Contents

Most recommended book

Robin Santra, *Einführung in den Lagrange- und Hamilton-Formalismus*, ISBN978-3-662-65358-6, e-book ISBN978-3-662-65359-3, Springer Spektrum Berlin, Heidelberg, DOI https://doi.org/10.1007/978-3-662-65359-3, link in the UHH library collection:
https://kataloge.uni-hamburg.de/DB=1/XMLPRS=N/PPN?PPN=1797763954.

General references on theoretical mechanics

[1] L. D. Landau and E. M. Lifshits, *Mechanics* (3rd ed.) 1976, Pergamon. Vol. 1 of the Course of Theoretical Physics.

[2] H. Goldstein, C. P. Poole and J. L. Safko, *Classical Mechanics*, Addison Wesley, 2002.

[3] V. I. Arnold, *Mathematical methods of classical mechanics*, Springer, 1989.

[4] W. Greiner, *Classical Mechanics*, Springer, 2010.

[5] W. Nolting, *Grundkurs Theoretische Physik 1 Klassische Mechanik und mathematische Vorbereitungen*, Springer, 2013, exists as e-book.

[6] F. Scheck, *Theoretische Physik 1: Mechanik*, Ausgabe 6 Springer, 2013, exists as e-book.

[7] T. Fliessbach, *Mechanik: Lehrbuch zur Theoretischen Physik I*, Springer 2020, exists as e-book.

General references on classical electrodynamics

[1] L. D. Landau and E. M. Lifshits, *Classical Theory of Fields* (3rd ed.) 1971, London: Pergamon. Vol. 2 of the Course of Theoretical Physics.

[2] J. D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, 1975.

[3] A. Sommerfeld, *Electrodynamics*, 1952, Acadenic Press INC., Publishers.

[4] J. Schwinger, L. L. DeRaad (jr.), K. A. Milton, W. Tsai, *Classical Electrodynamics*, 1998, Westview Press.

[5] D. J. Griffiths, *Introduction to Electrodynamics*, (3rd ed.), 2008, Pearson Education, Inc.

[6] W. Greiner, *Classical Electrodynamics*, Springer, 1998.

[7] A. Zangwill, *Modern Electrodynamics*, 2013, Cambridge University Press.

[8] A. Garg, *Classical Electromagnetism in a Nutshell*, 2012, Princeton University Press.

[9] L. D. Landau and E. M. Lifshits, *Electrodynamics of Continuous Media* (2nd ed.) 1984, Pergamon. Vol. 8 of the Course of Theoretical Physics.

[10] I. E. Tamm, *Fundamentals of the theory of electricity*, MIR (1979) (Translated from Russian).

[11] B. Thiedé, *Electromagnetic Field Theory*, Upsilon Books, Uppsala, Sweden.

[12] A. N. Vasil'ev, *Lectures on classical electrodynamics*, 2004. In Russian.

Organization of the course

The lectures[1] are on Mondays 8.30-10.00 and Thursdays 12.30-14.00 at Hörsaal II, Jungiusstraße 9. *The first lecture is on 17.10.2022.*

The schedule of classes is

1. Group A: Monday, 10.15-11.45, Seminar room 1
2. Group B: Monday, 10.15-11.45, Bibliothek INF
3. Group C: Monday, 12.00-13.30, Seminar room 1
4. Group D: Monday, 12.00-13.30, Hörsaal INF
5. Group E: Thursday, 09.00-10.30, Seminar room 1
6. Group F: Thursday, 10.45-12.15, Seminar room 1
7. Group G: Thursday, 10.45-12.15. Ju 11, A203
8. Group H: Thursday, 10.45-12.15. Ju 11, A20

The fist class for groups A, B, C, D is on 24.10.2022, for groups E, F, G, H is on 27.10.2022. In total there are 13 exercises classes.

The grading assistant for all groups: Victor Danescu - vdanescu@physnet.uni-hamburg.de

Seminar leaders:

Thore Posske	thore.posske@physik.uni-hamburg.de
Juan Miguel Nieto Garcia	juan.miguel.nieto.garcia@desy.de
Zhiguo He	zhiguo.he@desy.de
Alejandro Romero-Ros	alejandro.romero.ros@physnet.uni-hamburg.de
Kapil Goswami	kapil.goswami@physnet.uni-hamburg.de
Rohan Srikumar	rohan.srikumar@physnet.uni-hamburg.de

Students tutors:

Ali Bacha Mina	mina.ali.bacha@studium.uni-hamburg.de
Blum Peter	peter.blum@uni-hamburg.de
Burghart Philipp	philipp.burghart@studium.uni-hamburg.de
Ha Sven	sven.ha@studium.uni-hamburg.de
Maaß Robin	robin.maass@studium.uni-hamburg.de
Thaysen Torben	torben.thaysen@studium.uni-hamburg.de

The exam is scheduled on 17.02.2023 (Friday). The second exam is scheduled on 31.03.2021 (Friday).

[1] The number of lectures planned is 28.

Physical constants

Speed of light in vacuum	$299792458\,\mathrm{m/s} \approx 3 \times 10^8\,\mathrm{m/s}$
Planck's constant h	$6.62676 \times 10^{-34}\,\mathrm{kg \cdot m^2/s}$
Planck's constant $\hbar$	$1.05458 \times 10^{-34}\,\mathrm{kg \cdot m^2/s}$
Avogadro's number N_0	$6.02 \times 10^{23}\ \mathrm{mole^{-1}}$
Boltzmann constant k	$1.38065 \times 10^{-23}\,\mathrm{J/K}$
Gas constant $R = kN_0$	$8.314\,\mathrm{J/(K \cdot mole)}$
Charge of electron e	$-1.602177 \times 10^{-19}\,\mathrm{C}$
Rest mass of electron m_e	$9.109383 \times 10^{-31}\,\mathrm{kg} = 0.511$ Mev
Compton wave length of electron λ_e	$2.426310 \times 10^{-12}\,\mathrm{m}$
Rest mass of proton m_p	$1.67 \times 10^{-27}\,\mathrm{kg} = 938$ Mev
Gravitational constant G	$6.673 \times 10^{-11}\,\mathrm{m^3/(kg \cdot s^2)}$
Fine structure constant $\alpha = \frac{e^2}{\hbar c}$	$\approx \frac{1}{137}$

Decimal multiplets of units

Milli 10^{-3}	Kilo 10^3
Micro 10^{-6}	Mega 10^6
Nano 10^{-9}	Giga 10^9
Pico 10^{-12}	Terra 10^{12}
Femto 10^{-15}	

Important scales

Radius of the Earth	$R_\oplus = 6371$ km
Mass of the Earth	$M_\oplus = 5.972 \times 10^{24}$ kg
Age of the Earth	4.5 billion years (4.5×10^9 years)
Distance between the Earth and the Moon	384400 km
Age of the Universe	13.8 billion years

Part I

Theoretical Mechanics

Chapter 1

Newtonian mechanics

> "The initial state of a mechanical system (the totality of positions and velocities of its points at some moment of time) uniquely determines all of its motion.
> It is hard to doubt this fact, since we learn it very early. One can imagine a world in which to determine the future of a system one must also know the acceleration at the initial moment, but experience shows us that our world is not like this."
>
> *Vladimir Arnold "Mathematical Methods of Classical Mechanics"*

Newtonian mechanics studies the motion of a system of point masses in three-dimensional euclidean space. Newton's equations allow one to solve completely a series of important problems in mechanics, including the problem of motion in a central force field.

1.1 Conservative forces

Let us assume that a point particle moves under a force $\vec{F} = \vec{F}(\vec{r}(t))$ that depends on the radius-vector $\vec{r}(t)$ but does not depend on the velocity $\dot{\vec{r}}$.

Definition 1. Work which is done by a force to move particle along the trajectory γ from the position $\vec{r}_1$ to $\vec{r}_2$ is defined as

$$W_{12}(\gamma) = \int_{\vec{r}_1}^{\vec{r}_2} \vec{F} \cdot d\vec{r} = \int_{t_1}^{t_2} \vec{F}(\vec{r}(t)) \cdot \frac{d\vec{r}}{dt} dt \,.$$

Definition 2. A *force is called conservative* if W_{12} does not depend on the path connecting the initial and final points.

We have several characterisations of a conservative force.

- If a force is conservative then the work along any closed path vanishes: $W = 0$. Indeed, $W_{12}(\gamma_1) = W_{12}(\gamma_2) = -W_{21}(\gamma_2)$, so that the work over the closed path $W_{12}(\gamma_1)+W_{21}(\gamma_2) = 0$. This can be written as

$$\oint \vec{F} \cdot d\vec{r} = 0\,.$$

- A force is conservative if and only if there exists a scalar function $U(\vec{r})$ on $\mathbb{R}^3$ such that

$$\vec{F} = -\vec{\nabla} U.$$

  The function U is called *potential*. The minus sign in the above formula is chosen in such a way that the force is directed towards decreasing of the potential.

Let us prove an equivalence of these two definitions. Suppose first that U exists and $\vec{F} = -\vec{\nabla}U$. Let us show that work does not depend on a path. We have

$$W_{12} = \int_{t_1}^{t_2} \vec{F}(\vec{r}(t)) \cdot \frac{d\vec{r}}{dt} dt = -\int_{t_1}^{t_2} \vec{\nabla} U \cdot \frac{d\vec{r}}{dt} dt = -\int_{t_1}^{t_2} \frac{dU}{dt} dt = U(\vec{r}(t_1)) - U(\vec{r}(t_2))\,.$$

Thus, W_{12} does not depend on a path and, therefore, $\vec{F}$ is conservative. Oppositely, assume that W_{12} does not depend on a path. Then, the following function is well-defined

$$U(\vec{r}(t)) = -\int_{\vec{r}_1}^{\vec{r}(t)} \vec{F}(\vec{r}') \cdot d\vec{r}' = -\int_{t_1}^{t} \vec{F}(\vec{r}') \cdot \frac{d\vec{r}'}{dt} dt\,.$$

From here we get

$$\frac{dU}{dt} = \vec{\nabla} U \cdot \frac{d\vec{r}}{dt} = -\vec{F}(\vec{r}) \cdot \frac{d\vec{r}}{dt} \qquad \to \qquad \vec{F} = -\vec{\nabla} U\,.$$

Some remarks are in order. A potential is defined up to an additive constant c, as $\vec{\nabla}(U + c) = \vec{\nabla} U$. Finally, the work from moving at position $\vec{r}_1$ to a position $\vec{r}_2$ is given by a difference of potentials and does not depend on the form of a trajectory

$$W_{12} = -\int_{\vec{r}_1}^{\vec{r}_2} dU = U(\vec{r}_1) - U(\vec{r}_2)\,, \qquad dU = \vec{\nabla} U d\vec{r}.$$

There is a third definition of a conservative fourse.

- Force is conservative if and only if the curl of $\vec{F}$ vanishes

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

The proof goes as follows. If $\vec{F} = -\vec{\nabla}U$, then $\vec{\nabla} \times \vec{F} = -\vec{\nabla} \times \vec{\nabla} U = 0$ by the known formula of the mathematical analysis. Oppositely, taking any closed pass γ in a simply connected region where $\vec{\nabla} \times \vec{F} = 0$, we compute circulation of $\vec{F}$ over this path by using the Stokes theorem

$$\oint_\gamma \vec{F} \cdot d\vec{r} = \int_S \vec{\nabla} \times \vec{F}\, \mathrm{d}S = 0\,,$$

where S is any smooth, orientable surface whose boundary is γ. Thus, circulation vanishes for any closed path, *i.e.* $\vec{F}$ is conservative.

Examples of conservative forces

1. Homogeneous force $\vec{F} = \text{const}$. Then $\vec{F} = -\vec{\nabla}U$, $U = -\vec{F}\cdot\vec{r}$. For instance, the homogeneous gravitational field of the Earth $\vec{F} = -mg\vec{e}_z$, $U = mgz$.

2. Harmonic oscillator
$$\vec{F} = -k\vec{r}, \quad U = \frac{k}{2}\vec{r}\cdot\vec{r}.$$

3. Central field $\vec{F} = f(|\vec{r}|)\cdot\frac{\vec{r}}{|\vec{r}|}$, $\vec{F}||\vec{r}$ is always conservative. For instance, Newton's gravitational force
$$\vec{F} = \gamma\frac{m_1 m_2}{r^3}\,\vec{r}, \quad U = -\gamma\frac{m_1 m_2}{r}\,.$$

1.2 Kinetic and potential energy. Conservation of energy

Consider Newton's equations
$$m\ddot{\vec{r}} = \vec{F}(\vec{r})\,,$$
where we assume that $\vec{F}$ is a conservative force. We multiply this equation with $\dot{\vec{r}}$ and get
$$m\ddot{\vec{r}}\cdot\dot{\vec{r}} = \vec{F}(\vec{r})\cdot\dot{\vec{r}}.$$
The left hand side can obviously be represented in the form
$$\frac{m}{2}\frac{d}{dt}(\dot{\vec{r}}\cdot\dot{\vec{r}}) = \vec{F}\cdot\dot{\vec{r}}. \tag{I.1.1}$$

Definition 3. The quantity
$$T = \frac{m}{2}(\dot{\vec{r}}\cdot\dot{\vec{r}}) = \frac{m}{2}\dot{\vec{r}}^2 = \frac{m}{2}\vec{v}^2$$
is called *kinetic energy* of a particle.

We then integrate (I.1.1)
$$\int_{t_1}^{t_2} dt\frac{d}{dt}T = \int_{t_1}^{t_2}\vec{F}\cdot\dot{\vec{r}}dt\,,$$
so that
$$T(t_2) - T(t_1) = W$$
Since $\vec{F} = -\vec{\nabla}U$, we have
$$T(t_2) - T(t_1) = -\int_{t_1}^{t_2}\vec{\nabla}U\cdot\dot{\vec{r}}dt = -\int_{t_1}^{t_2}\frac{dU}{dt}dt = U(t_1) - U(t_2)\,.$$
From here we deduce that
$$T(t_1) + U(t_1) = T(t_2) + U(t_2)\,.$$
Thus, the quantity $E = T + U$ does not depend on time. It is called *the total energy of the system.* For a conservative force the total energy of a system is conserved along particle's trajectory. This can be checked directly by computing the time derivative of the total energy
$$\frac{dE}{dt} = \frac{dT}{dt} + \frac{dU}{dt} = \frac{m}{2}\frac{d}{dt}\dot{\vec{r}}^2 + \vec{\nabla}U\cdot\dot{r} = m\ddot{\vec{r}}\cdot\dot{r} - \vec{F}\cdot\dot{r} = \underbrace{(m\ddot{\vec{r}} - \vec{F})}_{\text{Newton}}\cdot\dot{r} = 0\,.$$

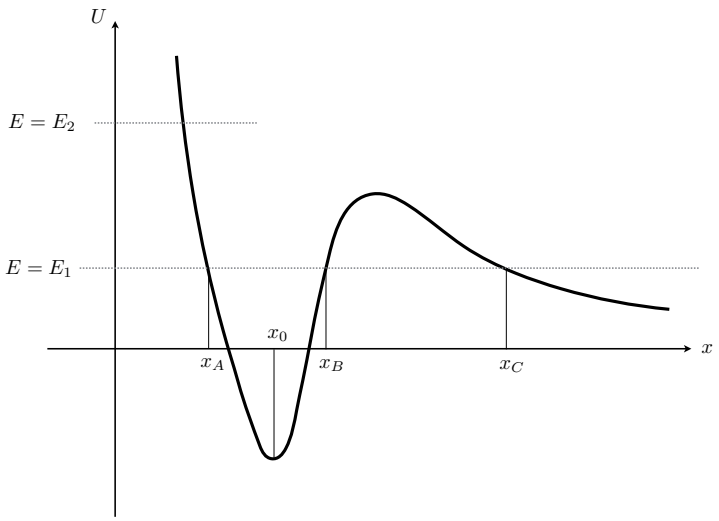

Figure 1.1: Motion in a typical potential.

1.3 One-dimensional systems

In most cases (for example, in the three-body problem) we can neither solve the system of differential equations nor completely describe the behavior of the solutions. Here we consider a few simple but important problems for which Newton's equations can be solved.

One example is provided by one-dimensional systems. For one-dimensional systems Newton's equation is

$$m\ddot{x} = F$$

and $F = F(x)$ is always conservative as there is always exists such U that $F = -\frac{\partial U}{\partial x}$. Energy

$$E = T + U = \frac{m}{2}\dot{x}^2 + U\,, \qquad \frac{dE}{dt} = 0\,.$$

Thus, fixing the value of E, we get

$$\dot{x} = \frac{dx}{dt} = \pm\sqrt{\frac{2}{m}(E-U)}\,.$$

We proceed by separating the differentials

$$\frac{dx}{\sqrt{\frac{2}{m}(E-U(x))}} = dt \qquad \to \qquad \int_{x_1(t_1)}^{x_2(t_2)} \frac{dx'}{\sqrt{\frac{2}{m}(E-U(x'))}} = \int_{t_1}^{t_2} dt = t_2 - t_1\,.$$

Consider a motion in a given potential, see Fig. 1.1. There should be

$$T \geq 0 \qquad \Longrightarrow \qquad E \geq U \qquad \Longrightarrow \qquad E > U(x_0)\,.$$

For $E = E_1$ the particle can only be in regions

$$x_A \leq x \leq x_B \qquad (1)$$

$$x_C \leq x \qquad (2)$$

(1) Periodic motion around equilibrium point x_0**.** Here the turning points are x_A and x_B. The position of equilibrium coinsides with extrema of the potential, *i.e.*

$$\frac{dU}{dx} = 0\,, \qquad \frac{d^2U}{dx^2} = \begin{cases} > 0 \text{ stable equilibrium} \\ < 0 \text{ unstable equilibrium} \end{cases}$$

Then U is minimal then T is maximal and at $x = x_0$ the particle has its highest velocity. On the other hand,

$$E_1 = U(x_A) = U(x_B) \quad \Longrightarrow \quad T = 0 \;\; \dot{x}(x = x_{A,B}) = 0\,.$$

The period of motion is

$$\tau = 2\int_{x_A}^{x_B} \frac{dx}{\sqrt{\frac{2}{m}(E - U(x))}}\,.$$

Consider small oscillations around equilibrium. Expand the potential in Teylor series

$$U(x) = \underbrace{U(x_0)}_{\text{const}} + \underbrace{\frac{dU}{dx}\Big|_{x=x_0}}_{=0}(x - x_0) + \frac{1}{2}\frac{d^2U}{dx^2}\Big|_{x=x_0}(x - x_0)^2 + o\Big((x - x_0)^2\Big).$$

We can always choose $U(x_0)$ to be zero as U is defined up to an additive constant. The second term is zero because the potential is expanded around minimum. Next we set

$$\frac{1}{2}\frac{d^2U}{dx^2}\Big|_{x=x_0}(x - x_0)^2 = \frac{1}{2}m\omega^2(x - x_0)^2$$

with

$$\omega^2 = \frac{1}{m}\frac{d^2U}{dx^2}\Big|_{x=x_0}$$

being the frequency of oscillations. Thus, the frequency of small oscillations around equilibrium is determined by the second derivative of the potential at the equilibrium point multiplied by the inverse of mass. Around equilibrium a potential can always be approximated by a parabola.

(2) Unbounded motion $E = E_2$. In this case particle runs at infinity $x = \infty$.

1.4 Motion of a system of N particles

So far we have considered just one particle and its Newton's equation

$$m\ddot{\vec{r}} = \vec{F}\,.$$

For a system of N particles we have

$$m_i\ddot{\vec{r}}_i = \vec{F}_i\,, \qquad i = 1, \ldots, N\,.$$

This is a system of $3N$ coupled differential equations and its solution requites $2{\cdot}3N = 6N$ integration constants, given by

$$\vec{r}_i(t = 0)\,, \quad \dot{\vec{r}}_i(t = 0)\,.$$

Definition 4. The kinetic energy of a system of N particles is

$$T = \sum_{i=1}^{N} \frac{m_i}{2}\vec{v}_i \cdot \vec{v}_i\,.$$

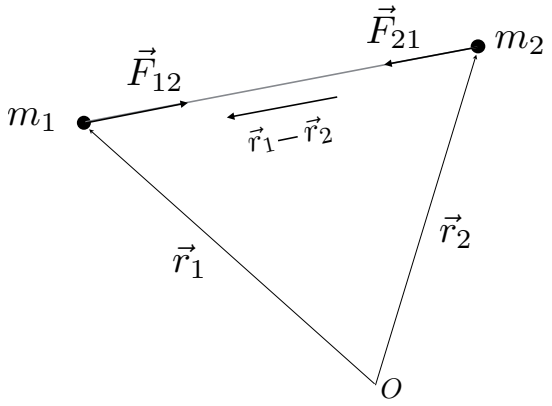

Figure 1.2: For two points central forces are equal in magnitude and act in opposite directions along the straight line joining the points.

We assume that forces $\vec{F}_i$ do not depend on velocities, there is no friction and no explicit time dependence

$$\vec{F}_i = \vec{F}_i(\vec{r}_1, \ldots, \vec{r}_N)\,.$$

Definition 5. Forces $\vec{F}_i$ are called conservative if the path integral

$$\sum_{i=1}^{N} \int_{t_1,\gamma_i}^{t_2} \vec{F}_i \frac{d\vec{r}_i}{dt} dt$$

does not depend on a path.

Forces $\vec{F}_i$ are conservative precisely when there exists a function $U(\vec{r}_1, \ldots, \vec{r}_N)$ such that

$$\vec{F}_i = -\vec{\nabla}_i U(\vec{r}_1, \ldots, \vec{r}_N)\,,$$

where

$$\vec{\nabla}_i = \vec{e}_x \frac{\partial}{\partial x_i} + \vec{e}_y \frac{\partial}{\partial y_i} + \vec{e}_z \frac{\partial}{\partial z_i}\,.$$

The proof is the same as for one particle. This means that all forces $\vec{F}_i$ can be expressed via one scalar function $U(\vec{r}_1, \ldots, \vec{r}_N)$.

Example. Central force between two particles. Newton's equations are

$$\begin{aligned} m_1 \ddot{\vec{r}}_1 &= \vec{F}_1 = -f(|\vec{r}_1 - \vec{r}_2|) \frac{\vec{r}_1 - \vec{r}_2}{|\vec{r}_1 - \vec{r}_2|}\,, \\ m_2 \ddot{\vec{r}}_2 &= \vec{F}_2 = -\vec{F}_1\,, \end{aligned}$$

where $\vec{F}_1 = \vec{F}_{12}$ is a force applied on mass m_1 and $\vec{F}_2 = \vec{F}_{21}$ is a force applied on mass m_2. Here

$$|\vec{r}_1 - \vec{r}_2| = \sqrt{(\vec{r}_1 - \vec{r}_2) \cdot (\vec{r}_1 - \vec{r}_2)} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}\,.$$

We look for the potential $U(|\vec{r}_1 - \vec{r}_2|)$. We have

$$\frac{\partial U}{\partial x_1} = U' \frac{x_1 - x_2}{|\vec{r}_1 - \vec{r}_2|}\,, \qquad \frac{\partial U}{\partial x_2} = -U' \frac{x_1 - x_2}{|\vec{r}_1 - \vec{r}_2|}$$

and similarly for other components $y_{1,2}$ and $z_{1,2}$. Thus, we see that $\vec{\nabla}_1 U = U' \frac{\vec{r}_1 - \vec{r}_2}{|\vec{r}_1 - \vec{r}_2|}$ and, therefore,

$$\vec{F}_1 = -\vec{\nabla}_1 U\,, \quad \text{provided} \quad U' = f\,.$$

Also,

$$\vec{F}_2 = -\vec{\nabla}_2 U = -\vec{F}_1\,.$$

The total energy is

$$E = \frac{m_1}{2}\dot{\vec{r}}_1^2 + \frac{m_2}{2}\dot{\vec{r}}_2^2 + U(|\vec{r}_1 - \vec{r}_2|)\,.$$

For instance for a gravitational field between two masses

$$f = \frac{\gamma m_1 m_2}{|\vec{r}_1 - \vec{r}_2|^2} \quad \Longrightarrow \quad U = -\frac{\gamma m_1 m_2}{|\vec{r}_1 - \vec{r}_2|}\,.$$

Central forces are always conservative.

Definition 6. Momentum of a ith particle is

$$\vec{p}_i = m\dot{\vec{r}}_i = m\vec{v}_i\,.$$

The total momentum is

$$\vec{P} = \sum_{i=1}^{N} \vec{p}_i\,.$$

Definition 7. Center of mass[1] (also known as center of inertia) is defined as the following vector

$$\vec{R} = \frac{\sum\limits_{i=1}^{N} m_i \vec{r}_i}{M}\,, \qquad M = \sum_{i=1}^{N} m_i\,,$$

where M is the total mass.

Some remarks are in order.

1. Evolution equation for the center of mass is determined by the total momentum

$$M\dot{\vec{R}} = \sum_{i=1}^{N} m_i \dot{\vec{r}}_i = \vec{P}\,.$$

2. With the notion of momentum Newton's equations can be written as

$$\dot{\vec{p}}_i = \vec{F}_i\,.$$

Definition 8. Internal forces are the forces that act between N particles. External forces are the forces that act on particles outside the system.

[1] It astronomy it is called *barycenter*.

Examples of internal forces constitute gravitation force and electromagnetic force. In general we have

$$\vec{F}_i = \vec{F}_i'(\vec{r}_i) + \sum_{j=1}^{N} \vec{F}_{ij}\,. \tag{I.1.2}$$

A force $\vec{F}_{ij}$ is the force that jth particle exerts in ith particle.

The total momentum is conserved if the sum of all external forces vanishes. Indeed,

$$\frac{d\vec{P}}{dt} = \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i = \sum_{i=1}^{N} \Big(\vec{F}_i' + \sum_{j=1}^{N} \vec{F}_{ij}\Big) = \sum_{i=1}^{N}\sum_{j=1}^{N} \vec{F}_{ij} = 0\,,$$

as $\vec{F}_{ij} = -\vec{F}_{ji}$.

Definition 9. Angular momentum of a particle is defined as

$$\vec{L} = \vec{r} \times \vec{p} = m\vec{r} \times \dot{\vec{r}}$$

and for $i$the particle

$$\vec{L}_i = m_i \vec{r}_i \times \dot{\vec{r}}_i\,.$$

The total angular momentum is

$$\vec{L} = \sum_{i=1}^{N} \vec{L}_i\,.$$

Definition 10. Torque $\vec{N}$ is defined as

$$\vec{N} = \frac{d\vec{L}}{dt}\,.$$

For a system of particles interacting by central forces the total momentum is conserved. The proof goes as follows. Consider the torque

$$\begin{aligned}
\frac{d\vec{L}}{dt} &= \sum_i m_i \frac{d}{dt}(\vec{r}_i \times \dot{\vec{r}}_i) = \sum_i m_i (\dot{\vec{r}}_i \times \dot{\vec{r}}_i) + \sum_i m_i (\vec{r}_i \times \ddot{\vec{r}}_i) = \sum_i \vec{r}_i \times \vec{F}_i \\
&= \sum_i \sum_j \vec{r}_i \times \vec{F}_{ij} = \frac{1}{2}\sum_i \sum_j \Big[\vec{r}_i \times \vec{F}_{ij} + \vec{r}_j \times \vec{F}_{ji}\Big] = \frac{1}{2}\sum_i \sum_j \Big[\vec{r}_i \times \vec{F}_{ij} - \vec{r}_j \times \vec{F}_{ij}\Big] \\
&= \frac{1}{2}\sum_i \sum_j (\vec{r}_i - \vec{r}_j) \times \vec{F}_{ij} = 0\,,
\end{aligned}$$

because the vector $\vec{r}_i - \vec{r}_j$ is collinear to the vector $\vec{F}_{ij}$, see Fig. 1.2.

1.5 Two-body problem

Two body problem: two particles interacting by means of internal forces.

1.5.1 General solution

Since we have only two particles, we can denote $\vec{F}_{12} = \vec{F}$ and write Newton's equations as

$$\begin{aligned} m_1\ddot{\vec{r}}_1 &= \vec{F}\,, \\ m_2\ddot{\vec{r}}_2 &= -\vec{F}\,. \end{aligned}$$

These formula reflect the *third Newton's law: actio = reactio*. Consider the sum and the difference of these two equations

$$\begin{aligned} m_1\ddot{\vec{r}}_1 + m_2\ddot{\vec{r}}_2 &= 0\,, \\ m_1\ddot{\vec{r}}_1 - m_2\ddot{\vec{r}}_2 &= 2\vec{F}\,. \end{aligned} \tag{I.1.3}$$

Introducing the center of mass

$$\vec{R} = \frac{1}{M}(m_1\vec{r}_1 + m_2\vec{r}_2)\,, \qquad M = m_1 + m_2\,,$$

we see that the first equation in (I.1.3) implies that $\ddot{\vec{R}} = 0$, so that the center of mass performs the uniform and rectilinear motion

$$\vec{R} = \vec{v}_0 t + \vec{R}_0\,.$$

Now we can describe a relative motion around the center of mass. Introduce

$$\vec{r} = \vec{r}_1 - \vec{r}_2\,, \qquad \mu = \frac{m_1 m_2}{m_1 + m_2}\,.$$

Here $\vec{r}$ is called *relative coordinate* and μ is *reduced mass*. We have

$$\ddot{\vec{r}} = \ddot{\vec{r}}_1 - \ddot{\vec{r}}_2 = \frac{\vec{F}}{m_1} + \frac{\vec{F}}{m_2} = \Big(\frac{1}{m_1} + \frac{1}{m_2}\Big)\vec{F} = \frac{m_1 + m_2}{m_1 m_2}\vec{F} = \frac{1}{\mu}\vec{F}\,,$$

that is

$$\mu\ddot{\vec{r}} = \vec{F}\,,$$

which is the only equation left to solve.

The total energy is

$$E = T + U = \frac{1}{2}m_1\dot{\vec{r}}_1^{\,2} + \frac{1}{2}m_1\dot{\vec{r}}_2^{\,2} + U(r)$$

The total angular momentum is

$$L = m_1\vec{r}_1 \times \dot{\vec{r}}_1 + m_2\vec{r}_2 \times \dot{\vec{r}}_2\,.$$

One can show that these quantities can be spit into separately conserved quantities associated with the center of mass and the ones corresponding to the relative motion. Namely,

1. $E = E_{\rm CM} + E_{\rm rel}\,,$ where $E_{\rm CM} = \frac{1}{2}M\dot{\vec{R}}^2\,, \quad E_{\rm rel} = \frac{1}{2}\mu\dot{\vec{r}}^{\,2} + U(r)\,.$
2. $\dot{E}_{\rm CM} = \dot{E}_{\rm rel} = 0\,.$
3. $\vec{L} = \vec{L}_{\rm CM} + \vec{L}_{\rm rel}\,,$ where $\vec{L}_{\rm CM} = M\vec{R} \times \dot{\vec{R}}\,, \quad \vec{L}_{\rm rel} = \mu\vec{r} \times \dot{\vec{r}}\,.$
4. $\dot{\vec{L}}_{\rm CM} = \dot{\vec{L}}_{\rm rel} = 0\,.$

The proof of these statements is in **Tutorial II**. What is especially important is that $\vec{L}_{\rm rel} = \mu\vec{r}\times\dot{\vec{r}}$ is conserved during the time evolution

$$\frac{\vec{L}_{\rm rel}}{dt} = 0\,,$$

i.e. $\vec{L}_{\rm rel}$ remains constant all the time. This means that the relative motion happens in the plane orthogonal to the vector $\vec{L}_{\rm rel}$. This follows from the fact that

$$\vec{L}_{\rm rel}\cdot\vec{r} = \mu(\vec{r}\times\dot{\vec{r}})\cdot\vec{r} = 0\,,$$
$$\vec{L}_{\rm rel}\cdot\dot{\vec{r}} = \mu(\vec{r}\times\dot{\vec{r}})\cdot\dot{\vec{r}} = 0\,,$$

and meaning that the vectors $\vec{r}$ and $\dot{\vec{r}}$ lie in the plane orthogonal to $\vec{L}_{\rm rel}$.[2] Thus, without loss of generality, we can choose a coordinate system such that $\vec{L}_{\rm rel}$ has only one non-trivial component $\vec{L}_{\rm rel} = \ell\vec{e}_z$, $\dot{\ell} = 0$ and, therefore $z = 0 = \dot{z}$. Then the motion happens in the two-dimensional plane (x, y), in particular

$$\vec{L}_{\rm rel} = \mu(x\dot{y} - y\dot{x})\vec{e}_z \quad\Longrightarrow\quad \ell = \mu(x\dot{y} - y\dot{x})\,.$$

As we have just discussed, ℓ is an integral of motion in addition to $E_{\rm rel}$.

On the two-dimensional plane we introduce polar coordinates

$$x = r\cos\varphi\,, \qquad y = r\sin\varphi\,,$$

and compute ℓ. We have

$$\frac{\ell}{\mu} = r\cos\varphi(\cancel{\dot{r}\sin\varphi} + r\cos\varphi\dot{\varphi}) - r\sin\varphi(\cancel{\dot{r}\cos\varphi} - r\sin\varphi\dot{\varphi}) = r^2\dot{\varphi}\,.$$

Thus, during the time evolution of our two-body system the quantity $\ell = \mu r^2\dot{\varphi}$ remains constant. This is nothing else but the conservation law of angular momentum, as originally discovered by Kepler through his observations of the motion of Mars. The quantity ℓ has a simple geometric meaning. Kepler introduced the *sectorial velocity* v_s:

$$v_s = \lim_{\Delta t\to 0}\frac{\Delta A}{\Delta t}\,,$$

where ΔA is the area of an infinitesimal sector swept by the radius-vector q for time Δt:

$$\Delta A = \frac{1}{2}r\cdot r\dot{\varphi}\Delta t + \mathcal{O}(\Delta t^2) \approx \frac{1}{2}r^2\dot{\varphi}\Delta t\,.$$

This equation expresses the second law by Kepler: *in equal times the radius vector sweeps out equal areas, hence the sectorial velocity is constant.* This is one of the formulations of the conservation law of angular momentum.

Now we rewrite the relative energy in polar coordinates. For simplicity we redefine the notation $E_{\rm rel} \equiv E$. We have

$$\begin{aligned} E_{\rm rel} = E &= \frac{1}{2}\mu\dot{\vec{r}}^2 + U(r) = \frac{1}{2}\mu(\dot{x}^2 + \dot{y}^2) + U(r) \\ &= \frac{1}{2}\mu\Big[(\dot{r}\cos\varphi - r\sin\varphi\dot{\varphi})^2 + (\dot{r}\sin\varphi + r\cos\varphi\dot{\varphi})^2\Big] + U(r)\,. \end{aligned}$$

From here we get

$$E = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) + U(r)\,.$$

[2] Three points (lying not on one line) uniquely define a plane that passes through them.

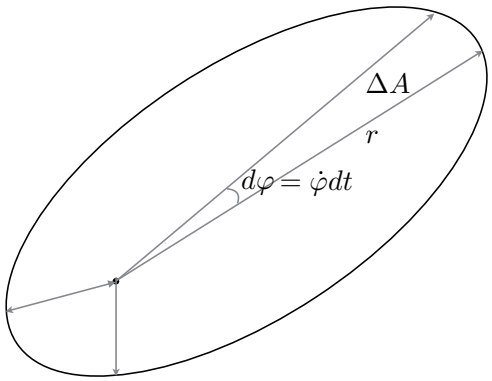

Figure 1.3: The 2nd Kepler law: the sectorial velocity $v_s = \frac{1}{2}\frac{\ell}{\mu}$ is constant.

Now the conservation law of ℓ can be used to exclude $\dot{\varphi}$, namely,

$$\dot{\varphi} = \frac{\ell}{\mu r^2}$$

so that the energy integral reads as

$$E = \frac{\mu}{2}\Big(\dot{r}^2 + \frac{\ell^2}{\mu^2 r^2}\Big) + U(r) = \frac{1}{2}\mu\dot{r}^2 + \frac{\ell^2}{2\mu r^2} + U(r)\,.$$

This can be written as

$$E = \frac{1}{2}\mu\dot{r}^2 + U_{\text{eff}}\,,$$

where we have introduced an effective potential

$$U_{\text{eff}} = \frac{\ell^2}{2\mu r^2} + U(r)\,.$$

Thus, the problem is reduced to a one-dimensional one! Initially, we started with 6 degrees of freedom of two bodies. Separating the motion of the center of mass, we reduced the number of degrees of freedom to 3. Conservation of angular momentum left us with two degrees of freedom on a plane and, simultaneously, allowed us to remove the angle variable leaving us with just one-dimensional problem for r.

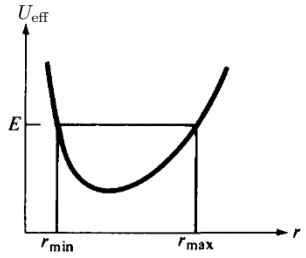

Figure 1.4: Effective potential.

Time evolution. Fixing the value of energy E, we separate the variables as was discussed above

$$dt = \pm\frac{dr}{\sqrt{\frac{2}{\mu}(E - U_{\text{eff}}(r))}}$$

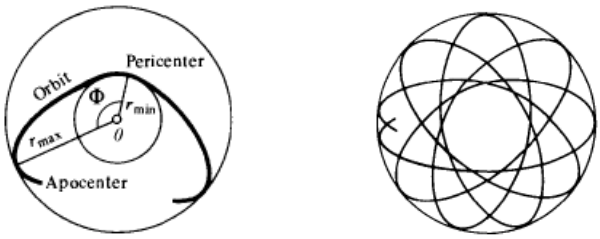

Figure 1.5: The left picture shows an orbit of a point in a central field. The right picture shows an orbit everywhere dense in an annulus.

so that the time evolution $r = r(t)$ will be determined from

$$t - t_0 = \pm \int_{r_0}^{r} \frac{dr'}{\sqrt{\frac{2}{\mu}(E - U_{\text{eff}}(r'))}} \,.$$

Form of the trajectory. One can also derive the form of the trajectory $r = r(\varphi)$ directly, *i.e.* without appealing to first finding $r(t)$ and $\varphi(t)$. Indeed, along the trajectory we have

$$\frac{dr(\varphi)}{dt} = \frac{dr}{d\varphi}\dot{\varphi} \,,$$

so that

$$\pm\sqrt{\frac{2}{\mu}(E - U_{\text{eff}}(r))} = \frac{dr}{d\varphi}\frac{\ell}{\mu r^2} \,.$$

From here

$$d\varphi = \pm \frac{dr}{\frac{\sqrt{2\mu}}{\ell}\, r^2 \sqrt{E - U_{\text{eff}}(r)}} \,,$$

that yields upon integration

$$\varphi - \varphi_0 = \pm \frac{\ell}{\sqrt{2\mu}} \int_{r_0}^{r} \frac{dr'}{r'^2 \sqrt{E - U_{\text{eff}}(r')}} \,.$$

This is an equation defining the trajectory for given values of E and ℓ.

These considerations complete the general solution of the two-body problem. Further progress relies on specification of the potential $U(r)$.

Qualitative behaviour of orbits. The qualitative behaviour of orbits can be understood from the graph of the effective potential U_{eff}, see Fig. 1.4. All orbits corresponding to the given E and ℓ

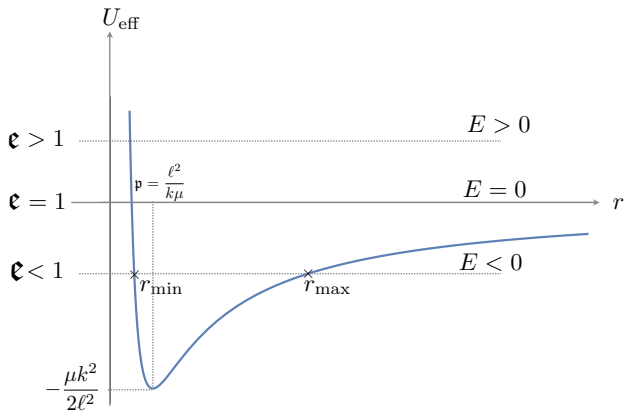

Figure 1.6: Effective potential of the Kepler problem.

are in the region $U_{\text{eff}}(r) \le E$. On the boundaries of this region $U_{\text{eff}} = E$, so that $\dot{r} = 0$. However, the point continues to move because at these points $\dot{\varphi} \neq 0$. The inequality $U_{\text{eff}}(r) \le E$ gives one or several annular regions in the plane

$$0 \le r_{\min} \le r \le r_{\max} \le \infty \,.$$

If $r_{\min} \le r \le r_{\max} < \infty$, then the motion is bounded and takes place inside the annulus between the circles with radius $r_{\min}$ and $r_{\max}$. The shape of the orbit is shown on the left picture of Fig. 1.5. The angle φ varies monotonically, while r oscillates between $r_{\min}$ and $r_{\max}$. The points where $r = r_{\min}$ are called pericentral, and where $r = r_{\max}$, apocentral (if the center is the earth – perigee and apogee; if it is the sun – perihelion and aphelion; if it is the moon – perilune and apolune).

In general the orbit is not closed. The angle Φ between successive pericenters and apocenters is given by

$$\Phi = \frac{\ell}{\sqrt{2\mu}} \int_{r_{\min}}^{r_{\max}} \frac{dr'}{r'^2 \sqrt{E - U_{\text{eff}}(r')}} \,.$$

The angle between two successive pericenters is twice big. The orbit is closed if 2Φ is commensurable (rationally comparable) with 2π, *i.e.* if $2\Phi = 2\pi \frac{m}{n}$, where m, n are integers. Then after repetition n times the period of time between reaching two successive apocenters, the radius-vector will make m full revolutions and return to its original position, *i.e.* the trajectory will close. If 2Φ is not commensurable with 2π then the orbit is everywhere dense on the annulus. If $r_{\min} = r_{\max}$, *i.e.* E is at minimum of U_{eff}, then the annulus degenerates into a circle, which is also the orbit.

Determination of a central potential for which *all* bounded orbits are closed is known as the *Bertrand problem*. Remarkably, it appears that there are only two potentials for which all bounded orbits are closed. They are

$$U = kr^2 \,, \quad k > 0$$

and

$$U = -\frac{k}{r} \,, \quad k > 0 \,.$$

The first is the spacial harmonic oscillator potential, the second one is Newton's gravitational potential. For $k > 0$ the potential is attractive. Motion in Newton's gravitational potential is the famous Kepler problem of planetary motion, which we now consider.

1.5.2 Kepler problem

Equation for trajectories. Consider Newton's potential $U(r) = -\frac{k}{r}$, $k > 0$. The equation defining the trajectory will be then

$$\varphi(r) = \frac{\ell}{\sqrt{2\mu}} \int^r \frac{dr'}{r'^2 \sqrt{E - U_{\text{eff}}}} = \frac{\ell}{\sqrt{2\mu}} \int^r \frac{dr'}{r'^2 \sqrt{E + \frac{k}{r'} - \frac{\ell^2}{2\mu r'^2}}} \,.$$

This integral can be computed explicitly. To this end, we make the change of variables

$$r' = \frac{1}{x} \,, \quad dr' = -\frac{dx}{x^2}$$

and get

$$\varphi(r) = -\int^{\frac{1}{r}} \frac{dx}{\sqrt{\frac{2\mu E}{\ell^2} + \frac{2\mu k}{\ell^2} x - x^2}} \,.$$

The right hand side here is the standard integral

$$\int \frac{dx}{\sqrt{\gamma + 2\beta x - x^2}} = -\arccos\left(\frac{x - \beta}{\sqrt{\beta^2 + \gamma}}\right), \qquad \beta^2 + \gamma > 0 \,.$$

We therefore identify

$$\beta = \frac{\mu k}{\ell^2} \,, \qquad \gamma = \frac{2\mu E}{\ell^2} \,.$$

In particular, $\beta^2 + \gamma = \frac{\mu^2 k^2}{\ell^4} + \frac{2\mu E}{\ell^2} > 0$, *i.e.* $E > -\frac{\mu k^2}{2\ell^2}$, where $-\frac{\mu k^2}{2\ell^2}$ is the minimum of the effective potential, see Fig. (1.6). In this way we get

$$\varphi(r) = \arccos\left(\frac{\frac{1}{r} - \frac{\mu k}{\ell^2}}{\sqrt{\frac{\mu^2 k^2}{\ell^4} + \frac{2\mu E}{\ell^2}}}\right) = \arccos\left(\frac{\frac{\ell^2}{\mu k r} - 1}{\sqrt{1 + \frac{2E\ell^2}{\mu k^2}}}\right). \tag{I.1.4}$$

Defining the following quantities

$$\mathfrak{p} = \frac{\ell^2}{k\mu} \,, \qquad \mathfrak{e} = \sqrt{1 + \frac{2E\ell^2}{\mu k^2}} \,,$$

we can write the equation for the orbit in the form

$$r = \frac{\mathfrak{p}}{1 + \mathfrak{e}\cos\varphi} \,.$$

This is the so-called *focal equation of a conic section*, see Fig. 1.7. When $\mathfrak{e} < 1$, *i.e.* $E < 0$, the conic section is an ellipse. The number $\mathfrak{p}$ is called the *parameter* of the ellipse and $\mathfrak{e}$ the *eccentricity*. The motion is bounded for $E < 0$, as is seen from the graph of the effective potential Fig. 1.6.

Note that in (I.1.4) the integration constant was set to zero which corresponds to the choice of the reference point for the angle φ at the pericenter: $\varphi(r_{\min}) = 0$, where

$$r_{\min} = \frac{k}{2E}\left(-1 + \sqrt{1 + \frac{2E\ell^2}{\mu k^2}}\right) = \frac{k}{2E}(\mathfrak{e} - 1) = \frac{k}{2E}\frac{(\mathfrak{e}^2 - 1)}{1 + \mathfrak{e}} = \frac{\mathfrak{p}}{1 + \mathfrak{e}} \,.$$

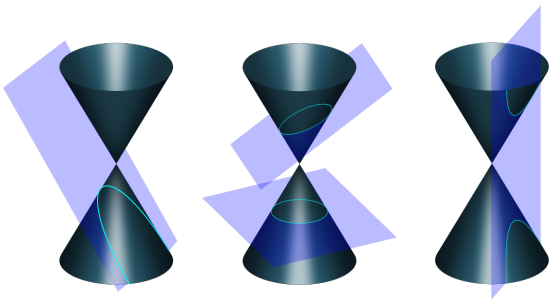

Figure 1.7: Three conic sections: parabola $e = 1$, ellipse $e < 1$ and hyperbola $e > 1$.

Elliptic orbits and Kepler's laws. For an elliptic orbit corresponding $\mathfrak{e} < 1$, we further notice three distinguished points

$$\begin{aligned} \varphi &= 0: \quad r = \frac{\mathfrak{p}}{1+\mathfrak{e}} = r_{\min}\,, \\ \varphi &= \frac{\pi}{2}: \quad r = \mathfrak{p}\,, \\ \varphi &= \pi: \quad r = \frac{\mathfrak{p}}{1-\mathfrak{e}} = r_{\max}\,, \end{aligned}$$

which give us certain geometric intuition about various quantities appearing in the description of elliptic orbits, see Fig. 1.8. In particular, the major semi-axis a is determined as

$$2a = \frac{\mathfrak{p}}{1-\mathfrak{e}} + \frac{\mathfrak{p}}{1+\mathfrak{e}} = \frac{2\mathfrak{p}}{1-\mathfrak{e}^2}\,.$$

We also have

$$c = a - \frac{\mathfrak{p}}{1+\mathfrak{e}} = \frac{\mathfrak{e}\mathfrak{p}}{1-\mathfrak{e}^2}\,.$$

From the last two equations the eccentricity can be determined via the major semi-axes as

$$\mathfrak{e} = \frac{c}{a} = \frac{\sqrt{a^2-b^2}}{a} = \sqrt{1-\frac{b^2}{a^2}}\,.$$

We can now formulate the Kepler laws:

1) The first law: planets describe ellipses with the Sun at one focus.

2) The second law: the sectorial velocity is constant.

3) The third law: the period of revolution around an elliptical orbit depends only on the size of the major semi-axes. The squares of the revolution periods of two planets on different elliptical orbits have the same ratio as the cubes of their major semi-axes.

