \frac{a_1}{\gamma(w)^3 D^3}, \qquad a_k' = \frac{Da_k + v_k\frac{a_1 w}{c^2}}{\gamma^2(w)D^3} \quad (k = 2, 3),$$

where we have set

$$D := 1 - \frac{v_1 w}{c^2}.$$

The vector $\mathbf{a}' = (a_1', a_2', a_3')$ is the particle's acceleration measured in an inertial frame S' (with axes parallel to those of S) moving relative to S with velocity $\mathbf{w} = w\mathbf{e}_1$. If we assume that at some instant t the particle is moving in the direction of the axis $\mathbf{e}_1$, that is if $\mathbf{v} = v\mathbf{e}_1$ at that time, taking $v_1 = v = w$ (and therefore $D = \gamma(v)^{-2}$) and $v_2 = v_3 = 0$ in the previous equations yields the particle's acceleration in its proper frame S' at that instant t:

$$a_1' = \gamma(v)^3 a_1, \qquad a_k' = \gamma(v)^2 a_k \qquad (k = 2, 3).$$

Obviously, in an arbitrary inertial frame (whose $\mathbf{e}_1$ axis need not coincide with the direction of the particle's velocity at time t) the previous formulas should be replaced by

$$a_\parallel' = \gamma(v)^3 a_{||}, \qquad a_\perp' = \gamma(v)^2 a_\perp,$$

where $a_\parallel'$ and $a_\perp'$ respectively denote the components of $\mathbf{a}'$ in the direction of the particle's velocity at time t and its perpendicular.

ii) If the particle moves at all times in the direction of the x_1 axis with velocity v (not necessarily constant) in a certain inertial frame S then $x_2 = x_3 = 0$ for all t, and therefore $a_2 = a_3 = 0$.

From the previous formulas it then follows that

$$a_1' = \gamma(v)^3 a_1 = \gamma(v)^3 \frac{\mathrm{d}v}{\mathrm{d}t}, \qquad a_2' = a_3' = 0.$$

Hence

$$\frac{\mathrm{d}u_1}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(\gamma(v)v) = (\gamma'(v)v + \gamma(v))\frac{\mathrm{d}v}{\mathrm{d}t} = \left(\gamma(v) + \gamma(v)^3 \frac{v^2}{c^2}\right)\frac{\mathrm{d}v}{\mathrm{d}t} = \gamma(v)^3 \frac{\mathrm{d}v}{\mathrm{d}t} = a_1',$$

as was to be shown. In particular, from the equation of motion under a constant force $\mathbf{F} = F\mathbf{e}_1$ it follows that

$$\frac{\mathbf{F}}{m} = \frac{1}{m}\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \mathbf{a}'.$$

Hence in the hyperbolic motion studied in this section the proper acceleration of the particle is constant and directed along the x_1 axis.

Index