The third law follows from the following considerations. Let T be a revolutionary period and A be the area swept out by the radius vector over the period. An ellipse with the semi-axes a and b encompasses the area

$$A = \pi ab = \pi a^2\sqrt{1-\mathfrak{e}^2} = \pi\frac{\mathfrak{p}^2}{(1-\mathfrak{e}^2)^2}\sqrt{1-\mathfrak{e}^2} = \pi\frac{\mathfrak{p}^2}{(1-\mathfrak{e}^2)^{\frac{3}{2}}} = \frac{\pi k\ell}{\mu^{1/2}(\sqrt{2|E|})^3}\,,$$

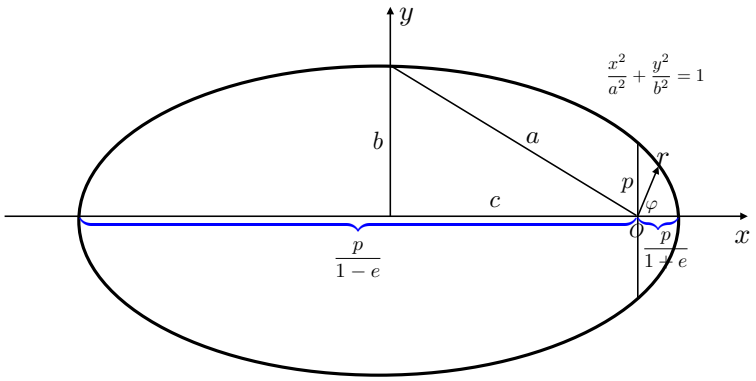

Figure 1.8: Keplerian ellipse: semi-axes a, b, parameter p and eccentricity e.

where we have taken into account that

$$a = \frac{\mathfrak{p}}{1-\mathfrak{e}^2} = \frac{k}{2|E|}\,. \tag{I.1.5}$$

On the other hand, since the sectorial velocity v_s is constant, we have

$$\int_0^T v_s = \int_0^T \mathrm{d}t \frac{dA}{dt} = A\,, \quad \to \qquad v_s T = \frac{\ell}{2\mu} T = A\,,$$

that is,

$$T = \frac{2\mu A}{\ell} = \frac{2\pi k \mu^{1/2}}{(\sqrt{2|E|})^3} = 2\pi \Big(\frac{\mu}{k}\Big)^{1/2} a^{3/2} \quad \Longrightarrow \quad T^2 \sim a^3\,.$$

It is interesting that according to (I.1.5) the total energy depends only on the major semi-axis a and it is the same for the whole set of elliptical orbits from a circle of radius a to a line segment of length $2a$. The value of the second semi-axis depends on the angular momentum.[3]

Kepler's three laws of planetary motion, published around 1610, were the result of his pioneering analysis of observations and laid the groundwork for Newton's great advances. The second law, the conservation of sectorial velocity, is a general theorem for central force motion. However, the first – that planets move in elliptical orbits around the Sun at one focus – and the third are restricted specifically to the inverse-square law of force. Eccentricities of planets in the solar system are rather small, see Table 1.11, so the planet's orbits are almost circular.

Hyperbolic orbits. For $E \geq 0$ the motion is infinite. If $E > 0$, them $\mathfrak{e} > 1$ and the trajectory is a hyperbola that snakes around the center of the field (focus). The distance to pericenter is

$$r_{\min} = \frac{\mathfrak{p}}{\mathfrak{e}+1} = a(\mathfrak{e}-1)\,,$$

where

$$a = \frac{\mathfrak{p}}{\mathfrak{e}^2-1} = \frac{k}{2E}\,, \qquad b = \sqrt{c^2-a^2} = a\sqrt{\mathfrak{e}^2-1} = \frac{\mathfrak{p}}{\sqrt{\mathfrak{e}^2-1}}$$

[3] An elementary derivation of the 3rd Kepler law can be done by assuming that the orbits are circular. In that case the centrifugal force mv^2/r, where r is the radius of the orbit and v is the linear velocity should be balanced by the gravitational force: $\frac{mv^2}{r} = \gamma\frac{mM}{r^2}$, *i.e.* $v = \sqrt{\gamma\frac{M}{r}}$. The period of revolution is then $T = \frac{2\pi r}{v} = \frac{2\pi}{\sqrt{\gamma M}} r^{3/2}$, hence the 3rd Kepler law $T \sim r^{3/2}$.

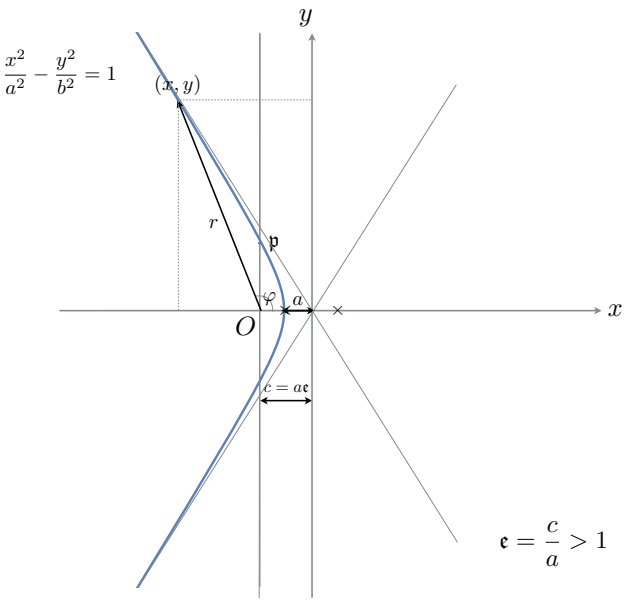

Figure 1.9: For $E \geq 0$ the motion goes over a hyperbolic trajectory. Here O is one of the two foci and is a force center.

are the semi-axes of the hyperbola (a is the major semi-axis). To demonstrate that we indeed obtain the hyperbola, we compute x and y and verify that the obey the canonical equation of the hyperbola

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\,.$$

From Fig. 1.9, we read off

$$x = -r\cos\varphi + a\mathfrak{e} = -\frac{\mathfrak{p}\cos\varphi}{1+\mathfrak{e}\cos\varphi} + \frac{\mathfrak{p}\mathfrak{e}}{\mathfrak{e}^2-1} = \frac{\mathfrak{p}}{\mathfrak{e}^2-1}\frac{\mathfrak{e}+\cos\varphi}{1+\mathfrak{e}\cos\varphi}\,.$$

and

$$y = r\sin\varphi = \frac{\mathfrak{p}\sin\varphi}{1+\mathfrak{e}\cos\varphi}\,.$$

Thus, after some algebraic manipulations we arrive at

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{(\mathfrak{e}+\cos\varphi)^2}{(1+\mathfrak{e}\cos\varphi)^2} - \frac{(\mathfrak{e}^2-1)\sin^2\varphi}{(1+\mathfrak{e}\cos\varphi)^2} = 1\,,$$

confirming that the trajectory is a hyperbola.

In the case $E = 0$ the eccentricity $\mathfrak{e} = 1$ and a particle moves over parabola with the distance to the pericentum $r_{\min} = \mathfrak{p}/2$. This case is realised when a particle starts its motion at infinite with zero initial velocity.

Laplace-Runge-Lenz vector. It turns out that the Kepler problem with its specific potential (with any sign of k) admits one more non-trivial conserved quantity that is absent for a generic central potential: *the Laplace-Runge-Lenz vector*[4] $\vec{A}$

$$\vec{A} = \dot{\vec{r}} \times \vec{L}_{\rm rel} - k\frac{\vec{r}}{r}\,.$$

Its conservation is proved in **Tutorial IV**. The existence of the Laplace-Runge-Lenz vector suggests that the Kepler problem has a hidden symmetry and this turns out to be $\mathfrak{so}(4) \simeq \mathfrak{su}(2) \oplus \mathfrak{su}(2)$.

[4]It was actually discovered by Jacob Hermann (16 July 1678–11 July 1733), a mathematician who worked on problems in classical mechanics.

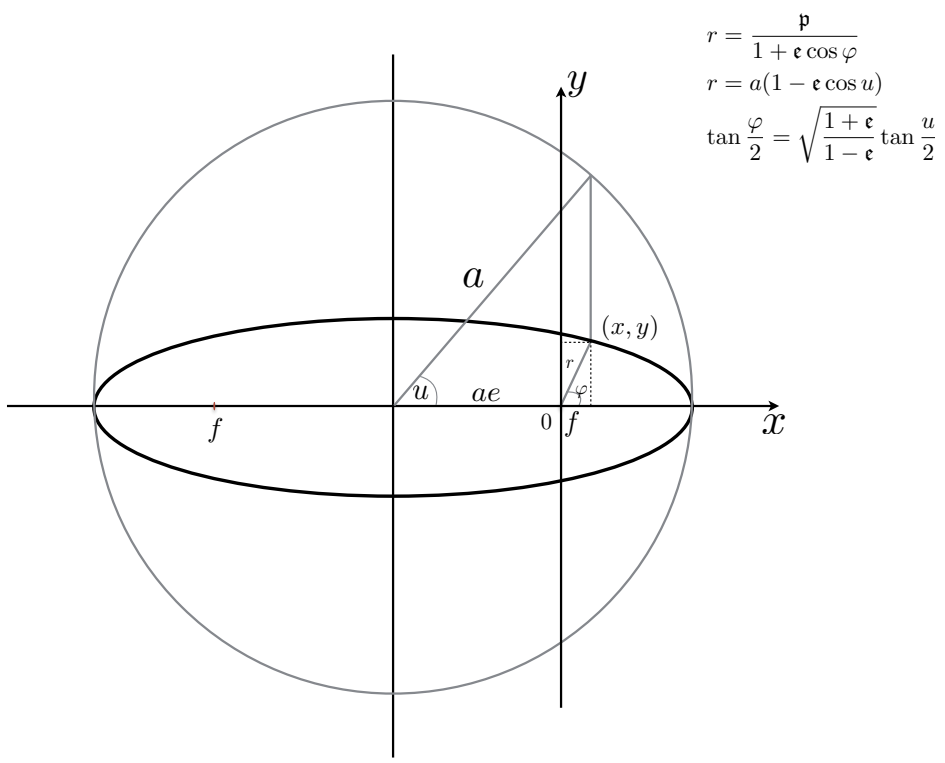

Figure 1.10: True anomaly φ, eccentric anomaly u and their relation.

Solving Kepler's equation. The last step in solving Kepler's problem is to determine the evolution laws along the already established elliptic orbit.[5] This is most easily done in terms of the so-called *eccentric anomaly* u, rather than in terms of the true anomaly φ. Looking at Fig. 1.10, it is easy to see that

$$r\cos\varphi + a\mathfrak{e} = a\cos u\,,$$

where $c = a\mathfrak{e}$ is half the distance between two foci. Expressing the product $r\cos\varphi$ from the equation $r = \mathfrak{p}/(1+\mathfrak{e}\cos\varphi)$, we get

$$r = a(1-\mathfrak{e}\cos u)\,.$$

Now we have two expressions for the radius r

$$\frac{\mathfrak{p}}{1+\mathfrak{e}\cos\varphi} = r = a(1-\mathfrak{e}\cos u)\,,$$

from which we can find the relation between the true and eccentric anomalies, namely

$$\frac{1-\mathfrak{e}^2}{1+\mathfrak{e}\cos\varphi} = 1-\mathfrak{e}\cos u \quad \rightarrow \quad \cos\varphi = \frac{\mathfrak{e}-\cos u}{\mathfrak{e}\cos u - 1}$$

and further,[6]

$$\tan\frac{\varphi}{2} = \sqrt{\frac{1+\mathfrak{e}}{1-\mathfrak{e}}}\tan\frac{u}{2}\,.$$

[5] We do it here for an elliptic orbit only.

[6] It is convenient to express $\cos\varphi$ and $\cos u$ with the tangent of the corresponding half-argument, *cf.* formulae on Fig. 1.10.

Introducing the coordinate system as shown in Fig. 1.10, we can express the cartesian coordinates (x, y) of a point on the orbit via the eccentric anomaly u. Namely,

$$x = a\cos u - a\mathfrak{e} = a(\cos u - \mathfrak{e})\,. \tag{I.1.6}$$

As to y, we have

$$y = r\sin\varphi = a(1-\mathfrak{e}\cos u)\frac{2\tan\frac{\varphi}{2}}{1+\tan^2\frac{\varphi}{2}} = a(1-\mathfrak{e}\cos u)\frac{2\sqrt{\frac{1+\mathfrak{e}}{1-\mathfrak{e}}}\tan\frac{u}{2}}{1+\frac{1+\mathfrak{e}}{1-\mathfrak{e}}\tan^2\frac{u}{2}}\,.$$

Further simplification of the right hand side of the last formula gives

$$y = a\sqrt{1-\mathfrak{e}^2}\,\sin u\,. \tag{I.1.7}$$

Now we recall that $\ell = \mu r^2\dot\phi$ is an integral of motion. We can rewrite it in cartesian coordinates

$$\ell = \mu r^2\dot\phi = \mu(x\dot y - y\dot x)\,.$$

Substituting here the formulae (I.1.6) and (I.1.7), we get the following differential equation

$$\ell = \mu a^2\sqrt{1-\mathfrak{e}^2}(1-\mathfrak{e}\cos u)\frac{du}{dt}\,. \tag{I.1.8}$$

Separating the variables

$$\frac{\ell}{\mu a^2\sqrt{1-\mathfrak{e}^2}}\,dt = du - \mathfrak{e}d\sin u$$

and integrating, one finds the famous Kepler's equation

$$u - \mathfrak{e}\sin u = n(t-t_0) = \zeta\,, \qquad n \equiv \frac{\ell}{\mu a^2\sqrt{1-\mathfrak{e}^2}}\,. \tag{I.1.9}$$

The function ζ depends linearly on time and is called *mean* anomaly.

Equation (I.1.9) can be solved in terms of power series and the solution is given by

$$u(\mathfrak{e},\zeta) = \zeta + 2\sum_{m=1}^{\infty}\frac{J_m(m\mathfrak{e})}{m}\sin m\zeta\,. \tag{I.1.10}$$

Here $J_m(z)$ is the Bessel function and the series converges for $\mathfrak{e} < 1$.

The derivation of (I.1.10) is carried out as follows. It is clear that the solution is a function $u \equiv u(\mathfrak{e},\zeta)$ and from the equation is clear that if we change $\zeta \to \zeta + 2\pi$ and simultaneously change $u \to u + 2\pi$ then the equation will remain invariant. Thus, $u(\zeta) - \zeta$ is periodic in ζ with period 2π, therefore it can be expanded in a Fourier series. Differentiating (I.1.9), we obtain

$$\frac{du}{d\zeta} = \frac{1}{1-\mathfrak{e}\cos u} \equiv f(\zeta)\,,$$

where $f(\zeta)$ must be periodic and, therefore, admits an expansion

$$f(\zeta) = \frac{1}{2}a_0 + \sum_{m=1}^{\infty}a_m\cos m\zeta + \sum_{m=1}^{\infty}b_m\sin m\zeta\,,$$

where the corresponding coefficients are

$$a_0 = \frac{1}{\pi}\int_0^{2\pi}f(\zeta)d\zeta\,, \qquad a_m = \frac{1}{\pi}\int_0^{2\pi}f(\zeta)\cos m\zeta d\zeta\,, \qquad b_m = \frac{1}{\pi}\int_0^{2\pi}f(\zeta)\sin m\zeta d\zeta\,.$$

Celestial Object	Mean Distance from Sun (million km)	Period of Revolution (d=days) (y=years)	Period of Rotation at Equator	Eccentricity of Orbit	Equatorial Diameter (km)	Mass (Earth = 1)	Density (g/cm^3)
SUN	—	—	27 d	—	1,392,000	333,000.00	1.4
MERCURY	57.9	88 d	59 d	0.206	4,879	0.06	5.4
VENUS	108.2	224.7 d	243 d	0.007	12,104	0.82	5.2
EARTH	149.6	365.26 d	23 h 56 min 4 s	0.017	12,756	1.00	5.5
MARS	227.9	687 d	24 h 37 min 23 s	0.093	6,794	0.11	3.9
JUPITER	778.4	11.9 y	9 h 50 min 30 s	0.048	142,984	317.83	1.3
SATURN	1,426.7	29.5 y	10 h 14 min	0.054	120,536	95.16	0.7
URANUS	2,871.0	84.0 y	17 h 14 min	0.047	51,118	14.54	1.3
NEPTUNE	4,498.3	164.8 y	16 h	0.009	49,528	17.15	1.8
EARTH'S MOON	149.6 (0.386 from Earth)	27.3 d	27.3 d	0.055	3,476	0.01	3.3

Figure 1.11: Solar system data.

From Kepler's equation one can see that $u(-\zeta) = -u(\zeta)$ and, therefore, the derivative $\frac{du}{d\zeta}$ is an even function of ζ. Consequently, the coefficients b_m vanish. Thus, we have

$$\begin{aligned}\frac{du}{d\zeta} &= \frac{1}{2\pi}\int_0^{2\pi}\frac{d\zeta}{1-\mathfrak{e}\cos u} + \sum_{m=1}^{\infty}\frac{\cos m\zeta}{\pi}\int_0^{2\pi}\frac{\cos m\zeta d\zeta}{1-\mathfrak{e}\cos u} \\ &= \frac{1}{2\pi}\int_0^{2\pi} du + \sum_{m=1}^{\infty}\frac{\cos m\zeta}{\pi}\int_0^{2\pi}\cos m(u-\mathfrak{e}\sin u)du\,.\end{aligned}$$

Now one needs to recall the definition of the Bessel function

$$J_m(z) = \frac{1}{2\pi}\int_0^{2\pi}\cos(mx - z\sin x)\mathrm{d}x = \sum_{k=0}^{\infty}\frac{(-1)^k(z/2)^{m+2k}}{k!(m+k)!}\,.$$

Therefore,

$$\frac{du}{d\zeta} = 1 + 2\sum_{m-1}^{\infty} J_m(m\mathfrak{e})\cos m\zeta\,.$$

It remains to integrate this formula with respect to ζ and get (I.1.10).

1.5.3 Scattering

So far we have discussed the behaviour of two particles that interacted through an attractive two-body central potential. We discussed the nature of bounded and unbounded orbits which might

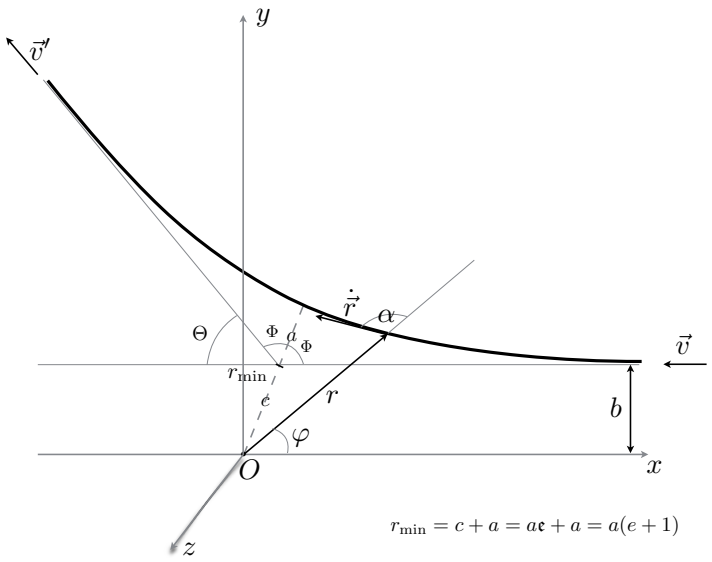

Figure 1.12: Scattering in a repulsive potential. The impact parameter b replaces angular momentum ℓ. The scattering angle Θ is a function of b.

appear, and quantitatively described the orbits in the case of the Kepler problem. However, we can also consider the case of a repulsive central potential. In this case, the two particles will not orbit each other - they will at most approach each other, before the repulsive potential causes them to move away from each other, and never meet again. Such a behaviour is typically referred to as scattering, see Fig. 1.12. This type of scenario is incredibly important in a wide range of physics, especially in condensed matter systems (where neutrons being scattered of a material reveal information about the microscopic details of the material) and in high energy physics (where scattering elementary particles against each other can reveal information about the existence of new fundamental particles). It is scattering of α particles bombarding a gold foli that led Rutherford to the discovery of the atomic kernel.

We consider elastic scattering which means that in the scattering process particles do not change. Denote by $\vec{v}_1, \vec{v}_2$ the particle velocities before the interaction and $\vec{v}_1', \vec{v}_2'$ after. We have

- Conservation of energy

$$\frac{1}{2}m_1\vec{v}_1^2 + \frac{1}{2}m_2\vec{v}_2^2 = \frac{1}{2}m_1\vec{v}_1'^2 + \frac{1}{2}m_2\vec{v}_2'^2\,.$$

- Conservation of momentum

$$m_1\vec{v}_1 + m_2\vec{v}_2 = m_1\vec{v}_1' + m_2\vec{v}_2'\,.$$

In total there are 4 equations for 6 unknowns, which are $\vec{v}_1'$ and $\vec{v}_2'$. This leaves us with $6 - 4 = 2$ unknowns. However, the angular momentum is conserved, therefore, scattering happens on a plane, this leaves only one unknown which is *scattering angle*.

To proceed, we separate the relative motion from the motion of the center of mass. We have

$$E_{\rm rel} = \frac{1}{2}\mu\dot{\vec{r}}^2 + U(r)\,, \qquad \dot{E}_{\rm rel} = 0\,, \quad \lim_{t\to\pm\infty} U(r(t)) = 0\,.$$

Requiring the potential to vanish at asymptotically infinite times, we have the following asymptotic condition on the energy

$$\lim_{t\pm\infty} E_{\rm rel} = \lim_{t\pm\infty}\frac{1}{2}\mu\dot{\vec{r}}^2 = \begin{cases} \frac{1}{2}\mu\vec{v}^2\,, & t\to-\infty\,, \\ \frac{1}{2}\mu\vec{v}'^2\,, & t\to+\infty\,. \end{cases}$$

From conservation of energy it is then follows that

$$|\vec{v}| = |\vec{v}'|\,.$$

Now our task is to obtain the scattering angle Θ, see Fig. 1.12.

Impact parameter. Consider the angular momentum $\vec{L}_{\rm rel} = \mu\vec{r}\times\dot{\vec{r}} = \ell\vec{e}_z$, where $\ell = \mu r^2\dot{\varphi}$. Evaluating the integral ℓ at a moment t, we have

$$\ell = \mu r(t)v(t)\sin\alpha(t) = \mu v(t)\big[r(t)\sin\alpha(t)\big]\,,$$

where $r(t)$ and $v(t)$ are lengths of vectors $\vec{r}(t)$ and $\vec{v}(t) = \dot{\vec{r}}(t)$, respectively, $\alpha(t)$ is an angle between $\vec{r}(t)$ and $\vec{v}(t)$. Asymptotically, as $t\to-\infty$, $v(t)$ tends to the initial velocity v, while $\alpha(t)$ tens to $\pi-\varphi(t)$. Thus, asymptotically,

$$r(t)\sin\alpha(t) \to r(t)\sin(\pi-\varphi(t)) = r(t)\sin(\varphi(t)) \to b\,,$$

as is seen from Fig. 1.12. Thus, in the limit $t\to-\infty$, we find

$$\ell = \mu v b\,.$$

This formula allows one to trade the angular momentum ℓ for the impact parameter b and the initial velocity v. We therefore have

$$\ell = \mu v b = \mu r^2\dot{\varphi} \qquad\Longrightarrow\qquad \dot{\varphi} = \frac{vb}{r^2} > 0\,.$$

The fact that $\dot{\varphi}$ is positive shows that φ monotonically grows starting from $\varphi = 0$ when time increases from $-\infty$ to $+\infty$.

Equation for the trajectory. Further, we have

$$d\varphi = \pm\frac{\ell}{\sqrt{2\mu}}\frac{dr}{r^2\sqrt{E-U_{\rm eff}(r)}}\,.$$

For the repulsive Coulomb potential

$$U = \frac{k}{r}\,,\quad k>0\,,$$

the effective potential

$$U_{\rm eff} = \frac{k}{r} + \frac{\ell^2}{2\mu r^2}$$

is a function that monotonically decreases from $+\infty$ to zero for r running from 0 to ∞. The energy of a particle can only be positive and the motion is always infinite (scattering).

The particle trajectory is symmetric with respect to the line that passes through the scattering center O and the point of the trajectory closest to O. On the right half of the trajectory, we have

$$d\varphi = -\frac{\ell}{\sqrt{2\mu}}\frac{dr}{r^2\sqrt{E-U_{\rm eff}(r)}}\,,$$

because decreasing of r corresponds to increasing of φ. We then integrate

$$\varphi = \int_0^\varphi d\varphi = -\frac{\ell}{\sqrt{2\mu}}\int_\infty^r \frac{dr'}{r'^2\sqrt{E-U_{\rm eff}(r')}} = \frac{\ell}{\sqrt{2\mu}}\int_r^\infty \frac{dr'}{r'^2\sqrt{E-U_{\rm eff}(r')}}\,.$$

Substituting here U_{eff} as well as $\ell = \mu v b$ and $E = \frac{1}{2}\mu v^2$, we get

$$\varphi = b\int_r^\infty \frac{dr'}{r'^2\sqrt{1-\frac{2k}{\mu v^2 r'}-\frac{b^2}{r'^2}}}\,,$$

The change of variables $u = 1/r$ yields

$$\varphi = -\int_{\frac{1}{r}}^0 \frac{bdu}{\sqrt{1-\frac{2k}{\mu v^2}u-b^2u^2}} = \int_0^{\frac{1}{r}} \frac{bdu}{\sqrt{1-\frac{2k}{\mu v^2}u-b^2u^2}}\,.$$

First we compute the indefinite integral[7]

$$\int \frac{bdu}{\sqrt{1-\frac{2k}{\mu v^2}u-b^2u^2}} = \arcsin\frac{bu+\frac{k}{b\mu v^2}}{\sqrt{1+\frac{k^2}{\mu^2 b^2 v^4}}}\,.$$

Then, for the definite one we will find the following answer

$$\varphi = \arcsin\frac{\frac{b}{r}+\frac{k}{b\mu v^2}}{\sqrt{1+\frac{k^2}{\mu^2 b^2 v^4}}} - \arcsin\frac{\frac{k}{b\mu v^2}}{\sqrt{1+\frac{k^2}{\mu^2 b^2 v^4}}}\,.$$

This formula can be further simplified by using the identity

$$\arcsin x - \arcsin y = \arccos\left(\sqrt{1-x^2}\sqrt{1-y^2}+xy\right), \qquad x > y\,.$$

Namely, we obtain the equation describing the orbit

$$\varphi = \arccos\frac{1+\frac{\mu b^2 v^2}{kr}+\frac{b^2\mu^2 v^4}{k^2}\sqrt{1-\frac{2k}{\mu v^2 r}-\frac{b^2}{r^2}}}{1+\frac{b^2\mu^2 v^4}{k^2}}\,,$$

or solving this equation for r

$$r(\varphi) = \frac{b}{2}\frac{\csc^2\frac{\varphi}{2}}{\cot\frac{\varphi}{2}-\frac{k}{\mu b v^2}}\,. \tag{I.1.11}$$

This solution satisfies correct boundary condition $\varphi|_{r=\infty} = 0$. The minimal distance (periapsis) from the force center to the orbit is given by the is a positive root of

$$1-\frac{2k}{\mu v^2 r}-\frac{b^2}{r^2} = 0\,,$$

which is

$$r_{\min} = k\frac{1+\sqrt{1+\frac{b^2\mu^2 v^4}{k^2}}}{\mu v^2} = \frac{k}{\mu v^2}(1+\mathfrak{e})\,,$$

where we have taken into account the expression for eccentricity

$$\mathfrak{e} = \sqrt{1+\frac{2E\ell}{\mu k^2}} = \sqrt{1+\frac{b^2\mu^2 v^4}{k^2}}\,. \tag{I.1.12}$$

[7]See *e.g.* I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series and products, Academic Press, 1965, page 81, integral 2.261.*

Impact parameter as a function of the scattering angle. From Fig. 1.12 we conclude that the angle Φ is then given by

$$\Phi = \varphi(r_{\min}) = \arccos \frac{1}{\sqrt{1+\frac{b^2\mu^2v^4}{k^2}}} = \arccos \frac{1}{\mathfrak{e}}\,.$$

The scattering angle is computed via Φ as

$$\Theta = \pi - 2\Phi\,.$$

Thus,

$$\sin\frac{\Theta}{2} = \sin\Big(\frac{\pi}{2}-\Phi\Big) = \cos\Phi = \frac{1}{\mathfrak{e}}$$

and from here it follows that

$$\cot^2\frac{\Theta}{2} = \mathfrak{e}^2 - 1 = \frac{b^2\mu^2v^4}{k^2}\,.$$

This allows us to express the impact parameter in terms of the scattering angle

$$b = \frac{k}{\mu v^2}\cot\frac{\Theta}{2}\,. \tag{I.1.13}$$

One remark is in order. Equation (I.1.11) for the trajectory $r = r(\varphi)$ exhibits a pole at φ_0 where

$$\cot\frac{\varphi_0}{2} = \frac{k}{\mu b v^2}$$

This pole corresponds to the angle $\varphi_0 = 2\Phi$ when $r \to \infty$ on the outgoing branch of the trajectory. Indeed, since $\cos\Phi = 1/\mathfrak{e}$, we find

$$\cot\frac{\varphi_0}{2} = \cot\Phi = \frac{\cos\Phi}{\sin\Phi} = \frac{1}{\mathfrak{e}}\frac{1}{\sqrt{1-\frac{1}{\mathfrak{e}^2}}} = \frac{1}{\sqrt{\mathfrak{e}^2-1}} = \frac{k}{\mu b v^2}\,,$$

where we have used (I.1.12).

Trajectory is a hyperbola. Equation (I.1.11), although looks complicated, still defines a hyperbola. To see this, we will use that $\cot\Phi = \frac{k}{\mu b v^2}$, so that

$$\begin{aligned} r(\varphi) &= \frac{b}{2}\frac{\csc^2\frac{\varphi}{2}}{\cot\frac{\varphi}{2}-\cot\Phi} = \frac{b}{2}\frac{1}{\sin^2\frac{\varphi}{2}\Big(\frac{1}{\sin\varphi}+\cot\varphi-\cot\Phi\Big)} = \frac{b}{2}\frac{1}{\sin^2\frac{\varphi}{2}\Big(\frac{1}{\sin\varphi}-\frac{\sin(\varphi-\Phi)}{\sin\varphi\sin\Phi}\Big)} \\ &= \frac{b}{2}\frac{\sin\Phi}{\frac{\sin^2\frac{\varphi}{2}}{\sin\varphi}\underbrace{(\sin\Phi-\sin(\varphi-\Phi))}_{2\cos\frac{\varphi}{2}\sin\frac{2\Phi-\varphi}{2}}} = \frac{b\sin\Phi}{2\sin\frac{\varphi}{2}\sin\frac{2\Phi-\varphi}{2}}\,. \end{aligned}$$

Thus, we finally obtain

$$r(\varphi) = \frac{b\sin\Phi}{\cos(\varphi-\Phi)-\cos\Phi} = \frac{b\tan\Phi}{-1+\frac{1}{\cos\Phi}\cos(\varphi-\Phi)}\,.$$

Recalling now that $1/\cos\Phi = \mathfrak{e}$ and $b\tan\Phi = \frac{\mu b^2v^2}{k} = \frac{\ell^2}{\mu k} = \mathfrak{p}$, we get for the equation describing the trajectory the following answer

$$r(\varphi) = \frac{\mathfrak{p}}{-1+\mathfrak{e}\cos(\varphi-\Phi)}\,.$$

This is an equation for a hyperbola passing outside the force center. Finally, we note that the impact parameter b coincides with what would be called the (non-major) "semi-axis". Indeed, the major semi-axis a and the parameter c are

$$a = \frac{r_{\min}}{1+e} = \frac{k}{\mu v^2}\,, \qquad c = a\mathfrak{e}\,,$$

so that

$$\sqrt{c^2 - a^2} = a\sqrt{\mathfrak{e}^2 - 1} = \frac{k}{\mu v^2}\frac{b\mu v^2}{k} = b\,.$$

Differential cross section. In physical applications one has to deal not with scattering of an individual particle but with scattering of a beam of the same particles falling on a target with the same velocity $\vec{v}$. Different particles in the beam have different impact parameters and, therefore, scatter under different angles Θ. Denote by dN the number of particles scattered per unit time into angles between Θ and $\Theta + d\Theta$. By itself, the number dN is not convenient to characterise the scattering process, because it depends on the density $\mathscr{L}$ of the incident beam. Therefore, one uses the quantity

$$d\sigma = \frac{dN}{\mathscr{L}}\,,$$

where $\mathscr{L}$ is the number of incident particles per unit area per unit time, called *luminosity*.[8] The quantity $d\sigma$ has dimension of area and is called *differential cross section*. It is determined by a force field that scatters and represents one of the most important characteristics of the scattering process.

We assume that the dependence of b on Θ is monotonic (this is not always so and depends on a scattering potential), *i.e.* the scattering angle is a monotonically decreasing function of the impact parameter. In this case particles that scatter in the interval between Θ and $\Theta + d\Theta$ are the ones that have the impact parameter between $b(\Theta)$ and $b(\Theta) + db(\Theta)$. The number of these particles is equal to the area of an annulus between circles with radius b and $b + db$ multiplied with $\mathscr{L}$, that is $dN = 2\pi b db \cdot \mathscr{L}$. Thus, for the differential cross section we get

$$d\sigma = 2\pi b db.$$

In order to find the dependence of the differential cross section on the scattering angle, we rewrite this formula in the form

$$d\sigma = 2\pi b \Big|\frac{db}{d\Theta}\Big| d\Theta\,,$$

where the modulus is taken because in general $\frac{db}{d\Theta}$ is negative (with Θ increasing b is decresing). Often it is convenient to consider $d\sigma$ not with respect to the flat angle Θ, but rather with respect to the solid angle $d\Omega = 2\pi \sin\Theta d\Theta$. Then the previous formula can be rewritten in the form

$$d\sigma = \frac{b}{\sin\Theta}\Big|\frac{db}{d\Theta}\Big| d\Omega\,,$$

or, in other words,

$$\frac{d\sigma}{d\Omega} = \frac{b}{\sin\Theta}\Big|\frac{db}{d\Theta}\Big|\,. \tag{I.1.14}$$

This is one of the main formulas of the scattering theory. The total cross section is defined as

$$\sigma_{\mathrm{t}} = \int_{S^2} \frac{d\sigma(\Omega)}{d\Omega} d\Omega = 2\pi \int_0^\pi \frac{d\sigma(\Theta)}{d\Theta} \sin\Theta d\Theta\,.$$

[8] We assume that the beam is homogeneous over its cross section.

Differential cross section for Coulomb scattering. Now we can compute the cross section which corresponds to scattering in the repulsive Coulomb field. Substituting (I.1.13) into (I.1.14), we find

$$\frac{d\sigma(\Theta)}{d\Omega} = \frac{k^2}{\mu^2 v^4} \frac{\cot\frac{\Theta}{2}}{\sin\Theta} \left| \frac{d\cot\frac{\Theta}{2}}{d\Theta} \right| = \frac{1}{4}\Big(\frac{k}{\mu v^2}\Big)^2 \frac{1}{\sin^4\frac{\Theta}{2}} = \frac{1}{4}\Big(\frac{k}{2E}\Big)^2 \frac{1}{\sin^4\frac{\Theta}{2}} \,.$$

This is the famous Rutherford formula, originally derived by Rutherford for the scattering of α particles by atomic nuclei. The total cross section is divergent as a consequence of a long-range force character of the Coulomb field. Note that the formula for the differential cross section does not depend on the sign of k and is valid for both the attractive and repulsive Coulomb potentials.

Chapter 2

Lagrangian and Hamiltonian mechanics

> "...the Lagrangian and Hamilton's principle together form a compact invariant way of obtaining mechanical equations of motion. This possibility is not reserved to mechanics only; in almost every field of physics variational principles can be used to express the "equations of motion," whether they be Newton's equations, Maxwell's equations, or the Schrödinder equation."
>
> *Herbert Goldstein "Classical Mechanics"*

Lagrangian mechanics describes motion in a mechanical system by means of the configuration space. A newtonian potential system is a particular case of a lagrangian system (the configuration space in this case is euclidean, and the lagrangian function is the difference between the kinetic and potential energies).

The lagrangian point of view allows us to solve completely a series of important mechanical problems, including problems in the theory of small oscillations and in the dynamics of a rigid body.

2.1 Lagrangian mechanics

Here we consider the formulation of the lagrangian mechanics. We start with the principle of the least action, introduce the action and derive the Euler-Lagrange equations. We also give a lagrangian description of the newtonian mechanics.

2.1.1 Principle of the least action

To determine a state of a system of N particles in space, one needs to specify N radius-vectors, *i.e.* $3N$ coordinates. In general, the number of independent quantities that are needed to be given in order to specify the position of a system is called *the number of degrees of freedom*. In the present

case it is $n = 3N$. In fact, these quantities should not necessarily be cartesian coordinates of the particles. Depending on the problem at hand, it might be more convenient to choose some other coordinates. Any n quantities $q_1, q_2, \ldots, q_n$ which fully characterise the position of the system (with n degrees of freedom) are called its *generalised coordinates*, while their derivatives $\dot{q}_i$ are called *generalised velocities*.

The most general formulation of laws of motion of mechanical systems is given by the so-called *principle of the least action*[1] (Maupertuis's-Hamilton's principle). According to this principle, a mechanical system is characterised by a certain function of generalised coordinates and velocities

$$L(q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n, t) \equiv L(q, \dot{q}, t)$$

and motion of the system satisfies a condition that we now describe.

Let at $t = t_1$ and $t = t_2$ the system be at positions characterised by two sets of coordinates, $q^{(1)}$ and $q^{(2)}$, respectively. Then between these two positions the systems moves in such a way that the integral

$$S = \int_{t_1}^{t_2} \mathrm{d}t\, L(q, \dot{q}, t) \tag{I.2.1}$$

will have its extremal value. Function L is called the lagrangian function or simply *lagrangian* and the integral S is called *action*. The fact that L depends on q and $\dot{q}$ only, but not on more higher derivatives $\ddot{q}, \ldots$, reflects the fact that a mechanical state of a system is fully determined by coordinates and velocities.

Now we derive differential equations which solve the problem about minimising (I.2.1). For simplicity we start with a system with one degree of freedom. Let $q = q(t)$ be a function for which S has a minimum. This means that S will take large values if we replace $q(t)$ by any function of the form

$$q(t) + \delta q(t)\,,$$

where δq is a function that is small on the whole interval from t_1 to t_2, that is $|\delta q(t)| < \epsilon$ and $|\dot{\delta q}(t)| < \epsilon$, this function is called *variation* of q. Since at t_1 and t_2 all comparable functions should have the one and the same values $q^{(1)}$ and $q^{(2)}$, we should have

$$\delta q(t_1) = \delta q(t_2) = 0\,. \tag{I.2.2}$$

The change of S under the replacement of q for $q + \delta q$ is given by

$$\int_{t_1}^{t_2} \mathrm{d}t\, L(q + \delta q, \dot{q} + \dot{\delta q}, t) - \int_{t_1}^{t_2} \mathrm{d}t\, L(q, \dot{q}, t)$$
$$= \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \dot{\delta q} + \ldots \right).$$

Expansion of this difference over powers of δq and $\dot{\delta q}$ under the integral starts for the linear terms. The necessary condition for the extremum of S is the vanishing of all these linear terms, the latter are called the *first variation (or simply variation)* δS of the integral. Thus, for the variation we have

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \dot{\delta q} \right) = \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \frac{d}{dt} \delta q \right).$$

[1] More accurately, the principle of stationary action.

The condition of the extremum of S is then

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \frac{d}{dt} \delta q \right) = 0 \,.$$

Here we integrate d/dt by parts and get

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q + \frac{\partial L}{\partial \dot{q}} \delta q \Big|_{t_1}^{t_2} = 0 \,.$$

But doe to (I.2.2), the boundary terms vanish and we get

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q = 0 \,.$$

Thus, we are left with the integral which should vanish for arbitrary δq. Due to the basic lemma of the variational calculus this is only possible if the integrand vanishes and we obtain the following equation

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \,.$$

If there are n degrees of freedom, one has to vary n functions $q_i(t)$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \,, \qquad i = 1, \ldots, n \,. \tag{I.2.3}$$

These are the differential equations which describe the actual motion of the system, they are called *Euler-Lagrange equations*. If the lagrangian of a given mechanical system is known, then these equations establish a relation between accelerations, velocities and coordinates, *i.e.* they are nothing else but the equations of motion of this system. From the mathematical point of view, equations (I.2.3) constitute a system of n differential equations of the second order for n unknown functions $q_i(t)$. The general solution of such a system contains $2n$ arbitrary integration constants. For their complete determination one needs to specify initial values, which characterise the system at a given moment of time, for instance, initial coordinates and velocities.

On comment is in order. It is clear that equations of motion are unchanged if we add to a lagrangian a total time derivative of a function which depends on the coordinates and time only:

$$L \to L + \frac{d}{dt} R(q, t) \,. \tag{I.2.4}$$

Indeed, the change of the action under the variation will be

$$\delta S \to \delta S' = \delta S + \int_{t_1}^{t_2} \mathrm{d}t \, \frac{d}{dt} \delta R(q, t) = \delta S + \frac{\partial R}{\partial q_i} \delta q_i |_{t=t_1}^{t=t_2} \,.$$

Since in deriving the equations of motion the variation is assumed to vanish at the initial and final moments of time, we see that $\delta S' = \delta S$ and the equations of motion are unchanged. The fact that modification (I.2.4) of the of lagrangian leaves the equations of motion untouched can be, of course, verified by a direct calculation. Denote the addition to the original lagrangian as

$$\Delta L = \frac{dR(q, t)}{dt} = \frac{\partial R}{\partial q_j} \dot{q}_j + \frac{\partial R}{\partial t} \,.$$

Then the contribution to the equations of motion caused by ΔL will be

$$\begin{aligned}\frac{d}{dt}\frac{\partial \Delta L}{\partial \dot{q}_i} - \frac{\partial \Delta L}{\partial q_i} &= \frac{d}{dt}\frac{\partial R}{\partial q_i} - \left(\frac{\partial^2 R}{\partial q_i \partial q_j}\dot{q}_j + \frac{\partial R}{\partial q_i \partial t}\right) \\ &= \left(\frac{\partial^2 R}{\partial q_i \partial q_j}\dot{q}_j + \frac{\partial R}{\partial q_i \partial t}\right) - \left(\frac{\partial^2 R}{\partial q_i \partial q_j}\dot{q}_j + \frac{\partial R}{\partial q_i \partial t}\right) = 0\,.\end{aligned}$$

Before we proceed, let us fix some terminology. The quantities

$$p_i = \frac{\partial L}{\partial \dot{q}_i}\,, \qquad p = (p_1, \ldots, p_n)\,,$$

are called *canonical momenta*. It is not always so that $p_i = m\vec{v}_i$; in general the expression for the canonical momentum corresponding to the generalised velocity $\dot{q}_i$ depends on the form of the concrete lagrangian. The quantities

$$F_i = \frac{\partial L}{\partial q_i}$$

are called generalised forces. Finally, if a given lagrangian does not depend on a particular generalised coordinate q_i, then the canonical momentum corresponding to this coordinate is conserved (integral of motion). Indeed, from (I.2.3) we get for this coordinate

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{dp_i}{dt} = 0\,, \qquad \frac{\partial L}{\partial q_i} = 0\,. \tag{I.2.5}$$

Such a coordinate is called *cyclic*.

2.1.2 Lagrangian for a system of N particles

Consider a system of N particles which interact by means of internal conservative forces only and denote by $U(\vec{r}_1, \ldots, \vec{r}_N)$ the corresponding potential. Consider the following lagrangian

$$L = T - U = \sum_{i=1}^{N} \frac{m_i \vec{v}_i^{\,2}}{2} - U(\vec{r}_1, \ldots, \vec{r}_N)\,.$$

Let us show that the Euler-Lagrange equations that follow from this lagrangian coincide with Newton's equations. The Euler-Lagrange equations are

$$0 = \frac{d}{dt}\frac{\partial L}{\partial \vec{v}_i} - \frac{\partial L}{\partial \vec{r}_i} = \frac{d}{dt} m_i \vec{v}_i + \frac{\partial U}{\partial \vec{r}_i}\,, \qquad i = 1, \ldots, N\,.$$

Since $\vec{v}_i = \dot{\vec{r}}_i$, the expression above is

$$m_i \ddot{\vec{r}}_i = -\frac{\partial U}{\partial \vec{r}_i} = \vec{F}_i\,,$$

that is we obtain the system of Newton's equations.

As was already mentioned, to describe a system we can use any kind of generalised coordinates, not necessarily the cartesian ones. If, for instance, $\vec{r}_i = \vec{r}_i(q_a)$, $\dot{\vec{r}}_i = \dot{\vec{r}}_i(q_a)$, where $a = 1, \ldots, n$, then

$$L = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\vec{r}}_i^{\,2}(q_a) - U(\vec{r}_i(q)) = \frac{1}{2}\sum_{a=1}^{n}\sum_{b=1}^{n} g_{ab}(q)\dot{q}_a\dot{q}_b - U(q)\,. \tag{I.2.6}$$

with

$$g_{ab} = \sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial q_a} \cdot \frac{\partial \vec{r}_i}{\partial q_b} \,.$$

The Euler-Lagrange equations are derived in the same way as before with respect to the generalised coordinates.

Example. Lagrangian for the relative motion in the two-body problem. We write the lagrangian in the polar coordinate system, namely,

$$L = T - U = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) - U(r) \,.$$

The coordinate φ does not enter the lagrangian and, therefore, it is cyclic. The corresponding canonical momentum, which is

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = \mu r^2 \dot{\varphi} \,,$$

is nothing else but the value of the conserved angular momentum. The Euler-Lagrange equations are

$$\begin{aligned} &\frac{d}{dt} p_\varphi = \frac{d}{dt}(\mu r^2 \dot{\varphi}) = 0 \,, \\ &\frac{d}{dt}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = \mu \ddot{r} - \mu r \dot{\varphi}^2 + \frac{\partial U}{\partial r} = 0 \,. \end{aligned}$$

In particular, the Lagrangian for the relative motion in the Kepler problem will be

$$L_{\text{Kepler}} = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) + \frac{k}{r} \,, \qquad k > 0 \,.$$

Often one has to deal with mechanical systems in which interactions between bodies (point particles) have a character of *constraints*, *i.e.* restrictions imposed on the mutual positions of bodies. In practice such constraints are realised by fasting the bodies by means of various rods, threads, hinges, etc. This brings a new factor, namely, the motion of the bodies is accompanied by friction at places where the bodies are in contact, so that, strictly speaking the problem goes beyond the framework of mechanics. However, in many cases friction appears to be so small that it can be completely neglected. If, in addition, one can neglect the masses of fasting elements, then the role of the latter simply reduced to diminishing the number of degrees of freedom n of a system (in comparison to $3N$). To determine the motion of such a system one can use the lagrangian (I.2.6) with the number of independent generalised coordinates equal to the number of actual degrees of freedom.

Example. Mathematical pendulum. Consider a mathematical pendulum. It moves in such a way that $x^2 + y^2 = l^2$. We draw the coordinate axes as in Fig. 2.1 and write the coordinates x and y via the angle φ

$$x = l\sin\varphi \,, \quad y = l\cos\varphi \,.$$

Write the Lagrangian[2]

$$L = T - U = \frac{1}{2} m\vec{v}^2 - U \,, \qquad U = -mgy \,.$$

Thus,

$$L = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2) + mgy$$

[2]Note that with the choice of the coordinate system as in Fig. 2.1, the potential decreases towards the Earth surface, while y increases. Therefore, the force acting on m is also directed towards the Earth surface, as it should be.

$$= \frac{m}{2}l^2(\cos\varphi^2\dot{\varphi}^2 + \sin\varphi^2\dot{\varphi}^2) + mgl\cos\varphi\,,$$

where we have used that

$$\dot{x} = l\dot{\varphi}\cos\varphi\,, \quad \dot{y} = -l\dot{\varphi}\sin\varphi\,.$$

Thus, we finally get

$$L = \frac{m}{2}l^2\dot{\varphi}^2 + mgl\cos\varphi\,.$$

The Euler-Lagrange equation is

$$\frac{d}{dt}\frac{\partial L}{\partial\dot{\varphi}} - \frac{\partial L}{\partial\varphi} = ml^2\ddot{\varphi} + mgl\sin\varphi = 0\,,$$

so that the pendulum moves according to the one-dimensional "sine-Gordon" equation

$$\ddot{\varphi} + \frac{g}{l}\sin\varphi = 0\,.$$

This equation can be solved analytically for small oscillations where $\sin\varphi \approx \varphi$, namely,

$$\ddot{\varphi} + \omega^2\varphi = 0\,, \quad \omega^2 = \frac{g}{l}$$

and $\varphi = A\cos\omega t + B\sin\omega t$.

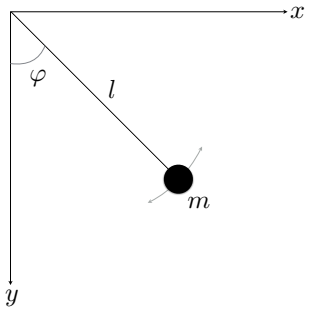

Figure 2.1: Mathematical Pendulum.

Example. Pendulum with a movable suspension point. Consider a mathematical pendulum with a movable suspension point. We introduce the generalised coordinates $q_1 = x(t)$ and $q_2 = \varphi(t)$. The Lagrangian $L = L(x, \dot{x}, \varphi, \dot{\varphi}) = T - U$. To find the kinetic energy

$$T = \frac{1}{2}m_1\vec{v}_1^2 + \frac{1}{2}m_2\vec{v}_2^2\,,$$

we first obtain

$$\begin{aligned}
\vec{v}_1 &= \dot{\vec{r}}_1\,, \quad \vec{r}_1 = x(t)\vec{e}_x\,, \quad \dot{\vec{r}}_1 = \dot{x}\vec{e}_x\,, \\
\vec{v}_2 &= \dot{\vec{r}}_2\,, \quad \vec{r}_2 = \vec{r}_1 + l\sin\varphi\vec{e}_x + l\cos\varphi\vec{e}_y\,, \quad \dot{\vec{r}}_2 = (\dot{x} + l\dot{\varphi}\cos\varphi)\vec{e}_x + (-l\dot{\varphi}\sin\varphi)\vec{e}_y\,.
\end{aligned}$$

Thus,

$$\begin{aligned}
T &= \frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2\Big((\dot{x} + l\dot{\varphi}\cos\varphi)^2 + l^2\dot{\varphi}^2\sin^2\varphi\Big) \\
&= \frac{1}{2}(m_1 + m_2)\dot{x}^2 + \frac{1}{2}m_2l^2\dot{\varphi}^2 + m_2l\dot{x}\dot{\varphi}\cos\varphi\,.
\end{aligned}$$

The potential energy is $U = -m_2gl\cos\varphi$. The Lagrangian is then

$$L(x, \dot{x}, \varphi, \dot{\varphi}) = \frac{1}{2}(m_1 + m_2)\dot{x}^2 + \frac{1}{2}m_2l^2\dot{\varphi}^2 + m_2l\dot{x}\dot{\varphi}\cos\varphi + m_2gl\cos\varphi\,.$$

The coordinate x is cyclic. The Euler-Lagrange equations are

$$\begin{aligned}
\frac{d}{dt}\frac{\partial L}{\partial\dot{x}} - \frac{\partial L}{\partial x} &= \frac{d}{dt}\Big[(m_1 + m_2)\dot{x} + m_2l\cos\varphi\dot{\varphi}\Big] \\
&= (m_1 + m_2)\ddot{x} + m_2l\ddot{\varphi}\cos\varphi - m_2l\dot{\varphi}^2\sin\varphi = 0\,,
\end{aligned}$$

$$\begin{aligned}\frac{d}{dt}\frac{\partial L}{\partial \dot\varphi} - \frac{\partial L}{\partial \varphi} &= \frac{d}{dt}\Big[m_2 l^2\dot\varphi + m_2 l\dot x\cos\varphi\Big] + m_2 l\dot x\dot\varphi\sin\varphi + m_2 gl\sin\varphi \\ &= m_2 l^2\ddot\varphi + m_2 l\ddot x\cos\varphi + m_2 gl\sin\varphi = 0\,.\end{aligned}$$

Thus, we obtained a system of coupled differential equations

$$\begin{aligned}&(m_1+m_2)\ddot x + m_2 l\cos\varphi\ddot\varphi - m_2 l\sin\varphi\dot\varphi^2 = 0\,,\\ &l\ddot\varphi + \cos\varphi\ddot x + g\sin\varphi = 0\,.\end{aligned}$$

For small oscillations $\sin\varphi\approx\varphi$, $\cos\varphi\approx 1$ and the system is approximated as

$$\begin{aligned}(m_1+m_2)\ddot x + m_2 l\ddot\varphi &= 0\,,\\ l\ddot\varphi + \ddot x + g\varphi &= 0\,.\end{aligned}\qquad\text{(I.2.7)}$$

From the second equation $\ddot x = -(l\ddot\varphi + g\varphi)$ and substituting this result in the first, we find

$$-(m_1+m_2)(l\ddot\varphi + g\varphi) + m_2 l\ddot\varphi = 0$$

or

$$\ddot\varphi + \omega^2\varphi = 0\,,\qquad \omega^2 = \frac{m_1+m_2}{m_1}\frac{g}{l}$$

and

$$\varphi = A\cos\omega t + B\sin\omega t\,.$$

As to x, from the first equation in (I.2.7) we have

$$\ddot x = -\frac{m_2 l}{m_1+m_2}\ddot\varphi\,.$$

Integrating this equation twice we find

$$x = -\frac{m_2 l}{m_1+m_2}\varphi + A't + B'\,.$$

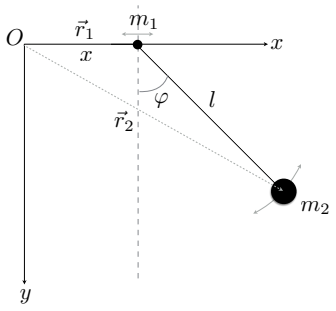

Figure 2.2: Pendulum with a movable suspension point.

Thus, we have obtained a generals solution for small oscillations which depends on 4 integration constants A, B, A', B'.

One can also make a progress towards an exact solution. Since x is cyclic, the momentum p_x is conserved, *i.e.*

$$p_x = (m_1+m_2)\dot x + m_2\ell\dot\varphi\cos\varphi = A' = \text{const}$$

Integrating this equation, we get

$$(m_1+m_2)x + m_2\ell\sin\varphi = A't + B'\,.$$

By properly choosing an inertial frame we can always adjust $A' = 0$. The cartesian coordinates of the mass m_2 are

$$\begin{aligned}x_2 &= x + l\sin\varphi = B' - \frac{m_2 l}{m_1+m_2}\sin\varphi + l\sin\varphi = B' + \frac{m_1 l}{m_1+m_2}\sin\varphi\,,\\ y_2 &= l\cos\varphi\,.\end{aligned}$$

From here we deduce that the trajectory of the mass m_2 on the xy-plane is an ellipse

$$\frac{(x-B')^2}{a^2}+\frac{y^2}{b^2}=1\,,$$

where

$$a=\frac{m_1 l}{m_1+m_2}\,,\quad b=l\,.$$

With the use of $\dot{x}=-\frac{m_2 l}{m_1+m_2}\dot{\varphi}\cos\varphi$, the total energy $E=T+U$ reads as

$$E=\frac{m_2 l^2\dot{\varphi}^2}{2}\Big(1-\frac{m_2}{m_1+m_2}\cos^2\varphi\Big)-m_2 gl\cos\varphi\,.$$

Separating the variables, we find the law of motion for the angle φ

$$t=l\sqrt{\frac{m_2}{2(m_1+m_2)}}\int\sqrt{\frac{m_1+m_2\sin^2\varphi}{E+m_2 gl\cos\varphi}}d\varphi\,.$$

2.2 Symmetries and conservation laws

We begin our discussion with Noether's theorem which is one of the most fundamental and general statements concerning the behaviour of dynamical systems. It relates symmetries of a theory with its conservation laws.

2.2.1 Noether's theorem

Noether's theorem. *For any continuous symmetry of the action there exists a quantity which is conserved due to the Euler-Lagrange equations. In other words, symmetries of the action yield conservation laws.*

Below we explain the notion of symmetry and provide a proof of Noether's theorem. Let an infinitezimal transformation $q_i\to q_i+\delta q_i$ depending on a continuous parameter(s) be such that the variation of the Lagrangian takes the form[3] of a total time derivative of some function F:

$$\delta L=\frac{dF}{dt}\,.$$

Transformation δq_i is called a symmetry of the action.

Proof. Now comes the proof of Noether's theorem. Suppose that $q_i'=q_i+\delta q_i$ is a symmetry of the action. Then[4]

$$\delta L=\frac{\partial L}{\partial q_i}\delta q_i+\frac{\partial L}{\partial\dot{q}_i}\delta\dot{q}_i=\frac{\partial L}{\partial q_i}\delta q_i+\frac{\partial L}{\partial\dot{q}_i}\frac{d}{dt}\delta q_i=\frac{dF}{dt}\,.$$

By the Euler-Lagrange equations, we get

$$\delta L=\frac{d}{dt}\Big(\frac{\partial L}{\partial\dot{q}_i}\Big)\delta q_i+\frac{\partial L}{\partial\dot{q}_i}\frac{d}{dt}\delta q_i=\frac{dF}{dt}\,.$$

This gives

$$\delta L=\frac{d}{dt}\Big(\frac{\partial L}{\partial\dot{q}_i}\delta q_i\Big)=\frac{dF}{dt}\,.$$

[3]Without usage of equations of motion! A variation of the lagrangian computed on the equations of motion is always a total derivative!

[4]Here we start to use the convention adopted in the literature that $\delta\dot{q}\equiv\dot{\delta q}=\frac{d}{dt}\delta q$.

As the result, we find the quantity which is conserved in time

$$\frac{dJ}{dt} \equiv \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\delta q_i - F\Big) = 0\,.$$

This quantity

$$J = \frac{\partial L}{\partial \dot{q}_i}\delta q_i - F$$

is called *Noether's charge*. As we have just shown, Noether's charge is a conserved quantity. Now we consider some important symmetries and corresponding conservation laws. Using the definition of the canonical momentum the expression for the conserved charge can also be written as

$$J = p_i \delta q_i - F\,.$$

This completes the proof of Noether's theorem. □

2.2.2 Fundamental conservation laws

The most important conservation laws include conservation of energy, momentum and angular momentum. As such, they reflect the fundamental symmetry properties of the space-time. Later on we will study another fundamental conservation law – conservation of electric charge.

1) Conservation of energy. Energy conservation is related to homogeneity of time, which shows up in the freedom of arbitrary choosing the reference point of time (one can perform an experiment today or after several years but its result will always be the same provided one use the same initial conditions).

We derive now the conservation law of energy in the framework of Noether's theorem. Suppose we make an infinitesimal time displacement $\delta t = \epsilon$, where ϵ is a small but otherwise arbitrary constant. The response of the lagrangian on this displacement is

$$\delta L = \frac{dL}{dt}\delta t = \frac{dL}{dt}\epsilon\,.$$

Thus, the displacement $\delta t = \epsilon$ is the symmetry because the lagrangian changed by the total derivative of the function $F = L\epsilon$. On the other hand,

$$\delta L = \frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i + \frac{\partial L}{\partial t}\delta t = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\frac{d}{dt}\delta q_i = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right)\,,$$

where we have used the Euler-Lagrange equations and assumed that L *does not explicitly depend* on time. Thus, we have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right) = \frac{dL}{dt}\epsilon\,.$$

Obviously, $\delta q^i = \dot{q}_i\epsilon$ and the above equation reduces to

$$\epsilon\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L\right) = 0\,.$$

Omitting ϵ and recalling that $\frac{\partial L}{\partial \dot{q}_i} = p_i$, we recover the corresponding conserved quantity, which we denote as H,

$$H = p_i\dot{q}_i - L\,, \qquad \frac{dH}{dt} = 0\,.$$

The quantity H is called *energy* of the system. We stress that conservation of energy tales place not only for closed systems but also for systems in a constant, *i.e.* time-independent, field. The only condition that was needed to prove conservation of energy was the absence of explicit time dependence of the lagrangian, the property that holds in a constant external field. Sometimes mechanical systems for which the energy is conserved are called *conservative*. The formula

$$H = \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L$$

relating L with H is called *the Legendre transformation*. For a mechanical system in cartesian coordinates $\vec{p}_i = m\vec{v}_i$ and

$$H = \sum_{i=1}^{N} m_i \dot{\vec{r}}_i^{\,2} - L = \sum_{i=1}^{N} m_i \dot{\vec{r}}_i^{\,2} - \left(\frac{1}{2}\sum_{i=1}^{N} m_i \dot{\vec{r}}_i^{\,2} - U(r)\right) = T + U \equiv E\,,$$

justifying that H coincides with the definition of energy we use in Newtonian's mechanics.

2) Momentum conservation. Momentum conservation is related to homogeneity of space. Due to this homogeneity the mechanical properties of a closed system do not change under any parallel translation of the system as a whole. This means that the lagrangian describing this system must be invariant under a shift of radius-vectors $\vec{r}_i \to \vec{r}_i + \vec{\epsilon}$, where $\vec{\epsilon}$ is an arbitrary constant vector. This means that $\delta\vec{r}_i = \vec{\epsilon}$ is the symmetry and, according to Noether's theorem, we have a conserved current

$$J = \vec{\epsilon}\cdot\sum_{i=1}^{N}\vec{p}_i = \vec{\epsilon}\cdot\vec{P}\,, \qquad \vec{P} = \sum_{i=1}^{N}\vec{p}_i\,,$$

where $\vec{P}$ is the total momentum. Since $\vec{\epsilon}$ is arbitrary and constant, conservation of the J implies conservation of the components of the total momentum

$$\frac{d\vec{P}}{dt} = 0\,.$$

The other way around, assuming the conservation law of the total momentum and summing up the Euler-Lagrange equations, we get

$$0 = \frac{d}{dt}\sum_{i=1}^{N}\frac{\partial L}{\partial \vec{v}_i} - \sum_{i=1}^{N}\frac{\partial L}{\partial \vec{r}_i} = \frac{d}{dt}\sum_{i=1}^{N} p_i - \sum_{i=1}^{N}\frac{\partial L}{\partial \vec{r}_i} = -\sum_{i=1}^{N}\frac{\partial L}{\partial \vec{r}_i}\,.$$

But $\frac{\partial L}{\partial \vec{r}_i} = -\frac{\partial U}{\partial \vec{r}_i} = \vec{F}_i$ is the force acting on i'th particle. Thus, if the total momentum is conserved, then the sum of forces acting on all particles of the closed system equals to zero

$$\sum_{i=1}^{N}\vec{F}_i = 0\,.$$

In particular, for a system of two particles the last relation turns into the third Newton's law: $\vec{F}_1 + \vec{F}_2 = 0$, *i.e.* the force acting on the first particle from the second one is equal in strength but appositely directed to the force with which the first particle acts on the second (*actio = reactio*).

Note that if the motion is described by generalised coordinates, individual generalised momenta p_i will be conserved provided the corresponding lagrangian is invariant under constant shifts $q_i \to q_i + \epsilon$. Indeed, in this case $\delta q_i = \epsilon$ is the symmetry, which implies that the action does not depend on the coordinate q_i, *i.e.* q_i is cyclic and, according to (I.2.5), p_i is conserved.

A lagrangian might be invariant under constant shifts only up to a total derivative, which also leads to an existence of the conservation law involving a non-trivial function F, as in the example below.

Example. Particle in a constant gravitational field. Consider the lagrangian

$$L = \frac{m}{2}\dot{z}^2 - mgz$$

and a shift $z \to z + \epsilon$, *i.e.* $\delta z = \epsilon$. We get $\delta L = -mg\epsilon = \frac{d}{dt}(-mg\epsilon t)$, so that $F = -mgt\epsilon$. Thus, according to Noether's theorem, the quantity

$$J = m\dot{z}\delta z - F = (m\dot{z} + mgt)\epsilon$$

is conserved. This is a conservation law of the initial velocity $\dot{z} + gt = \text{const}$.

3) Angular momentum conservation. Conservation of angular momentum is related to isotropy of space. Isotropy means that mechanical properties of a closed system do not change under any rotation of the system as a whole in space. From the point of view of the lagrangian dynamics, this means that the lagrangian of the system is invariant under rotations.

Let us first discuss rotations in more detail. Rotations are linear coordinate transformations which preserve the origin and keep the modulus $|\vec{r}|$ of a radius-vector $\vec{r} \in \mathbb{R}^3$ invariant. Let

$$\vec{r} = \sum_{a=1}^{3} x_a \vec{e}_a = x\vec{e}_x + y\vec{e}_y + z\vec{e}_z \,.$$

Any linear transformation can be represented by a 3×3 matrix $R = ||R_{ab}||$, $a, b = 1, \ldots, 3$. Such a transformation acts on coordinates as

$$x'_a = \sum_{b=1}^{3} R_{ab} x_b \,,$$

or in the matrix form

$$\vec{r}' = R\,\vec{r}$$

The length squared of the transformed vector $\vec{r}'$ is

$$\vec{r}'^t\vec{r}' = \sum_{a=1}^{3} x'_a x'_a = \sum_{a=1}^{3}\sum_{b=1}^{3}\sum_{c=1}^{3} R_{ab}x_b R_{ac}x_c = \sum_{b=1}^{3}\sum_{c=1}^{3} x_b \left(\sum_{a=1}^{3}(R^t)_{ba}R_{ac}\right) x_c \,,$$

where t stands for transposition. We require the preservation of length, *i.e.* $\vec{r}'^t\vec{r}' = \vec{r}^t\vec{r}$, which imposes on R the following relation

$$\sum_{a=1}^{3}(R^t)_{ba}R_{ac} = \delta_{bc} \,.$$

The last relation written in the matrix form is

$$R^t R = \mathbb{1} \,. \tag{I.2.8}$$

The matrices satisfying the condition (I.2.8) are called *orthogonal*. For such matrices we have

$$\det(R^t R) = (\det R)^2 \quad \Rightarrow \quad \det R = \pm 1 \,.$$

Matrices which represent rotations are those which have $\det R = 1$, such matrices are called *special*. All orthogonal matrices as well as orthogonal matrices with $\det R = 1$ constitute a *group*.

Definition. *A group G is a set equipped with a binary operation $\cdot$, called product, that combines any two elements to form a third element in such a way that four conditions called group axioms are satisfied. The group axioms are namely closure, associativity, identity and invertibility.*

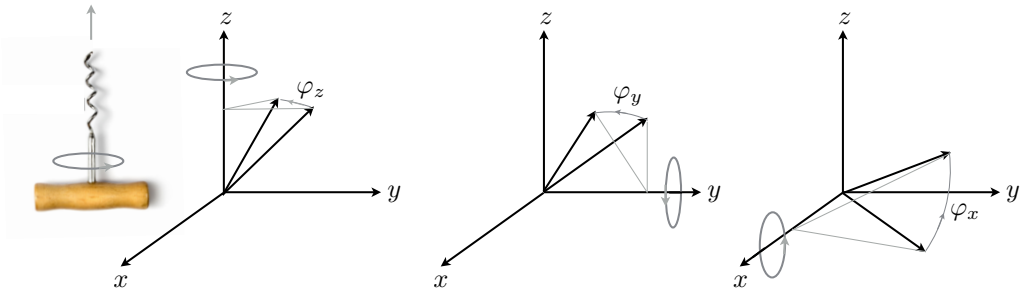

Figure 2.3: Rotations of the radius vector around x-, y- and z-axes. If a corkscrew downwards along one of the axes, then the direction of rotation of his handle driven by the right hand coincides with the direction of the (positive) rotation angle around this axis.

(a) *Closure*: $\forall\, g_1, g_2 \in G \quad \to \quad g_1 \cdot g_2 \in G$;

(b) *Associativity of the product*: $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$;

(c) *Identity*: $\exists\, e$, *such that* $e \cdot g = g \cdot e = g$;

(d) *Inverse: for any* $g \in G$ $\exists$ *the inverse* $g^{-1} \in G$ *such that* $g \cdot g^{-1} = g^{-1} \cdot g = e$.

Very often the sign of the product is omitted, so instead of $g_1 \cdot g_2$ one simply writes $g_1 g_2$.

The fact that orthogonal matrices form a group is easily verified. The product $\cdot$ is given by the usual matrix product. Let us show that if R_1 and R_2 are orthogonal, then their matrix product $R_1 R_2$ is also orthogonal. We have

$$(R_1 R_2)^t (R_1 R_2) = R_2^t R_1^t R_1 R_2 = R_2^t R_2 = \mathbb{1}\,.$$

Further, the matrix product is associative and the identity e is given by the identity matrix $\mathbb{1}$. Finally, any orthogonal matrix is invertible and its inverse is an orthogonal matrix. Indeed, from $R^t R = \mathbb{1}$ we get $R^t = R^{-1}$. Applying transposition to $RR^{-1} = \mathbb{1}$, we obtain $(R^{-1})^t R^t = \mathbb{1}$, where substituting $R^t = R^{-1}$, we find $(R^{-1})^t R^{-1} = \mathbb{1}$. If R_1 and R_2 have unit determinant, then their product $R_1 R_2$ also has a unit determinant. Matrices with determinant equal to -1 do not form a group.

The groups of orthogonal matrices and orthogonal matrices with unit determinant have the following notation

$$\begin{aligned} &\mathrm{O}(3) - \text{the group of orthogonal matrices}\,,\\ &\mathrm{SO}(3) - \text{the group of special orthogonal matrices}\,. \end{aligned}$$

Now we want to understand how parametrise special orthogonal matrices by rotation angles. Choosing the complex coordinates in the xy-plane, see Fig. 2.4, we write the result of rotation of a vector $\vec{r} = (x, y, z)$ around z-axis by angle φ_z in the complex form as

$$\begin{aligned} x' + iy' = (x + iy)e^{i\varphi_z} &= re^{i(\phi+\varphi_z)} = (x + iy)(\cos\varphi_z + i\sin\varphi_z)\\ &= x\cos\varphi_z - y\sin\varphi_z + i(x\sin\varphi_z + y\cos\varphi_z)\,,\\ z' &= z\,. \end{aligned}$$

In the matrix form we have

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\varphi_z & -\sin\varphi_z & 0 \\ \sin\varphi_z & \cos\varphi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \Rightarrow \quad R_z = \begin{pmatrix} \cos\varphi_z & -\sin\varphi_z & 0 \\ \sin\varphi_z & \cos\varphi_z & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Analogously, introducing complex coordinates in the zx-plane, for rotation around y-axis we find

$$\begin{aligned} z' + ix' = (z + ix)e^{i\varphi_y} &= re^{i(\phi+\varphi_y)} = (z + ix)(\cos\varphi_y + i\sin\varphi_y) \\ &= z\cos\varphi_y - x\sin\varphi_y + i(z\sin\varphi_y + x\cos\varphi_y)\,, \\ y' &= y\,. \end{aligned}$$

In the matrix form this yields

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\varphi_y & 0 & \sin\varphi_y \\ 0 & 1 & 0 \\ -\sin\varphi_y & 0 & \cos\varphi_y \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \Rightarrow \quad R_y = \begin{pmatrix} \cos\varphi_y & 0 & \sin\varphi_y \\ 0 & 1 & 0 \\ -\sin\varphi_y & 0 & \cos\varphi_y \end{pmatrix} .$$

Finally, introducing complex coordinates in the yz-plane, for rotation around x-axis we find

$$\begin{aligned} y' + iz' = (y + iz)e^{i\varphi_x} &= re^{i(\phi+\varphi_x)} = (y + iz)(\cos\varphi_x + i\sin\varphi_x) \\ &= y\cos\varphi_x - z\sin\varphi_x + i(y\sin\varphi_x + z\cos\varphi_x)\,, \\ x' &= x\,. \end{aligned}$$

The matrix realisation of these formulae is

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_x & -\sin\varphi_x \\ 0 & \sin\varphi_x & \cos\varphi_x \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \Rightarrow \quad R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_x & -\sin\varphi_x \\ 0 & \sin\varphi_x & \cos\varphi_x \end{pmatrix} .$$

One can check that all three matrices R_x, R_y, R_z are orthogonal and have unit determinant. They represent rotations around x-, y- and z-axes on a finite angle.

Consider now an infinitesimal rotation, *i.e.* a rotation on a small angle $\delta\varphi$. Then in the limit $\delta\varphi \to 0$ the matrices R_x, R_y, R_z turn to

$$\begin{aligned} R_x(\delta\varphi) &\approx \mathbb{1} + \delta\varphi A_1\,, \qquad A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\,, \\ R_y(\delta\varphi) &\approx \mathbb{1} + \delta\varphi A_2\,, \qquad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}\,, \\ R_z(\delta\varphi) &\approx \mathbb{1} + \delta\varphi A_3\,, \qquad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,. \end{aligned}$$

The matrices A_a are skew-symmetric, $A_a = -A_a^t$. Moreover, they form a basis in the space of all 3×3 real skew-symmetric matrices, *i.e.* any 3×3 skew-symmetric matrix A, $A = -A^t$, can be written as

$$A \equiv A(\vec{a}) = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} = a_1A_1 + a_2A_2 + a_3A_3\,,$$

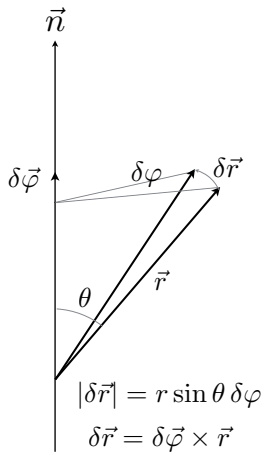

Figure 2.4: Rotation around an arbitrary axis $\vec{n}$ by an angle $\delta\varphi$.

The entries of these three matrices can be encoded in one formula

$$(A_a)_{bc} = -\epsilon_{abc}\,.$$

Infinitesimally, rotation of $\vec{r}$ is $\vec{r} \to \vec{r}' = \vec{r} + \delta r$, where

$$\begin{aligned}
&\text{around } x: \qquad \delta\vec{r} = \delta\varphi A_1\vec{r} = \delta\varphi(y\vec{e}_z - z\vec{e}_y) = \delta\varphi\,\vec{e}_x\times\vec{r}\,,\\
&\text{around } y: \qquad \delta\vec{r} = \delta\varphi A_2\vec{r} = \delta\varphi(z\vec{e}_x - x\vec{e}_z) = \delta\varphi\,\vec{e}_y\times\vec{r}\,,\\
&\text{around } z: \qquad \delta\vec{r} = \delta\varphi A_3\vec{r} = \delta\varphi(x\vec{e}_y - y\vec{e}_x) = \delta\varphi\,\vec{e}_z\times\vec{r}\,.
\end{aligned} \tag{I.2.9}$$

It is not difficult to write the displacement δr for an infinitesimal rotation of $\vec{r}$ on an angle $\delta\varphi$ around arbitrary direction specifies by a unit vector $\vec{n}$. Introducing an angle θ between $\vec{n}$ and $\vec{r}$, we have

$$|\delta\vec{r}| = r\sin\theta\delta\varphi\,.$$

The direction of vector $\delta\vec{r}$ is perpendicular to the plane passing through $\delta\vec{\varphi} \equiv \delta\varphi\,\vec{n}$ and $\vec{r}$. Therefore, we have

$$\delta\vec{r} = \delta\vec{\varphi}\times\vec{r}\,. \tag{I.2.10}$$

The formulae (I.2.9) are particular cases of this general formula.

Quite remarkably, the rotation matrices $R_x(\varphi)$, $R_y(\varphi)$ and $R_x(\varphi)$ on a finite angle φ can be uniquely restored via the infinitesimal matrices A_a. Namely the following formulae hold

$$\begin{aligned}
e^{\varphi A_1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} = R_x(\varphi)\,,\\
e^{\varphi A_2} &= \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} = R_y(\varphi)\,,\\
e^{\varphi A_3} &= \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = R_z(\varphi)\,.
\end{aligned} \tag{I.2.11}$$

These formulae will be proved in **Tutorial VI**.

We also point out that

$$\frac{\partial e^{\varphi A_1}}{\partial \varphi}\Big|_{\varphi=0} = A_1\,, \quad \frac{\partial e^{\varphi A_2}}{\partial \varphi}\Big|_{\varphi=0} = A_2\,, \quad \frac{\partial e^{\varphi A_3}}{\partial \varphi}\Big|_{\varphi=0} = A_3\,.$$

This property justifies to call A_1, A_2, A_3 the *infinitesimal generators* of the rotation group. It is not difficult to prove that a rotation on an angle φ around a unit vector $\vec{n}$ is given by

$$R(\varphi\vec{n}) = e^{\varphi\,\vec{n}\cdot\vec{A}}\,.$$

Here the 3×3 matrix $\vec{n}\cdot\vec{A}$ in the exponent has the matrix elements

$$(\vec{n}\cdot\vec{A})_{ab} = -n_c\epsilon_{cab} = -\epsilon_{abc}n_c\,.$$

Explicit computation reveals that R has the following matrix elelments

$$R_{ab}(\varphi\vec{n}) = \cos\varphi\,\delta_{ab} + (1-\cos\varphi)n_a n_b - \sin\varphi\,\epsilon_{abc}n_c\,, \quad a,b = 1,\dots 3\,. \tag{I.2.12}$$

Matrix R is orthogonal and it represents an element of the rotation group SO(3). Finally, we mention one important fact, namely, the tensor ϵ_{abc} is invariant under rotations, *i.e.*

$$R_{aa'}R_{bb'}R_{cc'}\epsilon_{a'b'c'} = \det R\,\epsilon_{abc} = \epsilon_{abc}\,.$$

Now we turn to the conservation law which arises due to the symmetry of the lagrangian with respect to simultaneous rotation of all $\vec{r}_i$ by an infinitesimal constant angle $\delta\varphi$ around arbitrary direction $\vec{n}$

$$\delta\vec{r}_i = \delta\varphi\,\vec{n}\times\vec{r}_i\,, \qquad \delta\varphi \equiv \epsilon\,.$$

According to Noether's theorem the following charge is conserved

$$J = \delta\varphi\sum_{i=1}^{N}(\vec{p}_i,\,\vec{n}\times\vec{r}_i) = -\delta\varphi\sum_{i=1}^{N}(\vec{p}_i,\,\vec{r}_i\times\vec{n}) = -\delta\varphi\sum_{i=1}^{N}(\vec{p}_i\times\vec{r}_i,\vec{n}) = \delta\varphi\sum_{i=1}^{N}(\vec{r}_i\times\vec{p}_i,\vec{n})\,.$$

Since $\delta\varphi$ and $\vec{n}$ are arbitrary, we conclude that the total angular momentum

$$\vec{L} = \sum_{i=1}^{N}\vec{r}_i\times\vec{p}_i$$

is conserved.

Since the definition of angular momentum includes $\vec{r}_i$'s, its definition does depend on the choice of the coordinate origin. Radius-vectors $\vec{r}_i$ and $\vec{r}\,'_i$ of the one and the same point but measured with respect to different coordinate origins are related as $\vec{r}_i = \vec{r}\,'_i + \vec{a}$, where $\vec{a}$ is a constant vector. Therefore,

$$\vec{L} = \sum_{i=1}^{N}\vec{r}_i\times\vec{p}_i = \sum_{i=1}^{N}\vec{r}\,'_i\times\vec{p}_i + \vec{a}\times\sum_{i=1}^{N}\vec{p}_i = \vec{L}' + \vec{a}\times\vec{P}\,.$$

As is clear form this formula, the angular momentum does not depend on the choice of the coordinate origin if and only if the system as a whole is at rest, *i.e.* $\vec{P}=0$. This ambiguity does not influence conservation of $\vec{L}$ because $\vec{P}$ is conserved.

Although conservation of all three components of angular momentum (with respect to an arbitrarily chosen origin of the reference frame) takes place only for a closed system, in a more restricted form this conservation law can sometimes be applied for systems in an external field. *What is conserved in this case is the projection of the angular momentum on the axis with respect to which this field is symmetric*, so that mechanical properties of the system do not change under any rotation around this axis. Of course, in such a situation the angular momentum should be defined with respect to a point (the origin of a reference frame) that lies on the same axis. One of the most important examples of this kind is provided by a field with the central symmetry, *i.e.* a field for which the potential energy depends only on the distance to a definite point (field center) in space. For motion in such a field the projection of the momentum on any axis passing through the center is conserved. In other words, the vector $\vec{L}$ is conserved when being defined not with respect to an arbitrary point in space but with respect to the field center.

This exhausts all *additive* integrals of motion. Thus, any closed system has 7 such integrals: energy, 3 components of momentum and 3 components of angular momentum. Finally, we note that for all the symmetry transformations we have considered so far the integration measure $\mathrm{d}t$ in the action did not transform (even for in the case of energy $\mathrm{d}t \to \mathrm{d}(t+\epsilon) = \mathrm{d}t$).

2.3 Oscillations

Here we study lagrangian systems that perform small coupled oscillations and show that they factorise into direct product of systems with one degree of freedom. To better appreciate the context of this discussion, we start with considering an example of coupled pendulums.

2.3.1 Coupled pendulums

Example. Coupled pendulums. Consider a system of two mathematical pendulums of equal length ℓ and of equal mass $m_1 = m_2 = m$ connected by a massless spiring with Hook's constant κ. We assume that in a state of equilibrium the length of the spring is constant and equals to d. We choose the angles φ_1 and φ_2 as generalised coordinates. Then cartesian coordinates of the masses are

$$\begin{aligned} x_1 &= l\sin\varphi_1\,, \qquad y_1 = l\cos\varphi_1\,, \\ x_2 &= d + l\sin\varphi_2\,, \qquad y_2 = l\cos\varphi_2\,. \end{aligned}$$

The kinetic energy is

$$T = \frac{m_1}{2}(\dot{x}_1^2 + \dot{y}_1^2) + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2) = \frac{ml^2}{2}(\dot{\varphi}_1^2 + \dot{\varphi}_2^2)\,.$$

The potential energy comes from two sources: the potential energy U_{gr} of masses in the gravitational field of the Earth and the potential energy U_{spring} stored in the spring. We have

$$U_{\mathrm{gr}} = -m_1 g y_1 - m_2 g y_2 = -mgl\cos\varphi_1 - mgl\cos\varphi_1 = -mgl(\cos\varphi_1 + \cos\varphi_2)\,.$$

To find the energy stored in the spring, we notice that the length of the spring in comparison to its equilibrium length is $x_2 - x_1 - d$ and, therefore,

$$U_{\mathrm{spring}} = \frac{\kappa}{2}(x_2 - x_1 - d)^2 = \frac{\kappa l^2}{2}(\sin\varphi_1 - \sin\varphi_2)^2\,.$$

The lagrangian of the system of coupled oscillators is then

$$L = T - U = \frac{ml^2}{2}(\dot{\varphi}_1^2 + \dot{\varphi}_2^2) + mgl(\cos\varphi_1 + \cos\varphi_2) - \frac{\kappa l^2}{2}(\sin\varphi_1 - \sin\varphi_2)^2\,.$$

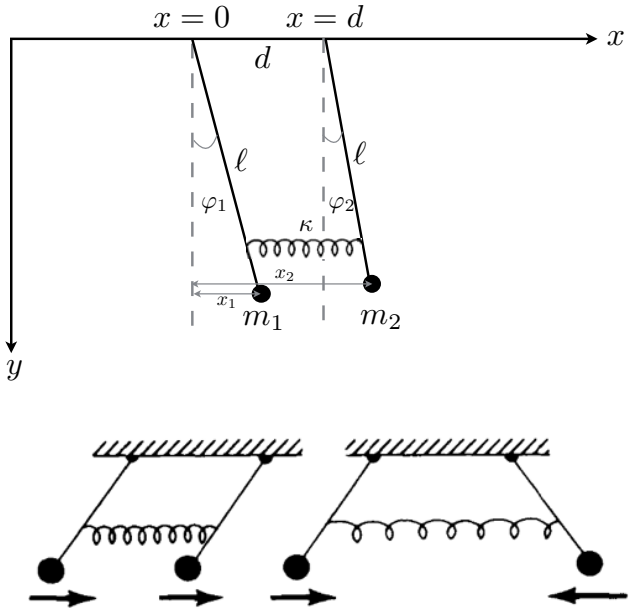

Figure 2.5: The upper figure shows coupled pendulums. The lower figure demonstrates characteristic oscillations of coupled pendulums.

For small oscillations, *i.e.* when φ_1 and φ_2 are small, we approximate $\cos\varphi_{1,2} \approx 1 - \frac{1}{2}\varphi_{1,2}^2$ and $\sin\varphi_{1,2} \approx \varphi_{1,2}$ and the lagrangian for small oscillations takes the form

$$L = T - U = \frac{ml^2}{2}(\dot\varphi_1^2 + \dot\varphi_2^2) - \frac{mgl}{2}(\varphi_1^2 + \varphi_2^2) - \frac{\kappa l^2}{2}(\varphi_1 - \varphi_2)^2\,, \tag{I.2.13}$$

where we have omitted an unessential constant term. The Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\dot\varphi_{1,2}} - \frac{\partial L}{\partial\varphi_{1,2}} = 0$$

are

$$\begin{aligned} ml^2\ddot\varphi_1 + mgl\varphi_1 + \kappa l^2(\varphi_1 - \varphi_2) = 0\,,\\ ml^2\ddot\varphi_2 + mgl\varphi_2 + \kappa l^2(\varphi_2 - \varphi_1) = 0\,. \end{aligned}$$

This is the system of two coupled 2nd order ordinary differential equations which we further brush up to

$$\begin{aligned} \ddot\varphi_1 + \Big(\frac{g}{l} + \frac{\kappa}{m}\Big)\varphi_1 - \frac{\kappa}{m}\varphi_2 = 0\,,\\ \ddot\varphi_2 + \Big(\frac{g}{l} + \frac{\kappa}{m}\Big)\varphi_2 - \frac{\kappa}{m}\varphi_1 = 0\,. \end{aligned} \tag{I.2.14}$$

It is convenient to rewrite this system in the matrix form

$$\left(\frac{d^2}{dt^2} + M\right)\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} = 0\,, \tag{I.2.15}$$

where a matrix M is

$$M = \begin{pmatrix} \frac{g}{l} + \frac{\kappa}{m} & -\frac{\kappa}{m}\\ -\frac{\kappa}{m} & \frac{g}{l} + \frac{\kappa}{m}\end{pmatrix}.$$

We will look for a solution of this system in the form

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \begin{pmatrix} \varphi_1^0 \\ \varphi_2^0 \end{pmatrix} e^{i\omega t}\,, \qquad \vec{\varphi}_0 = \begin{pmatrix} \varphi_1^0 \\ \varphi_2^0 \end{pmatrix}\,,$$

where φ_1^0 and φ_2^0 are constants. Plugging this ansatz into (I.2.15), we obtain the following matrix equation

$$M\vec{\varphi}_0 = \omega^2\vec{\varphi}_0\,.$$

This is an eigenvalue problem for the matrix M, where ω^2 is an eigenvalue and $\vec{\varphi}_0$ is the corresponding eigenvector. According to the standard procedure, eigenvalues are then found by solving the *characteristic equation*, namely,

$$\det(M - \omega^2\mathbb{1}) = 0\,.$$

In our present case this equation is

$$\begin{aligned} \det(M-\omega^2\mathbb{1}) &= \det\begin{pmatrix} \frac{g}{l}+\frac{\kappa}{m}-\omega^2 & -\frac{\kappa}{m} \\ -\frac{\kappa}{m} & \frac{g}{l}+\frac{\kappa}{m}-\omega^2 \end{pmatrix} \\ &= \frac{1}{lm^2}(g-l\omega^2)(2\kappa l + m(g-l\omega^2)) = 0\,. \end{aligned}$$

Solutions of this equation are

$$\omega_1^2 = \frac{g}{l}\,, \quad \& \quad \omega_2^2 = \frac{g}{l} + 2\frac{\kappa}{m}\,.$$

These ω's are called *characteristic or eigen frequencies*. Now we can determine the eigenvectors. We obtain

$$\begin{aligned} \omega_1^2: &\quad \begin{pmatrix} \frac{\kappa}{m} & -\frac{\kappa}{m} \\ -\frac{\kappa}{m} & \frac{\kappa}{m} \end{pmatrix}\begin{pmatrix} \varphi_1^0 \\ \varphi_2^0 \end{pmatrix} = 0 \quad \Rightarrow \quad \begin{pmatrix} \varphi_1^0 \\ \varphi_2^0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\,, \\ \omega_2^2: &\quad \begin{pmatrix} -\frac{\kappa}{m} & -\frac{\kappa}{m} \\ -\frac{\kappa}{m} & -\frac{\kappa}{m} \end{pmatrix}\begin{pmatrix} \varphi_1^0 \\ \varphi_2^0 \end{pmatrix} = 0 \quad \Rightarrow \quad \begin{pmatrix} \varphi_1^0 \\ \varphi_2^0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\,, \end{aligned}$$

where we have presented the normalised eigenvectors with the norm equal to one. Since the Euler-Lagrange equations are linear, the general solution is obtained by superposition of oscillations with characteristic frequencies and it reads as

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \frac{A}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{i\omega_1 t} + \frac{A^*}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-i\omega_1 t} + \frac{B}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega_2 t} + \frac{B^*}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-i\omega_2 t}\,,$$

i.e. there are 4 integration constants represented by two complex numbers A and B.

Special cases. There are two special cases.

1) $B = B^* = 0$. We have

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = \frac{A}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{i\omega_1 t} + \frac{A^*}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-i\omega_1 t}\,.$$

For this case, $\varphi_1 = \varphi_2$ for any time and $\omega_1^2 = \frac{g}{l}$. The spring is not stretched at all, see the lower left image on Fig. 2.5. The system behaves itself as a single pendulum.

2) $A = A^* = 0$. We have

$$\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = +\frac{B}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega_2 t} + \frac{B^*}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-i\omega_2 t} .$$

For this case, $\varphi_1 = -\varphi_2$ for any time and $\omega_1^2 = \frac{g}{l} + 2\frac{\kappa}{m}$. The spring reaches its maximal and minimal stretch, see the lower right image on Fig. 2.5.

These two special modes are called *eigen oscillations*. A general motion is an overlapping of these two modes. Let us now define the so-called normal coordinates

$$\begin{aligned} \Theta_1 &= \varphi_1 + \varphi_2 , \\ \Theta_2 &= \varphi_1 - \varphi_2 , \end{aligned}$$

so that

$$\begin{aligned} \varphi_1 &= \frac{1}{2}(\Theta_1 + \Theta_2) , \\ \varphi_2 &= \frac{1}{2}(\Theta_1 - \Theta_2) . \end{aligned}$$

Adding and subtracting equations (I.2.14), we then rewrite them via normal coordinates as

$$\begin{aligned} (1)+(2): &\quad \ddot{\Theta}_1 + \Big(\frac{g}{l} + \frac{\kappa}{m}\Big)\Theta_1 - \frac{\kappa}{m}\Theta_1 = 0 , \\ (1)-(2): &\quad \ddot{\Theta}_2 + \Big(\frac{g}{l} + \frac{\kappa}{m}\Big)\Theta_2 + \frac{\kappa}{m}\Theta_2 = 0 . \end{aligned}$$

Collecting similar terms, we bring the last two equations to the form

$$\begin{aligned} &\ddot{\Theta}_1 + \omega_1^2\Theta_1 = 0 , \quad \omega_1^2 = \frac{g}{l} , \\ &\ddot{\Theta}_2 + \omega_2^2\Theta_2 = 0 , \quad \omega_2^2 = \frac{g}{l} + 2\frac{\kappa}{m} . \end{aligned}$$

Thus, we observe that in the normal coordinates the differential equations decouple and describe two independent harmonic oscillators with frequencies ω_1 and ω_2. It remains to see how the lagrangian (I.2.13) looks in the normal coordinates. We have

$$L = \frac{ml^2}{2}\left(\frac{1}{4}(\dot{\Theta}_1 + \dot{\Theta}_2)^2 + \frac{1}{4}(\dot{\Theta}_1 - \dot{\Theta}_2)^2\right) - \frac{mgl}{2}\left(\frac{1}{4}(\Theta_1 + \Theta_2)^2 + \frac{1}{4}(\Theta_1 - \Theta_2)^2\right) - \frac{\kappa l^2}{2}Q_2^2 .$$

Simplifying, we obtain

$$L = \frac{ml^2}{4}\left[\dot{\Theta}_1^2 - \frac{g}{l}\Theta_1^2 + \dot{\Theta}_2^2 - \left(\frac{g}{l} + 2\frac{\kappa}{m}\right)\Theta_2^2\right] .$$

Thus, we observe that the original lagrangian factorises in the normal coordinates into the sum $L = L_1(\Theta_1) + L_2(\Theta_2)$, where

$$\begin{aligned} L_1 &= \frac{ml^2}{4}\left(\dot{\Theta}_1^2 - \omega_1^2\Theta_1^2\right) , \\ L_2 &= \frac{ml^2}{4}\left(\dot{\Theta}_1^2 - \omega_1^2\Theta_1^2\right) . \end{aligned}$$

This finishes our considerations of the example of coupled pendulums.

2.3.2 Oscillations with many degrees of freedom

The theory of free small oscillation is build analogously to the example of coupled oscillator considered above. Let potential energy U of a system considered as a function of generalised coordinates q_i, $i = 1, \ldots, n$, has minimum at $q_i = q_{i0}$. Introduce small displacements

$$x_i = q_i - q_{i0}$$

and expanding U over x_i up to quadratic order, we obtain the potential energy as the following positively definite quadratic form

$$U = \frac{1}{2} \sum_{ij} k_{ij} x_i x_j \,,$$

where we measure U from its minimum. Coefficients k_{ij} can be viewed as symmetric, *i.e.* $k_{ij} = k_{ji}$. The kinetic energy has in general the following form

$$T = \frac{1}{2} \sum_{ij} g_{ij}(q) \dot{q}_i \dot{q}_j \,.$$

Setting in $g_{ij}(q)$ the coordinates q_i to q_{i0} and denoting $g_{ij}(q_0) = m_{ij}$, we obtain T also as a positive definite quadratic form

$$T = \frac{1}{2} \sum_{ij} m_{ij} \dot{x}_i \dot{x}_j \,,$$

where the coefficients m_{ij} are also regarded as symmetric, $m_{ij} = m_{ji}$. Therefore, the lagrangian for a free system performing small oscillations is

$$L = \frac{1}{2} \sum_{ij} m_{ij} \dot{x}_i \dot{x}_j - \frac{1}{2} \sum_{ij} k_{ij} x_i x_j \,.$$

The Euler-Lagrange equations are

$$\sum_j m_{ij} \ddot{x}_j + \sum_j k_{ij} x_j = 0 \,, \quad i = 1, \ldots, n \,. \tag{I.2.16}$$

This is a system of n homogeneous differential equations with constant coefficients. We look for the general solution for n functions $x_j(t)$ in the form

$$x_j = A_j e^{i\omega t} \,, \tag{I.2.17}$$

where A_j are constant unknowns. Substituting these x_j into the differential equations and cancelling out $e^{i\omega t}$, we obtain a system of linear homogeneous algebraic equations for the unknowns A_j, namely,

$$\sum_j (k_{ij} - \omega^2 m_{ij}) A_j = 0 \,, \tag{I.2.18}$$

or in the matrix form

$$(k - \omega^2 m) A = 0 \,, \quad k = ||k_{ij}|| \,, \quad m = ||m_{ij}|| \,.$$

For this system to have non-vanishing solutions, the determinant of the matrix $k - \omega^2 m$ must vanish

$$\det(k - \omega^2 m) = 0 \,.$$

This is a characteristic equation of order n with respect to ω^2. In general it has n different real positive roots ω_α^2, $\alpha = 1, \ldots, n$. These roots are called *eigen or normal frequences*. Reality and positivity of roots can be deduced from the following argument. We multiply equation (I.2.18) by A_i^* and sum over i

$$\sum_{ij} A_i^*(k_{ij} - \omega^2 m_{ij})A_j = 0\,,$$

from where we express ω^2 as

$$\omega^2 = \frac{\sum_{ij} A_i^* k_{ij} A_j}{\sum_{ij} A_i^* m_{ij} A_j}\,.$$

Since k and m are real and symmetric matrices, the numerator and denominator are both real, moreover, positive definiteness of k and m implies that the numerator and denominator are both positive.

After ω_α^2 are found, substituting each of them into (I.2.18) we can find the corresponding A_j. If all the roots of the characteristic equation are different, then the coefficients A_k are proportional to cofactors (adjuncts) of the matrix $k - \omega^2 m$, where ω^2 is replaced with ω_α^2.[5] Denoting these cofactors as $\Delta_{\alpha j}$, we obtain a particular solution of the Euler-Lagrange equations

$$x_j = \Delta_{\alpha j} C_\alpha e^{i\omega_\alpha t}\,,$$

where C_α is an arbitrary complex constant. The general solution is then the sum of all particular solutions. Passing to the real part, we have

$$x_j = \mathrm{Re}\left\{ \sum_{\alpha=1}^{n} \Delta_{\alpha j} C_\alpha e^{i\omega_\alpha t} \right\} = \sum_{\alpha=1}^{n} \Delta_{\alpha j} \Theta_\alpha\,, \tag{I.2.19}$$

where we introduced

$$\Theta_\alpha = \mathrm{Re}\Big\{ C_\alpha e^{i\omega_\alpha t} \Big\}\,.$$

Therefore, motion of each of the coordinates in time represents an overlaying of n simple periodic oscillations $\Theta_1, \ldots, \Theta_n$ with arbitrary amplitudes and phases but with definite frequencies. The quantities $\Theta_1, \ldots, \Theta_n$ can be conveniently taken as new generalised coordinates, these new coordinates are precisely the *normal coordinates*. From their definition it follows that they satisfy

$$\ddot{\Theta}_\alpha + \omega_\alpha^2 \Theta_\alpha = 0\,,$$

i.e. in the normal coordinates the equations of motion factorise into n independent equations. In other words, normal oscillations are fully independent.

[5] For a symmetric matrix $\Delta_{\alpha j} = \Delta_{j\alpha}$. The proof that $A_k = \Delta_{\alpha k}$ up a an arbitrary simultaneous rescaling of all A_k goes as follows. Assume that ω does not coincide with any of the roots ω_α, $\alpha = 1, \ldots, n$. Then the matrix $k - \omega^2 m$ is invertible and by the general rule of computing the inverse via its cofactors $\Delta_{ij}(\omega)$ we have

$$(k - \omega^2 m)^{-1}_{ik} = \frac{\Delta_{ki}(\omega)}{\Delta(\omega)}\,, \quad \Delta(\omega) = \det(k - \omega^2 m)\,.$$

Thus, with Einstein's convention for summation of indices,

$$(k - \omega^2 m)_{ik}(k - \omega^2 m)^{-1}_{kj} = (k - \omega^2 m)_{ik} \frac{\Delta_{jk}(\omega)}{\Delta(\omega)} = \delta_{ij}\,,$$

or

$$(k - \omega^2 m)_{ik} \Delta_{jk}(\omega) = \Delta(\omega)\delta_{ij} \qquad \forall\, i, j\,.$$

Sending here $\omega \to \omega_\alpha$ and taking into account that $\Delta(\omega_\alpha) = 0$, we obtain that $(k - \omega_\alpha^2 m)_{ik}\Delta_{jk}(\omega_\alpha) = 0$ for any i and j. Taking further $j = \alpha$ we obtain the proof of the statement.

It follows from this discussion that the lagrangian expressed via normal coordinates turns into a sum of expressions, each of which corresponds to one-dimensional oscillations with frequency ω_α and, therefore, it has the form

$$L = \sum_\alpha \frac{m_\alpha}{2}\left(\dot\Theta_\alpha^2 - \omega_\alpha^2\Theta_\alpha^2\right),$$

where m_α are positive constants. From the mathematical point of view this means that transformation (I.2.19) diagonalises both quadratic forms corresponding to the kinetic and potential energy.

Usually, normal coordinates are chosen in such a way that the coefficients of velocities squared in the lagrangian are equal to 1/2. For this it is enough to define the formal coordinates as $Q_\alpha = \sqrt{m_\alpha}\Theta_\alpha$, so that

$$L = \sum_\alpha \frac{1}{2}\left(\dot Q_\alpha^2 - \omega_\alpha^2 Q_\alpha^2\right).$$

If there are multiple (degenerate) frequencies, then the number of the normal coordinates corresponding to each degenerate frequency coincides with the order of degeneration. However, the choice of these normal coordinates is not unique. Since normal coordinates enter in the kinetic and potential energy (with equal ω_α) enter as equally transformed quantities $\sum \dot Q_\alpha^2$ and $\sum Q_\alpha^2$, they can be simultaneously transformed by any transformation which leaves the sum of squares invariant.

Finally, we note that the case of a degenerate matrix k defining the potential energy will be considered in the next subsection and in **Tutorial VII**.

2.3.3 Periodic chain of coupled oscillators[6]

As another application of the techniques discussed above we consider a linear chain of coupled oscillators. This picture of a linear chain of coupled oscillators and its three-dimensional generalisation is used in solid state physics to model the vibrational motion of atoms in a solid. The masses represent the atomic nuclei that make up the solid and the spacing between the masses is the atomic separation. The "springs" coupling the masses represent a harmonic approximation to the forces binding the nuclei into the solid. In the context of applications to solid state physics the *normal modes are identified with phonons.* After incorporating quantum mechanics, this phonon picture of vibrational modes of a solid is used to describe thermal conductivity, specific heat, propagation of sound, and other properties of the solid.

In fact, here we consider a chain of oscillators with periodic boundary conditions. These boundary conditions will be bring a new complication, namely, the quadratic form defining the potential energy appears to be degenerate, and we have to modify out treatment of normal modes to account for this feature.

Example. Periodic linear chain of oscillators coupled via springs. We consider $N \in 2\mathbb{N}$ particles of equal mass m coupled by springs with equal Hooke's constant κ and impose on generalised coordinates the periodicity condition $q_{N+1} = q_1$. The generalised coordinates are introduced as displacements

$$x_i = q_i - (i-1)d\,.$$

The lagrangian is

$$L = T - U \quad = \quad \frac{m}{2}\sum_{i=1}^N \dot x_i^2 - \frac{\kappa}{2}\sum_{i=1}^N (x_{i+1} - x_i)^2 = \frac{1}{2}\sum_{ij} m_{ij}\dot x_i \dot x_j - \frac{1}{2}\sum_{ij} k_{ij} x_i x_j\,.$$

[6]This subsection is optional.

where $m_{ij} = m\delta_{ij}$ and

$$||k_{ij}|| = \kappa \cdot \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & -1 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & -1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -1 & 0 \\ 0 & \dots & 0 & -1 & 2 & -1 \\ -1 & 0 & \dots & 0 & -1 & 2 \end{pmatrix} .$$

The Euler-Lagrange equations are

$$m\ddot{x}_i + \sum_j k_{ij} x_j = 0\,, \quad i = 1, \dots, N\,. \tag{I.2.20}$$

Before we proceed with finding their general solution, we note that the matrix k is degenerate, $\det k = 0$, and it has a null vector

$$A^{(0)} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} ,$$

that is $kA^{(0)} = 0$. The presence of the null vector leads to an existence of a special solution of (I.2.20), namely, the solution which is a vector $x^{(0)}$ obeying two conditions

$$\ddot{x} = 0\,, \qquad kx = 0\,.$$

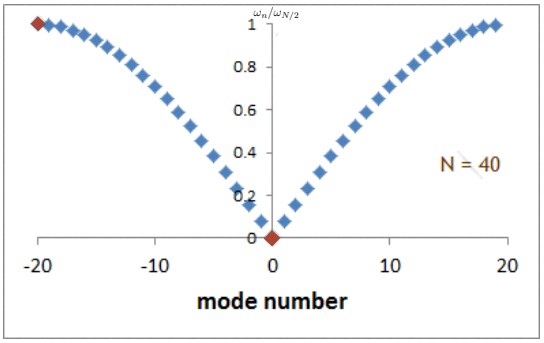

Figure 2.6: Normal modes.

A general solution of these conditions is then

$$x^{(0)} = (a + bt)A^{(0)}\,, \tag{I.2.21}$$

where $a, b \in \mathbb{R}$ are two real arbitrary constants. This special solution does not have an oscillating character, rather it describes a linear motion with a constant velocity.

To find other solutions of (I.2.20), we employ (I.2.17). The eigenvectors satisfy

$$\sum_{k=1}^{N} (k_{jk} - m\omega^2 \delta_{jk}) A_k = 0\,.$$

To solve this equation, we consider an ansatz $A_k = e^{ikp}$ with $p = \text{const}$. Then for a fixed index j

$$\begin{aligned} \sum_{k=1}^{N} k_{jk} A_k &= \kappa(-e^{i(j-1)p} + 2e^{ijp} - e^{i(j+1)p}) \\ &= \kappa e^{ijp}(2 - e^{-ip} - e^{ip}) = 2\kappa A_j(1 - \cos p) = 4\kappa A_j \sin^2 \frac{p}{2}\,. \end{aligned}$$

Thus,

$$\sum_{k=1}^{N} (k_{jk} - m\omega^2 \delta_{jk}) A_k = \left(4k \sin^2 \frac{p}{2} - m\omega^2\right) A_j = 0$$

from where we find

$$\omega_p^2 = \frac{4\kappa}{m}\sin^2\frac{p}{2}\,.$$

In general, the dependence of frequency ω on the wave number n (or the phonon momentum p) is called *dispersion relation*. Now we have to take into account the boundary condition $x_{N+1} = x_1$, *i.e.* $A_{N+1} = A_1$ which results into

$$e^{i(N+1)p} = e^{ip} \quad \Rightarrow \quad e^{iNp} = 1 \quad \Rightarrow \quad p = \frac{2\pi n}{N}\,, \quad n \in \mathbb{Z}\,.$$

Thus,

$$A_j^{(n)} = e^{\frac{2\pi i j n}{N}}\,, \qquad \omega_n^2 = \frac{4\kappa}{m}\sin^2\frac{\pi n}{N}\,,$$

so that

$$\omega_n = \sqrt{\frac{4\kappa}{m}}\Big|\sin\frac{\pi n}{N}\Big|\,.$$

It is clear that shifting $n \to n+N$ does not give any new A because $A_j^{(n)} = A_j^{(n+N)}$ and it does not change $\omega_n:\ \omega_{n+N} = \omega_n$. Therefore, we can restrict n to lie in the interval

$$-\frac{N}{2} \le n \le \frac{N}{2} - 1\,, \tag{I.2.22}$$

where we recall that according to our assumption N is even. Thus, a particular solution labeled by n from (I.2.22)

$$x_j^{(n)}(t) = \mathrm{Re}\left[C_n e^{\frac{2\pi i j n}{N}} e^{i\omega_n t}\right]\,. \tag{I.2.23}$$

For $n = 0$ we see that $\omega_0 = 0$ and the corresponding constant solution should be replaced by a more general solution $x_j^{(0)}$ given by (I.2.21). The general solution is then the sum of particular solutions[7]

$$x_j(t) = \frac{2}{\sqrt{N}}\sum_{\substack{n=-N/2\\ n\neq 0}}^{N/2-1}\mathrm{Re}\left[C_n e^{\frac{2\pi i j n}{N}} e^{i\omega_n t}\right] + \frac{1}{\sqrt{N}}x_j^{(0)}(t)\,. \tag{I.2.24}$$

Our next goal will be to represent this solution as a superposition of the normal modes. To this end, we consider two terms in the above corresponding to $n \neq -N/2$ and $-n$ and manipulate it in the following way

$$\begin{aligned}
&2\mathrm{Re}\left[C_n e^{\frac{2\pi i j n}{N}} e^{i\omega_n t}\right] + 2\mathrm{Re}\left[C_{-n} e^{-\frac{2\pi i j n}{N}} e^{i\omega_n t}\right]\\
&= C_n e^{\frac{2\pi i j n}{N}} e^{i\omega_n t} + C_n^* e^{-\frac{2\pi i j n}{N}} e^{-i\omega_n t} + C_{-n} e^{-\frac{2\pi i j n}{N}} e^{i\omega_n t} + C_{-n}^* e^{\frac{2\pi i j n}{N}} e^{-i\omega_n t}\\
&= e^{\frac{2\pi i j n}{N}}\Theta_n(t) + e^{-\frac{2\pi i j n}{N}}\Theta_{-n}(t)\,,
\end{aligned}$$

where we have introduced the normal modes Θ_n

$$\begin{aligned}
\Theta_n(t) &= C_n e^{i\omega_n t} + C_{-n}^* e^{-i\omega_n t}\,,\\
\Theta_{-n}(t) &= C_n^* e^{-i\omega_n t} + C_{-n} e^{i\omega_n t}\,.
\end{aligned}$$

By construction, we see that these modes obey the following reality condition

$$\Theta_n^* = \Theta_{-n}\,. \tag{I.2.25}$$

[7]The overall normalisation $1/\sqrt{N}$ is introduced for further convenience.

The term in the sum $n = -N/2$ requires a separate treatment. We have

$$2\mathrm{Re}\left[C_{-N/2}e^{-\pi ij}e^{i\omega_{N/2}t}\right] = (-1)^j\left(C_{-N/2}e^{i\omega_{N/2}t} + C^*_{-N/2}e^{-i\omega_{N/2}t}\right) = e^{i\pi j}\Theta_{N/2,j} + e^{-i\pi j}\Theta_{-N/2,j}\,,$$

where we have introduced the components of the modes $Q_{N/2}$ and $Q_{-N/2}$

$$\begin{aligned}\Theta_{N/2} &\equiv \frac{1}{2}\left(C_{-N/2}e^{i\omega_{N/2}t} + C^*_{-N/2}e^{-i\omega_{N/2}t}\right),\\ \Theta_{-N/2} &\equiv \frac{1}{2}\left(C_{-N/2}e^{i\omega_{N/2}t} + C^*_{-N/2}e^{-i\omega_{N/2}t}\right).\end{aligned}$$

Clearly, the modes $\Theta_{N/2}$ and $\Theta_{-N/2}$ are both real and equal to each other

$$\Theta^*_{N/2} = \Theta_{N/2} = \Theta_{-N/2} = \Theta^*_{-N/2}\,. \tag{I.2.26}$$

Finally, introducing a real mode $\Theta_0(t) = x^{(0)}(t)$, we write solution (I.2.24) as an expansion over normal modes

$$x_j(t) = \frac{1}{\sqrt{N}}\sum_{n=-N/2}^{N/2} e^{\frac{2\pi ijn}{N}}\Theta_n(t)\,. \tag{I.2.27}$$

In such a form reality of x_j follows from the reality conditions (I.2.25) and (I.2.26).

Let us count the number of integration constants. There are two real constants in Θ_0 and, since $C_{-N/2}$ is complex, $\Theta_{N/2}$ also depends on two real constants. The rest of independent modes is delivered by Θ_n with $n = 1, \ldots, N/2-1$, each of Θ_n depends on 4 real constants (two complex C_n and C^*_n). Thus, the total number of real integration constants is

$$2 + 2 + 4(N/2 - 1) = 2N\,,$$

as it should be for a system with N degrees of freedom.

Let us now perform an exercise of rewriting the original lagrangian in terms of the normal modes. To this end, we will use formula (I.2.27) which we simply consider as a change of variables from x to Θ in the lagrangian.[8] First we compute

$$\sum_{j=1}^{N}\dot{x}_j\dot{x}_j = \frac{1}{N}\sum_{n=-N/2}^{N/2}\sum_{m=-N/2}^{N/2}\dot{\Theta}_n\dot{\Theta}_m\sum_{j=1}^{N}e^{\frac{2\pi ij(n+m)}{N}}\,,$$

where we need to evaluate the following geometric sum with $q = e^{\frac{2\pi i(n+m)}{N}}$

$$\sum_{j=1}^{N}e^{\frac{2\pi ij(n+m)}{N}} = \sum_{j=1}^{N}q^j = \frac{q(q^N-1)}{q-1} = \frac{e^{\frac{2\pi i(n+m)}{N}}(e^{2\pi i(n+m)}-1)}{e^{\frac{2\pi i(n+m)}{N}}-1} = \begin{cases} 0 \text{ if } n+m \neq 0 \\ N \text{ if } n+m = 0 \end{cases}.$$

Thus, the kinetic energy is

$$T = \frac{m}{2}\sum_{j=1}^{N}\dot{x}_j\dot{x}_j = \frac{1}{N}\sum_{n=-N/2}^{N/2}\sum_{m=-N/2}^{N/2}\dot{\Theta}_n\dot{\Theta}_m N\delta_{m,-n} = \frac{m}{2}\sum_{n=-N/2}^{N/2}|\dot{\Theta}_n|^2\,,$$

[8] We do not use our knowledge of the solution for Θ's.

where we have used the reality conditions (I.2.25) and (I.2.26). Analogously, for the potential energy one finds

$$\begin{aligned} U &= \frac{1}{2}\sum_{ij} k_{ij}x_i x_j = \frac{1}{2N}\sum_{n=-N/2}^{N/2}\sum_{s=-N/2}^{N/2}\sum_i A_i^{(n)} \underbrace{\Big[\sum_j k_{ij}A_j^{(s)}\Big]}_{m\omega_s^2 A_i^{(s)}} \Theta_n\Theta_s \\ &= \frac{m}{2N}\sum_{n=-N/2}^{N/2}\sum_{s=-N/2}^{N/2}\omega_s^2\Theta_n\Theta_s \underbrace{\sum_{i=1}^{N} A_i^{(n)}A_i^{(s)}}_{N\delta_{n+s,0}} = \frac{m}{2}\sum_{n=-N/2}^{N/2}\omega_n^2|\Theta_n|^2\,, \end{aligned}$$

where we again used the geometric sum and the reality conditions. Thus, the lagrangian is indeed factorised into a sum of the individual normal mode contributions

$$L = \frac{m}{2}\sum_{n=-N/2}^{N/2}\Big(|\dot\Theta_n|^2 - \omega_n^2|\Theta_n|^2\Big)\,.$$

With the reality conditions taken into account, this lagrangian can be more explicitly written as

$$\begin{aligned} L &= \frac{m}{2}\dot\Theta_0^2 + m\Big(\dot\Theta_{N/2}^2 - \omega_{N/2}^2\Theta_{N/2}^2\Big) \\ &+ m\sum_{n=1}^{N/2-1}\left[\Big(\mathrm{Re}\dot\Theta_n\Big)^2 - \omega_n^2\Big(\mathrm{Re}\Theta_n\Big)^2 + \Big(\mathrm{Im}\dot\Theta_n\Big)^2 - \omega_n^2\Big(\mathrm{Im}\Theta_n\Big)^2\right]. \end{aligned}$$

The physical content of this theory constitutes two real non-degenerate normal models (one for for the linear motion and another for oscillations with the maximal frequency $\omega_{N/2}$) and $2(N/2-1)$ doubly degenerate modes, see Fig. 2.6. Double degeneracy means that for one ω_n there are two different oscillatory modes corresponding to $\mathrm{Re}\Theta_n$ and $\mathrm{Im}\Theta_n$. The total number of real normal modes entering the lagrangian is $1+1+2(N/2-1) = N$, $N-1$ of which are oscillatory and one corresponds to free motion.

2.3.4 Lagrangians for continuous systems

So far our discussion concerned a dynamical system with a finite number of degrees of freedom. To describe continuous systems, such as vibrating solid, a transition to an infinite number of degrees of freedom is necessary. Indeed, one has to specify the position coordinates of all the points which are infinite in number.

The continuum case can be reached by taking the appropriate limit of a system with a finite number of discrete coordinates. Our first example is an elastic rod of fixed length ℓ which undergoes small longitudinal vibrations. We approximate the rod by a system of equal mass m particles spaced a distance Δa apart and connected by uniform massless springs having the force (Hooke's) constant κ. The total length of the system is $\ell = (n+1)\Delta a$. We describe the displacement of the ith particle from its equilibrium position by the coordinate ϕ_i. Then the kinetic energy of the particles is

$$T = \sum_{i=1}^{n}\frac{m}{2}\dot\phi_i^2\,.$$

The potential energy is stored into springs and it is given by the sum

$$U = \frac{1}{2}\kappa\sum_{i=0}^{n}(\phi_{i+1}-\phi_i)^2\,.$$

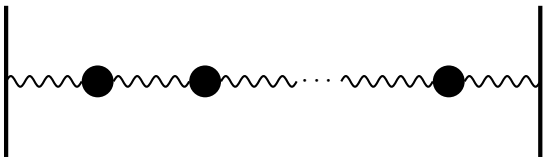

Figure 2.7: Particles connected by springs. Particles are enumerated from the left to the right as $1, 2, \ldots, n$, while springs as $0, 1, \ldots, n$. The distance between two neighbouring particles in the equilibrium position is Δa and the length between the boundaries is $\ell = (n+1)\Delta a$.

Here we associate $\phi_0 = 0 = \phi_{n+1}$ with the end points of the interval which do not move. The force acting on ith particle is $F_i = -\frac{\partial U}{\partial \phi_i}$:

$$F_i = \kappa(\phi_{i+1} + \phi_{i-1} - 2\phi_i)\,.$$

This formula shows that the force exerted by the spring on the right of the ith particle equals to $\kappa(\phi_{i+1} - \phi_i)$, while the force exerted from the left is $\kappa(\phi_i - \phi_{i-1})$. The Lagrangian is

$$L = T - U = \sum_{i=1}^{n} \frac{m}{2}\dot{\phi}_i^2 - \frac{1}{2}\kappa \sum_{i=0}^{n} (\phi_{i+1} - \phi_i)^2\,.$$

At this stage we can take a continuum limit by sending $n \to \infty$ and $\Delta a \to 0$ so that $\ell = (n+1)\Delta a$ is kept fixed. Increasing the number of particles we will be increasing the total mass of a system. To keep the total mass finite, we assume that the ratio $m/\Delta a \to \mu$, where μ is a finite mass density. To keep the force between the particles finite, we assume that in the large particle limit $\kappa \Delta a \to Y$, where Y is a finite quantity. Thus, we have

$$L = T - U = \frac{1}{2}\sum_{i=1}^{n} \Delta a \Big(\frac{m}{\Delta a}\Big)\dot{\phi}_i^2 - \frac{1}{2}\sum_{i=0}^{n} \Delta a(\kappa\,\Delta a)\Big(\frac{\phi_{i+1} - \phi_i}{\Delta a}\Big)^2\,.$$

Taking the limit, we replace the discrete index i by a continuum variable x. As a result, $\phi_i \to \phi(x)$. Also

$$\frac{\phi_{i+1} - \phi_i}{\Delta a} \to \frac{\phi(x + \Delta a) - \phi(x)}{\Delta a} \to \partial_x \phi(x)\,.$$

Thus, taking the limit we find

$$L = \frac{1}{2}\int_0^\ell \mathrm{d}x \left[\mu\dot{\phi}^2 - Y(\partial_x \phi)^2\right].$$

Also equations of motion can be obtained by the limiting procedure. Starting from

$$\frac{m}{\Delta a}\ddot{\phi}_i - k\Delta a \frac{\phi_{i+1} + \phi_{i-1} - 2\phi_i}{\Delta a^2} = 0,$$

and using

$$\lim_{\Delta a \to 0} \frac{\phi_{i+1} + \phi_{i-1} - 2\phi_i}{\Delta a^2} = \frac{\partial^2 \phi}{\partial x^2} \equiv \partial_{xx}\phi$$

we obtain the equation of motion

$$\mu\ddot{\phi} - Y\partial_{xx}\phi = 0\,.$$

Just as there is a generalized coordinate ϕ_i for each i, there is a generalized coordinate $\phi(x)$ for each x. Thus, the finite number of coordinates ϕ_i has been replaced by a function of x. Since ϕ

depends also on time, we are dealing with the function of two variables $\phi(x,t)$ which is called the *displacement field*. The Lagrangian is an integral over x of the *Lagrangian density*

$$\mathscr{L} = \frac{1}{2}\mu\dot{\phi}^2 - \frac{1}{2}Y(\partial_x\phi)^2\,.$$

The action is a functional of $\phi(x,t)$:

$$S[\phi] = \int_{t_1}^{t_2} \mathrm{d}t \int_0^\ell \mathrm{d}x\, \mathscr{L}(\phi(x,t), \dot{\phi}(x,t), \partial_x\phi(x,t))\,.$$

It is possible to obtain the equations of motion for the field $\phi(x,t)$ directly from the continuum Lagrangian. One has to understand how the action changes under an infinitesimal change of the field

$$\phi(x,t) \to \phi(x,t) + \delta\phi(x,t)\,. \tag{I.2.28}$$

The derivatives change accordingly,

$$\begin{aligned}
\frac{\partial}{\partial t}\phi(x,t) &\to \frac{\partial}{\partial t}\phi(x,t) + \frac{\partial}{\partial t}\delta\phi(x,t)\,,\\
\frac{\partial}{\partial x}\phi(x,t) &\to \frac{\partial}{\partial x}\phi(x,t) + \frac{\partial}{\partial x}\delta\phi(x,t)\,.
\end{aligned}$$

This gives

$$\delta S[\phi] = S[\phi+\delta\phi] - S[\phi] = \int_{t_1}^{t_2} \mathrm{d}t \int_0^\ell \mathrm{d}x \Big[\frac{\partial\mathscr{L}}{\partial\phi}\delta\phi + \frac{\partial\mathscr{L}}{\partial\dot{\phi}}\partial_t\delta\phi + \frac{\partial\mathscr{L}}{\partial(\partial_x\phi)}\partial_x\delta\phi\Big]\,.$$

Integrating by parts, we find

$$\begin{aligned}
\delta S[\phi] &= \int_{t_1}^{t_2} \mathrm{d}t \int_0^\ell \mathrm{d}x \Big[\frac{\partial\mathscr{L}}{\partial\phi} - \partial_t\frac{\partial\mathscr{L}}{\partial\dot{\phi}} - \partial_x\frac{\partial\mathscr{L}}{\partial(\partial_x\phi)}\Big]\delta\phi\\
&+ \int_0^\ell \mathrm{d}x \frac{\partial\mathscr{L}}{\partial(\partial_t\phi)}\delta\phi\big|_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2} \mathrm{d}t \frac{\partial\mathscr{L}}{\partial(\partial_x\phi)}\delta\phi\big|_{x=0}^{x=\ell}\,.
\end{aligned} \tag{I.2.29}$$

The action principle requires that the action principle be stationary with respect to infinitesimal variations of the fields that leave the field values at the initial and finite time unaffected, *i.e.*

$$\delta\phi(x,t_1) = \delta\phi(x,t_2) = 0\,.$$

On the other hand, since the rod is clamped, the displacement at the end points must be zero, *i.e.*

$$\delta\phi(0,t) = \delta\phi(\ell,t) = 0\,.$$

Under these circumstances we derive the Euler-Lagrange equations for our continuum system

$$\frac{\partial}{\partial t}\Big(\frac{\partial\mathscr{L}}{\partial(\partial_t\phi)}\Big) + \frac{\partial}{\partial x}\Big(\frac{\partial\mathscr{L}}{\partial(\partial_x\phi)}\Big) - \frac{\partial\mathscr{L}}{\partial\phi} = 0\,.$$

Let us now discuss the solution of the field equation

$$\ddot{\phi} - c^2\partial_{xx}\phi = 0\,, \qquad c = \sqrt{\frac{Y}{\mu}}\,,$$

where c is the propagation velocity of vibrations through the rod. This equation is linear and, for this reason, its solutions satisfy the superposition principle. Take an ansatz

$$\phi(x,t) = e^{ikx}a_k(t) + e^{-ikx}b_k(t)\,.$$

If we impose $\phi(0,t)=0$, then $b_k(t) = -a_k(t)$ and we can refine the ansatz as

$$\phi(x,t) = a_k(t)\sin kx\,.$$

Requiring that $\phi(\ell,t)=0$ we get $\sin k\ell = 0$, i.e. $k \equiv k_n = \frac{\pi n}{\ell}$. Coefficients $a_k(t)$ then obey

$$\ddot{a}_k + c^2k^2a_k(t) = 0 \qquad \to \qquad a_k(t) = e^{i\omega_k t}a_k\,,$$

where $\omega_k = \pm ck$ is the dispersion relation. Thus, the general solution is

$$\phi(x,t) = \sum_n \sin k_n x\Big(A_n\cos\omega_n t + B_n\sin\omega_n t\Big)\,, \qquad \omega_n = ck_n\,,$$

and the constants A_n, B_n are fixed by the initial conditions, which is an initial profile $\phi(x,0)$ and an initial velocity $\dot{\phi}(x,0)$.

The generalisation to continuous systems in more space dimensions is now straightforward. In two dimensions one can start with two-dimensional lattice of springs. The displacement of a particle at the site (i,j) is measured by the quantity $\vec{\phi}_{ij}$, which is a two-dimensional vector. In the limit when we go to a continuum, this becomes a displacement field $\vec{\phi}(x,y,t)$ of a membrane subjected to small vibrations in the (x,y)-plane. In three dimensions we get a vector $\vec{\phi}_{ijk}$. The continuous limit yields a three-dimensional displacement field $\vec{\phi}(x,y,z,t)$ of a continuous solid vibrating in the x,y,z directions with eoms of a partial differential equation type:

$$\ddot{\vec{\phi}} - c_1\partial_{xx}\vec{\phi} - c_2\partial_{yy}\vec{\phi} - c_3\partial_{zz}\vec{\phi} - c_4\partial_{xy}\vec{\phi} - c_5\partial_{yz}\vec{\phi} - c_6\partial_{xz}\vec{\phi} = 0\,,$$

the coefficients c_i encode the properties of the solid.

2.4 Rigid body

A rigid body is a system of point masses, constrained by holonomic relations expressed by the fact that the distance between points is constant.[9] Historically, the description of the motion of rigid bodies was amongst the first problems of analytic mechanics. Of special interest, since 18th century, remains the dynamics of spinning tops, where the cases of Euler, Lagrange and Kowalevski provided prominent examples of completely integrable systems.

Below we outline the general approach to describe the motion of a rigid body and further consider the solvable case of Euler's top.

2.4.1 Angular velocity

To describe a motion of a rigid body, we introduce two orthogonal coordinate systems: one which is stationary (immovable) and, therefore, inertial with coordinate axes XYZ, and another one (moving) which is rigidly fixed to a body and which participates in all its motions. We denote the axes of this moving coordinate system as $x_1 = x$, $x_2 = y$, $x_3 = z$ and fix its origin at some point O which should nor necessarily coincide with the center of mass.

[9] Holonomic constraints are constraints which are expressible as a function of the coordinates x_j and time t, *i.e.* they do not involve velocities and higher derivatives of coordinates.

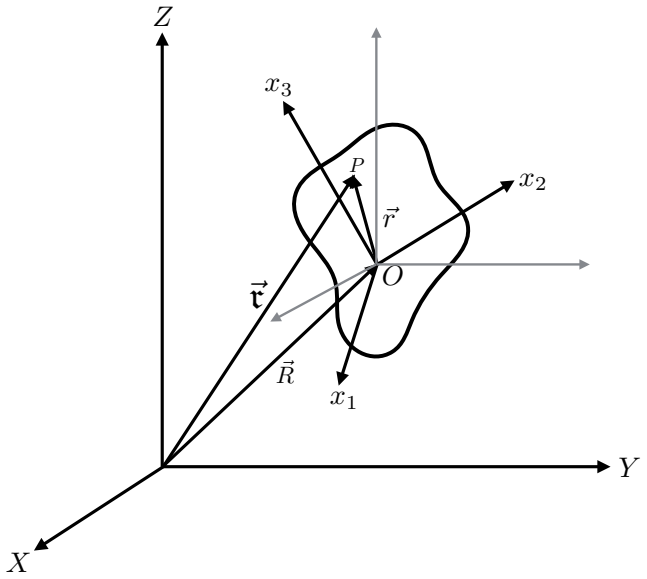

Figure 2.8: Relation between the radius-vectors of a point of the rigid body in the stationary and moving coordinate systems.

Consider an arbitrary infinitesimal displacement of a rigid body. An infinitesimal displacement $d\vec{\mathfrak{r}}$ of a point P of the body for a time dt will be given by the sum of the displacement $d\vec{R}$ of the center of the moving coordinate system and the shift $\delta\vec{r} = d\vec{\varphi} \times \vec{r}$ with is due to rotation of this system by an angle $d\vec{\varphi}$

$$d\vec{\mathfrak{r}} = d\vec{R} + d\vec{\varphi} \times \vec{r}. \tag{I.2.30}$$

Dividing this expression by dt, we obtain

$$\vec{v} = \vec{V} + \vec{\omega} \times \vec{r}. \tag{I.2.31}$$

Here $\vec{v} = \frac{d\vec{\mathfrak{r}}}{dt}$ is the velocity of P with respect to the stationary coordinate system, $\vec{V} = \frac{d\vec{R}}{dt}$ is the velocity of O with respect to this system and the vector

$$\vec{\omega} = \frac{d\vec{\varphi}}{dt}$$

is called *angular velocity* of the rigid body. The definition of the angular velocity does not depend on the choice of the moving coordinate system and is associated to the rigid body as the whole. In general, $\vec{\omega}$ depends on t.

It should be pointed out that formulae (I.2.30) and (I.2.31) are essentially based on the

Eulers's theorem. *The general displacement of a rigid body with one fixed point is a rotation about some axis.*

Equation (I.2.30) is an infinitesimal version of the global formula describing an arbitrary motion of the rigid body in time, namely,

$$\vec{\mathfrak{r}}(t) = \vec{R}(t) + B_t\vec{\mathrm{r}}, \tag{I.2.32}$$

where $\vec{\mathrm{r}}$ is the radius-vector of the point P measured with respect to the axes of the *moving* coordinate system and B_t is an orthogonal transformation that encodes rotation of the moving system in time with respect to the stationary system (the XYZ-system translated to O).[10] Would the translational

[10]We stress that $\vec{\mathrm{r}}$ is time-independent as the body point P does not move with respect to the moving system.

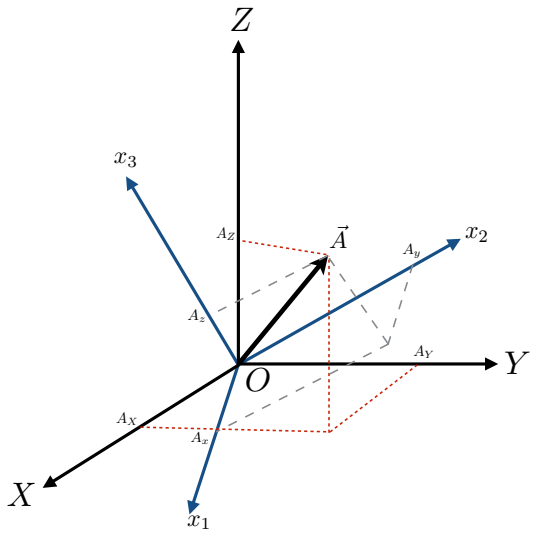

Figure 2.9: Coordinates of an arbitrary vector $\vec{A}$ in the stationary XYZ-frame and in the rotating xyz-frame. Coordinates do depend on the choice of the coordinate system. Two sets of coordinates are related by means of an orthogonal transformation B. This a *passive* point of view on rotations: the position vector $\vec{A}$ is a spectator fixed in space while the coordinate system transforms.

motion be absent and the origins of the stationary and moving coordinates systems coincide, then $\vec{\mathrm{r}} = \vec{r}$ and formula (I.2.32) would take the form

$$\vec{r}(t) = B_t \vec{\mathrm{r}}, \tag{I.2.33}$$

In fact, for any fixed time, $\vec{r}$ and $\vec{\mathrm{r}}$ represent coordinates of the one and the same geometric point P, but with respect to two different coordinate systems, the stationary and the rotating, respectively, see Fig. 3.6. The orthogonal matrix B_t connects these coordinates at any moment of time, mathematically realising the physical picture of rotation.

To make a connection between (I.2.31) and (I.2.32), we differentiate (I.2.32) in time to get

$$\vec{v} = \vec{V} + \dot{B}_t \vec{\mathrm{r}}$$

Regarding (I.2.33) as the relation between the coordinates of the one and the same vector with respect to the stationary and moving frames *sharing the same origin*, we use it to write $\vec{\mathrm{r}} = B_t^{-1} r$ and, substituting this expression for $\vec{\mathrm{r}}$ into the previous equation, we obtain[11]

$$\vec{v} = \vec{V} + \dot{B} B^{-1} r\,. \tag{I.2.34}$$

Here $\dot{B}B^{-1} = \dot{B}B^t$ appears a skew-symmetric 3×3-matrix, where the upper index t stands for transposition.[12] Indeed, differentiating over time the orthogonality condition $BB^t = \mathbb{1}$, we get

$$\dot{B}B^t + B\dot{B}^t = \dot{B}B^t + (\dot{B}B^t)^t = 0\,.$$

It is known that the result of the action of a skew-symmetric matrix 3×3 on an arbitrary vector r, can be realised as cross-product by some fixed vector $\vec{\omega}$, so that $\dot{B}B^{-1} r = \vec{\omega} \times \vec{r}$. In this way we have shown how (I.2.34) reproduces (I.2.31), as well as found the relation between $\vec{\omega}$ and B. This relation can be explicitly written as

$$\omega_i = -\frac{1}{2} \epsilon_{ijk} (\dot{B}B^t)_{jk}\,, \tag{I.2.35}$$

[11] We omit the subscript "t", as it is clear from the context that B is t-dependent.

[12] We note by passing that $\dot{B}B^{-1}$ is an element of the so(3) Lie algebra.

where ω_i are components of $\vec{\omega}$ with respect to the stationary coordinate system.

As we have already seen, any orthogonal matrix describing rotations depends on three parameters. For instance, in subsection 2.2.2 we have chosen a unit vector n_a and an angle φ as such a set of parameters, see (I.2.12) for an explicit form of the corresponding rotation matrix. Different choices of parameters are possible, but the one most convenient for describing the dynamics of the rigid body is proved to be in terms of *eulearin angles*. We postpone an introduction of these angles for later, dwelling here on the general idea that three generalised angles are needed to specify the position of the rotating coordinate system with respect to the stationary one. Together with three components of the position vector $R(t)$ describing the motion of the center O of the moving system, these generalised angles constitute 6 degrees of freedom of the rigid body.

What concerns rotations, the information about the position of the body in space relative to the axes of the stationary system is contained in the matrix B_t, which time evolution becomes the main problem of the rigid body dynamics. According to (I.2.35), components of the angular velocity are functions of generalised angles and their time derivatives; as such, they can be understood as generalised velocities corresponding to generalised angles. As is also clear from (I.2.35), the components do not depend on the choice of the moving coordinate system – neither on the position of its origin nor on the orientation of its axes, provided all these orientations are related to each other by orthogonal transformations. Indeed, under any such constant (time-independent) orthogonal transformation g of the moving system, the matrix B will transform as $B \to Bg$, which leaves the combination $\dot{B}B^t$ invariant.

We also point out the following fact important for later considerations. In fact, equation (I.2.33) is valid for any vector $\vec{A}$, not necessarily constant. In this case we have

$$\vec{A}(t) = B_t \vec{\mathrm{A}}(t)\,, \tag{I.2.36}$$

where $\vec{\mathrm{A}}$ and $\vec{A}$ are components of the one and the same vector but in the rotating and stationary frames, respectively. Taking time derivative, we get

$$\dot{\vec{A}} = B_t \dot{\vec{\mathrm{A}}} + \dot{B}_t \vec{\mathrm{A}} = B_t \dot{\vec{\mathrm{A}}} + \dot{B}_t B_t^{-1} \vec{A} = B_t \dot{\vec{\mathrm{A}}} + \omega \times \vec{A}\,. \tag{I.2.37}$$

It will be further convenient to define the angular momentum velocity $\vec{\Omega}$ measured with respect to the body axes and, therefore, related to ω as (I.2.36), that is $\vec{\omega} = B_t \vec{\Omega}$. Then (I.2.37) can be written as

$$\dot{\vec{A}} = B_t \dot{\vec{\mathrm{A}}} + B_t \vec{\Omega} \times B_t \vec{\mathrm{A}}\,. \tag{I.2.38}$$

It remains to recall the following fact from geometry. For any rotation matrix B, *i.e.* an orthogonal matrix with the unit determinant, the following equivariance property of the cross product holds

$$B(u \times v) = Bu \times Bv\,, \qquad \forall u, v \in \mathbb{R}^3\,. \tag{I.2.39}$$

This property allows one to rewrite (I.2.38) in the form

$$\dot{\vec{A}} = B_t \left[\dot{\vec{\mathrm{A}}} + \vec{\Omega} \times \vec{\mathrm{A}} \right]\,. \tag{I.2.40}$$

Let us stress that we deal with the one and the same vector $\vec{A}$. The expression in the brackets is evaluated with respect to the coordinate axes of the moving system, the vector is evaluated in the stationary frame. We will essentially use (I.2.40) when deriving Euler's equations. Finally, we give an expression for the angular velocity $\vec{\Omega}$ inside the body

$$\Omega_i = -\frac{1}{2} \epsilon_{ijk} (B^t \dot{B})_{jk}\,. \tag{I.2.41}$$

This formula follows from (I.2.35) together with the equivariance property of the Levi-Civita tensor ϵ_{ijk} under orthogonal transformations, *cf.* (I.2.39). Under $B \to Bg$, one has $\vec{\Omega} \to g^{-1}\vec{\Omega}$.

2.4.2 Lagrangian

To compute the kinetic energy of a rigid body, one can consider it as a discrete system of material particles and then sum (integrate in the case of continuum) their kinetic energies

$$T = \sum \frac{m}{2}\vec{v}^2 \,.$$

Using (I.2.31), we obtain

$$T = \sum \frac{m}{2}(\vec{V} + \vec{\omega} \times \vec{r})^2 = \frac{1}{2}\vec{V}^2 \sum m + \Big(\vec{V}, \vec{\omega} \times \sum m\vec{r}\Big) + \frac{1}{2}\sum m(\vec{\omega} \times \vec{r}, \vec{\omega} \times \vec{r}) \,.$$

From this formula it is clear that it is convenient to put the origin of the moving system in the center of mass, where $\sum m\vec{r} = 0$. Denoting the total mass as $M = \sum m$, we will then have

$$T = \frac{M}{2}\vec{V}^2 + \frac{1}{2}\sum m(\vec{\omega} \times \vec{r}, \vec{\omega} \times \vec{r}) \,.$$

Here the first term is the kinetic energy of the translational motion and it looks like as if the whole mass of the body would be concentrated in its center of mass. The second term is the kinetic energy of the rotational motion around an axis passing through the center of mass. The possibility to split the energy in these two parts is due to the choice of the origin of the body reference frame in its center of mass.

Let U be the potential energy of the rigid body in an external field. In general, the potential energy is a function of six variables defining the position of the rigid body: 3 coordinates $(X, Y, Z) = \vec{R}$ of the center of mass and three angles $\vec{\varphi}$, defining the orientation of the body axes with respect to the stationary coordinate system. The lagrangian of the body is then $L = T - U$, that is

$$L\big(\vec{R}, \vec{\varphi}, \vec{V}, \vec{\omega}\big) = \frac{M}{2}\vec{V}^2 + \frac{1}{2}\sum m(\vec{\omega} \times \vec{r}, \vec{\omega} \times \vec{r}) - U(\vec{R}, \vec{\varphi}) \,. \tag{I.2.42}$$

This lagrangian is a function of generalised coordinates $\vec{R}$ and $\vec{\varphi}$, and the corresponding generalised velocities $\vec{V}$ and $\vec{\omega}$. The first set of the Euler-Lagrange equations is

$$\frac{d}{dt}\frac{\partial L}{\partial \vec{V}} - \frac{\partial L}{\partial \vec{R}} = \frac{d}{dt}(M\vec{V}) + \frac{\partial U}{\partial \vec{R}} = \frac{d\vec{P}}{dt} + \frac{\partial U}{\partial \vec{R}} = 0 \,,$$

where we recall that $\vec{P} = M\vec{V}$ is the total momentum. Here

$$\vec{F} = -\frac{\partial U}{\partial \vec{R}}$$

is the force which is equal to the sum of all forces acting on each particle of the body. In fact, $\vec{F}$ is equal to the sum of external forces because all internal forces cancel. Thus, the first set of Euler-Lagrange equations is

$$\frac{d\vec{P}}{dt} = \vec{F} \,. \tag{I.2.43}$$

The second set of the Euler-Lagrange equations is

$$\frac{d}{dt}\frac{\partial L}{\partial \vec{\omega}} - \frac{\partial L}{\partial \vec{\varphi}} = \frac{d}{dt}\Big(\sum m\vec{r} \times (\vec{\omega} \times \vec{r})\Big) + \frac{\partial U}{\partial \vec{\varphi}} = 0 \,,$$

where to vary the lagrangian over $\vec{\omega}$, we made use of the formula

$$(\vec{\omega} \times \vec{r}, \vec{\omega} \times \vec{r}) = (\vec{r} \times (\vec{\omega} \times \vec{r}), \vec{\omega}) \,.$$

Further, we note that

$$\sum m\vec{r}\times(\vec{\omega}\times\vec{r}) = \sum \vec{r}\times m\vec{v} = \sum \vec{r}\times\vec{p} = \vec{L}\,, \tag{I.2.44}$$

where $\vec{L}$ is the total angular momentum computed with respect to the stationary coordinate system with the origin at the center of mass of the rigid body. With this observation, the second set of Euler-Lagrange equations take the form

$$\frac{d\vec{L}}{dt} = -\frac{\partial U}{\partial\vec{\varphi}} = \vec{N}\,, \tag{I.2.45}$$

where $\vec{N}$ is the torque. Equations (I.2.43) and (I.2.45) are equations of motion for the rigid body. In the absence of external forces these equations simply reduce to conservation laws of the total linear and angular momenta

$$\frac{d\vec{P}}{dt} = 0\,, \qquad \frac{d\vec{L}}{dt} = 0\,.$$

We stress that $\vec{P}$ and $\vec{L}$ here are computed with respect to a stationary, *i.e.* inertial frame, in particular, for $\vec{L}$ with the coordinate origin in the center of mass of the rigid body.

2.4.3 Inertia tensor

In the following we assume that translational motion is absent. Now we would like to change the coordinates in the kinetic energy to that of the moving system. We have

$$\begin{aligned} T &= \frac{1}{2}\sum m(\vec{\omega}\times\vec{r},\vec{\omega}\times\vec{r}) = \frac{1}{2}\sum m\left(B_t\vec{\Omega}\times B_t\vec{\mathrm{r}}, B_t\vec{\Omega}\times B_t\vec{\mathrm{r}}\right) \\ &= \frac{1}{2}\sum m\left(B_t(\vec{\Omega}\times\vec{\mathrm{r}}), B_t(\vec{\Omega}\times\vec{\mathrm{r}})\right) = \frac{1}{2}\sum m\big(\vec{\Omega}\times\vec{\mathrm{r}},\vec{\Omega}\times\vec{\mathrm{r}}\big)\,. \end{aligned}$$

Here in the first step we used property (I.2.39) of the cross product and in the second step, the invariance of the scalar product with respect to orthogonal transformations. Thus, the kinetic energy of the rotating body expressed via quantities of the moving system is

$$T = \frac{1}{2}\sum m\big(\vec{\Omega}\times\vec{\mathrm{r}},\vec{\Omega}\times\vec{\mathrm{r}}\big)\,,$$

where we recall that the position vector $\vec{\mathrm{r}}$ of mass m is time-independent.

Let x_i and Ω_i denote the components of $\vec{\mathrm{r}}$ and $\vec{\Omega}$, respectively. In terms of these components the kinetic energy is

$$T = \frac{1}{2}\Omega_i\Omega_j\sum m(\delta_{ij}\mathrm{x}^2 - \mathrm{x}_i\mathrm{x}_j)\,, \qquad \mathrm{x}^2 \equiv \vec{\mathrm{r}}^2\,. \tag{I.2.46}$$

Introducing a tensor

$$I_{ij} = \sum m(\delta_{ij}\mathrm{x}^2 - \mathrm{x}_i\mathrm{x}_j)\,, \tag{I.2.47}$$

we rewrite this kinetic energy in the form

$$T = \frac{1}{2}I_{ij}\Omega_i\Omega_j = \frac{1}{2}\vec{\Omega}^t I\vec{\Omega}\,,$$

where the last term is written in the matrix form. As is clear from its definition, tensor I_{ij} is symmetric $I_{ij} = I_{ji}$. This second rank tensor is called the *moment of inertia tensor* or simply the *inertia tensor*. In the continuum case the inertia tensor is given by

$$I_{ij} = \int \rho(\vec{\mathrm{r}})(\mathrm{x}^2\delta_{ij} - \mathrm{x}_i\mathrm{x}_j)dV\,,$$

where $\rho(\vec{\mathrm{r}})$ is the mass density.

In the matrix form the inertia tensor looks as

$$I = \begin{pmatrix} \sum m(\mathrm{x}_2^2 + \mathrm{x}_3^2) & -\sum m\mathrm{x}_1\mathrm{x}_2 & -\sum m\mathrm{x}_1\mathrm{x}_3 \\ -\sum m\mathrm{x}_1\mathrm{x}_2 & \sum m(\mathrm{x}_1^2 + \mathrm{x}_3^2) & -\sum m\mathrm{x}_2\mathrm{x}_3 \\ -\sum m\mathrm{x}_1\mathrm{x}_3 & -\sum m\mathrm{x}_2\mathrm{x}_3 & \sum m(\mathrm{x}_1^2 + \mathrm{x}_2^2) \end{pmatrix}.$$

As any real symmetric matrix, the inertia tensor can be brought to the diagonal form by the corresponding choice of directions of the axes x_1, x_2, x_3. These directions are then called *the principal inertia axes*. The diagonal components I_1, I_2, I_3 are known as *principal moments of inertia*. Via the principle moments of intertia the rotation energy is expressed as

$$T_{\mathrm{rot}} = \frac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2)\,.$$

With respect to the principle moments of inertia, rigid bodies are characterised as

- a body for which all three moments are different is called an *asymetric top*;
- a body for which $I_1 = I_2 \neq I_3$ is called *symmetric top*. In this case the choice of principal axes in the x_1x_2-plane is arbitrary;
- a body for which $I_1 = I_2 = I_3$ is called *spherical top*. A choice of any three principal axes is arbitrary: any three perpendicular axes do the job;
- a body for which $I_1 = I_2$ and $I_3 = 0$ is called *rotator*. This is the case of a system of particles lying on the same x_3-axis. Its peculiarity that it has not three but two rotational degrees of freedom corresponding to rotations around the axes x_1 and x_2.

Rotator is a rigid stick

Now we explain the relationship between the total angular momentum and the angular velocity. Choosing the center of a stationary coordinate system in the center of mass of the rigid body, for the angular momentum we have

$$\vec{L} = \sum m\vec{r} \times (\vec{\omega} \times \vec{r}) = B_t \sum m\vec{\mathrm{r}} \times (\vec{\Omega} \times \vec{\mathrm{r}}) = B_t\vec{M}\,,$$

where we have introduced the angular momentum $\vec{M}$ inside the body

$$\vec{M} = \sum m\vec{\mathrm{r}} \times (\vec{\Omega} \times \vec{\mathrm{r}})\,. \tag{I.2.48}$$

We further have

$$\Big[\vec{\mathrm{r}} \times (\vec{\Omega} \times \vec{\mathrm{r}})\Big]_i = \Omega_i\mathrm{x}^2 - \mathrm{x}_i\mathrm{x}_j\Omega_j = (\mathrm{x}^2\delta_{ij} - \mathrm{x}_i\mathrm{x}_j)\Omega_j\,.$$

Thus, the components of the angular momentum in the moving frame are related to the components of the angular velocity in this frame as

$$M_i = \sum m(\mathrm{x}^2\delta_{ij} - \mathrm{x}_i\mathrm{x}_j)\Omega_j = I_{ij}\Omega_j\,.$$

Regarding I as a 3×3 matrix, we can write the last relation in the form

$$\vec{M} = I\vec{\Omega}\,. \tag{I.2.49}$$

If the axes x_1, x_2, x_3 are directed along the principal inertia axes, then the above formula yields

$$M_1 = I_1\Omega_1\,, \qquad M_2 = I_2\Omega_2\,, \qquad M_3 = I_3\Omega_3\,. \tag{I.2.50}$$

2.4.4 Euler's top

Here we study the motion of a rigid body attached to one point in the absence of external forces.

Definition. *The Euler top is a rigid body without any particular symmetry that rotates in the absence of any external forces around a fixed point that coincides with its center of mass.*

For the Euler top $\vec{P} = 0$ and the Euler-Lagrange equation coincides with the conservation law of the total angular momentum

$$\frac{d\vec{L}}{dt} = 0\,. \tag{I.2.51}$$

This equation is written with respect to a stationary coordinate system, *i.e.* with respect to this system $\vec{L}$ is a constant vector. On the other hand, relation (I.2.50) between the angular momentum and the angular velocity arises in the moving coordinate system oriented along the principal inertia axis. To exploit this relation, we therefore need to transform equation (I.2.51) in the moving coordinate system. This is done with the help of equation (I.2.40), according to which

$$\frac{d\vec{L}}{dt} = B_t\left[\frac{d\vec{M}}{dt} + \Omega \times \vec{M}\right].$$

Here $\vec{L}$ and $\vec{M}$ are components of the angular momentum with respect to the stationary and moving coordinate systems, respectively. Thus, with respect to the rotating system equation of motion (I.2.51) takes the form

$$\frac{d\vec{M}}{dt} + \vec{\Omega} \times \vec{M} = 0\,. \tag{I.2.52}$$

Introducing the orts $\vec{e}_x$, $\vec{e}_y$, $\vec{e}_z$ of the rotating system with coordinates $x_1 = x$, $x_2 = y$, $x_3 = z$ directed along the principal inertia axes, we first compute

$$\vec{\Omega} \times \vec{M} = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ \Omega_1 & \Omega_2 & \Omega_3 \\ I_1\Omega_1 & I_2\Omega_2 & I_3\Omega_3 \end{vmatrix} = \vec{e}_x(I_3 - I_2)\Omega_2\Omega_3 + \vec{e}_y(I_1 - I_3)\Omega_1\Omega_3 + \vec{e}_z(I_2 - I_1)\Omega_2\Omega_3\,.$$

Thus, for the components of the angular velocity in the moving frame we find the following equations

$$\begin{aligned} I_1\frac{d\Omega_1}{dt} + (I_3 - I_2)\Omega_2\Omega_3 &= 0\,, \\ I_2\frac{d\Omega_2}{dt} + (I_1 - I_3)\Omega_3\Omega_1 &= 0\,, \\ I_3\frac{d\Omega_3}{dt} + (I_2 - I_1)\Omega_1\Omega_2 &= 0. \end{aligned} \tag{I.2.53}$$

These are the so-called *Euler's equations*.

Conservations laws. Euler's equations have two integrals of motion. The first one is the total energy E which just coincides with the kinetic energy

$$E = \frac{1}{2}\Omega_i I_{ij}\Omega_j = \frac{1}{2}I_1\Omega_1^2 + \frac{1}{2}I_2\Omega_2^2 + \frac{1}{2}I_3\Omega_3^2\,. \tag{I.2.54}$$

It is conserved due to Euler's equations. Indeed,

$$\frac{dE}{dt} = I_1\dot{\Omega}_1\Omega_1 + I_2\dot{\Omega}_2\Omega_2 + I_3\dot{\Omega}_3\Omega_3$$

$$= \quad (I_2 - I_3)\Omega_1\Omega_2\Omega_3 + (I_3 - I_1)\Omega_1\Omega_2\Omega_3 + (I_1 - I_2)\Omega_1\Omega_2\Omega_3 = 0\,.$$

The second integral is the length of the angular momentum. We have

$$L^2 \equiv \vec{L}^2 = (B_t\vec{M}, B_t\vec{M}) = \vec{M}^2 = M_1^2 + M_2^2 + M_3^2 = I_1^2\Omega_1^2 + I_2^2\Omega_2^2 + I_3^2\Omega_3^2\,. \tag{I.2.55}$$

Its conservation can also be proved by an explicit calculation

$$\begin{aligned}\frac{1}{2}\frac{dL^2}{dt} &= I_1^2\dot{\Omega}_1\Omega_1 + I_2^2\dot{\Omega}_2\Omega_2 + I_3^2\dot{\Omega}_3\Omega_3 \\ &= I_1(I_2 - I_3)\Omega_1\Omega_2\Omega_3 + I_2(I_3 - I_1)\Omega_1\Omega_2\Omega_3 + I_3(I_1 - I_2)\Omega_1\Omega_2\Omega_3 = 0\,.\end{aligned}$$

Note that nether E nor L^2 involve time derivatives of Ω's.

Thus, we have proven that Euler's equations have two quadratic integrals – the energy and L^2. As a consequence, $\vec{M}$ lies on the intersection of an ellipsoid and a sphere:

$$2E = \frac{M_1^2}{I_1} + \frac{M_2^2}{I_2} + \frac{M_3^2}{I_3}\,, \qquad L^2 = M_1^2 + M_2^2 + M_3^2\,. \tag{I.2.56}$$

One can further study the structure of the curves of intersection by fixing the ellipsoid $E > 0$ and changing the radius of the sphere.

Integrating Euler's equations. From the conservation laws (I.2.54) and (I.2.55) we can express two angular velocities, for instance, Ω_1 and Ω_3,

$$\begin{aligned}\Omega_1^2 &= \frac{1}{I_1(I_3 - I_1)}\Big((2EI_3 - L^2) - I_2(I_3 - I_2)\Omega_2^2\Big)\,, \\ \Omega_3^2 &= \frac{1}{I_3(I_3 - I_1)}\Big((L^2 - 2EI_1) - I_2(I_2 - I_1)\Omega_2^2\Big)\,.\end{aligned}$$

Then plugging these expressions into the Euler equation for Ω_2, we obtain

$$\frac{d\Omega_2}{dt} = \frac{1}{I_2\sqrt{I_1I_3}}\sqrt{\Big((2EI_3 - L^2) - I_2(I_3 - I_2)\Omega_2^2\Big)\Big((L^2 - 2EI_1) - I_2(I_2 - I_1)\Omega_2^2\Big)}\,.$$

For definiteness we assume that $I_3 > I_2 > I_1$ and also that $L^2 > 2EI_2$. Then making the substitutions

$$\tau = t\sqrt{\frac{(I_3 - I_2)(L^2 - 2EI_1)}{I_1I_2I_3}}\,, \qquad s = \Omega_2\sqrt{\frac{I_2(I_3 - I_2)}{2EI_3 - L^2}}$$

and introducing the positive parameter $k^2 < 1$ by[13]

$$k^2 = \frac{(I_2 - I_1)(2EI_3 - L^2)}{(I_3 - I_2)(L^2 - 2EI_1)}\,,$$

we obtain

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

The initial time $\tau = 0$ is chosen such that for $s = 0$ one has $\Omega_2 = 0$. Inverting the last integral, one gets the Jacobi elliptic function

$$s = \mathrm{sn}\,\tau\,.$$

Using two other elliptic functions

$$\mathrm{cn}^2\,\tau + \mathrm{sn}^2\,\tau = 1\,, \qquad \mathrm{dn}^2\,\tau + k^2\,\mathrm{sn}^2\,\tau = 1\,,$$

[13]For a solution to exist the values of L^2 must be bounded: $2EI_1 < L^2 < 2EI_3$.

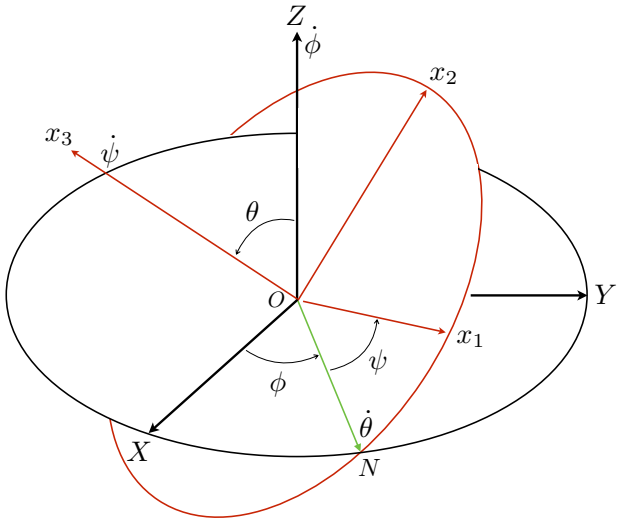

Figure 2.10: The eulerian angles. Here XYZ and $x_1x_2x_3$ are the stationary and moving coordinate systems, respectively. The line ON is called the line of nodes.

we obtain the solution

$$\Omega_1 = \sqrt{\frac{2EI_3 - L^2}{I_1(I_3 - I_1)}}\,\mathrm{cn}\,\tau\,, \qquad \Omega_2 = \sqrt{\frac{2EI_3 - L^2}{I_2(I_3 - I_1)}}\,\mathrm{sn}\,\tau\,, \qquad \Omega_3 = \sqrt{\frac{L^2 - 2EI_1}{I_3(I_3 - I_1)}}\,\mathrm{dn}\,\tau\,.$$

The real period of all these three elliptic functions is given by 4K, where K is the complete elliptic integral of the first kind:

$$\mathrm{K} = \int_0^1 \frac{\mathrm{d}s}{\sqrt{(1-s^2)(1-k^2s^2)}}\,.$$

The period T in time t is, therefore,

$$T = 4\mathrm{K}\sqrt{\frac{I_1I_2I_3}{(I_3 - I_2)(L^2 - 2EI_1)}}\,.$$

After this time both $\vec\Omega$ and $\vec M$ will return to their original values. Thus, $\vec\Omega$ and $\vec M$ perform a strictly periodic motion. What is remarkable, is that the top itself *does not return* to its original position with respect to the stationary coordinate system, as we now explain.

Eulerian angles. Our primary interest is to understand how the rigid body rotates in space rather than to know the time evolution of the angular velocity or the angular momentum in the moving frame. This information is encoded in the matrix B_t that is parametrised by three parameters, which we choose to be the eulerian angles.

Let XYZ and $x_1x_2x_3$ be the the stationary and moving coordinate frames, respectively, see Fig. 2.10. Denote by ON the line of nodes, that is the line of intersection of the planes XOY and x_1Ox_2. The angles ϕ, θ, ψ in Fig. 2.10 are called *eulerian angles*. The angle θ takes values from 0 to π, and ϕ and ψ from 0 to 2π.

Eulerian angles parametrise the following orthogonal matrix B

$$B = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & \sin\theta\sin\phi \\ \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & -\sin\theta\cos\phi \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{pmatrix} . \tag{I.2.57}$$

Matrix B relates the coordinates $\vec{\mathrm{A}} = (x, y, z)^t$ of any vector in the moving frame with axes $x_1x_2x_3$ with the coordinates $\vec{A} = (X, Y, Z)^t$ of the same vector in the stationary frame with axes XYZ[14]

$$\vec{A} = B\vec{\mathrm{A}} .$$

With an explicit parametrisation of B_t in term of eulerian angles, we can now apply (I.2.35) to read off the components of the angular velocity $\vec{\omega}$ in the stationary frame

$$\begin{aligned} \omega_1 &= \dot{\psi}\sin\theta\sin\phi + \dot{\theta}\cos\phi , \\ \omega_2 &= -\dot{\psi}\sin\theta\cos\phi + \dot{\theta}\sin\phi , \\ \omega_3 &= \dot{\psi}\cos\theta + \dot{\phi} . \end{aligned} \tag{I.2.58}$$

Analogously, applying (I.2.41) we find the components of the angular velocity $\vec{\Omega}$ inside the body

$$\begin{aligned} \Omega_1 &= \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi , \\ \Omega_2 &= \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi , \\ \Omega_3 &= \dot{\phi}\cos\theta + \dot{\psi} . \end{aligned} \tag{I.2.59}$$

As a remark, both (I.2.58) and (I.2.59) can also be found geometrically, just by inspecting Fig. 2.10. For instance, let us show this for (I.2.59). The angular velocity of the system is compounded of angular velocities $\dot{\theta}$ about ON, $\dot{\phi}$ about OZ and $\dot{\psi}$ about Ox_3. First we find the projections of these velocities on the coordinate axes of the moving frame. For the projections of $\dot{\theta}$ we have

$$\dot{\theta}_1 = \dot{\theta}\cos\psi , \quad \dot{\theta}_2 = -\dot{\theta}\sin\psi , \quad \dot{\theta}_3 = 0 .$$

The velocity $\dot{\phi}$ is directed along the axis Z of the stationary coordinate system. Its projections on the axes of the moving frame are

$$\dot{\phi}_1 = \dot{\phi}\sin\theta\sin\psi , \quad \dot{\phi}_2 = \dot{\phi}\sin\theta\cos\psi , \quad \dot{\phi}_3 = \dot{\phi}\cos\theta .$$

The velocity $\dot{\psi}$ is directed along x_3. Collecting components along each axis, we obtain (I.2.59).

Lagrangian. Substituting (I.2.59) into the expression for the kinetic energy $T = \frac{1}{2}I_i\Omega_i^2$, we obtain T in terms of the eulerian angles and their derivatives

$$T = \frac{I_1}{2}(\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi)^2 + \frac{I_2}{2}(\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)^2 + \frac{I_3}{2}(\dot{\phi}\cos\theta + \dot{\psi})^2 . \tag{I.2.60}$$

Note that this expression simplifies for a symmetric top $I_1 = I_2$

$$T_{\mathrm{sym}} = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\phi}\cos\theta + \dot{\psi})^2 . \tag{I.2.61}$$

The form (I.2.60) fits the general expression for the kinetic energy in generalised coordinates

$$T = \frac{1}{2}\sum_{i,j=1}^{3} g_{ij}(\varphi)\dot{\varphi}_i\dot{\varphi}_j , \qquad \vec{\varphi} = (\phi, \theta, \psi) , \tag{I.2.62}$$

[14]The proof of this fact can be found, for instance, in Goldstein, Poole and Safko, *Classical Mechanics*, Addison Wesley, 2002.

where $g_{ij}(\varphi)$ is the corresponding metric. Regarding (I.2.60) as the lagrangian $L = T$, we can derive the Euler-Lagrange equations for $\vec{\varphi}$,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = 0\,, \qquad \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = 0\,, \qquad \frac{d}{dt}\frac{\partial L}{\partial \dot{\psi}} - \frac{\partial L}{\partial \psi} = 0\,. \tag{I.2.63}$$

The reader can verify that these 2nd order differential equations coincide with the 1st order Euler's equations (I.2.53) upon we substitute in them the expressions (I.2.59) for the components of the angular velocity. Note also that ϕ is a cyclic coordinate, so that the corresponding momentum

$$p_\phi = M_1 \sin\theta \sin\psi + M_2 \sin\theta \cos\psi + M_3 \cos\theta = L_Z$$

is conserved. It is now clear how integration of the very complicated Euler-Lagrange equations (I.2.63) proceeds. First, by a clever choice of variables Ω_i as in (I.2.59) one reduces the 2nd order equations to the 1st order Euler's equations, which can be further integrated by quadrature due to the existence of two integrals of motion: E and L^2. After Eulers's equations for Ω_i are solved, one needs to come back to (I.2.59) and solve these equations for (ϕ, θ, ψ).

Motion of Euler's top in space. There is a simpler way to solve for for (ϕ, θ, ψ) than to try to directly integrate (I.2.59). By using eulerian angles we can relate the angular momenta in the moving and the stationary coordinate systems. Using yet another arbitrariness in the choice of the orientation of the stationary coordinate system XYZ, we choose it such that the angular momentum $\vec{L}$ will directed along the Z axis. For the momentum in the moving coordinate system we then get

$$\vec{M} = B^{-1}\vec{L} = B^t \begin{pmatrix} 0 \\ 0 \\ L \end{pmatrix}\,, \tag{I.2.64}$$

where B is given by (I.2.57). Explicitly,

$$\begin{aligned} L\sin\theta\sin\psi &= I_1\Omega_1\,, \\ L\sin\theta\cos\psi &= I_2\Omega_2\,, \\ L\cos\theta &= I_3\Omega_3\,, \end{aligned}$$

where $L = |\vec{L}|$ is the length of $\vec{L}$. From here

$$\cos\theta = \frac{I_3\Omega_3}{L}\,, \qquad \tan\psi = \frac{I_1\Omega_1}{I_2\Omega_2}\,.$$

Substituting here the solution for Ω_i allows one to find

$$\begin{aligned} \cos\theta &= \sqrt{\frac{I_3(L^2 - 2EI_1)}{L^2(I_3 - I_1)}}\,\mathrm{dn}\,\tau\,, \\ \tan\psi &= \sqrt{\frac{I_1(I_3 - I_2)}{I_2(I_3 - I_1)}}\,\frac{\mathrm{cn}\,\tau}{\mathrm{sn}\,\tau}\,. \end{aligned}$$

Thus, both angles θ and ψ are periodic functions of time with the period T (the same period as for $\vec{\Omega}$). However, the angle ϕ *does not appear* in the formulas relating the angular momenta in the moving and the stationary coordinate systems. We can find it from the first two equations of (I.2.59), namely,

$$\dot{\phi} = \frac{\Omega_1\sin\psi + \Omega_2\cos\psi}{\sin\theta}\,.$$

This yields the following differential equation

$$\frac{d\phi}{d\tau} = L\frac{I_1\Omega_1^2 + I_2\Omega_2^2}{I_1^2\Omega_1^2 + I_2^2\Omega_2^2}\,.$$

The solution of this equation is found by integration but the integrand appears a rather involved combination of elliptic functions. Indeed, substituting the solution for Ω's in the last formula, we find

$$\frac{d\phi}{d\tau} = \frac{L}{I_1 \operatorname{cn}^2\tau + I_2 \operatorname{sn}^2\tau} = \frac{L}{I_1}\frac{1}{1-\alpha^2\operatorname{sn}^2\tau}\,,$$

where we introduced $\alpha^2 \equiv 1 - I_2/I_1 < 0$, since by our assumptions $I_2 > I_1$. Thus,

$$\phi(\tau) = \phi_0 + \frac{L}{I_1}\int_0^\tau \frac{\mathrm{d}\tau'}{1-\alpha^2\operatorname{sn}^2\tau'} = \phi_0 + \frac{L}{I_1}\Pi\Big(\operatorname{am}(\tau,k),\alpha^2,k\Big)\,,$$

where $\Pi(\varphi,\alpha^2,k)$ is the *incomplete elliptic integral of the third kind*. Here $\varphi = \operatorname{am}(\tau,k)$ is the Jacobi amplitude. By using this explicit expression, one can show that the period of ϕ, which is denoted by T', is not commensurable with T. Consequently, this implies that the top never returns to its original state.

Matrix form of Euler's equations and generalisations. It represents a theoretical interest to evaluate the lagrangian $L = T$ for a matrix B without specifying its parametrisation in terms of eulerian angles. We thus start from $T = 1/2 I_i \Omega_i^2$ and substitute here Ω_i given by (I.2.41). Further calculation proceeds with the use of the formula (III.7.3). Introducing a skew-symmetric matrix $S = B^{-1}\dot{B}$, we find that the kinetic energy is

$$L = \frac{1}{4}\Big[2\mathrm{Tr}(IS^2) - \mathrm{Tr}I\,\mathrm{Tr}S^2\Big] = \frac{1}{2}\mathrm{Tr}(\mathcal{I}S^2)\,, \tag{I.2.65}$$

Here I is the inertia tensor regarded as a 3×3 symmetric matrix which we consider to be diagonal and further introduce $\mathcal{I} = I - 1/2\mathrm{Tr}I\;\mathbb{1}$. Let us derive equations of motion that follow from this Lagrangian. First we note that

$$\delta L = \mathrm{Tr}(\delta S \Lambda)\,,$$

where we have introduces a skew-symmetric matrix

$$\Lambda = \frac{1}{2}\big(S\mathcal{I} + \mathcal{I}S\big)\,. \tag{I.2.66}$$

Next, we have

$$\delta S = \delta(B^{-1}\dot{B}) = -B^{-1}\delta B S + B^{-1}\delta\dot{B}\,.$$

Plugging this into the variation of the lagrangian and integrating by parts, we will get

$$\delta L = -\mathrm{Tr}\left[B^{-1}\delta B\big(\dot{\Lambda} + [S,\Lambda]\big)\right]\,,$$

from where we deduce the equations of motion

$$\dot{\Lambda} + [S,\Lambda] = 0\,, \tag{I.2.67}$$

where Λ is given by (I.2.66).

Further we note that this Lagrangian is invariant under $B \to hB$, where h is a constant orthogonal matrix. According to Noether's theorem there exists the conserved Noether charge J corresponding to this symmetry. It is given by

$$J = B\Lambda B^{-1}\,. \tag{I.2.68}$$

Computing the time derivative of J one can verify that it vanishes due to equations of motion (I.2.67). This charge is nothing else but the conserved angular momentum. Although the lagrangian is invariant under $B \to hB$, the Noether charge is not, it transforms as $J \to hJh^{-1}$. Obviously, S

can be identified with the matrix form of $\vec{\Omega}$, Λ is the angular momentum in the moving coordinate system,[15] J is the angular momentum in the stationary frame and (I.2.67) are Euler's equations. These equations admit straightforward generalisation to an arbitrary $\mathfrak{so}(n)$ Lie algebra where they are known as *Euler-Arnold* equations.

2.4.5 Lagrange's top[16]

When a rigid body fixed at a stationary point O is in a gravitational field of the Earth, its weight has to be taken into account. The problem of the motion of such a "heavy" rigid body has not yet been solved in the general case and in some sense is unsolvable. One special solvable case, found by Lagrange, is when two inertia moments are equal, for example $I_1 = I_2$ and the center of mass is located at a position $(x_1 = 0, x_2 = 0, x_3 = l)$ with respect to the rotating frame. This situation is achieved when the top has an axis of symmetry (around the third axis) and is attached to a point on this axis.

Definition. *Lagrange's top is a heavy axially symmetric rigid body fixed at a stationary point on its symmetry axis in a uniform gravitational field.*

We assume that the stationary point O coincides with the apex of the top, as in Fig. 2.11. Let both the stationary coordinate system XYZ and the rotating system $x_1x_2x_3$ have the origin at O and let the axis Z of the system XYZ be directed vertically. The fixed point O does not coincide with the center of mass, the latter is lying at the distance l from the origin along the x_3-axis. Denote by I_{ij} the inertia tensor of the body with respect to the center of mass. The axes x_1, x_2, x_3 coincide with the principal inertia axes and, therefore, I_{ij} is diagonal, $I_{ij} = I_i\delta_{ij}$. However, to describe the kinetic energy of the body fixed at O, we need to use the inertia tensor with respect to O, rather then with respect to the center of mass. In the rotating system the center of mass is away from the origin by the vector $\vec{a} = l\vec{e}_z$, where $\vec{e}_z$ is the ort of x_3. The inertia tensor I'_{ij} with respect to O is then obtained from I_{ij} by using the formula (proved in **Tutorial VIII**)

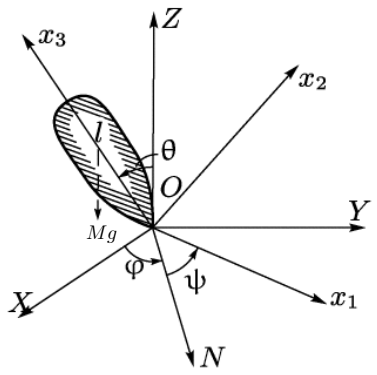

Figure 2.11: Lagrange's top spinning with its apex being a fixed point.

$$I'_{ij} = I_{ij} + M(a^2\delta_{ij} - a_ia_j)\,,$$

where M is the mass of the top. From here we see that the principal moments of inertia with respect to O are shifted according to

$$I'_1 = I_1 + Ml^2\,, \qquad I'_2 = I_2 + lM^2\,, \qquad I'_3 = I_3\,,$$

so that $I'_1 = I'_2$, because for Lagrange's top $I_1 = I_2$. Using the eulerian angles, we write the lagrangian $L = T - U$ by taking into account the expression (I.2.61) for the kinetic energy of the symmetric top where we replace I_1 by I'_1, namely,

$$L = \frac{I'_1}{2}(\dot\theta^2 + \dot\phi^2\sin^2\theta) + \frac{I_3}{2}(\dot\phi\cos\theta + \dot\psi)^2 - Mgl\cos\theta\,. \tag{I.2.69}$$

[15] The exact relations are $\Omega_i = -\frac{1}{2}\epsilon_{ijk}S_{jk}$ and $M_i = \epsilon_{ijk}\Lambda_{jk}$, where M_i is given by (I.2.50). The inverse relations are $S_{ij} = -\epsilon_{ijk}\Omega_k$ and $\Lambda_{ij} = \frac{1}{2}\epsilon_{ijk}M_k$.

[16] This subsection is optional.

We observe that in spite of the presence of the gravitational potential, the coordinates ϕ and ψ remain cyclic and, therefore, their conjugate momenta are conserved. These momenta are

$$\begin{aligned} p_\psi &= \frac{\partial L}{\partial \dot\psi} = I_3(\dot\phi\cos\theta + \dot\psi)\,, \\ p_\phi &= \frac{\partial L}{\partial \dot\phi} = (I_1'\sin^2\theta + I_3\cos^2\theta)\dot\phi + I_3\dot\psi\cos\theta\,. \end{aligned}$$

One can see that p_ψ and p_ψ coincide with the components of the angular momentum L_3 (in the rotation frame) and L_Z (in the stationary frame), respectively. Indeed, from (I.2.59) we get the components of the angular momentum $\vec{L}$ in the moving frame

$$\begin{aligned} L_1 &= I_1'\omega_1 = I_1'(\dot\phi\sin\theta\sin\psi + \dot\theta\cos\psi)\,, \\ L_2 &= I_1'\omega_2 = I_1'(\dot\phi\sin\theta\cos\psi - \dot\theta\sin\psi)\,, \\ L_3 &= I_3\omega_3 = I_3(\dot\phi\cos\theta + \dot\psi)\,. \end{aligned} \tag{I.2.70}$$

Therefore,

$$p_\psi = L_3 \tag{I.2.71}$$

By using (I.2.57), we find the component L_Z of the angular momentum in the stationary frame through its components (L_1, L_2, L_3) in the rotating frame

$$L_Z = L_1\sin\theta\sin\psi + L_2\sin\theta\cos\psi + L_3\cos\theta\,.$$

Upon substituting here (I.2.70), we find that

$$p_\phi = L_Z\,. \tag{I.2.72}$$

Of course, conservation of L_Z and L_3 also follows from Noether's theorem. Lagrangian (I.2.69) is invariant under infinitesimal constant rotations $\delta\phi = \epsilon$ and $\delta\psi = \epsilon$, so that the corresponding components of the angular momentum, L_z and L_3 must be conserved.

In addition to L_3 and L_Z, there is on more conserved quantity is the energy

$$E = \frac{I_1'}{2}(\dot\theta^2 + \dot\phi^2\sin^2\theta) + \frac{I_3}{2}(\dot\phi\cos\theta + \dot\psi)^2 + Mgl\cos\theta\,. \tag{I.2.73}$$

Thus, Lagrange's top has three integrals of motion (L_Z, L_3, E) and nor we use them to solve the equations of motion.

From (I.2.71) and (I.2.72) we the find

$$\begin{aligned} \dot\phi &= \frac{L_Z - L_3\cos\theta}{I_1'\sin^2\theta}\,, \\ \dot\psi &= \frac{L_3}{I_3} - \cos\theta\frac{L_Z - L_3\cos\theta}{I_1'\sin^2\theta}\,. \end{aligned} \tag{I.2.74}$$

Substituting these expressions in the formula for the energy, we obtain the one-dimensional problem

$$E' = \frac{I_1'}{2}\dot\theta^2 + U_{\text{eff}}(\theta)\,, \tag{I.2.75}$$

where we have introduced

$$E' = E - \frac{L_3^2}{2I_3}\,, \qquad U_{\text{eff}} = \frac{(L_Z - L_3\cos\theta)^2}{2I_1'\sin^2\theta} + Mgl\cos\theta\,.$$

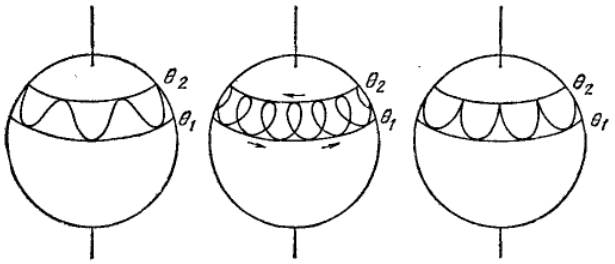

Figure 2.12: The possible shapes for the locus of the top axis on the unit sphere with center at the apex of the top. The axis monotonically *precesses* around the vertical and simultaneously *nutates* up and down.

Expressing from (I.2.75) the variable $\dot\theta$ and separating the variables, we find

$$t = \int \frac{d\theta}{\sqrt{\frac{2}{I_1'}(E' - U_{\rm eff}(\theta))}}\,. \tag{I.2.76}$$

It is not so difficult to understand quantitatively some general features of the motion of Lagrange's top. To this end, we rewrite (I.2.75) as

$$\frac{I_1'}{2}\dot\theta^2 = E' - \frac{(L_Z - L_3\cos\theta)^2}{2I_1'\sin^2\theta} - Mgl\cos\theta\,.$$

Introducing a new variable $x = \cos\theta$, the last equation can be written in the form

$$(I_1')^2\dot x^2 = 2I_1'E'(1-x^2) - (L_Z - L_3x)^2 - 2I_1'Mgl(x - x^3)\,. \tag{I.2.77}$$

On the right hand side we have a cubic polynomial, which is negative for both $x = 1$ and $x = -1$ and it tends to $+\infty$ for $x \to +\infty$. For some real values of x between -1 and 1 the cubic polynomial must be positive because the left hand side of equation (I.2.77) is positive. Thus, the polynomial should have two real roots in between -1 and 1, the third root is also real and bigger than 1. If we order two real roots $x_1 = \cos\theta_1$ and $x_2 = \cos\theta_2$ between 1 and -1 as $x_1 > x_2$, then they define two turning points $\theta_1 < \theta_2$, so that θ oscillates between θ_1 and θ_2. Under these oscillations the sign of $\dot\phi$ remains constant or changes, depending on the behaviour of the sign on the difference $L_Z - L_3\cos\theta$. In the first case the axis of the top *precesses* around the vertical monotonically, simultaneously performing *nutation* up and down, see Fig. 2.12, first picture. In the second case the direction of *precession* is opposite at two turning points, so that the top axis moves by making loops, see the second picture of Fig. 2.12. Finally, if one of θ_1, θ_2 coincides with zero of $L_Z - L_3\cos\theta$, then on the corresponding limiting circle $\dot\varphi$ and $\dot\theta$ turn to zero simultaneously, so that the axis describes the trajectory as in the third picture of Fig. 2.12.

The integral (I.2.76) can be computed in terms of the Weierstrass elliptic function $\wp$. Further, making a substitution

$$x(t) = \frac{2I_1'}{Mgl}z(t) + \frac{2E'I_1' + L_3^2}{6I_1'Mgl}\,,$$

we bring (I.2.77) to the form

$$\dot z^2 = 4z^3 - g_2z - g_3\,, \tag{I.2.78}$$

where

$$g_2 \;=\; \frac{1}{12I_1'}\Big((2E'I_1 + L_3^2)^2 - 12I_1'L_3L_ZMgl + 12(I_1'Mgl)^2\Big)$$

$$g_3 \;=\; \frac{1}{216I_1'}\Big((2E'I_1+L_3^2)^3-18I_1'L_3L_Z(2E'I_1+L_3^2)Mlg+18(I_1'Mgl)^2(L_3^2+3L_Z^2-4E'I_1')\Big)\,.$$

Equation (I.2.78) is nothing else but the differential equation satisfied by the Weierstrass elliptic function

$$\dot{\wp}^2 = 4\wp^3 - g_2\wp - g_3\,, \tag{I.2.79}$$

so that the general solution is $z=\wp(t+\delta)$, where δ is an integration constant. Thus,

$$\cos\theta(t) = \frac{2I_1'}{Mgl}\wp(t+\delta)+\frac{2E'I_1'+L_3^2}{6I_1'Mgl}\,.$$

2.5 Motion in noninertial system

Here we would like to address a general question how the equations of motion look like in a noninertial coordinate system. As a starting point we can again employ the principle of the least action which applicability is not restricted by any choice of the coordinate system. This principle implies the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \vec{v}} - \frac{\partial L}{\partial \vec{v}} = 0.$$

However, now the lagrangian will be different from the lagrangian in an inertial system k where the latter for a single particle has the form

$$L_{\rm inert} = \frac{m\vec{\mathfrak{v}}^2}{2} - U\,, \qquad \vec{\mathfrak{v}} = \frac{d\vec{\mathfrak{r}}}{dt}\,, \tag{I.2.80}$$

where we decided to denote the quantities related to the inertial frame k by gothic letters. We transform $L_{\rm inert}$ into a noninertial frame in two steps.

First we introduce a noninertial frame K' which performs with respect to k a translational motion with velocity $\vec{V}(t)$. Then the velocity of a particle in k and K' are related as

$$\vec{\mathfrak{v}} = \vec{v}' + \vec{V}(t)\,.$$

Substituting this into $L_{\rm inert}$, we get

$$L = \frac{m\vec{v}'^2}{2} + m(\vec{v}',\vec{V}) + \frac{m}{2}\vec{V}^2 - U\,.$$

Here $\vec{V}^2(t)$ is some given function of time, it can be represented as a total time derivative of some other function and, therefore, the third term in the expression above can be thrown away. Next, we have

$$m\vec{v}'\cdot\vec{V} = m\Big(\frac{d\vec{r}'}{dt},\vec{V}\Big) = -m\Big(\vec{r}',\frac{d\vec{V}}{dt}\Big) + \frac{d}{dt}(m\vec{r}',\vec{V})\,.$$

Omitting here the last term which is also the total time derivative, we get the following lagrangian

$$L = \frac{m\vec{v}'^2}{2} - m(\vec{W},\vec{r}') - U\,, \tag{I.2.81}$$

where $\vec{W}=\frac{d\vec{V}}{dt}$ is an acceleration of the translational motion of K'. From this lagrangian we derive the following Euler-Lagrange equation

$$m\frac{d\vec{v}'}{dt} = -\frac{\partial U}{\partial \vec{r}'} - m\vec{W}(t) = \vec{F} - m\vec{W}(t)\,. \tag{I.2.82}$$

Thus, the influence of translational motion of the coordinate system leads to an appearance of an additional homogenous force field $-m\vec{W}$, where $\vec{W}$ is the acceleration of the origin of the coordinate system K'. This force is called the *inertial force of translations* and it is directed opposite to $\vec{W}$.

Example. At the moment of takeoff, a rocket has acceleration $\ddot{r}$ directed upward. Thus, the coordinate system K' connected to the rocket is not inertial, and an observer inside can detect the existence of a force field mW and measure the inertial force, for example, by means of weighted springs. In this case the inertial force is called *overload*.

Example. When jumping from a loft, a person has acceleration g, directed downwards. Thus, the sum of the inertial force and the force of gravity is equal to zero: weighted springs show that the weight of any object is equal to zero, so such a state is called *weightlessness*. In exactly the same way, weightlessness is observed in the free ballistic flight of a satellite since the force of inertia is opposite to the gravitational force of the Earth.

As the second step, we introduce one more coordinate system K, which has the common origin with K' but rotates with respect to K with angular velocity $\vec{\omega}(t)$. The velocity $\vec{v}'$ of the particle with respect to K is the sum of its velocity $\vec{v}$ with respect to K' and the transferred velocity $\vec{\omega} \times \vec{r}$ of its rotation with the system K

$$\vec{v}' = \vec{v} + \vec{\omega} \times \vec{r}\,,$$

where we have taken into account that $\vec{r}'$ and $\vec{r}$ in the systems K' and K coincide. Substituting this into (I.2.81), we get

$$L = \frac{m\vec{v}^2}{2} + m(\vec{v}, \vec{\omega} \times \vec{r}) + \frac{m}{2}(\vec{\omega} \times \vec{r}, \vec{\omega} \times \vec{r}) - m(\vec{W}, \vec{r}) - U\,. \tag{I.2.83}$$

This is the general form of the lagrangian of a particle in an arbitrary noninertial frame. We then derive the Euler-Lagrange equations

$$\frac{d}{dt}\left(m\vec{v} + m\vec{\omega} \times \vec{r}\right) - \left(m\vec{v} \times \vec{\omega} - m\vec{\omega} \times (\vec{\omega} \times \vec{r}) - m\vec{W} - \frac{\partial U}{\partial r}\right) = 0\,.$$

These equations are then brought to the form of Newton's equations

$$m\frac{d\vec{v}}{dt} = \vec{F} - m\vec{W} - m\dot{\vec{\omega}} \times \vec{r} - 2m\vec{\omega} \times \vec{v} - m\vec{\omega} \times (\vec{\omega} \times \vec{r})\,. \tag{I.2.84}$$

Thus, the effect of rotation brings three new force fields which have the following names

1) the inertial force of rotation: $-m\dot{\vec{\omega}} \times \vec{r}$;

2) the Coriolis force: $-2m\vec{\omega} \times \vec{v}$;

3) the centrifugal force: $-m\vec{\omega} \times (\vec{\omega} \times \vec{r})$.

The last two forces are present for uniform rotation. The Coriolis force is different from other forces considered so far that it depends on particle velocity. The centrifugal force It lies in the plane passing through $\vec{\omega}$ and $\vec{r}$ and is perpendicular to the rotation axis (*i.e.* the direction of $\vec{\omega}$) and is directed away from the axis. Quantitatively, this force is equal to $m\rho\omega^2$, where ρ is the distance from the particle to the rotation axis.

Consider separately the case of a uniformly rotating system which does not have translational acceleration. Putting in (I.2.85) $\vec{\omega} = \text{const}$ and $\vec{W} = 0$, one gets the lagrangian

$$L = \frac{m\vec{v}^2}{2} + m(\vec{v}, \vec{\omega} \times \vec{r}) + \frac{m}{2}(\vec{\omega} \times \vec{r}, \vec{\omega} \times \vec{r}) - U \tag{I.2.85}$$

and the Euler-Lagrange equations

$$m\frac{d\vec{v}}{dt} = \vec{F} - 2m\vec{\omega}\times\vec{v} - m\vec{\omega}\times(\vec{\omega}\times\vec{r})\,. \tag{I.2.86}$$

Compute the energy of the particle in this case. We have the canonical momentum

$$\vec{p} = \frac{\partial L}{\partial \vec{v}} = m\vec{v} + m\vec{\omega}\times\vec{r}\,.$$

Substituting it in the expression for the energy $E = \vec{p}\dot{\vec{v}} - L$, we find

$$E = \frac{m\vec{v}^2}{2} - \frac{m}{2}(\vec{\omega}\times\vec{r})^2 + U\,.$$

The energy does not contain the term linear in velocity. The effect of rotation boils down to adding a term depending on coordinates of a particle and proportional to the square of the angular velocity. This additional energy $-\frac{m}{2}(\vec{\omega}\times\vec{r})^2$ is called *centrifugal*.

2.6 Hamiltonian mechanics

Formulation of the lagrangian mechanics assumes the description of a mechanical system by specifying its generalised coordinates and velocities. Such a description is not the only possible one. In many cases, especially for study general questions of mechanics, it is more convenient to describe a system with the help of generalised coordinated and momenta. The corresponding description is known as *hamiltonian formalism*.

To develop a certain culture of working with indices in the situations when the metric is not euclidean, in this section we use for coordinates the standard notation where they come with upper indices, while the corresponding momenta (co-vectors) naturally carry the lower indices. Of course, for the euclidean metric on the phase space there is no distinction between upper and lower indices.

2.6.1 Hamilton's equations

The main object of the hamiltonian description of mechanics is *hamiltonian* H, which is the energy of the system expressed in terms of canonical (generalised) coordinates and momenta, $H(p,q) \equiv H(p_j, q_j)$. For a system with n degrees of freedom, instead of the n second-order Euler-Lagrange equations, in the hamiltonian formalism one has $2n$ first-order differential equations for the canonical coordinates and momenta which have the following remarkably symmetrical form

$$\begin{aligned} \dot{q}_j &= \quad \frac{\partial H}{\partial p_j}\,, \\ \dot{p}_j &= -\frac{\partial H}{\partial q_j}\,. \end{aligned} \tag{I.2.87}$$

These are *Hamilton's equations*. Because of their beauty and symmetry these equations are called *canonical* equations of mechanics, and the variables - coordinates and momenta - *canonical variables*. Here the function $H(p_j, q_j)$ is related to the Lagrangian of the system by the so-called Legendre transformation

$$H(p,q) = \sum_{j=1}^{n} p_j\dot{q}_j - L(q,\dot{q},t)\Big|_{\dot{q}^i\to p_i}\,, \qquad p_i = \frac{\partial L}{\partial \dot{q}_i}\,. \tag{I.2.88}$$

This means that we have to invert the equations $p_i = \frac{\partial L}{\partial \dot{q}_i}$ and obtain $\dot{q}_i = \dot{q}_i(p,q)$ and then substitute these expressions for the velocities into the Legendre transform, obtaining thereby the Hamiltonian as a function of the canonical coordinates and momenta.

The Hamilton's equations can be obtained starting form the lagrangian description. as follows. We take the full differential of the lagrangian[17]

$$dL = \frac{\partial L}{\partial q_i} dq_i + \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i = \dot{p}_i dq_i + p_i d\dot{q}_i = \dot{p}_i dq_i + d(p_i \dot{q}_i) - \dot{q}_i dp_i \,,$$

where we have used the definition of the canonical momentum and the Euler-Lagrange equations. From here we find

$$d(p_i \dot{q}_i - L) = \dot{q}_i dp_i - \dot{p}_i dq_i \,. \tag{I.2.89}$$

Defining the hamiltonian by means of the Legendre transform

$$H(p,q) = p_i \dot{q}_i - L(q,\dot{q})\Big|_{\dot{q}_i \to p_i} \,,$$

we then deduce from (I.2.89) that the differential of the hamiltonian considered as the function of p_i and q_i reads as

$$dH = \dot{q}_i dp_i - \dot{p}_i dq_i \,.$$

From this expression, the differential Hamilton's equations (I.2.87) follow. The inverse Legendre transform allows one to reconstruct the lagrangian from a given hamiltonian and, therefore, in the non-singular situation[18] these descriptions are perfectly equivalent.

The total derivative of the hamiltonian reads

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_{i=1}^{n} \frac{\partial H}{\partial q_i} \dot{q}_i + \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} \dot{p}_i \,.$$

Substituting here $\dot{q}_i$ and $\dot{p}_i$ from equations (I.2.87), we see that the last two terms cancel and we get

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} \,.$$

In particular, if the hamiltonian does not explicitly depends on time, then $dH/dt = 0$ and we again obtain the conservation law of energy.

Hamilton's equations can also be obtained by means of the variational principle. The corresponding action has the form

$$S[p,q] = \int_{t_1}^{t_2} \big(p_i \dot{q}^i - H(p,q)\big) \mathrm{d}t \,. \tag{I.2.90}$$

Varying this action with respect to p and q, *considered as independent variables*, one obtains the hamiltonian equations.

Example. Newtonian mechanics. We start form the following lagrangian for a system on N particles interacting by means of potential forces

$$L = \sum_{i=1}^{N} \frac{m_i \vec{v}_i^{\,2}}{2} - U(\vec{r}_1, \ldots, \vec{r}_N) \,.$$

[17] To simplify our consideration, we use here Einstein's convention for the summation of indices.

[18] A singular situation arises when one cannot solve equations $p_i = \frac{\partial L}{\partial \dot{q}_i}$. In this case one has to modify the hamiltonian formalism, how to do this was explained by P.M. Dirac. The corresponding discussion goes beyond the scope of the present course.

The canonical momenta are

$$\vec{p}_i = \frac{\partial L}{\partial \vec{v}_i} = m_i \vec{v}_i \quad \rightarrow \quad \vec{v}_i = \frac{\vec{p}_i}{m_i} \,,$$

so that for the corresponding hamiltonian we find

$$H = \sum_{i=1}^{N} \vec{p}_i \cdot \vec{v}_i - L = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m_i} + U(\vec{r}_1, \ldots, \vec{r}_N) = T + U \,.$$

Hamilton's equations are

$$\begin{aligned} \dot{\vec{r}}_i &= \frac{\partial H}{\partial \vec{p}_i} = \frac{\vec{p}_i}{m_i} \,, \\ \dot{\vec{p}}_i &= -\frac{\partial H}{\partial \vec{r}_i} = -\frac{\partial U}{\partial \vec{r}_i} \,. \end{aligned}$$

These equations are equivalent to Newton's equations. Indeed, we differentiate the first equation over t, multiply it by m_i and use the second equation to get

$$m_i \ddot{\vec{r}}_i = \dot{\vec{p}}_i = -\frac{\partial U}{\partial \vec{r}_i} = \vec{F}_i \,,$$

which are nothing else but Newton's equations.

Hamilton's equations can be represented in the form of a single equation. Introduce two $2n$-dimensional vectors

$$\vec{x} = \begin{pmatrix} q \\ p \end{pmatrix} \,, \qquad \vec{\nabla} H = \begin{pmatrix} \frac{\partial H}{\partial q_j} \\ \frac{\partial H}{\partial p_j} \end{pmatrix} \tag{I.2.91}$$

and $2n \times 2n$ matrix J:

$$J = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \,, \tag{I.2.92}$$

where $\mathbb{1}$ is the $n \times n$ unit matrix. Then (I.2.87) are concisely written as

$$\dot{\vec{x}} = -J \cdot \vec{\nabla} H \,, \qquad \text{or} \qquad J \cdot \dot{\vec{x}} = \vec{\nabla} H \,. \tag{I.2.93}$$

In this form the Hamiltonian equations were written for the first time by Lagrange in 1808.

The vector $x = (x^1, \ldots, x^{2n})$ defines a state of a dynamical system in classical mechanics. The set of all states forms the *phase space* $\mathscr{P} = \{x\}$ of the system which in the present case is the $2n$-dimensional space with the euclidean metric $(x, y) = \sum_{i=1}^{2n} x^i y^i$. Solving Hamilton's equations with given initial conditions (p_0, q_0) representing a point in the phase space, we obtain a phase space curve

$$p \equiv p(t; p_0, q_0) \,, \qquad q \equiv q(t; p_0, q_0)$$

passing through this point. As follows from the uniqueness theorem for ordinary differential equations, there is one and only one phase space curve through every phase space point. The tangent vector to the phase space curve is called *the phase velocity vector or the Hamiltonian vector field*. By construction, it is determined by the Hamiltonian equations. The phase curve can consist of

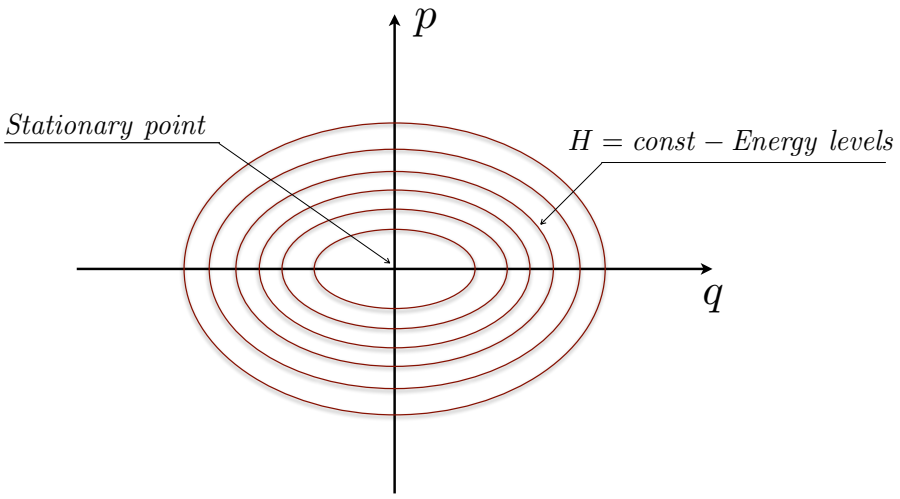

Figure 2.13: Phase space trajectories of the harmonic oscillator.

only one point. Such a point is called an *equilibrium position*. The Hamiltonian vector field at an equilibrium position vanishes.

Example. Phase space of the one-dimensional harmonic oscillator. To get more familiar with the concept of a phase space, consider as an example the one-dimensional the harmonic oscillator. The potential is $U(q) = \frac{m\omega^2}{2}q^2$. The Hamiltonian

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2\,.$$

Hamilton's equations of motion are given by *ordinary* differential equations:

$$\dot{q} = \frac{p}{m}\,, \quad \dot{p} = -m\omega^2\, q \quad \Longrightarrow \quad \ddot{q} = -\omega^2\, q\,.$$

The law of conservation of energy allows one to find the phase curves easily. On each phase curve the value of the total energy $E = H$ is constant. Therefore, each phase curve lies entirely in one energy level set $H(p,q) = h$. For the harmonic oscillator

$$\frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 = h \quad \rightarrow \quad \frac{p^2}{2mh} + \frac{q^2}{\frac{2h}{m\omega^2}} = 1\,.$$

and the phase space curves are ellipses and the origin, see Fig. 2.13. The Hamiltonian vector field is

$$X_H = \dot{q}\frac{\partial}{\partial q} + \dot{p}\frac{\partial}{\partial p} = \frac{p}{m}\frac{\partial}{\partial q} - m\omega^2\, q\frac{\partial}{\partial p}\,, \quad X_H \cdot H = 0\,.$$

2.6.2 Poisson brackets

Let $\mathcal{F}(\mathscr{P})$ be the space of smooth real-valued functions on $\mathscr{P}$. This space carries the structure of an algebra with respect to the point-wise multiplication and its elements are called *observables*. Using the matrix J, one can define on $\mathcal{F}(\mathscr{P})$ the following *Poisson bracket*

$$\{f,g\}(x) = J^{ij}\partial_i f \partial_j g = \sum_{i=1}^{n}\left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}\right)$$

for any $f, g \in \mathcal{F}(\mathscr{P})$. The Poisson bracket is a map $\mathcal{F}(\mathscr{P}) \times \mathcal{F}(\mathscr{P}) \to \mathcal{F}(\mathscr{P})$ which has the following properties

1) Linearity $\{f + \alpha h, g\} = \{f, g\} + \alpha\{h, g\}$;

2) Skew-symmetry $\{f, g\} = -\{g, f\}$;

3) Jacobi identity $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$;

4) Leibniz rule $\{f, gh\} = \{f, g\}h + g\{f, h\}$

for arbitrary functions $f, g, h \in \mathcal{F}(\mathscr{P})$ and $\alpha \in \mathbb{R}$. The first three properties imply that the Poisson bracket introduces on $\mathcal{F}(\mathscr{P})$ the structure of an infinite-dimensional Lie algebra, while the Leibniz rule expresses the compatibility of the bracket with multiplication in $\mathcal{F}(\mathscr{P})$. Due to this rule, the bracket is fully determined by its values on the coordinate functions x^i for which $\{x^i, x^j\} = J^{ij}$ or, explicitly,

$$\{q_i, q_j\} = 0\,, \quad \{p_i, p_j\} = 0\,, \quad \{p_i, q_j\} = \delta_{ij}\,. \tag{I.2.94}$$

Using the Poisson bracket, Hamilton's equations for the coordinate functions can be rephrased in the following concise form

$$\dot{x}^j = \{H, x^j\} \qquad \Leftrightarrow \qquad \dot{x} = \{H, x\} = X_H \cdot x\,,$$

where X_H is the hamiltonian vector field. As a consequence, evolution of any function $f = f(q, p, t)$ on the phase space is governed by the equation

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{H, f\}\,.$$

In particular, if f does not explicitly depend in time,

$$\frac{df}{dt} = \{H, f\} = X_H \cdot f\,.$$

Due to the skew-symmetry property of the Poisson bracket, this form of Hamilton's equations makes the conservation law for H obvious

$$\frac{dH}{dt} = \{H, H\} = 0\,.$$

It follows from Jacobi identity that the Poisson bracket of two integrals of motion is again an integral of motion (*Poisson theorem*). The Leibniz rule implies that a product of two integrals of motion is also an integral of motion. The algebra of integrals of motion represents an important characteristic of a Hamiltonian system and it is closely related to the existence of a symmetry group.

Poisson manifolds. The phase space is not always Euclidean $\mathbb{R}^{2n}$. The generic situation is that the phase space is a manifold and the the Poisson tensor J is different from (I.2.92). The properties $1) - 4)$ provide a general definition of the Poisson bracket for an arbitrary smooth manifold $\mathscr{P}$. Any Poisson bracket is described by a skew-symmetric tensor J on $\mathscr{P}$ satisfying the Jacoby identity. In local coordinates this identity takes the form

$$\sum_{(i,l,m)} J^{ik}\partial_k J^{lm} = 0\,,$$

where the sum is over the cyclic permutation of indices. A manifold endowed with a Poisson bracket is called *Poisson*.

In general, the rank r of the matrix J is less than or equal to the dimension $\dim \mathscr{P}$ of a manifold and it might change from point to point. In the case when $r = \dim \mathscr{P}$ at every point, the matrix J is invertible and the corresponding Poisson bracket is called non-degenerate. This is only possible if $\dim \mathscr{P}$ is even. Indeed, since $J^t = -J$, one has

$$\det J = \det(-J) = (-1)^{\dim \mathscr{P}} \det J \,,$$

so that $(-1)^{\dim \mathscr{P}} = 1$ since $\det J \neq 0$.

Example (*optional*). As a non-trivial example of the hamiltonian dynamics we consider the motion of a rigid body. Starting from the rigid body lagrangian

$$L = \frac{1}{2}\mathrm{Tr}(\mathcal{I}S^2)\,, \qquad S = B^{-1}\dot{B}\,,$$

see (I.2.65), we derive the Poisson structure and Hamilton's equations, as well as verify that they are perfectly match our findings in the lagrangian approach. We start with computing the canonical momentum

$$P_{ji} = \frac{\partial L}{\partial \dot{B}_{ij}} = \frac{1}{2}\frac{\partial L}{\partial \dot{B}_{ij}}\big(\mathcal{I}B^{-1}\dot{B}B^{-1}\dot{B}\big)_{kk} = \frac{1}{2}(B^{-1}\dot{B}\mathcal{I}B^{-1})_{ji} + \frac{1}{2}(\mathcal{I}B^{-1}\dot{B}B^{-1})_{ji}\,.$$

Here, by definition, we understand P_{ji} as the canonical momentum conjugate to the coordinate B_{ij}. The above formula for the canonical momentum can be written in the matrix form

$$P = \frac{1}{2}(S\mathcal{I} + \mathcal{I}S)B^{-1}\,, \tag{I.2.95}$$

so that

$$PB = \frac{1}{2}\big(S\mathcal{I} + \mathcal{I}S\big) = \Lambda\,, \tag{I.2.96}$$

where Λ is the angular momentum in the moving frame, *cf.* (I.2.66). By using the Legendre transform, we determine the hamiltonian

$$\begin{aligned} H = P_{ij}\dot{B}_{ji} - L &= \mathrm{Tr}(P\dot{B}) - \frac{1}{2}\mathrm{Tr}(\mathcal{I}S^2) = \mathrm{Tr}(P\dot{B}) - \frac{1}{4}\mathrm{Tr}(\mathcal{I}S + S\mathcal{I})S \\ &= \mathrm{Tr}(P\dot{B}) - \frac{1}{2}\mathrm{Tr}(\Lambda S) = \mathrm{Tr}(P\dot{B}) - \frac{1}{2}\mathrm{Tr}(PBS) = \frac{1}{2}\mathrm{Tr}(P\dot{B})\,, \end{aligned}$$

where we have used (I.2.96). The hamiltonian appears to coincide with the lagrangian because the lagrangian is just the kinetic energy. This is not yet the final answer because we still have to express H in terms of coordinates and momenta. To solve for $\dot{B}$ in terms of P, we write (I.2.96) in the matrix form

$$2(PB)_{ij} = (S\mathcal{I} + \mathcal{I}S)_{ij} = S_{ij}\mathcal{I}_j + \mathcal{I}_i S_{ij} = (\mathcal{I}_i + \mathcal{I}_j)S_{ij}\,,$$

from where we find

$$S_{ij} = (B^{-1}\dot{B})_{ij} = \frac{2(PB)_{ij}}{\mathcal{I}_i + \mathcal{I}_j} = \frac{2\Lambda_{ij}}{\mathcal{I}_i + \mathcal{I}_j}\,. \tag{I.2.97}$$

Thus, the hamiltonian is the following function of the canonical coordinates and momenta

$$H = \frac{1}{2}(PB)_{ij}S_{ji} = \sum_{ij}\frac{(PB)_{ij}(PB)_{ji}}{\mathcal{I}_i + \mathcal{I}_j}\,. \tag{I.2.98}$$

Determination of the Poisson structure between the canonical variables is not straightforward. This is because we deal with a constrained system. Although we do not use the method of lagrangian multipliers, we should have in mind the orthogonality constraint

$$C_1 = B^t B - \mathbb{1} = 0\,. \tag{I.2.99}$$

Differentiating this constraint in time,

$$\dot{C}_1 = \dot{B}^t B + B^t \dot{B} = (B^t \dot{B})^t + B^t \dot{B} = (B^{-1}\dot{B})^t + B^{-1}\dot{B} = S^t + S\,.$$

Thus, if we want to keep $C_1 = 0$ for all the times, we have to impose a condition $S^t + S = 0$ meaning that S should be a skew-symmetric matrix. Then, definition (I.2.96) of the canonical momentum leads to another hamiltonian constraint

$$C_2 = (PB)^t + PB = 0\,. \tag{I.2.100}$$

This constraint implies that not all the components of the canonical momentum are independent, rather from (I.2.100) together with (I.2.99) one gets that

$$P^t = -BPB\,.$$

Note by passing that the above relation also implies that

$$(BP)^t + BP = 0\,. \tag{I.2.101}$$

No new constraints further arise; if we differentiate C_2 then the corresponding equation will be satisfied due to the equations of motion.

It is now clear that the Poisson bracket between the conjugate coordinates and momenta cannot be the canonical bracket $\{P_{ij}, B_{kl}\} = \delta_{il}\delta_{kl}$, as the latter is not compatible with constraints (I.2.99) and (I.2.100). An educated guess for the Poisson bracket compatible with constraints is

$$\begin{aligned} &\{B_{ij}, B_{kl}\} = 0\,,\\ &\{P_{ij}, B_{kl}\} = \tfrac{1}{2}\big(\delta_{il}\delta_{jk} - B_{ki}B_{jl}\big)\,,\\ &\{P_{ij}, P_{kl}\} = \tfrac{1}{2}\big(\delta_{ik}(BP)_{jl} - \delta_{jl}(PB)_{ik}\big)\,. \end{aligned} \tag{I.2.102}$$

At the end of this example we provide (in small print) a straightforward verification of the Jacobi identity, as well as show the compatibility of (I.2.102) with constraints (I.2.99) and (I.2.100). One way to derive (I.2.102) from the canonical Poisson bracket $\{P_{ij}, B_{kl}\} = \delta_{il}\delta_{kl}$ is to use the Dirac bracket construction known in the theory of constrained hamiltonian systems. Another way is to note that (I.2.102) is equivalent to the canonical Poisson structure of the cotangent bundle $\mathrm{T}^*\mathrm{SO}(n)$ of the orthogonal group $\mathrm{SO}(n)$.

More generally, the cotangent bundle T^*G to a Lie group G is a manifold isomorphic to the product $\mathrm{T}^*G \simeq G \times \mathfrak{g}^*$, where $\mathfrak{g}^*$ is the dual space to the Lie algebra $\mathfrak{g}$ of G. If the space $\mathfrak{g}$ is supplied with a non-degenerate bilinear form, we can use this form to identify $\mathfrak{g}^*$ with $\mathfrak{g}$: $\mathfrak{g}^* \simeq \mathfrak{g}$, so that the cotangent bundle is isomorphic to $G \times \mathfrak{g}$. As such, it can be parametrised by elements (g, ℓ), where $g \in G$ and $\ell \in \mathfrak{g}$. Matrix elements of the defining representations of G in $\mathrm{GL}(n, \mathbb{C})$ and $\mathfrak{g}$ in $\mathrm{Mat}(n, \mathbb{C})$ can be regarded as coordinate functions on the cotangent bundle; we denote these coordinate functions as g_{ij} and ℓ_{ij}, respectively, where $i, j = 1, \ldots, n$. Finally, the cotangent bundle of G is a Poisson manifold with the following Poisson bracket which we write as the set of brackets between the coordinate functions

$$\begin{aligned} &\{g_1, g_2\} = 0\,,\\ &\{\ell_1, g_2\} = g_2 C_{12}\,,\\ &\{\ell_1, \ell_2\} = [C_{12}, \ell_1] = \tfrac{1}{2}[C_{12}, \ell_1 - \ell_2]\,. \end{aligned} \tag{I.2.103}$$

Here subscript 1 and 2 stands as a concise notation for the matrix indices ij and kl, respectively, and $C_{12} \in \mathfrak{g} \otimes \mathfrak{g}$ is the so-called *split Casimir*. For any $A \in \mathfrak{g}$ the spilt Casimir has the property that

$$A_1 = \mathrm{Tr}_2 C_{12} A_2 \,, \tag{I.2.104}$$

which in components means $A_{ij} = C_{ij,kl} A_{lk}$, *i.e.* C is the identity operator in $\mathfrak{g}$.

Now we specify the group G and the Poisson structure of its cotangent bundle to the case of interest $G = \mathrm{SO}(n)$. For the orthogonal group the corresponding split Casimir is

$$C_{ij,kl} = \frac{1}{2}(\delta_{il}\delta_{jk} - \delta_{ik}\delta_{jl}) \,.$$

It is skew-symmetric with respect to the interchange $i \leftrightarrow j$ and separately with respect to $k \leftrightarrow l$ and it fulfils (I.2.104) for any skew-symmetric matrix A. For the Poisson structure (I.2.103) we then find in components

$$\begin{aligned} &\{g_{ij}, g_{kl}\} = 0 \,, \\ &\{\ell_{ij}, g_{kl}\} = g_{km} C_{ij,ml} = \tfrac{1}{2}\big(\delta_{il} g_{kj} - \delta_{jl} g_{ki}\big) \,, \\ &\{\ell_{ij}, \ell_{kl}\} = C_{im,kl}\ell_{mj} - \ell_{im} C_{mj,kl} = \tfrac{1}{2}(\delta_{il}\ell_{kj} - \delta_{ik}\ell_{lj} - \delta_{jk}\ell_{il} + \delta_{jl}\ell_{ik}) \,. \end{aligned} \tag{I.2.105}$$

Now one can verify that with an identification

$$g = B \,, \qquad \ell = PB = \Lambda \tag{I.2.106}$$

the Poisson structure (I.2.102) for B and P precisely yields the structure (I.2.105) for g and ℓ. To prove this result, upon evaluation of the brackets by the Leibniz rule one has to use the constraints (I.2.99) and (I.2.100). This shows that a generalised Euler's top can be understood as a dynamical system on the cotangent bundle of the orthogonal group $\mathrm{SO}(n)$. The structure (I.2.103) is called the Poisson structure in the *left* parametrisation. The *right* parametrisation amounts to replacing ℓ with an element

$$m = g\ell g^{-1} = B(PB)B^{-1} = BP = B\Lambda B^{-1} = J \,,$$

which physically coincide with the angular momentum in the stationary frame, see (I.2.68). By using either (I.2.102) or (I.2.103), one can show that angular momenta in the moving and stationary frames Poisson commute $\{\ell_{ij}, m_{kl}\} = \{\Lambda_{ij}, J_{kl}\} = 0$.

Let us now derive Hamilton's equations. We have

$$\dot{B}_{ij} = \{H, B_{ij}\} = \sum_l B_{il} \frac{(PB)_{lj}}{\mathcal{I}_j + \mathcal{I}_l} = \sum_l B_{il} S_{lj} = (BS)_{ij} \,,$$

where we have used (I.2.99) and (I.2.100). Note that the result we obtained agrees with our original definition of S as $S = B^{-1}\dot{B}$. Analogously,

$$\dot{P}_{ij} = \{H, P_{ij}\} = -\sum_l \frac{(PB)_{ik}}{\mathcal{I}_k + \mathcal{I}_i} P_{kj} = -(SP)_{ij} \,,$$

where again (I.2.99) and (I.2.100) have been used. Hence, Hamilton's equations written in the matrix form are

$$\dot{B} = BS \,, \qquad \dot{P} = -SP \,, \tag{I.2.107}$$

where S is understood as the matrix with entries given by (I.2.97). We can now find the evolution equation for $\Lambda = PB$

$$\dot{\Lambda} = \dot{P}B + P\dot{B} = -SPB + PBS = -S\Lambda + \Lambda S = -[S, \Lambda] \,,$$

which obviously coincide with (I.2.67) found from the lagrangian approach.

Now we make the following interesting observation. According to the last line in (I.2.105), the Poisson bracket between the components of Λ is closed

$$\{\Lambda_{ij},\Lambda_{kl}\} \quad = \quad \tfrac{1}{2}(\delta_{il}\Lambda_{kj}-\delta_{ik}\Lambda_{lj}-\delta_{kj}\Lambda_{il}+\delta_{jl}\Lambda_{ik})\,. \tag{I.2.108}$$

The matrix Λ is skew-symmetric and the above relations coincide with the defining relations of the Lie algebra so(n) for the case of general n. The hamiltonian can also be expressed via Λ only

$$H=\sum_{ij}\frac{\Lambda_{ij}\Lambda_{ji}}{\mathcal{I}_i+\mathcal{I}_j}=\sum_{i\neq j}\frac{\Lambda_{ij}\Lambda_{ji}}{\mathcal{I}_i+\mathcal{I}_j}\,. \tag{I.2.109}$$

Note that the actual summation here runs over $i\neq j$ because $\Lambda_{ii}=0$.[19] Now we can also show that Euler's equations are hamiltonian with respect to the Poisson structure (I.2.108) and hamiltonian (I.2.109). Indeed,

$$\begin{aligned}\dot\Lambda_{kl}=\{H,\Lambda_{kl}\} \quad &= \quad 2\sum_{ij}\frac{\Lambda_{ji}}{\mathcal{I}_i+\mathcal{I}_j}\{\Lambda_{ij},\Lambda_{kl}\}=\sum_{ij}\frac{\Lambda_{ji}}{\mathcal{I}_i+\mathcal{I}_j}(\delta_{il}\Lambda_{kj}-\delta_{ik}\Lambda_{lj}-\delta_{kj}\Lambda_{il}+\delta_{jl}\Lambda_{ik})\\ &= \quad \sum_j\Lambda_{kj}\frac{2\Lambda_{jl}}{\mathcal{I}_l+\mathcal{I}_j}-\sum_i\frac{2\Lambda_{ki}}{\mathcal{I}_i+\mathcal{I}_k}\Lambda_{il}=(\Lambda S)_{kl}-(S\Lambda)_{kl}\,.\end{aligned}$$

An algebraic variety generated by the matrix elements Λ_{ij} of a skew-symmetric matrix Λ supplied with the Poisson bracket (I.2.108) also gives an example of Poisson manifold. This manifold is the dual space $\mathfrak{g}^*$ to the Lie algebra $\mathfrak{g}$ and the corresponding Poisson bracket (I.2.108) is the so-called *Kirillov-Kostant bracket*. In opposite the canonical bracket on the cotangent bundle, bracket (I.2.108) is degenerate and has $C_k=\mathrm{Tr}\Lambda^k$ as Casimir functions.

For the reader who is not familiar with the notion of the cotangent bundle and its Poisson structure, here we proved the proof of the Jacobi identity for the Poisson brackets (I.2.102) by straightforward calculation. Because of the structure of the brackets, the Jacobi identities involving three B or two B and one P are trivially satisfied. The first non-trivial Jacobi identity is

$$\{\{P_{ij},P_{kl}\},B_{mn}\}+\{\{P_{kl},B_{mn}\},P_{ij}\}+\{\{B_{mn},P_{ij}\},P_{kl}\}=0\,.$$

We then compute the brackets involved here one by one. We start with

$$\begin{aligned}\{\{P_{ij},P_{kl}\},B_{mn}\} \quad &= \quad \{\delta_{ik}(BP)_{jl}-\delta_{jl}(PB)_{ik},B_{mn}\}=\delta_{is}B_{js}\{P_{sl},B_{mn}\}-\delta_{jl}\{P_{is},B_{mn}\}B_{sk}\\ &= \quad \delta_{ik}B_{js}(\delta_{sn}\delta_{lm}-B_{ms}B_{ln})-\delta_{jl}(\delta_{in}\delta_{ms}-B_{mi}B_{sn})B_{sk}\\ &= \quad \delta_{ik}\delta_{lm}B_{jn}-\delta_{ik}\delta_{jm}B_{ln}-\delta_{jl}\delta_{in}B_{mk}+\delta_{jl}\delta_{nk}B_{mi}\,,\end{aligned}$$

where constraint (I.2.99) has been implemented. next,

$$\begin{aligned}\{\{P_{kl},B_{mn}\},P_{ij}\} \quad &= \quad \{\delta_{kn}\delta_{ml}-B_{mk}B_{ln},P_{ij}\}=\{P_{ij},B_{mk}\}B_{ln}+B_{mk}\{P_{ij},B_{ln}\}\\ &= \quad (\delta_{ik}\delta_{jm}-B_{mi}B_{jk})B_{ln}+B_{mk}(\delta_{in}\delta_{jl}-B_{li}B_{jn})\,.\end{aligned}$$

Analogously,

$$\{\{B_{mn},P_{ij}\},P_{kl}\}=-(\delta_{ki}\delta_{ml}B_{jn}-B_{mk}B_{li}B_{jn})-B_{mi}(\delta_{kn}\delta_{jl}-B_{jk}B_{ln})\,.$$

Now adding up these three pieces we obtain zero, *i.e.* the Jacobi identity at hand is satisfied.

The last Jacobi identity to check is

$$\{\{P_{ij},P_{kl}\},P_{mn}\}+\{\{P_{mn},P_{ij}\},P_{kl}\}+\{\{P_{kl},P_{mn}\},P_{ij}\}=0\,.$$

We have

$$\{\{P_{ij},P_{kl}\},P_{mn}\} \quad = \quad \{\delta_{ik}(BP)_{jl}-\delta_{jl}(PB)_{ik},P_{mn}\}$$

[19] For the usual case when $i,j,k=1,2,3$ one has $\mathcal{I}_i+\mathcal{I}_j=-I_k<0$, where $k\neq i$ and $k\neq j$. Since $\Lambda_{ij}\Lambda_{ji}=-\Lambda_{ij}^2<0$, the hamiltonian H is positive-definite.

$$
\begin{aligned}
&= \delta_{ik}\{B_{js},P_{mn}\}P_{sl} + \delta_{ik}B_{js}\{P_{sl},P_{mn}\} - \delta_{jl}\{P_{is},P_{mn}\}B_{sk} - \delta_{jl}P_{is}\{B_{sk},P_{mn}\}\\
&= -\delta_{ik}(\delta_{ms}\delta_{nj} - B_{jm}B_{ns})P_{sl} + \delta_{ik}B_{js}(\delta_{sm}(BP)_{ln} - \delta_{ln}(PB)_{sn})\\
&\quad -\delta_{jl}(\delta_{im}(BP)_{sn} - \delta_{sn}(PB)_{im})B_{sk} + \delta_{jl}P_{is}(\delta_{mk}\delta_{ns} - B_{sm}B_{nk})\,.
\end{aligned}
$$

Hence, we get

$$
\begin{aligned}
\{\{P_{ij},P_{kl}\},P_{mn}\} &= -\delta_{ik}\delta_{nj}P_{ml} + \cancel{\delta_{ik}B_{jm}(BP)_{nl}} + \cancel{\delta_{ik}B_{jm}(BP)_{ln}} + \delta_{ik}\delta_{ln}P_{mj}\\
&= -\delta_{jl}\delta_{im}P_{kn} + \cancel{\delta_{jl}(PB)_{im}B_{nk}} + \delta_{jl}\delta_{mk}P_{in} - \cancel{\delta_{jl}(PB)_{im}B_{nk}}\,,
\end{aligned}
$$

where the cancellation on the first line happen due to the constraint (I.2.101). Thus,

$$
\{\{P_{ij},P_{kl}\},P_{mn}\} = -\delta_{ik}\delta_{nj}P_{ml} + \delta_{ik}\delta_{ln}P_{mj} - \delta_{jl}\delta_{im}P_{kn} + \delta_{jl}\delta_{mk}P_{in}\,.
$$

Making here the cyclic permutations of pairs of indices we get the other two terms in the left hand side of the Jacobi identity and adding all them up we find zero. Thus, we conclude that on the constraint surface given by equations (I.2.99) and (I.2.100), the Jacobi identity is satisfied.

Now we check the compatibility of constraints with the Poisson structure. To this end, we compute the Poisson brackets of the constraints (I.2.99) and (I.2.100) with coordinates and momenta and show that on the constraint surface given by (I.2.99) and (I.2.100) these brackets vanish. We start with (I.2.99) and notice that its Poisson bracket with B_{ij} trivially vanishes. Then, we have to compute

$$
\begin{aligned}
\{(B^tB)_{ij},P_{kl}\} &= \{B_{si}B_{sj},P_{kl}\} = -\{P_{kl},B_{si}\}B_{sj} - B_{si}\{P_{kl},B_{sj}\}\\
&= -(\delta_{ki}\delta_{ls} - B_{sk}B_{li})B_{sj} - B_{si}(\delta_{kj}\delta_{ls} - B_{sk}B_{lj})\\
&= -\delta_{ki} + \delta_{kj}B_{li} - \delta_{kj}B_{li} + \delta_{ik}B_{lj} = 0\,,
\end{aligned}
$$

where constraint (I.2.99) have been used. Next, we look at the brackets of (I.2.100). First,

$$
\begin{aligned}
\{(PB)_{ij} + (PB)_{ji},B_{kl}\} &= \{P_{is},B_{kl}\}B_{sj} + i\leftrightarrow j = (\delta_{il}\delta_{ks} - B_{ki}B_{sl})B_{sj} + i\leftrightarrow j\\
&= \delta_{il}B_{kj} - \delta_{lj}B_{ki} + \delta_{jl}B_{ki} - \delta_{li}B_{kj} = 0\,.
\end{aligned}
$$

Second,

$$
\begin{aligned}
\{(PB)_{ij} + (PB)_{ji},P_{kl}\} &= \{P_{is},P_{kl}\}B_{sj} + P_{is}\{B_{sj},P_{kl}\} + i\leftrightarrow j\\
&= (\delta_{ik}(BP)_{sl} - \delta_{sl}(PB)_{ik})B_{sj} - P_{is}(\delta_{kj}\delta_{sl} - B_{sk}B_{lj}) + i\leftrightarrow j\\
&= \delta_{ik}P_{jl} - \cancel{(PB)_{ik}B_{lj}} - P_{il}\delta_{kj} + \cancel{(PB)_{ik}B_{lj}} + i\leftrightarrow j = 0\,.
\end{aligned}
$$

This completes the proof of compatibility of the constraints with the Poisson structure.

2.6.3 Canonical transformations

Consider a generic change of coordinates on a phase space

$$
Q_i = Q_i(q,p,t)\,,\qquad P_i = P_i(q,p,t) \tag{I.2.110}
$$

and ask which of these transformations preserve the form (I.2.87) of Hamilton's equations with a new hamiltonian $H' = H'(P,Q)$, that is

$$
\dot{Q}_i = \frac{\partial H'}{\partial P_i}\,,\qquad \dot{P}_i = -\frac{\partial H'}{\partial Q_i}\,.
$$

To derive the transformations that preserve the form of Hamilton's equations, we can appeal to the variational principle according to which Hamilton's equations are obtained as

$$
\delta\int(p_idq_i - Hdt) = 0\,,
$$

where the independent variables to vary are coordinates q_i and momenta p_i. In order for new variables P_i, Q_i to also satisfy Hamilton's equations, they must also obey the principle of the least action

$$
\delta\int(P_idQ_i - H'dt) = 0\,.
$$

The equivalence of two variational proinciples can be achieved provided the difference between the integrands is a total differential of an arbitrary function F of coordinates, momenta and time

$$p_i dq_i - H dt = P_i dQ_i - H' dt + dF\,.$$

Transformations satisfying this property are called *canonical*.[20] Any canonical transformation is characterised by its function F, called the generating function of canonical transformation. Rewriting the last relation in the form

$$dF = p_i dq_i - P_i dQ_i + (H' - H)dt\,, \tag{I.2.111}$$

one sees that

$$p_i = \frac{\partial F}{\partial q_i}\,, \quad P_i = -\frac{\partial F}{\partial Q_i}\,, \quad H' = H + \frac{\partial F}{\partial t}\,, \tag{I.2.112}$$

where F is considered as a function of and and new coordinates and time: $F = F(q, Q, t)$. Equation

$$p_i = \frac{\partial F}{\partial q_i}(q, Q, t) \tag{I.2.113}$$

should yield $Q_i = Q_i(p, q, t)$, *i.e.* to express the new coordinates in terms of the old coordinates and momenta.[21] Substituting these esperssions for Q_i into the equation for P_i

$$P_i = -\frac{\partial F}{\partial Q_i}(q, Q(p, q), t)\,, \tag{I.2.114}$$

we obtain $P_i = P_i(p, q, t)$, *i.e.* the expression for the new momenta in terms of old coordinates and momenta. This would give the canonical transformation generated by a function $F = F(q, Q)$ in the form (I.2.110).

It might be convenient to express the generating function not via q and Q, but rather via q and P. For that one needs to rewrite (I.2.111) in the form

$$d(F + P_i Q_i) = p_i dq_i + Q_i dP_i + (H' - H)dt\,.$$

On the left hand side under the differential one has a new generating function $\Phi = F + P_i Q_i = \Phi(q, P, t)$ and the last relation implies that

$$p_i = \frac{\partial \Phi}{\partial q_i}\,, \quad Q_i = \frac{\partial \Phi}{\partial P_i}\,, \quad H' = H + \frac{\partial \Phi}{\partial t}\,.$$

Analogously, one can obtain the formulae for canonical transformations expressed via generating functions depending on p and Q, or on p and P. If any of these generating functions does not depend on t explicitly, $H' = H$, in other words, to get H' one needs to substitute in $H(p, q)$ the variables p and q expressed via P and Q.

Example. Consider a system with one degree of freedom and the phase space (p, q). Consider a family of transformations which has a form of rotation on the phase plane by an angle α

$$Q = \cos\alpha\, q + \sin\alpha\, p\,, \quad P = -\sin\alpha\, q + \cos\alpha\, p\,. \tag{I.2.115}$$

[20] Note that Hamilton's equations also preserve their form if the integrands are different by a constant multiplier. An example is given by transformations of the form $P_i = a p_i$, $Q_i = q_i$, $H' = aH$.

[21] This can only be done if $\frac{D(p_1,\ldots,p_n)}{D(Q_1,\ldots,Q_n)} = ||\frac{\partial^2 F}{\partial q_i \partial Q_j}|| \neq 0$.

We show that these transformations are canonical by finding the corresponding generating function. From the expression for Q we find $p = p(q, Q)$ and then substituting the result into the equation for P, we find $P = P(q, Q)$. Explicitly,

$$\begin{aligned} p &= \frac{Q}{\sin\alpha} - \frac{\cos\alpha}{\sin\alpha} q\,, \\ P &= \frac{\cos\alpha}{\sin\alpha} Q - \frac{q}{\sin\alpha}\,. \end{aligned}$$

Thus, according to (I.2.112), the generating function must be found from the system of equations

$$\begin{aligned} \frac{\partial F}{\partial q} &= \frac{Q}{\sin\alpha} - \frac{\cos\alpha}{\sin\alpha} q\,, \\ \frac{\partial F}{\partial Q} &= -\frac{\cos\alpha}{\sin\alpha} Q + \frac{q}{\sin\alpha}\,. \end{aligned}$$

Integrating the first equation, we obtain

$$F = \frac{Qq}{\sin\alpha} - \frac{\cos\alpha}{\sin\alpha}\frac{q^2}{2} + f(Q)\,,$$

where $f(Q)$ is yet an unknown function of Q. Substituting this result into the second equation gives

$$f'(Q) = -\frac{\cos\alpha}{\sin\alpha} Q\,, \quad \Rightarrow \quad f(Q) = -\frac{\cos\alpha}{\sin\alpha}\frac{Q^2}{2}\,.$$

Hence, for the generating function we find the following result

$$F = \frac{Qq}{\sin\alpha} - \frac{\cos\alpha}{\sin\alpha}\frac{q^2}{2} - \frac{\cos\alpha}{\sin\alpha}\frac{Q^2}{2}\,.$$

Note that for $\alpha = 0$ this expression is divergent. Note that an arbitrary transformation (I.2.110) may not be canonical, to be canonical there must exist a generating function, for instance the one satisfying the system of differential equations

$$\begin{aligned} \frac{\partial F}{\partial q_i} &= p_i\,, \\ \frac{\partial F}{\partial Q_i} &= -P_i\,. \end{aligned}$$

For a number of degrees of freedom bigger than one this system is not necessarily compatible, *i.e.* the generating function does not exists.

The existence and large variety of canonical transformations, deprives to a large extent the generalised coordinates and momenta of their original meaning. This is clearly seen on the example of the transformation $Q_i = p_i$, $P_i = -q_i$, which does not affect Hamilton's equations and amounts calling coordinates momenta and vice versa. The usefulness of canonical transformations shows up in the fact that in many cases they allow to simplify the hamiltonian and, therefore, to simplify and hopefully solve Hamilton's equation.

On account of this arbitrariness of nomenclature, the variables p and q are are often called in the hamiltonian treatment *canonically conjugate variables*. The conditions related such quantities can be expressed in terms of the Poisson brackets and the following theorem

Theorem. *Poisson brackets are invariant under canonical transformations, i.e. for any two functions f and g on the phase space*

$$\{f, g\}_{p,q} = \{f, g\}_{P,Q}\,. \tag{I.2.116}$$

This property can be put as an independent definition of canonical transformations, see below.

Proof. We fist show that to satisfy (I.2.116) for any functions f and g, it is enough to require the fulfilment of this relation for the coordinate functions of new variables, namely,

$$\{Q_i, Q_j\}_{p,q} = \{Q_i, Q_j\}_{Q,P} = 0\,, \quad \{P_i, P_j\}_{p,q} = \{P_i, P_j\}_{Q,P} = 0\,, \quad \{P_i, Q_j\}_{p,q} = \{P_i, Q_j\}_{Q,P} = \delta_{ij}\,.$$

Consider

$$\begin{aligned}
\{f,g\}_{p,q} &= \sum_{i=1}^{n}\left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}\right) \\
&= \sum_{i,j,l=1}^{n}\left(\frac{\partial f}{\partial P_j}\frac{\partial P_j}{\partial p_i} + \frac{\partial f}{\partial Q_j}\frac{\partial Q_j}{\partial p_i}\right)\left(\frac{\partial g}{\partial P_l}\frac{\partial P_l}{\partial q_i} + \frac{\partial g}{\partial Q_l}\frac{\partial Q_l}{\partial q_i}\right) - \left(\frac{\partial f}{\partial P_j}\frac{\partial P_j}{\partial q_i} + \frac{\partial f}{\partial Q_j}\frac{\partial Q_j}{\partial q_i}\right)\left(\frac{\partial g}{\partial P_l}\frac{\partial P_l}{\partial p_i} + \frac{\partial g}{\partial Q_l}\frac{\partial Q_l}{\partial p_i}\right) \\
&= \sum_{i,j,l=1}^{n}\frac{\partial f}{\partial P_j}\frac{\partial g}{\partial P_l}\left(\frac{\partial P_j}{\partial p_i}\frac{\partial P_l}{\partial q_i} - \frac{\partial P_j}{\partial q_i}\frac{\partial P_l}{\partial p_i}\right) + \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial Q_l}\left(\frac{\partial P_j}{\partial p_i}\frac{\partial Q_l}{\partial q_i} - \frac{\partial P_j}{\partial q_i}\frac{\partial Q_l}{\partial p_i}\right) \\
&\quad + \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial P_l}\left(\frac{\partial Q_j}{\partial p_i}\frac{\partial P_l}{\partial q_i} - \frac{\partial Q_j}{\partial q_i}\frac{\partial P_l}{\partial p_i}\right) + \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial Q_l}\left(\frac{\partial Q_j}{\partial p_i}\frac{\partial Q_l}{\partial q_i} - \frac{\partial Q_j}{\partial q_i}\frac{\partial Q_l}{\partial p_i}\right)\,.
\end{aligned}$$

Hence, we see that

$$\{f,g\}_{p,q} = \sum_{j,l=1}^{n}\frac{\partial f}{\partial P_j}\frac{\partial g}{\partial P_l}\{P_j, P_l\}_{p,q} + \frac{\partial f}{\partial P_j}\frac{\partial g}{\partial Q_l}\{P_j, Q_l\}_{p,q} + \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial P_l}\{Q_j, P_l\}_{p,q} + \frac{\partial f}{\partial Q_j}\frac{\partial g}{\partial Q_l}\{Q_j, Q_l\}_{p,q}\,.$$

It is clear from this formula that if the Poisson brackets of new coordinates and momenta evaluated with respect to the old variables are

$$\{P_j, P_l\}_{p,q} = 0 = \{Q_j, Q_l\}\,, \qquad \{P_j, Q_l\}_{p,q} = \delta_{jl} = -\{Q_j, P_l\}_{p,q}\,, \tag{I.2.117}$$

then equations (I.2.117) represent the sufficient condition that (I.2.116) will be satisfied. This is also the necessary condition, as is seen by substituting in (I.2.116) as f and g all pairs P_j, P_l; Q_j, Q_l; P_j, Q_l and recalling that the Poisson brackets of variables computed with respect to themselves have the canonical form.

Now we evaluate the Poisson brackets between the new coordinate functions by assuming that the canonical transformation at hand is generated by a function $F = F(q, Q)$. From equation (I.2.113) we determine $Q_i = Q_i(q, p)$ and then we substitute it back into (I.2.113), so that (I.2.113) becomes an identity

$$p_i = \frac{\partial F}{\partial q_i}(q, Q(p, q))\,. \tag{I.2.118}$$

We then differentiate this identity with respect to q_j and p_j and get

$$\begin{aligned}
0 = \frac{\partial p_i}{\partial q_j} &= \frac{\partial^2 F}{\partial q_i \partial q_j} + \frac{\partial^2 F}{\partial q_i \partial Q_m}\frac{\partial Q_m}{\partial q_j}\,, \\
\delta_{ij} = \frac{\partial p_i}{\partial p_j} &= \frac{\partial^2 F}{\partial q_i \partial Q_m}\frac{\partial Q_m}{\partial p_j}\,.
\end{aligned}$$

From these matrix equations, we find

$$\frac{\partial Q_i}{\partial q_j} = -\left(\frac{\partial^2 F}{\partial q \partial Q}\right)^{-1}_{im}\frac{\partial^2 F}{\partial q_m \partial q_j}\,, \tag{I.2.119}$$

$$\frac{\partial Q_i}{\partial p_j} = \left(\frac{\partial^2 F}{\partial q \partial Q}\right)^{-1}_{ij}\,. \tag{I.2.120}$$

First we compute

$$\begin{aligned}\{Q_i,Q_j\}_{p,q} &= \frac{\partial Q_i}{\partial p_k}\frac{\partial Q_j}{\partial q_k}-\frac{\partial Q_i}{\partial q_k}\frac{\partial Q_j}{\partial p_k}\\ &= -\left(\frac{\partial^2 F}{\partial q\partial Q}\right)^{-1}_{ik}\left(\frac{\partial^2 F}{\partial q\partial Q}\right)^{-1}_{jm}\frac{\partial^2 F}{\partial q_m\partial q_k}+\left(\frac{\partial^2 F}{\partial q\partial Q}\right)^{-1}_{jk}\left(\frac{\partial^2 F}{\partial q\partial Q}\right)^{-1}_{im}\frac{\partial^2 F}{\partial q_m\partial q_k}=0\,,\end{aligned}$$

as the second derivative is symmetric

$$\frac{\partial^2 F}{\partial q_m\partial q_k}=\frac{\partial^2 F}{\partial q_k\partial q_m}\,.$$

Next, taking into account that $F=F(q,Q(q,p))$, we get

$$\begin{aligned}\{P_i,Q_j\}_{p,q} &= \frac{\partial P_i}{\partial p_k}\frac{\partial Q_j}{\partial q_k}-\frac{\partial P_i}{\partial q_k}\frac{\partial Q_j}{\partial p_k}=-\frac{\partial}{\partial p_k}\left(\frac{\partial F}{\partial Q_i}\right)\frac{\partial Q_j}{\partial q_k}+\frac{\partial}{\partial q_k}\left(\frac{\partial F}{\partial Q_i}\right)\frac{\partial Q_j}{\partial p_k}\\ &= -\frac{\partial^2 F}{\partial Q_i\partial Q_m}\frac{\partial Q_m}{\partial p_k}\frac{\partial Q_j}{\partial q_k}+\frac{\partial^2 F}{\partial Q_i\partial Q_m}\frac{\partial Q_m}{\partial q_k}\frac{\partial Q_j}{\partial p_k}+\frac{\partial^2 F}{\partial q_k\partial Q_i}\frac{\partial Q_j}{\partial p_k}\\ &= -\frac{\partial^2 F}{\partial Q_i\partial Q_m}\left[\frac{\partial Q_m}{\partial p_k}\frac{\partial Q_j}{\partial q_k}-\frac{\partial Q_m}{\partial q_k}\frac{\partial Q_j}{\partial p_k}\right]+\frac{\partial^2 F}{\partial q_k\partial Q_i}\frac{\partial Q_j}{\partial p_k}\,.\end{aligned}$$

Here the expression in the brackets coincide with $\{Q_m,Q_j\}_{pq}$ and, therefore, it vanishes, according to what has been already proved. Thus, we are left with

$$\{P_i,Q_j\}_{p,q} = \frac{\partial^2 F}{\partial q_k\partial Q_i}\frac{\partial Q_j}{\partial p_k}=\frac{\partial^2 F}{\partial q_k\partial Q_i}\left(\frac{\partial^2 F}{\partial q\partial Q}\right)^{-1}_{jk}=\left(\frac{\partial^2 F}{\partial q\partial Q}\right)^{-1}_{jk}\left(\frac{\partial^2 F}{\partial q\partial Q}\right)_{ki}=\delta_{ij}\,,$$

where we have applied (I.2.120). It remains to consider

$$\begin{aligned}\{P_i,P_j\}_{p,q} &= \frac{\partial P_i}{\partial p_k}\frac{\partial P_j}{\partial q_k}-\frac{\partial P_i}{\partial q_k}\frac{\partial P_j}{\partial p_k}=\frac{\partial}{\partial p_k}\left(\frac{\partial F}{\partial Q_i}\right)\frac{\partial}{\partial q_k}\left(\frac{\partial F}{\partial Q_j}\right)-\frac{\partial}{\partial q_k}\left(\frac{\partial F}{\partial Q_i}\right)\frac{\partial}{\partial p_k}\left(\frac{\partial F}{\partial Q_j}\right)\\ &= \frac{\partial^2 F}{\partial Q_i\partial Q_m}\frac{\partial Q_m}{\partial p_k}\left[\frac{\partial^2 F}{\partial q_k\partial Q_j}+\frac{\partial^2 F}{\partial Q_j\partial Q_n}\frac{\partial Q_n}{\partial q_k}\right]-\left[\frac{\partial^2 F}{\partial q_k\partial Q_i}+\frac{\partial^2 F}{\partial Q_i\partial Q_m}\frac{\partial Q_m}{\partial q_k}\right]\frac{\partial^2 F}{\partial Q_j\partial Q_n}\frac{\partial Q_n}{\partial p_k}\\ &= \frac{\partial^2 F}{\partial Q_i\partial Q_m}\frac{\partial^2 F}{\partial Q_j\partial Q_n}\left[\frac{\partial Q_m}{\partial p_k}\frac{\partial Q_n}{\partial q_k}-\frac{\partial Q_m}{\partial q_k}\frac{\partial Q_n}{\partial p_k}\right]\\ &\qquad +\frac{\partial^2 F}{\partial Q_i\partial Q_m}\frac{\partial Q_m}{\partial p_k}\frac{\partial^2 F}{\partial q_k\partial Q_j}-\frac{\partial^2 F}{\partial Q_j\partial Q_m}\frac{\partial Q_m}{\partial p_k}\frac{\partial^2 F}{\partial q_k\partial Q_i}\,.\end{aligned}$$

The expression in the brackets is $\{Q_m,Q_n\}_{p,q}$ and, therefore, vanishes. Thus, taking into account (I.2.120), we obtain

$$\begin{aligned}\{P_i,P_j\}_{p,q} &= \frac{\partial^2 F}{\partial Q_i\partial Q_m}\frac{\partial Q_m}{\partial p_k}\frac{\partial^2 F}{\partial q_k\partial Q_j}-\frac{\partial^2 F}{\partial Q_j\partial Q_m}\frac{\partial Q_m}{\partial p_k}\frac{\partial^2 F}{\partial q_k\partial Q_i}\\ &= \frac{\partial^2 F}{\partial Q_i\partial Q_m}\left(\frac{\partial^2 F}{\partial q\partial Q}\right)^{-1}_{mk}\left(\frac{\partial^2 F}{\partial q\partial Q}\right)_{kj}-\frac{\partial^2 F}{\partial Q_j\partial Q_m}\left(\frac{\partial^2 F}{\partial q\partial Q}\right)^{-1}_{mk}\left(\frac{\partial^2 F}{\partial q\partial Q}\right)_{ki}\\ &= \frac{\partial^2 F}{\partial Q_i\partial Q_j}-\frac{\partial^2 F}{\partial Q_j\partial Q_i}=0\,.\end{aligned}$$

The proof for other types of generating functions is analogous. □

Now we demonstrate preservation of Poisson brackets by canonical transformations directly from considering transformations of Hamilton's equations under coordinate changes. Thus, imagine we

perform a change of variables[22] $y^j = f^j(x^k)$, where the vector $\vec{x} = (x^k)$ as in (I.2.91). Then

$$\dot{y}^j = \underbrace{\frac{\partial y^j}{\partial x^k}}_{A^j_k} \dot{x}^k = A^j_k J^{km} \nabla^x_m H = A^j_k J^{km} \frac{\partial y^p}{\partial x^m} \nabla^y_p H'$$

or in the matrix form

$$\dot{y} = AJA^t \cdot \nabla_y H' \,.$$

The new equations for y are Hamiltonian with the new Hamiltonian $H'(y) = H(f^{-1}(y)) = H(x)$, if and only if

$$AJA^t = J \,.$$

Hence, this construction motivates the following definition.

Definition. *Transformations of the phase space which satisfy the condition*

$$AJA^t = J$$

are called canonical.[23]

This is an alternative but equivalent definition of canonical transformations.

Example. Consider transformation (I.2.115) and show that it preserves the Poisson brackets. Since the Poisson brackets are skew-symmetric, the only bracket we have to check is

$$\begin{aligned} \{P, Q\} &= \{-\sin\alpha\, q + \cos\alpha\, p, \cos\alpha\, q + \sin\alpha\, p\} \\ &= -\sin^2\alpha\{q,p\} + \cos^2\alpha\{p,q\} = (\cos^2\alpha + \sin^2\alpha)\{p,q\} = 1 \,. \end{aligned}$$

In many cases to check that a given transformation is canonical it is much simpler to compute the Poisson brackets (differentiation) and check their canonicity than to obtain the corresponding generating function (integration).

[22] For simplicity we restrict our consideration to transformations that do not explicitly depend on time.

[23] In the case when A does not depend on x, the set of all such matrices form a Lie group known as the real symplectic group $\mathrm{Sp}(2n, \mathbb{R})$. The term "symplectic group" was introduced by Herman Weyl. The geometry of the phase space which is invariant under the action of the symplectic group is called *symplectic geometry*.

Part II

Electrodynamics

Chapter 3

Electrostatics

"The problem of finding the solution to any electrostatic problem is equivalent to finding a solution of Laplace's equation throughout the space not occupied by conductors."

Sir James Jeans (1925)

Classical electrodynamics is a theory of electric and magnetic fields caused by distributions of electric charges and currents. This theory is entirely based on Maxwell's equations

$$\begin{aligned} \vec{\nabla}\cdot\vec{H} &= 0\,, & \frac{\partial\vec{H}}{\partial t} &= -c\,\vec{\nabla}\times\vec{E}\,, \\ \vec{\nabla}\cdot\vec{E} &= 4\pi\rho\,, & \frac{\partial\vec{E}}{\partial t} &= c\,\vec{\nabla}\times\vec{H} - 4\pi\vec{j}\,. \end{aligned} \tag{II.3.1}$$

Within the field of electrodynamics, one can study electromagnetic fields under certain static conditions leading to electrostatics (electric fields independent of time) and magnetostatics (magnetic fields independent of time). First, we focus on the laws of electrostatics.

3.1 Laws of electrostatics

Electrostatics studies electric fields produced by static electric charges. Historically, it stems from the discovery of Coulomb's law (1785). This law determines the force that two electrically charged bodies (point charges) exert on each other

$$\vec{F} = k\,q_1 q_2 \frac{\vec{x}_1 - \vec{x}_2}{|\vec{x}_1 - \vec{x}_2|^3}\,. \tag{II.3.2}$$

Here q_1 and q_2 are the magnitudes of charges located at positions $\vec{x}_1$ and $\vec{x}_2$, as presented in Fig. 3.1. The coefficient k is Coulomb's constant and its value, as well as physical dimension depends on the system of units used. More precisely, the situation is as follows.

1. In SI units, the force is measured in *newtons* and charge in *coulombs*. Coulomb's constant is

$$k = \frac{1}{4\pi\epsilon_0}.$$

Here ϵ_0 is the so-called vacuum permittivity and its numerical value is $\epsilon_0 = 8.8542 \cdot 10^{-12} F/m$, where F stands for *farad*, a unit of capacitance being equal to one coulomb per volt.

2. In the Gaussian system of units the vacuum permittivity $\epsilon_0 = \frac{1}{4\pi}$, so that $k = 1$. From Coulomb's law we can deduce the physical dimension of the electric charge in this system, namely,
$$[q] = [F \cdot \ell^2]^{1/2} = [E \cdot \ell]^{1/2} = m^{1/2} \ell^{3/2} t^{-1} = [\hbar \cdot c]^{1/2} ,$$
where the last term expresses the physical dimension of electric charge via dimensions of two fundamental constants, the Planck constant h and the speed of light c. The Gaussian system is more suitable for theoretical studies.

3. The Heaviside-Lorentz system is a modification of the Gaussian system which differs from the latter only by factors of 4π. In particular, in this system $\epsilon_0 = 1$, so that $k = 1/4\pi$.

There are other system of units which we will not discuss here. In the following we adopt the *Gaussian system*. The Coulomb's law then reads

$$\vec{F}_{12} = q_1 q_2 \frac{\vec{x}_1 - \vec{x}_2}{|\vec{x}_1 - \vec{x}_2|^3} , \tag{II.3.3}$$

where $\vec{F}_{12}$ is the force that the second charge q_2 experts on the first charge q_1, *cf.* the general definition of F_{ij} after (I.1.2).

One can introduce the concept of an electric field $\vec{E}$ as the force experienced by a point-like charge q_1 in the limit of vanishing q_1

$$\vec{E}(\vec{x}) = \lim_{q_1 \to 0} \frac{\vec{F}_{12}(\vec{x})}{q_1} = q_2 \frac{\vec{x} - \vec{x}_2}{|\vec{x} - \vec{x}_2|^3} , \qquad \vec{x} = \vec{x}_1 .$$

We have used the limiting procedure to introduce a test charge such that it will only measure the electric field at a certain point and not create its own field. Hence, using Coulomb's law, we obtained an expression for the electric field of a point charge q located at $\vec{x}'$

$$\vec{E}(\vec{x}) = q \frac{\vec{x} - \vec{x}'}{|\vec{x} - \vec{x}'|^3} .$$

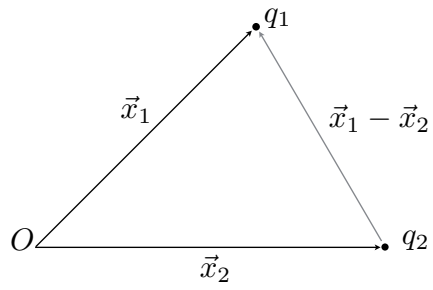

Figure 3.1: Two charges q_1 and q_2 and their respective position vectors $\vec{x}_1$ and $\vec{x}_2$. The charges exert an electric force on one another.

Since $\vec{E}$ is a vector quantity, to find $\vec{E}$ produced by many charges we can apply the superposition principle. Consequently, the field strength will simply be a sum of all of the contributions, which we can write as

$$\vec{E}(\vec{x}) = \sum_{i=1}^{N} q_i \frac{\vec{x} - \vec{x}_i}{|\vec{x} - \vec{x}_i|^3} . \tag{II.3.4}$$

Introducing an electric charge density $\rho(\vec{x})$, the electric field for a continuous distribution of charge is given by

$$\vec{E}(\vec{x}) = \int \rho(\vec{x}') \frac{\vec{x} - \vec{x}'}{|\vec{x} - \vec{x}'|^3} \mathrm{d}^3 x' . \tag{II.3.5}$$

The Dirac delta-function (distribution) allows one to write down the electric charge density which corresponds to local charges

$$\rho(\vec{x}) = \sum_{i=1}^{N} q_i \delta(\vec{x} - \vec{x}_i) . \tag{II.3.6}$$

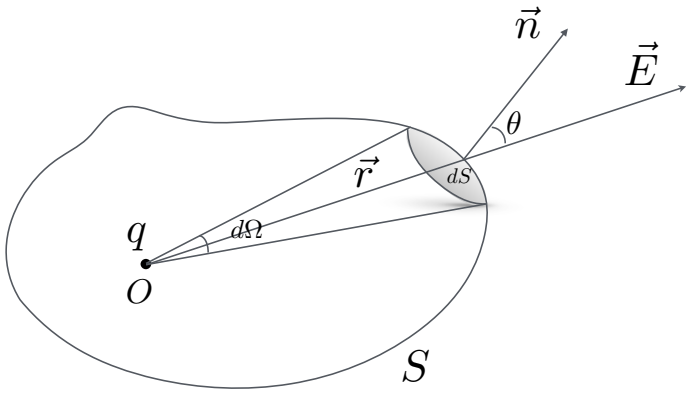

Figure 3.2: The electric flux through a surface, which is proportional to the charge within the surface.

Substituting this formula into eq.(II.3.5), one recovers eq.(II.3.4).

Gauss theorem (integral form). However, eq.(II.3.5) is not very convenient for finding the electric field. For this purpose, one typically turns to another integral relation known as the Gauss theorem, which states that the flux through an arbitrary surface is proportional to the charge contained inside it. Let us consider the flux of $\vec{E}$ through a small region of surface $\mathrm{d}S$, represented graphically in Fig. 3.2,

$$\mathrm{d}N = \big(\vec{E}\cdot\vec{n}\big)\mathrm{d}S = \frac{q}{r^3}\,(\vec{r}\cdot\vec{n})\,\mathrm{d}S = \frac{q}{r^2}\cos(\vec{r},\vec{n})\,\mathrm{d}S = \frac{q}{r^2}\mathrm{d}S'\,,$$

where on the first step we have used that $\vec{E} = q\frac{\vec{r}}{r^3}$. Here $(\vec{r},\vec{n}) \equiv \theta$ is an angle between $\vec{r}$ and the normal $\vec{n}$ to the surface. By the definition of $\mathrm{d}S'$, we observe that it is positive for an angle θ between $\vec{E}$ and $\vec{n}$ less than $\frac{\pi}{2}$, and negative otherwise. Numerically, the product $\cos(\vec{r},\vec{n})\,\mathrm{d}S$ equals to the projection of the area dS on a surface perpendicular to $\vec{r}$. We introduce a solid angle $\mathrm{d}\Omega$

$$\mathrm{d}\Omega = \frac{\mathrm{d}S'}{r^2}\,. \tag{II.3.7}$$

Plugging this relation into eq.(II.3.7) leaves us with the following expression for the flux

$$\mathrm{d}N = q\cdot\mathrm{d}\Omega\,. \tag{II.3.8}$$

By integrating eq.(II.3.8), we obtain the following equation for the flux N

$$\oint_S \big(\vec{E}\cdot\vec{n}\big)\mathrm{d}S = \begin{cases} 4\pi q & \text{if } q \text{ is inside the surface} \\ 0 & \text{otherwise} \end{cases}$$

Equivalently, using the fact that the integral of the charge distribution over volume V is equal to the total charge enclosed in the volume, i.e. $q = \int_V \rho(x)\,\mathrm{d}^3x$, one finds a similar expression

$$\oint_S \big(\vec{E}\cdot\vec{n}\big)\,\mathrm{d}S = 4\pi\int \rho(x)\,\mathrm{d}^3x\,.$$

The above formulae convey the essence of the Gauss theorem.

Gauss theorem. *In an arbitrary electrostatic field the flux of electric field $\vec{E}$ through an arbitrary closed surface equals to the total amount of charge multiplied by 4π inside this surface.*

Gauss theorem (differential form). By using of the Gauss-Ostrogradsky theorem, one may rewrite the surface integral in terms of the volume integral of the divergence of the vector field $\vec{E}$

$$\oint_S (\vec{E}\cdot\vec{n})\mathrm{d}S \;\; = \;\; \int_V \mathrm{div}\,\vec{E}\,(\vec{x})\,\mathrm{d}^3x\,.$$

Recalling that the left hand side is equal to $4\pi q$, a relation between the divergence of the electric field and the charge density arises

$$0 = \int_V \Big[\mathrm{div}\,\vec{E}\,(\vec{x}) - 4\pi\rho\,(\vec{x})\,\Big]\mathrm{d}^3x\,.$$

Since the relation holds for any chosen volume, then the expression inside the integral must equal to zero. The resulting equation is then

$$\mathrm{div}\,\vec{E}\,(\vec{x}) = 4\pi\rho\,(\vec{x})\;.$$

This is known as *the differential form of the Gauss (law) theorem for electrostatics*. This is the first equation from the set of four Maxwell's equations, the latter being the essence of electrodynamics.

Irrotationality of electric field. The Gauss theorem is not enough, however, to determine all the components of $\vec{E}$. A vector field $\vec{A}$ is known if its divergence and its curl, denoted as $\mathrm{div}\vec{A}$ and $\mathrm{rot}\vec{A}$ respectively, are known.[1] Hence, some information is necessary about the curl of electric field. This is in fact given by the second equation of electrostatics

$$\vec{\nabla}\times\vec{E} = 0\,. \tag{II.3.9}$$

The second equation of electrostatics is known as Faraday's law in the absence of time-varying magnetic fields, which are obviously not present in electrostatics (since we required all fields to be time independent).

We will argue about (II.3.9) in the following way. Starting from the definition of the electric field (Coulomb's law) given by equation (II.3.5), we rewrite it in terms of a gradient and pull the differential operator outside of the integral

$$\vec{E}\,(\vec{x}) \;\; = \;\; \int \rho\,(\vec{x}')\,\frac{\vec{x}-\vec{x}'}{|\vec{x}-\vec{x}'|^3}\mathrm{d}^3x' = -\int \rho\,(\vec{x}')\,\vec{\nabla}_x\frac{1}{|\vec{x}-\vec{x}'|}\mathrm{d}^3x' = -\vec{\nabla}\int\frac{\rho\,(\vec{x}')}{|\vec{x}-\vec{x}'|}\mathrm{d}^3x'\,.$$

The curl of the gradient is zero, so that

$$\vec{\nabla}\times\vec{\nabla}\varphi = 0 \quad\;\; \Rightarrow \quad\;\; \vec{\nabla}\times\vec{E} = 0\,.$$

Scalar potential. This derivation shows that the vanishing of $\vec{\nabla}\times\vec{E}$ is not related to the inverse square law. It also shows that the electric field is the minus gradient of a *scalar potential* φ

$$\vec{E} = -\vec{\nabla}\,\varphi\,.$$

From the above, it then follows that this scalar potential is given by

$$\varphi(\vec{x}) = \int\frac{\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}\mathrm{d}^3x'\,, \tag{II.3.10}$$

where the integration is carried out over the entire space. Obviously, the scalar potential is defined up to an additive constant; adding any constant to a given $\varphi(\vec{x})$ does not change the corresponding electric field $\vec{E}$.

[1] This is Helmholtz theorem, see section 7.7.

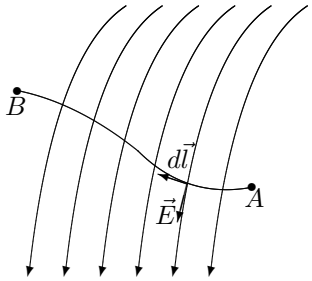

Figure 3.3: The work that has to be done over a charged particle to move it along the path from A to B through an electric field $\vec{E}$.

Before we proceed, we also mention the scalar potential for N point changes. This can be obtained, for instance, by substituting in (II.3.10) the charge density (II.3.6)

$$\varphi(\vec{x}) = \int \sum_{i=1}^{N} \frac{q_i \delta(\vec{x}' - \vec{x}_i)}{|\vec{x} - \vec{x}'|} d^3x' = \sum_{i=1}^{N} \frac{q_i}{|\vec{x} - \vec{x}_i|} \,.$$

What is the physical interpretation of $\varphi(x)$? Consider the work which has to be done to move a test charge along a path from point A to B through an electric field $\vec{E}$

$$W \;=\; -\int_A^B \vec{F} \cdot \mathrm{d}\vec{l} = -q \int_A^B \vec{E} \cdot \mathrm{d}\vec{l}.$$

The minus sign represents the fact that the test charge does work against the electric forces. By associating the electric field as the gradient of a scalar potential, one obtains

$$\begin{aligned} W &= q \int_A^B \vec{\nabla}\varphi \cdot \mathrm{d}\vec{l} = q \int_A^B \frac{\partial \varphi}{\partial x}\mathrm{d}x + \frac{\partial \varphi}{\partial y}\mathrm{d}y + \frac{\partial \varphi}{\partial z}\mathrm{d}z \\ &= q \int_{t_A}^{t_B} \Big(\frac{\partial \varphi}{\partial x}\frac{dx}{dt} + \frac{\partial \varphi}{\partial y}\frac{dy}{dt} + \frac{\partial \varphi}{\partial z}\frac{dz}{dt} \Big) dt = q \int_{t_A}^{t_B} \frac{d\varphi}{dt} dt = q\,(\varphi_B - \varphi_A) \;, \end{aligned}$$

where we have parametrized the path as $(x(t), y(t), z(t))$. The result is just a difference between the potentials at the end points of the path. This implies that the potential energy of a test charge is given by

$$V = q\,\varphi \,.$$

In other words, the potential energy does not depend on the choice of path (hence, the electric force is a conservative force). If a path is chosen such that it is closed, i.e. $A = B$, the integral reduces to zero

$$\oint \vec{E} \cdot \mathrm{d}\vec{l} = 0 \,.$$

This result can also be obtained from Stokes' theorem

$$\oint \big(\vec{E} \cdot \mathrm{d}\vec{l} \big) = \oint_S \mathrm{rot}\, \vec{E} \cdot \mathrm{d}\vec{S} = 0 \,,$$

where we have used the fact that rot $\vec{E} = 0$.

Matching condition. One important application of the Gauss theorem is determination of the electric field of charged surfaces. Analogously to ρ, one can define the surface charge density $\sigma = dq/dS$, where dq is an amount of charge spread over an infinitesimal surface dS. Consider no an arbitrary charged surface S. Choose a normal $\vec{n}$ to S in an arbitrary way and agreed to call 1 and 2 the quantities that refer to internal and external sides (with respect to $\vec{n}$) of the surface.

Theorem. *The normal component of electric has a discontinuity $4\pi\sigma$ passing through a charged surface, independently of the shape of this surface.*

$$E_{2n} - E_{1n} = 4\pi\sigma\,. \qquad \text{(II.3.11)}$$

Here E_{1n} and E_{2n} are normal components of the electric field on the internal and external sides of the surface. Formula (II.3.11) is called *matching condition*. The discontinuity of E_n is explained by the fact that the electric field of surface charges on different sides of the surface has opposite directions: from the surface if it is charged positively and towards the surface if it is charged negatively.

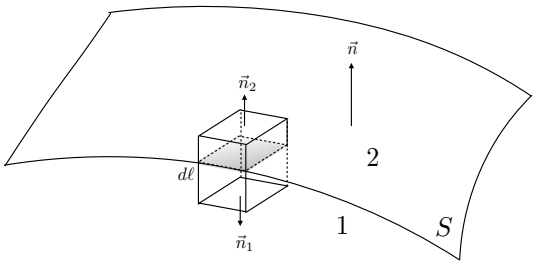

Figure 3.4: Discontinuity of E_n.

We can use the matching condition to immediately determine the electric field on the surface of a conductor. Since inside a conductor $\vec{E} = 0$, and, in particular, $E_{1n} = 0$, the matching condition yields $E_n = 4\pi\sigma$. The electric field is always normal to the surface of the conductor, as follows from electrostatic equilibrium.

Summary. To summarize, we have derived two laws of electrostatics in the differential form

$$\vec{\nabla}\cdot\vec{E} = 4\pi\rho\,, \qquad \text{(II.3.12)}$$

$$\vec{\nabla}\times\vec{E} = 0\,. \qquad \text{(II.3.13)}$$

The main problem of electrostatics is, therefore, giving the charge density ρ, to solve the above equations and find the corresponding electric field. The electrostatics of conductors amounts to determining the electric field in the vacuum outside the conductors and the distribution of charges on their surfaces.

3.2 Laplace and Poisson equations

In the previous section it was shown that the curl of the electric field is equal to zero, thus the field is simply the gradient of some scalar function, which can be written as

$$\vec{\nabla}\times\vec{E} = 0 \quad\Rightarrow\quad \vec{E}\,(\vec{x}) = -\vec{\nabla}\varphi\,(\vec{x})\ .$$

Substituting the right hand side of this expression into equation (II.3.12), we obtain

$$\operatorname{div}\vec{\nabla}\varphi\,(\vec{x}) = -4\pi\rho\,(\vec{x})$$

that is

$$\Delta\varphi\,(\vec{x}) = -4\pi\rho\,(\vec{x})\ , \qquad \text{(II.3.14)}$$

where $\vec{\nabla}^2 \equiv \Delta$ is the Laplace operator. Equation (II.3.14) is known as the *Poisson equation*.

In the case $\rho(\vec{x}) = 0$, i.e. in a region of no charge, the left hand side of (II.3.14) is zero, which is known as the *Laplace equation*. Functions satisfying $\Delta\varphi = 0$ are called *harmonic functions*. The potential solving the Laplace equation has the following remarkable property

Earnshaw's theorem. *The function $\varphi(x, y, z)$ can take maximum and minimum values only at the boundaries of the region where there is a field. As a result, a test charge q introduced into the field cannot be in stable equilibrium, since there is no point at which its potential energy $q\varphi$ would have a minimum.*

The proof is as follows. Let us suppose, for example, that the potential has a maximum at some point A not on the boundary of a region where there is a field. Then the point A can be surrounded by a small closed surface on which the the normal derivative $\vec{n} \cdot \vec{\nabla}\varphi < 0$. Consequently, the integral over this surface

$$\oint_S (\vec{n} \cdot \vec{\nabla}\varphi) dS < 0\,.$$

Using the Gauss-Ostrogradsky theorem, we get an equivalent statement in terms of an integral over the volume V that bounds S

$$\oint_S (\vec{n} \cdot \vec{\nabla}\varphi) dS = \int_V (\vec{\nabla} \cdot \vec{\nabla}\varphi) dV = \int_V \Delta\varphi \, dV < 0\,,$$

giving a contradiction because the last integral vanishes because $\Delta\varphi = 0$.

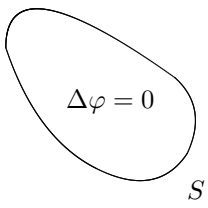

Figure 3.5: The field $\varphi(\vec{x})$, which obeys the Laplace equation, has no maximum or minimum inside a region S.

Another way to understand this remarkable property is to note that for an extremum to exist one needs to have $\frac{\partial\varphi}{\partial x_i} = 0$ together with either $\frac{\partial^2\varphi}{\partial x_i^2} > 0$ or $\frac{\partial^2\varphi}{\partial x_i^2} < 0$ for $\forall i$. The latter condition is impossible to satisfy because $\Delta\varphi = 0$.

Let us now see why and how the scalar potential (II.3.10) satisfies the Poisson equation. Substituting (II.3.10) into (II.3.14), we get

$$\Delta\varphi(\vec{x}) \quad = \quad \Delta \int \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|} \mathrm{d}^3x' = \int \mathrm{d}^3x' \, \rho(\vec{x}') \Delta \frac{1}{|\vec{x} - \vec{x}'|}\,.$$

Without loss of generality we can take $x' = 0$, which is equivalent to choosing the origin of our coordinate system. By switching to spherical coordinates, we can show that

$$\Delta \frac{1}{|\vec{x} - \vec{x}'|} = \Delta \frac{1}{r} = \frac{1}{r} \frac{d^2}{dr^2} \left(r \, \frac{1}{r} \right) = 0\,.$$

This is true everywhere except for $r = 0$, for which the expression above is undetermined. To determine its value at $r = 0$ we can use the following trick. Integrating over volume V, using the Gauss-Ostrogradsky theorem and the fact that $\vec{\nabla} r = \vec{n}$, one obtains

$$\begin{aligned}\int\limits_V \Delta\left(\frac{1}{r}\right) \mathrm{d}^3x &= \int\limits_V \operatorname{div}\vec{\nabla}\left(\frac{1}{r}\right) \mathrm{d}^3x = \oint\limits_S \vec{n}\cdot\vec{\nabla}\frac{1}{r}\,\mathrm{d}S \\ &= \oint\limits_S \vec{n}\cdot\frac{\partial}{\partial r}\left(\frac{1}{r}\right)\vec{n}\,\mathrm{d}S = \oint\limits_S \frac{\partial}{\partial r}\left(\frac{1}{r}\right)\underbrace{r^2\mathrm{d}\Omega}_{\mathrm{d}S} = -4\pi\,.\end{aligned}$$

Therefore,

$$\Delta\,\frac{1}{r} = -4\pi\delta(\vec{x})\ , \tag{II.3.15}$$

or

$$\Delta_x \frac{1}{|\vec{x}-\vec{x}'|} = -4\pi\delta\left(\vec{x}-\vec{x}'\right)\,.$$

Thus, we find

$$\Delta\varphi = \int \rho(\vec{x}')\big(-4\pi\delta(\vec{x}-\vec{x}')\big)\mathrm{d}^3x' = -4\pi\rho(\vec{x})\,.$$

As (II.3.15) shows, we have proved that $\frac{1}{r}$ solves the Poisson equation with the point-like charge source represented by delta-function.

3.3 Solution of the boundary-value problems

If in electrostatics we would always deal with discrete or continuous distributions of charges without any boundary surfaces, then the general expression (where one integrates over all of space)

$$\varphi(\vec{x}) = \int \mathrm{d}^3x' \frac{\rho(\vec{x}')}{|\vec{x}-\vec{x}'|} \tag{II.3.16}$$

would be the most convenient and straightforward solution of the problem. In other words, given some distribution of charge, one can find the corresponding potential and, hence, the electric field $\vec{E} = -\vec{\nabla}\varphi$.

In reality, most of the problems deals with finite regions of space (containing or not containing the charges), on the boundaries of which definite boundary conditions are assumed. These boundary conditions can be created by a specially chosen distribution of charges outside the region in question. In this situation our general formula (II.3.16) cannot be applied with the exception of some particular cases (as in the method of images).

3.3.1 Green's theorems and functions

To handle the boundary conditions it is necessary to invoke the identities or theorems due to George Green.

Green's theorems

Consider an arbitrary vector field[2] $\vec{A}$. We have

$$\int_V \operatorname{div} \vec{A}\, \mathrm{d}^3x = \oint_S \left(\vec{A}\cdot\vec{n}\right)\mathrm{d}S\,. \tag{II.3.17}$$

Let us assume that $\vec{A}$ has the following specific form

$$\vec{A} = \varphi\cdot\vec{\nabla}\psi\,,$$

where ψ and φ are arbitrary functions. Then

$$\operatorname{div}\vec{A} \;=\; \operatorname{div}\left(\varphi\cdot\vec{\nabla}\psi\right) = \frac{\partial}{\partial x^i}\left(\varphi\frac{\partial\psi}{\partial x_i}\right) = \vec{\nabla}\varphi\cdot\vec{\nabla}\psi + \varphi\Delta\psi\,.$$

Substituting this back into eq.(II.3.17), we get

$$\int_V \left(\vec{\nabla}\varphi\cdot\vec{\nabla}\psi + \varphi\Delta\psi\right)\mathrm{d}^3x = \oint_S \varphi\cdot\left(\vec{\nabla}\psi\cdot\vec{n}\right)\mathrm{d}S = \oint_S \varphi\left(\frac{\partial\psi}{\partial n}\right)\mathrm{d}S\,, \tag{II.3.18}$$

where $\partial/\partial n$ is the normal derivative to the surface S. Equation (II.3.18) is known as the *first Green formula*. When we interchange φ for ψ in the above expression and take a difference of these two we obtain *the second Green formula*

$$\int_V \left(\varphi\Delta\psi - \psi\Delta\varphi\right)\mathrm{d}^3x = \oint_S \left(\varphi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\varphi}{\partial n}\right)\mathrm{d}S\,. \tag{II.3.19}$$

By using this formula, the differential Poisson equation can be reduced to an integral equation. Indeed, consider a function ψ such that

$$\psi \equiv \frac{1}{R} = \frac{1}{|\vec{x}-\vec{x}'|} \quad\Rightarrow\quad \Delta\psi = -4\pi\delta\left(\vec{x}-\vec{x}'\right)\,. \tag{II.3.20}$$

Substituting it into the second Green formula (II.3.19) and assuming x is inside the space V integrated over, one gets

$$\int_V \left(-4\pi\varphi(\vec{x}')\delta\left(\vec{x}-\vec{x}'\right) + \frac{4\pi\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}\right)\mathrm{d}^3x' = \oint_{S'}\left[\varphi\frac{\partial}{\partial n'}\left(\frac{1}{R}\right) - \frac{1}{R}\frac{\partial\varphi}{\partial n'}\right]dS'\,.$$

Here we have chosen $\varphi\left(\vec{x}'\right)$ to satisfy the Poisson equation $\Delta\varphi\left(\vec{x}'\right) = -4\pi\rho\left(\vec{x}'\right)$. Taking into account that $\int_V \varphi\left(\vec{x}'\right)\delta\left(\vec{x}-\vec{x}'\right) = \varphi\left(\vec{x}\right)$, the expression above allows one to express $\varphi(\vec{x})$ as

$$\varphi\left(\vec{x}\right) = \int_V \frac{\rho\left(\vec{x}'\right)}{R}\mathrm{d}^3x' + \frac{1}{4\pi}\oint_S\left[\frac{1}{R}\frac{\partial\varphi}{\partial n'} - \varphi\frac{\partial}{\partial n'}\left(\frac{1}{R}\right)\right]\mathrm{d}S'\,, \tag{II.3.21}$$

which is the general solution for the scalar potential. The terms inside the integrals are equal to zero if x lies outside of V.

Consider the following two special cases:

1) If S goes to ∞ and the electric field vanishes on it faster than $\frac{1}{R}$, then the surface integral turns to zero and $\varphi(\vec{x})$ turns into our general solution given by eq.(II.3.16).

[2]Now introduced for mathematical convenience, but it will later prove to be of greater importance.

2) For a volume which does not contain charges, the potential at any point (which gives a solution of the Laplace equation) is expressed in terms of the potential and its normal derivative on the surface enclosing the volume. This result, however, does not give a solution of the boundary problem, rather it represents an integral equation, because given φ and $\frac{\partial\varphi}{\partial n}$ (Cauchy boundary conditions) we overdetermined the problem.

Therefore, the question arises which boundary conditions should be imposed to guarantee a unique solution to the Laplace and Poisson equations. Experience shows that given a potential on a closed surface uniquely defines the potential inside (e.g. a system of conductors on which one maintains different potentials). Giving a potential on a closed surface corresponds to the *Dirichlet boundary conditions*.

Analogously, given an electric field (i.e. normal derivative of a potential) or likewise the surface charge distribution ($E \sim 4\pi\sigma$) also defines a unique solution. These are the *Neumann boundary conditions*[3].

Uniqueness of solution with Dirichlet ot Neumann boundary conditions

One can prove, with the help of the first Green formula, that the Poisson equation

$$\Delta\varphi = -4\pi\rho\,,$$

in a volume V has a unique solution under the Dirichlet or the Neumann conditions given on a surface S enclosing V. To do so, assume there exist two different solutions φ_1 and φ_2 which both have the same boundary conditions. Consider

$$U = \varphi_2 - \varphi_1\,.$$

It solves $\Delta U = 0$ inside V and has either $U = 0$ on S (Dirichlet) or $\frac{\partial U}{\partial n} = 0$ on S (Neumann). In the first Green formula one plugs $\varphi = \psi = U$, so that

$$\int\limits_V \left(\left|\vec{\nabla} U\right|^2 + U\Delta U \right) \mathrm{d}^3x = \oint\limits_S U \left(\frac{\partial U}{\partial n}\right) \mathrm{d}S\,. \tag{II.3.22}$$

Here the second term in the integral vanishes as $\Delta U = 0$ by virtue of being the solution to the Laplace equation and the right hand side is equal to zero, since we have assumed that the value of the potential (Dirichlet) or its derivative (Neumann) vanish at the boundary. Therefore,

$$\begin{aligned} \int_V |\vec{\nabla} U|^2 = 0 \quad &\longrightarrow \quad |\vec{\nabla} U| = 0 \\ &\longrightarrow \quad \vec{\nabla} U = 0 \end{aligned} \tag{II.3.23}$$

Thus, inside V the function U is constant everywhere. For Dirichlet boundary conditions $U = 0$ on the boundary and so it is zero uniformly, such that $\varphi_1 = \varphi_2$ everywhere, i.e. there is only one solution. Similarly, the solution under Neumann boundary conditions is also unique up to unessential boundary terms.

Method of Green's functions

This method is used to find solutions of many second order differential equations and provides a formal solution to the boundary-value problems. The method is based on an impulse from a source,

[3]Note that both Dirichlet as well as Neumann boundary conditions are not only limited to electrodynamics, but are more general and appear in the theory of ordinary or partial differential equations.

which is later integrated with the source function over entire space. Recall

$$\Delta \frac{1}{|\vec{x}-\vec{x}'|} = -4\pi\delta\left(\vec{x}-\vec{x}'\right) . \tag{II.3.24}$$

However, the function $\frac{1}{|\vec{x}-\vec{x}'|}$ is just one of many functions which obeys $\Delta\psi = -4\pi\delta\left(\vec{x}-\vec{x}'\right)$. The functions that are solutions of this second order differential equation are known as *Green's functions*. In general,

$$\Delta G\left(\vec{x},\vec{x}'\right) = -4\pi\delta\left(\vec{x}-\vec{x}'\right) , \tag{II.3.25}$$

where

$$G\left(\vec{x},\vec{x}'\right) = \frac{1}{|\vec{x}-\vec{x}'|} + F\left(\vec{x},\vec{x}'\right) ,$$

so that $\Delta F\left(\vec{x},\vec{x}'\right) = 0$, *i.e.* it obeys the Laplace equation inside V.

The point is now to find such $F\left(\vec{x},\vec{x}'\right)$, that gets rid of one of the terms in the integral equation (II.3.21) we had for $\varphi\left(\vec{x}\right)$. Letting $\varphi = \varphi\left(\vec{x}\right)$ and $\psi = G\left(\vec{x},\vec{x}'\right)$, we then get

$$\varphi\left(\vec{x}\right) = \int_V \rho\left(\vec{x}'\right) G\left(\vec{x},\vec{x}'\right) \mathrm{d}^3x' + \frac{1}{4\pi}\oint_S \left[G\left(\vec{x},\vec{x}'\right)\frac{\partial\varphi\left(\vec{x}'\right)}{\partial n'} - \varphi\left(\vec{x}'\right)\frac{\partial G\left(\vec{x},\vec{x}'\right)}{\partial n'}\right]\mathrm{d}S' .$$

By using the arbitrariness in the definition of the Green function we can leave in the surface integral a chosen type of boundary conditions.

1) For the Dirichlet case we can choose $G_D\left(\vec{x},\vec{x}'\right) = 0$ when $\vec{x}' \in S$, then $\varphi(\vec{x})$ simplifies to

$$\varphi\left(\vec{x}\right) = \int_V \rho\left(\vec{x}'\right) G_D\left(\vec{x},\vec{x}'\right) \mathrm{d}^3x' - \frac{1}{4\pi}\oint_S \varphi\left(\vec{x}'\right)\frac{\partial G_D\left(\vec{x},\vec{x}'\right)}{\partial n'}\mathrm{d}S' ,$$

where $G_D\left(\vec{x},\vec{x}'\right)$ is referred to as the *bulk-to-bulk propagator* and $\frac{\partial G_D\left(\vec{x},\vec{x}'\right)}{\partial n'}$ is the *bulk-to-boundary propagator*.

2) For the Neumann case we could try to choose $\frac{\partial G_N\left(\vec{x},\vec{x}'\right)}{\partial n'} = 0$ when $\vec{x}' \in S$. However, one has

$$\begin{aligned}\oint_S \frac{\partial G_N\left(\vec{x},\vec{x}'\right)}{\partial n'}\mathrm{d}S' &= \oint_S \left(\vec{\nabla}G_N\cdot\vec{n}\right)\mathrm{d}S' = \int_V \mathrm{div}\vec{\nabla}G_N\,\mathrm{d}^3x' = \int_V \Delta G_N\,\mathrm{d}^3x' \\ &= -4\pi\int\delta(\vec{x}-\vec{x}')\,\mathrm{d}^3x' = -4\pi .\end{aligned}$$

For this reason we can not demand $\frac{\partial G_N\left(\vec{x},\vec{x}'\right)}{\partial n'} = 0$. Instead, one chooses another simple condition $\frac{\partial G_N\left(\vec{x},\vec{x}'\right)}{\partial n'} = -\frac{4\pi}{S}$, where S is the total surface area, and the left hand side of the equation is referred to as the Neumann Green function. Using this condition, we get

$$\varphi\left(\vec{x}\right) = \int_V \rho\left(\vec{x}'\right) G_N\left(x,x'\right)\mathrm{d}^3x' + \frac{1}{4\pi}\oint_S G_N\left(\vec{x},\vec{x}'\right)\frac{\partial\varphi\left(\vec{x}'\right)}{\partial n'}\mathrm{d}S' + \frac{1}{S}\oint_S \varphi\left(\vec{x}'\right)\mathrm{d}S' .$$

The last term represents $\langle\varphi\rangle$, the averaged value of the potential on S. If S is infinite this term vanishes. In any case, the extra term $\frac{1}{S}\oint_S \varphi\left(\vec{x}'\right)\mathrm{d}S'$ is just a constant (does not depend on x) and, therefore, does not contribute to the electric field $\vec{E} = -\vec{\nabla}\varphi$.

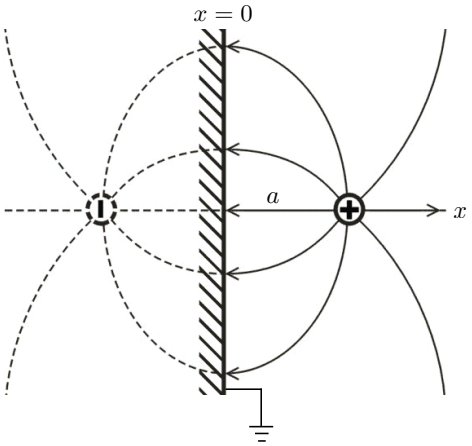

Figure 3.6: Charge placed in front of a metallic grounded plate. An image charge is put in the left half-space to emulate the boundary condition $\varphi = 0$ all along the plate.

3.3.2 Method of images

The method of images concerns itself with the problem of one or more point charges in the presence of boundary surfaces, for example, conductors either grounded or held at fixed potentials. Under favorable conditions it is possible to infer from the geometry of the situation that a small number of suitably placed charges of appropriate magnitude , external to the region of interest, can simulate the required boundary conditions. These charges are called *image changes* and the replacement of actual problem with boundaries by an enlarged region with image charges but no boundaries is called the *method of images*.

Example. Consider a point charge q located at a distance $\vec{x}_1 = a\vec{e}_x$ of an infinite grounded metallic plate. In the absence of the plate the scalar potential would be

$$\varphi = \frac{q}{|\vec{x} - \vec{x}_1|} = \frac{q}{\sqrt{y^2 + z^2 + (x-a)^2}}\,.$$

Although it satisfies Laplace's equation, this solution cannot persist in the presence of the plate. Since the plate is grounded, one should have $\varphi = 0$ all along the plate, which is not the case for the solution above. The problem can be cured by adding an image charge of opposite value precisely on the mirror side of the plate. Since the potentials are additive, the potential by the charge and its mirror is

$$\varphi = \frac{q}{|\vec{x} - \vec{x}_1|} - \frac{q}{|\vec{x} + \vec{x}_1|} = \frac{q}{\sqrt{y^2 + z^2 + (x-a)^2}} - \frac{q}{\sqrt{y^2 + z^2 + (x+a)^2}}, \quad x \geq 0\,. \qquad \text{(II.3.26)}$$

We should consider this potential above only in the region $x \geq 0$. Obviously $\varphi = 0$ along the plate and the boundary condition is satisfied. The solution we found satisfies the same Poisson equation in the right half-space. Indeed,

$$\Delta\varphi = -4\pi q\,\delta(\vec{x} - \vec{x}_1) + \cancel{4\pi q\,\delta(\vec{x} + \vec{x}_1)}\,.$$

The second 3d-delta-function never contributes in the region $x \geq 0$, as it contains the factor $\delta(x+a)$ which is always zero in this region. Thus, (II.3.26) is the desired solution – it solves the Poisson equation of the charge q in the right half-space and vanishes all along the plate $x = 0$.

We can now compute the other physical quantities, in first turn the the electric field

$$\vec{E} = -\vec{\nabla}\varphi = q\left[\frac{\vec{x}-\vec{x}_1}{|\vec{x}-\vec{x}_1|^3} - \frac{\vec{x}+\vec{x}_1}{|\vec{x}+\vec{x}_1|^3}\right].$$

Knowing $\vec{E}$, it is straightforward to find the surface charge distribution σ, namely,

$$\sigma = \frac{1}{4\pi}\vec{E}\cdot\vec{n}\Big|_{x=0} = \frac{q}{4\pi}\left[\frac{x-a}{((x-a)^2+y^2+z^2)^{3/2}} - \frac{x+a}{((x+a)^2+y^2+z^2)^{3/2}}\right]_{x=0},$$

so that

$$\sigma = -\frac{q}{2\pi}\frac{a}{(a^2+y^2+z^2)^{3/2}}.$$

This allows us to determine the total (induced) charge of the plate

$$Q = \iint_{-\infty}^{+\infty}\sigma dydz = -\frac{qa}{2\pi}2\pi\int_0^\infty \frac{rdr}{(a^2+r^2)^{3/2}} = -q.$$

Another example of application of the method of images will be considered in Tutorial X.

3.3.3 Separation of variables

Many problems of potential theory have a symmetry that makes a particular coordinate system most natural for its description. In an orthogonal coordinate system where (u, v, w) labels a point in space, the method of *separation of variables* assumes a product solution of the form

$$\varphi(u, v, w) = A(u)B(v)C(w).$$

This ansatz separates Laplace's partial differential equation into three ordinary, second-order differential equation in 13 different coordinate systems. The separation process generates three undetermined "separation" constants (not all independent) which serve as parameters in differential equations and label their solutions. If α is a sole separation constant in the differential equation for $A(u)$, the most general solution for the latter is an arbitrary sum of two linearly independent solutions $a_\alpha^{(1)}A_\alpha^{(1)}(u) + a_\alpha^{(2)}A_\alpha^{(2)}(u)$ and similarly for the separation constants β and γ governing the differential equations for $B(v)$ and $C(w)$. The most general solution of Laplace's equation is a sum over all possible values of separation constants of products of the corresponding solutions

$$\varphi(u,v,w) = \sum_{\alpha\beta\gamma}[a_\alpha^{(1)}A_\alpha^{(1)}(u) + a_\alpha^{(2)}A_\alpha^{(2)}(u)][b_\beta^{(1)}B_\beta^{(1)}(v) + b_\beta^{(2)}B_\beta^{(2)}(v)][c_\gamma^{(1)}C_\gamma^{(1)}(w) + c_\gamma^{(2)}C_\gamma^{(2)}(w)].$$

The separation constants and coefficients must be chosen in such a way as to satisfy the boundary conditions and the symmetry of the problem, as well as to retain the potential finite through the solution volume.

Electrostatic problems with cartesian symmetry

For potential problems with natural rectangular boundaries, the trial solution

$$\varphi = X(x)Y(y)Z(z)$$

converts Laplace's equation

$$\Delta\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2}$$

into

$$\frac{X''(x)}{X(x)} + \frac{Y''(y)}{Y(y)} + \frac{Z''(z)}{Z(z)} = 0\,.$$

For this to be true one has to have

$$X'' = \alpha^2 X\,, \qquad Y'' = \beta^2 Y\,, \qquad Z'' = \gamma^2 Z$$

with real separation constants α^2, β^2 and γ^2 satisfying

$$\alpha^2 + \beta^2 + \gamma^2 = 0\,. \tag{II.3.27}$$

Elementary solutions are

$$\begin{aligned} X_\alpha(x) &= \begin{cases} A_0 + B_0 x, & \alpha = 0\,, \\ A_\alpha e^{\alpha x} + B_\alpha e^{-\alpha x}\,, & \alpha \neq 0\,, \end{cases} \\ Y_\beta(y) &= \begin{cases} C_0 + B_0 y, & \beta = 0\,, \\ C_\beta e^{\beta y} + D_\beta e^{-\beta y}\,, & \beta \neq 0\,, \end{cases} \\ Z_\gamma(z) &= \begin{cases} E_0 + F_0 z, & \gamma = 0\,, \\ E_\gamma e^{\gamma z} + F_\gamma e^{-\gamma z}\,, & \gamma \neq 0\,. \end{cases} \end{aligned}$$

The linearity of Laplace's equation allows one to superpose the products of elementary solutions and produce a general solution in the form

$$\varphi(x,y,z) = \sum_\alpha \sum_\beta \sum_\gamma X_\alpha(x) Y_\beta(y) Z_\gamma(z) \delta(\alpha^2 + \beta^2 + \gamma^2)\,. \tag{II.3.28}$$

Example. Let us use the above formula to find the potential inside the rectangular box shown in Fig. 3.7. We assume that all the walls are fixed at zero potential except for the $z = 0$ wall, where the potential takes a given value $V(x,y)$. The homogeneous boundary conditions on the vertical sides of the wall can be easily satisfied if we take α and β to be purely imaginary, *i.e.* $\alpha \to i\alpha$ and $\beta \to i\beta$. T We can then choose the new $X_\alpha(x)$ and $Y_\beta(y)$ to be sine functions which vanish at $x = a$ and $y = b$, respectively. Having in mind that constraint (II.3.27) gives $\gamma^2 = \alpha^2 + \beta^2$, the general ansatz (II.3.28) boils down to

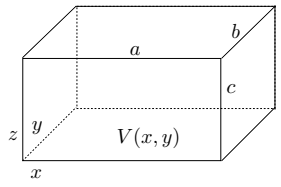

Figure 3.7: Empty box with the $z = 0$ bottom wall maintained at non-zero potential $V(x,y)$.

$$\varphi(x,y,z) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \sin\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \left[E_{mn} e^{\gamma_{mn} z} + F_{mn} e^{-\gamma_{mn} z}\right],$$

where

$$\gamma_{mn}^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2\,.$$

The next task is to choose E_{mn} and F_{mn} so the potential vanishes at $z = c$. If V_{mn} are coefficients still to be determined, a convenient way to write the result is

$$\varphi(x,y,z) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} V_{mn} \sin\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \frac{\sinh\gamma_{mn}(c-z)}{\sinh\gamma_{mn} c}\,.$$

It remains only to impose the final boundary condition that $\varphi(x, y, 0) = V(x, y)$. This gives

$$V(x, y) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} V_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} ,$$

which is a double Fourier sine representation of $V(x, y)$. To find the coefficients V_{mn}, multiply both sides of the last formula by $\sin(m'\pi x/a) \sin(n'\pi y/b)$ and integrate over the intervals $0 \leq x \leq a$ and $0 \leq y \leq b$, and use the orthogonality relation

$$\int_0^{\pi} du \sin(mu) \sin(nu) = \frac{\pi}{2} \delta_{mn} .$$

This yields

$$V_{mn} = \frac{4}{ab} \int_0^a \int_0^b dxdy V(x, y) \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} .$$

This completes the solution of the problem. Another problem on separation of variables in cartesian coordinates will be considered in Tutorial X.

The example raises the question of how to arrange a complete set of eigenvalues if we had specified non-zero potentials on any (or all) of the vertical walls. The solution is to superpose the separated-valuable solutions to several independent potential problems, each like the one in the present example but with a different wall held at a non-zero potential. This general approach works for other coordinate systems also.

Electrostatic problems with spherical symmetry

Frequently, when dealing with electrostatics, one encounters the problems exhibiting spherical symmetry. As an example, take Coulomb's law (II.3.2), which depends on the radial distance only and has no angular dependence. When encountering a symmetry of that sort, one often chooses a set of convenient coordinates which greatly simplifies the corresponding problem.

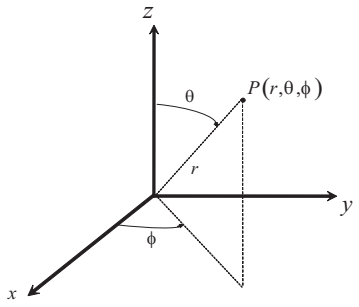

Figure 3.8: Spherical coordinate system.

It is no surprise that in this case, we will be making use of spherical coordinates, which in terms of the Cartesian coordinates, are given by

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} , \\ \theta &= \arccos \left(\frac{z}{\sqrt{x^2 + y^2 + z^2}} \right) , \end{aligned} \qquad \text{(II.3.29)}$$

$$\phi \;=\; \arctan\left(\frac{y}{x}\right),$$

To obtain the Cartesian coordinates from the spherical ones, we use

$$\begin{aligned} x &= r\sin\theta\cos\phi\,, \\ y &= r\sin\theta\sin\phi\,, \\ z &= r\cos\theta\,. \end{aligned} \tag{II.3.30}$$

In spherical coordinates Laplace's operator looks as

$$\Delta = \frac{1}{r^2}\left(\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\left(\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\,.$$

Hence, in these coordinates the Laplace equation reads as

$$\Delta\varphi = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\varphi) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\varphi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\varphi}{\partial\phi^2} = 0\,.$$

The separation ansatz is

$$\varphi(r,\theta,\phi) = \frac{U(r)}{r}P(\theta)\,Q(\phi)\,.$$

Upon substituting this into the Laplace equation and multiplying both sides by $\frac{r^3\sin^2\theta}{U(r)P(\theta)Q(\phi)}$, one obtains

$$r^2\sin^2\theta\left[\left(\frac{1}{U}\frac{\partial^2 U}{\partial r^2}\right) + \frac{1}{r^2\sin\theta P}\left(\frac{\partial}{\partial\theta}\sin\theta\frac{\partial P}{\partial\theta}\right)\right] + \frac{1}{Q}\frac{\partial^2 Q}{\partial\phi^2} = 0\,.$$

Since we only have ϕ dependence in the last term we can state that, since there are no other terms with ϕ, then this term has to be constant (chosen here for convenience with anticipation of the solution)

$$\frac{1}{Q}\frac{\partial^2 Q}{\partial\phi^2} = -m^2\,.$$

Hence the solution is $Q_m = e^{\pm im\phi}$, where m is an integer such that Q_m is single valued. This leaves us with two separated equations. For $P(\theta)$ the equation simplifies to

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\sin\theta\frac{\mathrm{d}P}{\mathrm{d}\theta}\right) + \left[l(l+1) - \frac{m^2}{\sin^2\theta}\right]P = 0\,,$$

and for $U(r)$ one obtains

$$\frac{\mathrm{d}^2U}{\mathrm{d}r^2} - \frac{l(l+1)}{r^2}U = 0\,,$$

where we have just again conveniently picked $l(l+1)$ to be the integration constant such that in our solution it will appear in a convenient form. It is easy to verify that the solution to the equation for $U(r)$ is given by

$$U(r) = Ar^{l+1} + Br^{-l}\,,$$

where l is assumed to be non-negative and A and B are arbitrary constants. The combination $U(r)/r$ entering the potential is then

$$\frac{U(r)}{r} = Ar^l + Br^{-(l+1)}\,.$$

Legendre polynomials. The second equation, on the other hand, is a bit more complicated and upon substitution $\cos\theta = x$ it transforms into

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\left(1-x^2\right)\frac{\mathrm{d}P}{\mathrm{d}x}\right] + \left[l(l+1) - \frac{m^2}{1-x^2}\right]P = 0\,,$$

which one can recognize as the so-called *generalized Legendre equation*. Its solutions are the associated Legendre functions. For $m^2 = 0$, we obtain the Legendre equation

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\frac{\mathrm{d}P}{\mathrm{d}x}\right] + l(l+1)P = 0\,. \tag{II.3.31}$$

The solutions to this equation are referred to as *the Legendre polynomials*. In order for our solution to have physical meaning, it must be finite and continuous on the interval $-1 \le x \le 1$. We try as a solution the following power series (*the Frobenius method*)

$$P\left(x\right) = x^{\alpha}\sum_{j=0}^{\infty} a_j x^j\,, \tag{II.3.32}$$

where α is unknown. Substituting our trial solution (II.3.32) into the Legendre equation (II.3.31), we obtain

$$\sum_{j=0}^{\infty}\left[\left(\alpha+j\right)\left(\alpha+j-1\right)a_j x^{\alpha+j-2} - \left[\left(\alpha+j\right)\left(\alpha+j+1\right) - l\left(l+1\right)\right]a_j x^{\alpha+j}\right] = 0\,.$$

For $j = 0$ and $j = 1$, the first term will have $x^{\alpha-2}$ and $x^{\alpha-1}$ and the second term will have x^{α} and $x^{\alpha+1}$ respectively, which will never make the left hand side to vanish unless

- $a_0 \neq 0$, then $\alpha\left(\alpha-1\right) = 0$ so that (A) $\alpha = 0$ or $\alpha = 1$
- $a_1 \neq 0$, then $\alpha\left(\alpha+1\right) = 0$ so that (B) $\alpha = 0$ or $\alpha = -1$

For other j, one obtains a recurrence relation

$$a_{j+2} = \frac{\left(\alpha+j\right)\left(\alpha+j+1\right) - l\left(l+1\right)}{\left(\alpha+j+1\right)\left(\alpha+j+2\right)}a_j\,.$$

Cases (A) and (B) are actually equivalent. We will consider case (A) for which $\alpha = 0$ or 1. The expansion contains only even powers of x for $\alpha = 0$ and only odd powers of x for $\alpha = 1$. We note two properties of this series:

1. The series is convergent for $x^2 < 1$ for any l.
2. The series is divergent at $x = \pm 1$ unless it is truncated.

It is obvious from the recurrent formula that the series is truncated in the case that l is a non-negative integer. The corresponding polynomials are normalized in such a way that they are all equal to identity at $x = 1$. These are the Legendre polynomials $P_l(x)$:

$$\begin{aligned}
P_0\left(x\right) &= 1\,;\\
P_1\left(x\right) &= x\,;\\
P_2\left(x\right) &= \frac{1}{2}\left(3x^2-1\right)\,;\\
P_3\left(x\right) &= \frac{1}{3}\left(5x^3-2x\right)\,;
\end{aligned}$$

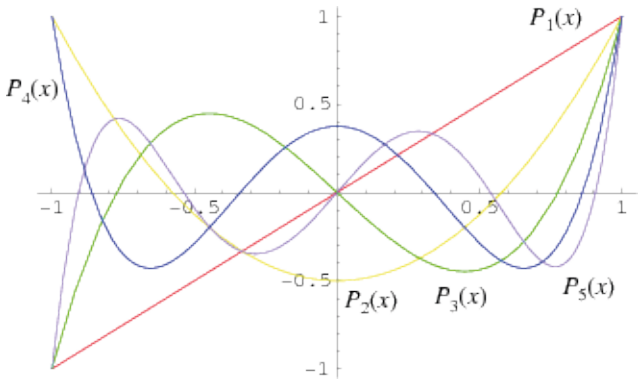

Figure 3.9: Profiles of a few Legendre polynomials.

$$\begin{aligned} &\cdots \\ P_l(x) &= \frac{1}{2^l l!}\frac{d^l}{dx^l}\left(x^2-1\right)^l . \end{aligned}$$

The general expression given in the last line is also known as *the Rodrigues formula.*

The Legendre polynomials form a complete system of orthogonal functions on $-1 \le x \le 1$. To check whether they are indeed orthogonal, one takes the differential equation for P_l, multiplies it by $P_{l'}$, and then integrates

$$\int_{-1}^{1} P_{l'}\left[\frac{\mathrm{d}}{\mathrm{d}x}(1-x^2)\frac{\mathrm{d}P_l}{\mathrm{d}x} + l(l+1)P_l\right]\mathrm{d}x = 0\,,$$

or

$$\int_{-1}^{1}\left[(x^2-1)\frac{\mathrm{d}P_{l'}}{\mathrm{d}x}\frac{\mathrm{d}P_l}{\mathrm{d}x} + l(l+1)P_{l'}P_l)\right]\mathrm{d}x = 0\,.$$

Now subtract the same equation, but with the interchange of l and l', such that the following expression is left

$$\left[l'(l'+1) - l(l+1)\right]\int_{-1}^{1} P_{l'}P_l = 0\,.$$

The equation above shows that for $l \neq l'$ the polynomials are orthogonal

$$\int_{-1}^{1} P_{l'}P_l = 0\,.$$

By using the Rodrigues formula, one can get the following formula

$$\int_{-1}^{1} P_{l'}(x)P_l(x)\mathrm{d}x = \frac{2}{2l+1}\delta_{l',l}\,. \qquad \text{(II.3.33)}$$

Let us prove this formula. Since we have already proved that Legendre polynomials with different l are orthogonal, to prove (II.3.33), we can put $l = l'$

$$\int_{-1}^{1} P_l(x)P_l(x)\mathrm{d}x = \frac{1}{(2^l l!)^2}\int_{-1}^{1}\mathrm{d}x\left[\frac{d^l}{dx^l}\left(x^2-1\right)^l\right]\left[\frac{d^l}{dx^l}\left(x^2-1\right)^l\right],$$

where we used the Rodrigues formula. Now we integrate all l derivatives in the first bracket by parts. All arising boundary terms vanish because

$$\frac{d^k}{dx^k}(x^2-1)^l$$

vanish for $k<l$ at both $x=1$ and $x=-1$. Hence, we obtain

$$\int_{-1}^{1} P_l(x)P_l(x)\mathrm{d}x = \frac{(-1)^l}{(2^l l!)^2}\int_{-1}^{1}\mathrm{d}x\left(x^2-1\right)^l\left[\frac{d^{2l}}{dx^{2l}}\left(x^2-1\right)^l\right].$$

By using, for instance, the binomial expansion of $(x^2-1)^l$, one can see that

$$\frac{d^{2l}}{dx^{2l}}\left(x^2-1\right)^l = \frac{d^{2l}}{dx^{2l}}x^{2l} = (2l)!\,,$$

and, as a result, we are left with

$$\int_{-1}^{1} P_l(x)P_l(x)\mathrm{d}x = \frac{(-1)^l(2l)!}{(2^l l!)^2}\int_{-1}^{1}\mathrm{d}x\left(x^2-1\right)^l = \frac{(2l)!}{(2^{2l}l!)^2}\int_0^{\pi}\mathrm{d}\theta\sin^{2l+1}\theta = \frac{(2l)!}{(2^{2l}l!)^2}\frac{\sqrt{\pi}\Gamma(l+1)}{\Gamma(l+3/2)}\,,$$

where a substitution $x=\cos\theta$ was made and we have used a formula

$$\int_0^{\pi}\mathrm{d}\theta\sin^k\theta = \sqrt{\pi}\frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k+2}{2})}\,,$$

where $\Gamma(z)$ is the so-called gamma-function, its integral representation is given by (II.3.34). Taking into account that $\Gamma(l+1)=l!$ and

$$\Gamma(l+3/2) = (l+1/2)\Gamma(l+1/2) = (l+1/2)\sqrt{\pi}\frac{(2l)!}{4^l l!}\,,$$

we find

$$\int_{-1}^{1} P_l(x)P_l(x)\mathrm{d}x = \frac{(2l)!}{(2^l l!)^2}\frac{\sqrt{\pi}\,l!\times 4^l l!}{(l+1/2)\sqrt{\pi}(2l)!} = \frac{1}{l+1/2} = \frac{2}{2l+1}\,,$$

which completes the proof.

For any function defined on $-1\le x\le 1$

$$f(x) \quad = \sum_{l=0}^{\infty} A_l P_l(x)\,,$$

where the coefficients A_l are found by using the orthogonality relation (II.3.33) for Legendre polynomials

$$A_l \quad = \frac{2l+1}{2}\int_{-1}^{1} f(x)P_l(x)\mathrm{d}x\,.$$

Note that this expansion and its coefficients is not different to any other set of orthogonal functions in the function space. In situations where there is an azimuthal symmetry, one can take $m=0$. Thus,

$$\varphi(r,\theta) = \sum_{l=0}^{\infty}\left(A_l r^l + B_l r^{-(l+1)}\right)P_l(\cos\theta)\,.$$

If charge is absent anywhere in the vicinity of the coordinate system, one should take $B_l=0$.

Example. Determine a potential inside the sphere of radius R if on its surface the potential $V(\theta)$ is held. We look for the potential in the form

$$\varphi(r,\theta) = \sum_{l=0}^{\infty} A_l r^l P_l(\cos\theta)$$

that is regular solution of $\Delta\varphi = 0$ inside the sphere. On its surface the potential is

$$V(\theta) \;=\; \sum_{l=0}^{\infty} A_l R^l P_l(\cos\theta)$$

so that

$$A_l \;=\; \frac{2l+1}{2R^l}\int_0^{\pi} V(\theta) P_l(\cos\theta)\sin\theta \mathrm{d}\theta\,.$$

There is a non-trivial electric field inside the sphere $\vec{E} = -\vec{\nabla}\varphi$. This electric field is divergenceless $\vec{\nabla}\cdot\vec{E} = -\Delta\varphi = 0$ because there is no charge inside. In particular, if $V(\theta) = 0$ (this is the case when the sphere is grounded), then all $A_l = 0$ so that $\varphi = 0$ and $\vec{E} = 0$ everywhere inside the sphere.

Example. Let us find the potential of an empty sphere of radius $r = R$ which has two semi-spheres with separate potentials $V(\theta)$, such that the potential is equal to V for $0 \le \theta < \frac{\pi}{2}$ and equal to $-V$ for $\frac{\pi}{2} < \theta \le \pi$. For such a system, the scalar potential is given by

$$\begin{aligned}\varphi(r,\theta) \;&=\; \frac{V}{\sqrt{\pi}}\sum_{j=1}^{\infty}(-1)^{j-1}\frac{(2j-\frac{1}{2})\Gamma(j-\frac{1}{2})}{j!}\left(\frac{r}{R}\right)^{2j} P_{2j-1}(\cos\theta)\\ &=\; V\left[\frac{3}{2}\left(\frac{r}{R}\right)P_1(\cos\theta) - \frac{7}{8}\left(\frac{r}{R}\right)^3 P_3(\cos\theta) + \frac{11}{16}\left(\frac{r}{R}\right)^5 P_5(\cos\theta) - \ldots\right].\end{aligned}$$

Here $\Gamma(z)$ for $\Re(z) > 0$ is defined as

$$\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}\mathrm{d}t\,. \tag{II.3.34}$$

Associated Legendre polynomials. Now we come back to the general case when azimuthal symmetry is absent. In this case we have an equation

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\left(1-x^2\right)\frac{\mathrm{d}P}{\mathrm{d}x}\right] + \left[l(l+1) - \frac{m^2}{1-x^2}\right]P = 0\,,$$

whose solutions are *associated Legendre polynomials* which can be also written explicitly with the help of the Rodriges formula

$$P_l^m = \frac{(-1)^m}{2^l l!}(1-x^2)^{\frac{m}{2}}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l\,.$$

As in the case of Legendre polynomials, one can show that finiteness of the solution on $-1 \le x \le 1$ requires m to be an integer running $-l, -(l-1), \ldots, 0, \ldots, l-1, l$. It can be further shown that

$$P_l^{-m}(x) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(x)\,.$$

Spherical harmonics. Solutions of the Laplace equation are represented as the product of three terms depending on r, θ and ϕ respectively. It is convenient to combine an angular dependence and construct a complete system of orthogonal functions on a sphere. Such functions $Y_{lm} \sim P_l^m Q_m$ are called *spherical harmonics*, explicitly,

$$Y_{lm}(\theta,\phi) = \left(\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\right)^{\frac{1}{2}} P_l^m(\cos\theta)e^{im\phi}$$

and they obey $Y_{l\,-m}(\theta,\phi) = (-1)^m Y^*_{lm}(\theta,\phi)$. Introducing the Laplace operator on the *unit sphere*

$$\Delta_\Omega = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\,,$$

we note that spherical harmonics are eigenstates of this operator

$$\Delta_\Omega Y_{lm} = -l(l+1)Y_{lm}\,.$$

Spherical harmonics form an orthogonal system of functions on the unit sphere S^2 with the measure $\mathrm{d}\Omega = \sin\theta\mathrm{d}\theta\mathrm{d}\phi$

$$\int_{S^2}\mathrm{d}\Omega\, Y^*_{lm}(\Omega)Y_{l'm'}(\Omega) \equiv \int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\mathrm{d}\theta\,\sin\theta\, Y^*_{lm}(\theta,\phi)Y_{l'm'}(\theta,\phi) = \delta_{ll'}\delta_{mm'}$$

and satisfy the completeness condition

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_{lm}(\theta,\phi)Y^*_{lm}(\theta',\phi') = \delta(\cos\theta-\cos\theta')\delta(\phi-\phi')\,.$$

As a consequence of completeness of an orthogonal system given by $\{Y_{lm}(\theta,\phi)\}$, an arbitrary function $f(\theta,\phi)$ on a sphere can be expanded in a series over spherical harmonics

$$f(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{m=l} A_{lm}Y_{lm}(\theta,\phi)\,.$$

Coefficients A_{lm} are found by using orthogonality condition for spherical harmonics. This completes our discussion of solving the Laplace equation in spherical coordinates.[4]

Thus, combining our results on the separation of variables for the Laplace operator in spherical coordinates, we write down a general solution of the Laplace equation

$$\varphi(r,\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left(A_{lm}r^l + B_{lm}r^{-(l+1)}\right)Y_{lm}(\theta,\phi)\,.$$

In the case when there is an azimuthal symmetry, *i.e.* the potential does not depend on ϕ, the expansion simplifies to

$$\varphi(r,\theta) = \sum_{l=0}^{\infty}\left(A_l r^l + B_l r^{-(l+1)}\right)P_l(\cos\theta)\,.$$

Finally, if a problem at hand has a spherical symmetry, *i.e.* the potential depends neither on θ nor on ϕ, we are left with a very simple expression, namely, the $l=0$ term in the previous expansions

$$\varphi(r) = A + \frac{B}{r}\,.$$

One concrete problem of solving the Laplace equation in spherical coordinates will be discussed in Tutorial X.

The Legendre equation is of the second order. Therefore, it must have another independent solution Q. It can be found in the following way. Consider

$$\frac{d}{dx}(1-x^2)P' + l(l+1)P \quad = \quad 0$$

[4] Analogously, one can treat the case of cylindrical, elliptical or other orthogonal coordinate systems.

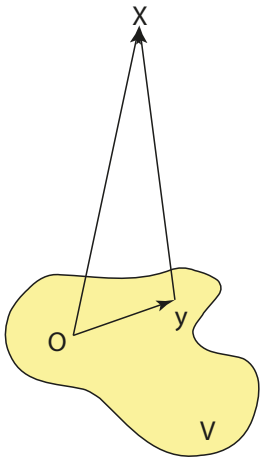

Figure 3.10: Multipole expansion is an expansion of the exact expression for the scalar potential on distances that are large in comparison with a region of charge localization.

$$\frac{d}{dx}(1-x^2)Q' + l(l+1)Q \quad = \quad 0\,.$$

Multiply the first equation by Q and another by P and subtract one from the other. We get

$$\frac{d}{dx}\Big[(1-x^2)(PQ' - QP')\Big] = 0\,.$$

Integration gives

$$(1-x^2)(PQ' - QP') = C\,,$$

where C is an integration constant. This can be brought to the form

$$\frac{d}{dx}\left(\frac{Q}{P}\right) = \frac{C}{(1-x^2)P^2}\,.$$

Integration gives

$$Q(x) = P(x)\int_\infty^x \frac{dy}{(1-y^2)P^2(y)}\,,$$

where normalization has been chosen such that $Q(\infty) = 0$. For n integer

$$Q_n(x) = P_n(x)\int_\infty^x \frac{dy}{(1-y^2)P_n^2(y)}\,,$$

the functions $Q_n(x)$ are not polynomials because the integrand above exhibits logarithmic singularities at $y = \pm 1$. The functions $Q_n(x)$ are called the *Legendre functions of the second kind.*

3.4 Multipole expansion of scalar potential

Let us assumed that electric charge is localized with the local charge density $\rho(\vec{x})$ inside a bounded region V. We chose an origin of a coordinate system somewhere inside V. Let us call $\max|\vec{y}| = L$, where $\vec{y}$ is an arbitrary point in V, "the size" of our system of charges.

It is interesting to know the scalar potential $\varphi(\vec{x})$ outside V, that is in the region $r \equiv |\vec{x}| \geq L$. Clearly, on large distances one can treat the system of charges as a point-like charge q that creates the potential $\varphi = q/r$. *The multipole expansion* is a representation of the exact answer

$$\varphi(\vec{x}) = \int_V \mathrm{d}y \frac{\rho(\vec{y})}{|\vec{x}-\vec{y}|}$$

in the form of a power series, which contains all the corrections to the simplest approximation $\varphi = q/r$. To build up the multipole expansion, we simply expand $|\vec{x} - \vec{y}|^{-1}$ into Taylor series in coordinates y_i of $\vec{y}$:

$$\frac{1}{|\vec{x} - \vec{y}|} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} y_{i_1} \cdots y_{i_n} \, \partial_{i_1} \cdots \partial_{i_n} \frac{1}{r} \,,$$

where $|\vec{y}| < |\vec{x}| = r$. Substituting this expansion into the expression for the potential, we get

$$\varphi(\vec{x}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \, T_{i_1 \ldots i_n} \, \partial_{i_1} \cdots \partial_{i_n} \frac{1}{r} \,,$$

where

$$T_{i_1 \ldots i_n} = \int \mathrm{d}^3 y \, \rho(\vec{y}) \, y_{i_1} \cdots y_{i_n} \,.$$

This is a multipole expansion and $T_{i_1 \ldots i_n}$ are called the multipole momenta. The first ones are

$$\begin{aligned} Q &= \int \mathrm{d}^3 y \, \rho(\vec{y}) \quad - \quad \text{monopole moment (total electric charge)} \\ d_i &= \int \mathrm{d}^3 y \, \rho(\vec{y}) \, y_i \quad - \quad \text{dipole moment} \\ T_{ij} &= \int \mathrm{d}^3 y \, \rho(\vec{y}) \, y_i y_j \quad - \quad \text{quadrupole moment} \end{aligned} \tag{II.3.35}$$

The multipole momenta have the following properties:

- Symmetry with respect to permutation of indices $i_1 \ldots i_n$.
- They are tensors with respect to the action of the orthogonal group.
- Transformation properties with respect to shifts of the origin: $y_i \to y'_i = y_i + a$. Since $d^3 y' = d^3 y$, one gets[5]

  $$T'_{i_1 \ldots i_n} = \int \mathrm{d}^3 y \, \rho(\vec{y}) (y_{i_1} + a_i) \cdots (y_{i_n} + a_{i_n})$$

  that upon opening the brackets give 2^n terms. The first term is the tensor $T_{i_1 \ldots i_n}$ itself, while all the other terms will contain a multiplied by multipole momenta of lower rank than n, *i.e.*;

  $$T'_{i_1 \ldots i_n} = T_{i_1 \ldots i_n} + \text{contributions of lower } T \,.$$

  Thus, $T_{i_1 \ldots i_n}$ do not depend on the choice of the origin of the coordinate system if and only if all lower multipole moments vanish. In other words, only the first non-trivial moment is invariant with respect to shifts of the origin. The first moment which is a total charge is always invariant under shifts. The second moment, which is the dipole moment, is invariant only if the total charge q is equal to zero.[6]

[5] Here we have taken into account that the charge density $\rho(\vec{y})$ is a scalar, which means that under general coordinate transformations $y_i \to y'_i(y)$ it transforms as $\rho'(\vec{y}') = \rho(\vec{y})$.

[6] For a discrete system of charges the arguing is very similar. The dipole moment is $\vec{d} = \sum_{i=1}^{N} e_i \vec{x}_i$, where e_i is the magnitude of a charge at some distance R_i taken from an arbitrary point, in this case chosen to be the origin. For neutral system $\sum_{i=1}^{N} d_i = 0$. Thus, shifting all $\vec{R}_i \to \vec{R}_i - \vec{a}$ gives

$$\vec{d}_{\vec{a}} = \sum_{i=1}^{N} e_i \, (\vec{x}_i - \vec{a}) = \sum_{i=1}^{N} e_i \vec{x}_i - \vec{a} \sum_{i=1}^{N} e_i = \sum_{i=1}^{N} e_i \vec{x}_i = \vec{d}.$$

Now we discuss how to construct the multipole expansion in terms of irreducible moments. Recall that a tensor is called irreducible if being contracted on any pair of two indices it gives zero. *Irreducibility means that that from a given tensor one cannot construct by contacting indices a simpler object – a tensor of lower rank.* Any tensor can be reduced to its irreducible component by adding proper terms containing Kronecker's delta. For, instance, for a second rank tensor one finds that its irreducible component is

$$\mathscr{T}_{ij} = T_{ij} - \frac{\delta_{ij}}{3} T_{kk}\,,$$

so that the irreducible tensor of quadrupole moment is

$$\mathscr{T}_{ij} = \int \mathrm{d}^3 y\, \rho(\vec{y}) \left(y_i y_j - \frac{y^2}{3}\delta_{ij} \right) .$$

It turns out that the multipole expansion is unchanged if one replaces all multipole momenta for their irreducible components. This follows from the fact that

$$\delta_{ij}\partial_i\partial_j \frac{1}{r} = \Delta \frac{1}{r} = 0\,,$$

as there is no charge located at $\vec{x}$. Thus, the multipole expansion can be written as

$$\varphi(\vec{x}) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\, \mathscr{T}_{i_1\ldots i_n}\, \partial_{i_1}\cdots\partial_{i_n} \frac{1}{r}\,.$$

We further notice that

$$\begin{aligned} \partial_i \frac{1}{r} &= -\frac{x_i}{r^3}\,, \\ \partial_i\partial_j \frac{1}{r} &= -\frac{\delta_{ij}}{r^3} + (-1)(-3)\frac{x^i x^j}{r^5} \end{aligned} \tag{II.3.36}$$

and so on. In general,

$$\partial_{i_1}\cdots\partial_{i_n} \frac{1}{r} = (-1)^n (2n-1)!!\, \frac{x_{i_1}\cdots x_{i_n}}{r^{2n+1}} + \ldots\,,$$

where $\ldots$ stand for all the terms containing Kronecker's delta. Since all such terms drop out when being contracted with irreducible tensors, one finds that the multipole expansion takes the form

$$\varphi(\vec{x}) = \sum_{n=0}^{\infty} \frac{(2n-1)!!}{n!}\, \mathscr{T}_{i_1\ldots i_n}\, \frac{x_{i_1}\cdots x_{i_n}}{r^{2n+1}}\,.$$

Explicitly,

$$\varphi(\vec{x}) = \frac{q}{r} + \frac{d_i x_i}{r^3} + \frac{3\mathscr{T}_{ij} x_i x_j}{r^5} + \ldots$$

The first term vanishes as $1/r$ as $r \to \infty$, the second one representing the dipole moment as $1/r^2$, the third term as $1/r^3$ and so on. Thus, if a potential vanishes faster than $1/r$, its first several moments must be zero. For instance, if $\varphi \sim 1/r^3$, then the total charge and the dipole moment must be zero, while the quadrupole moment must not.

If one knows an expansion of $\varphi(\vec{x})$ in power series in $1/r$, then one can restore all irreducible moments $\mathscr{T}_{i_1\ldots i_n}$ and vice versa, knowing all $\mathscr{T}_{i_1\ldots i_n}$ one can restore the potential. That is there is a one-to-one map between a set of multiple moments and the corresponding potential. Knowing $T_{i_1\ldots i_n}$one can also uniquely restore the potential, but the inverse is not true.

Thus, for the potential we find

$$\varphi = \frac{q}{r} + \frac{\vec{x}\cdot\vec{d}}{r^3} + \dots .$$

If the system is neutral, then $q = 0$ and the leading term in the expansion of φ is the dipole term. The electric field corresponding to the dipole potential is

$$\vec{E} = -\vec{\nabla}\frac{(\vec{x}\cdot\vec{d})}{r^3} = \frac{3\vec{n}(\vec{n}\cdot\vec{d}) - \vec{d}}{r^3}\,. \tag{II.3.37}$$

Thus, for a neutral system the electric field at large distances from this system behaves itself as $1/r^3$, rather than $1/r^2$!

Field of an idealised point dipole. Formula (II.3.37) for $\vec{E}$ of a dipole is valid everywhere outside the origin. To find the modification of this formula at the origin, one can use the following statement. If a sphere of radius R completely encloses the charge density giving rise to the dipole moment $\vec{d}$, then the integral of the electric field $\vec{E}$ over the sphere is

$$\int_{r<R} \vec{E}(\vec{x})\, d^3x = -\frac{4\pi}{3}\vec{d}.$$

Note that this volume integral is independent on the size of the spherical region of integration provided all the charge is inside. To be consistent with this formula, the dipole field must be written as

$$\vec{E} = \frac{3\vec{n}(\vec{n}\cdot\vec{d}) - \vec{d}}{r^3} - \frac{4\pi}{3}\vec{d}\,\delta(\vec{r})\,. \tag{II.3.38}$$

The delta-function does not contribute away from the origin. Its only purpose to yield the required volume integral, with the convention that the spherically symmetric (around the origin) volume integral of the first term is zero from the angular integration, the singularity at the origin causing an otherwise ambiguous result. This formula treat dipoles as idealised point dipoles.

Addition theorem for spherical harmonics. Consider an expression

$$\frac{1}{|\vec{x}-\vec{x}'|} = \frac{1}{\sqrt{r^2 + r'^2 - 2(\vec{x}\cdot\vec{x}')}} = \frac{1}{r}\frac{1}{\sqrt{1 - 2(\vec{n}\cdot\vec{n}')\frac{r'}{r} + \frac{r'^2}{r^2}}}\,, \tag{II.3.39}$$

where $r = |x|$, $r' = |x'|$, $\vec{n} = \frac{\vec{x}}{r}$ and $\vec{n}' = \frac{\vec{x}'}{r'}$ are the unit vectors in the directions of $\vec{x}$ and $\vec{x}'$, respectively. We assume that $r' < r$ and construct an expansion of (II.3.39) in powers of r'/r. To this end, it is convenient to remember the generating function for the Legendre polynomials, namely,

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{l=0}^{\infty} t^l P_l(x)\,. \tag{II.3.40}$$

Taking $t = r'/r$ and $\vec{n}\cdot\vec{n}' = \cos\gamma$, we compare (II.3.39) with the generating function (II.3.40), and find

$$\frac{1}{|\vec{x}-\vec{x}'|} = \frac{1}{r}\sum_{\ell=0}^{\infty}\left(\frac{r'}{r}\right)^l P_l(\cos\gamma)\,, \qquad r' < r\,. \tag{II.3.41}$$

Here $\cos\gamma$ is a bit awkward variable, as γ is an angle between $\vec{x}$ and $\vec{x}'$. It would be much more convenient if the right hand side of the above expression were given in terms of the spherical coordinates $\vec{x} = (r,\theta,\phi)$ and $\vec{x}' = (r',\theta',\phi')$ defined with respect to a fixed set of coordinate axes in

space. This goal is precisely achieved with the help of the *spherical harmonic addition theorem*, which states that

$$P_l(\cos\gamma) = \frac{4\pi}{2l+1} \sum_{m=-l}^{m=l} Y^*_{lm}(\theta',\phi') Y_{lm}(\theta,\phi)\,. \tag{II.3.42}$$

Substituting this expansion into (II.3.41) produces a spherical expansion for the inverse distance referred to a single origin of coordinates

$$\frac{1}{|\vec{x}-\vec{x}'|} = \frac{1}{r}\sum_{l=0}^{\infty} \frac{4\pi}{2l+1}\left(\frac{r'}{r}\right)^l \sum_{m=-l}^{m=l} Y^*_{lm}(\theta',\phi') Y_{lm}(\theta,\phi)\,, \qquad r' < r\,. \tag{II.3.43}$$

We observe that the right hand side of the last formula has the form

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm} r^{-l} Y_{lm}(\theta,\varphi)\,,$$

which is a solution of the Laplace equation. This is to be expected, because the inverse distance $1/|\vec{x}-\vec{x}'|$ solves the Laplace equation as soon as $\vec{x}' \neq \vec{x}$ and, for this reason, admits an expansion in spherical harmonics. Equation (II.3.43) gives the scalar potential in a completely factorised form in coordinates $\vec{x}$ and $\vec{x}'$. This is useful in any integrations over charge densities, etc., where one variable is the variable of integration and the other is the coordinates of the observation point, see the corresponding problem in Tutorial XII.

Finally, we point out one consequence of (II.3.42). Since $P_l(1) = 1$ for any l, setting in (II.3.42) $\theta' = \theta$ and $\phi' = \phi$, so that $\gamma = 0$, we get

$$\sum_{m=-l}^{m=l} |Y_{lm}(\theta,\phi)|^2 = \frac{2l+1}{4\pi}\,.$$

Chapter 4

Magnetostatics

"I have constructed an expression for the attraction of two infinitely small currents which was, in truth, only a hypothesis, but the simplest one that could be adopted and, consequently, the one that should be tried first."

André-Marie Ampère (1820)

4.1 Laws of magnetostatics

In the case when electric field is static, *i.e.* it does not depend on time, the second pair of the Maxwell equations (II.3.1) takes the form

$$\operatorname{div}\vec{H} = 0\,, \qquad \operatorname{rot}\vec{H} = \frac{4\pi}{c}\vec{j}\,.$$

The first equation allows one to write

$$\vec{H} = \operatorname{rot}\vec{A}\,.$$

Substituting this in the second equation, we obtain

$$\operatorname{grad}\operatorname{div}\vec{A} - \Delta\vec{A} = \frac{4\pi}{c}\vec{j}\,.$$

Because of gauge invariance the vector potential is not uniquely defined, therefore, we can subject it to one additional constraint, which will chose to be

$$\operatorname{div}\vec{A} = 0\,.$$

Then, the equation defining the vector potential of time-independent magnetic field takes the form

$$\Delta\vec{A} = -\frac{4\pi}{c}\vec{j}\,.$$

Obviously, this is the Poisson equation, very similar to the equation for the electrostatic potential. Therefore, the solution reads as

$$\vec{A}(\vec{x}) = \frac{1}{c}\int \mathrm{d}^3x'\,\frac{\vec{j}(\vec{x}')}{|\vec{x}-\vec{x}'|}\,.$$

Now we can determine the corresponding magnetic field

$$\vec{H} = \mathrm{rot}\,\vec{A} = \frac{1}{c}\int \mathrm{d}^3x' \left[\vec{\nabla}\frac{1}{|\vec{x}-\vec{x}'|}\times\vec{j}(x')\right] = \frac{1}{c}\int \mathrm{d}^3x' \frac{\vec{j}(\vec{x}')\times(\vec{x}-\vec{x}')}{|\vec{x}-\vec{x}'|^3}\,,$$

where the bracket means the vector product[1]. This is the *Biot-Savart law*. It describes the magnetic field produced by time-independent currents.

The integral form of Maxwell's equation $\mathrm{rot}\,\vec{H} = \frac{4\pi}{c}\vec{j}$ is called *Ampère's law*. To derive it, consider a surface S enclosed by a contour C. The flux of both sides of the last equation through S is

$$\int_S (\mathrm{rot}\,\vec{H}\cdot\vec{n})\mathrm{d}S = \frac{4\pi}{c}\int_S (\vec{j}\cdot\vec{n})\mathrm{d}S\,.$$

Application of the Stoks theorem gives

$$\oint_C \vec{H}\cdot\mathrm{d}\vec{\ell} = \frac{4\pi}{c}\int_S (\vec{j}\cdot\vec{n})\,\mathrm{d}S = \frac{4\pi}{c}I\,, \tag{II.4.1}$$

where

$$I = \int_S (\vec{j}\cdot\vec{n})\,\mathrm{d}S \tag{II.4.2}$$

is the full current through the surface S. Formula (II.4.1) is Ampère's law.

4.2 Continuity equation

The charge density ρ and the current density $\vec{j}$ are related by the so-called *continuity equation*. In fact, this equation is a compatibility condition of Maxwell's equations. We take the Gauss law $\mathrm{div}\,\vec{E} = 4\pi\rho$ and differentiate it with respect to time

$$4\pi\frac{\partial\rho}{\partial t} = \mathrm{div}\,\frac{\partial\vec{E}}{\partial t} = \mathrm{div}\left(c\,\vec{\nabla}\times\vec{H} - 4\pi\vec{j}\right) = -4\pi\mathrm{div}\,\vec{j}\,,$$

because $\vec{\nabla}\cdot(\vec{\nabla}\times\vec{H}) = 0$. In this way we obtain the continuity equation

$$\frac{\partial\rho}{\partial t} + \mathrm{div}\,\vec{j} = 0\,. \tag{II.4.3}$$

Physically, the continuity equation expresses the *conservation law of electric charge*. First of all, the total current that passes through a finite surface S is

$$I = \frac{dQ}{dt} = \int_S (\vec{j}\cdot\vec{n})\,\mathrm{d}S\,. \tag{II.4.4}$$

If we take the surface S to be closed surrounding the volume V, then we can apply the Gauss-Ostrogradsky theorem and write

$$\int_S (\vec{j}\cdot\vec{n})\,\mathrm{d}S = \int_V d^3x\,\mathrm{div}\,\vec{j}\,. \tag{II.4.5}$$

[1]Here we have used the formula $\mathrm{rot}\,f\vec{A} = f\mathrm{rot}\,\vec{A} + [\mathrm{grad}\,f, \vec{A}]$.

Because the vector $\vec{n}$ here points outwards from V, equation (II.4.4) written for the charge Q inside the volume should describe its decrease and, for this reason, must be written as

$$-\frac{dQ}{dt} = -\frac{d}{dt}\int_V d^3x\,\rho = -\int_V d^3x\frac{\partial\rho}{\partial t}\,. \tag{II.4.6}$$

Thus, combining (II.4.5) and (II.4.6), we obtain

$$-\int_V d^3x\frac{\partial\rho}{\partial t} = \int_V d^3x\,\mathrm{div}\,\vec{j} \quad \longrightarrow \quad \int d^3x\left[\frac{\partial\rho}{\partial t} + \mathrm{div}\,\vec{j}\right] = 0\,,$$

from where, because of arbitrariness of V, we derive a local statement of charge conservation, which is the continuity equation (II.4.3).

Let N point charges q_i follow the trajectories $\vec{x}_i(t)$. Then the charge density is

$$\rho(t,\vec{x}) = \sum_{i=1}^{N} q_i\delta(\vec{x}-\vec{x}_i(t))\,. \tag{II.4.7}$$

Let us find the current density produced by the flow of these charges. We have

$$\frac{\partial\rho}{\partial t} = \sum_{i=1}^{N} q_i\frac{\partial}{\partial x_{i,k}}\delta(\vec{x}-\vec{x}_i(t))\frac{dx_{i,k}}{dt} = -\sum_{i=1}^{N} q_i\frac{\partial}{\partial x_k}\delta(\vec{x}-\vec{x}_i(t))v_{i,k}\,,$$

where $x_{i,k}$ id the k-th component of $\vec{x}_i$ and $v_{i,k} = dx_{i,k}/dt$ is the k-th component of the velocity $\vec{v}_i$ of i-th particle. The last formula can be written as

$$\frac{\partial\rho}{\partial t} = -\sum_{i=1}^{N} q_i\vec{v}_i\cdot\vec{\nabla}\delta(\vec{x}-\vec{x}_i) = -\vec{\nabla}\cdot\sum_{i=1}^{N} q_i\vec{v}_i\delta(\vec{x}-\vec{x}_i)\,.$$

Comparing this equation with the continuity equation (II.4.3), we conclude that

$$\vec{j}(t,\vec{x}) = \sum_{i=1}^{N} q_i\vec{v}_i(t)\delta(\vec{x}-\vec{x}_i(t))\,, \qquad \vec{v}_i = \frac{d\vec{x}_i}{dt}\,. \tag{II.4.8}$$

Thus, equations (II.4.7) and (II.4.8) represent the charge and current density for a system of moving charges satisfying the continuity equation.

4.3 Magnetic dipole moment

Free magnetic charges do not exist. The really existing object which plays the basic role[2] in study of magnetic phenomena is the so-called *magnetic dipole*. A small magnetic dipole is a magnetic arrow (like the compass arrow) which aligns along the direction of an external magnetic field.

Consider the magnetic field created by a system of stationary moving charges on distances large in comparison with the size of this system. We choose a center of a reference frame somewhere inside the system of moving charges. Then $x' \ll x$ and we can expand

$$\frac{1}{|\vec{x}-\vec{x}'|} = \frac{1}{|\vec{x}|} + \frac{(\vec{x}\cdot\vec{x}')}{|\vec{x}|^3} + \dots$$

[2]The same role as elementary electric charge in electrostatics.

Therefore, for the vector potential we get

$$A_i(\vec{x}) = \frac{1}{c|\vec{x}|}\int j_i(\vec{x}')\mathrm{d}^3x' + \frac{1}{c|\vec{x}|^3}\cdot\int j_i(\vec{x}')(\vec{x}\cdot\vec{x}')\mathrm{d}^3x' + \cdots$$

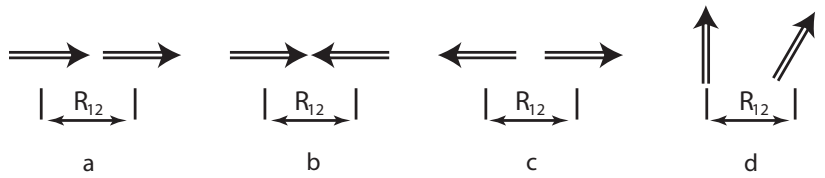

Figure 4.1: Force between magnetic dipoles depends not only on the distance between them but also on their mutual orientation: a) magnetic dipoles attract ($U_M < 0$), b) and c) magnetic dipoles repeal $U_M > 0$), d) the sign of energy U_M is determined by the general formula $U_M = \frac{(\vec{M}_1\cdot\vec{M}_2)-3(\vec{M}_1\cdot\vec{n}_{12})(\vec{M}_2\cdot\vec{n}_{12})}{R_{12}^3}$, $\vec{n}_{12} = \frac{\vec{R}_{12}}{R_{12}}$.

From the continuity equation $\frac{\partial\rho}{\partial t} + \operatorname{div}\vec{j} = 0$ we have $\operatorname{div}\vec{j} = 0$. Taking this into account, for *any* function $f(x)$ we can write

$$0 = \int f(x')\operatorname{div}\vec{j}\,\mathrm{d}^3x' = -\int(\vec{\nabla}f\cdot\vec{j})\,\mathrm{d}^3x'\,,$$

where we have integrated by parts. Picking now $f = x_i$, we get $(\vec{\nabla}x_i)_j = \delta_{ij}$, so that $(\vec{\nabla}x_i\cdot\vec{j}) = j_i$. Thus, we arrive at

$$\int j_i(x')\mathrm{d}^3x' = 0 \quad \text{for any } i\,.$$

This is also intuitively clear, because the current is assumed to have vanishing normal components everywhere on the surface S – the current is concentrated in the volume surrounded by S and never flows out through S. Hence, the leading term of the vector potential is

$$\vec{A}(\vec{x}) = \frac{1}{c|\vec{x}|^3}\cdot\int\vec{j}(x')(\vec{x}\cdot\vec{x}')\,\mathrm{d}^3x'$$

To make further progress, we recall an identity

$$[\vec{a},[\vec{b},\vec{c}]] = (\vec{a}\cdot\vec{c})\vec{b} - (\vec{a}\cdot\vec{b})\vec{c},$$

which allows one to write

$$(\vec{x}\cdot\vec{x}')\vec{j} = (\vec{x}\cdot\vec{j})\vec{x}' - \vec{x}\times(\vec{x}'\times j)\,.$$

It turns out that the integral from $(\vec{x}\cdot\vec{x}')\vec{j}$ is equal up to the minus sign to the integral from $(\vec{x}\cdot\vec{j})\vec{x}'$. Indeed, since $\operatorname{div}\vec{j} = 0$, we have

$$\int\mathrm{d}^3x'\,j_kx_i' = \int\mathrm{d}^3x'\operatorname{div}(x_k'\vec{j})x_i' \overset{\text{by parts}}{=} -\int\mathrm{d}^3x'\,x_k'(\vec{j}\cdot\operatorname{grad}')x_i' = -\int\mathrm{d}^3x'\,x_k'j_i\,.$$

From here we deduce that

$$\int\mathrm{d}^3x'\,(\vec{x}\cdot\vec{j})\,x_i' = -\int\mathrm{d}^3x'\,(\vec{x}\cdot\vec{x'})\,j_i\,,$$

or, in the vector form,

$$\int\mathrm{d}^3x'\,(\vec{x}\cdot\vec{j})\,\vec{x}' = -\int\mathrm{d}^3x'\,(\vec{x}\cdot\vec{x'})\,\vec{j}\,.$$

Therefore, we arrive at

$$\vec{A} = -\frac{\vec{x}}{|\vec{x}|^3} \times \frac{1}{2c} \int \mathrm{d}^3x' \, \vec{x}' \times \vec{j}(\vec{x}') \, .$$

Define the density of the magnetic moment as

$$\vec{\mathcal{M}} = \frac{1}{2c} \vec{x}' \times \vec{j}(x')$$

and the magnetic moment as

$$\vec{M} = \int \mathrm{d}^3x' \, \vec{\mathcal{M}}(x') = \frac{1}{2c} \int \mathrm{d}^3x' \, \vec{x}' \times \vec{j}(\vec{x}') \, .$$

We, therefore, find

$$\vec{A}(\vec{x}) = \frac{\vec{M} \times \vec{x}}{|\vec{x}|^3} \, .$$

This is the leading term in the expansion of the vector potential for a bounded stationary current distribution. As a result, the magnetic field of a magnetic dipole is

$$\vec{H} = \mathrm{rot} \, \vec{A} = \frac{3\vec{n}(\vec{n} \cdot \vec{M}) - \vec{M}}{|\vec{x}|^3} \, ,$$

where $\vec{n}$ is the unit vector in the direction of $\vec{x}$. This expression for the magnetic field coincides with the formula for the electric field of an electric dipole.

4.4 Gyromagnetic ratio. Magnetic moment of electron.

Suppose that the current I flows over a closed flat loop C on an arbitrary shape. For the magnetic moment we have

$$\vec{M} = \int \mathrm{d}^3x' \, \vec{\mathcal{M}}(x') = \frac{1}{2c} \int \mathrm{d}^3x' \, \vec{x}' \times \vec{j}(x') = \frac{1}{2c} \int \mathrm{d}S' \mathrm{d}\ell \; \vec{x}' \times \vec{j}(x') \, ,$$

where $\mathrm{d}S'$ is an area differential corresponding the transverse section of the (thin) loop C. Since the current I is defined as

$$I = \int_S (\vec{j} \cdot \vec{n}) \mathrm{d}S \, ,$$

we have

$$\vec{M} = \frac{1}{2c} \int \mathrm{d}S' \; \vec{x}' \times (\vec{j}(x') \cdot \vec{n}) \mathrm{d}\vec{\ell}$$

so that the magnetic moment can be written in the form

$$\vec{M} = \frac{I}{2c} \oint_C \vec{x} \times \mathrm{d}\vec{\ell}.$$

Since $\vec{x} \times \mathrm{d}\vec{\ell} = 2 \, \mathrm{d}\vec{S}$, where $\mathrm{d}\vec{S}$ is the area of an elementary triangle formed by the radii drawn from the origin of the coordinate system to the end points of the element $\mathrm{d}\vec{\ell}$, the integral above is equal to the total area S enclosed by the current loop C. Therefore,

$$|\vec{M}| = \frac{IS}{c}$$

independently of the shape of the contour. Here $|\vec{M}|$ is a magnitude of the magnetic dipole moment of a current loop.

If the current is formed by particles of masses m_i with charges e_i moving with velocities $\vec{v}_i \ll c$, then the magnetic moment can be expressed via the angular momentum. We have

$$\vec{j}(x) = \sum_i e_i \vec{v}_i \delta(\vec{x} - \vec{x}_i)\,,$$

where $\vec{x}_i$ is the radius-vector of i'th particle. In this case the magnetic moment is

$$\vec{M} = \frac{1}{2c}\sum_i e_i(\vec{x}_i \times \vec{v}_i) = \sum_i \frac{e_i}{2c\,m_i}(\vec{x}_i \times m_i\vec{v}_i) = \sum_i \frac{e_i}{2c\,m_i}\underbrace{[\vec{x}_i, \vec{p}_i]}_{\vec{L}_i},$$

where $\vec{L}_i = [\vec{x}_i, \vec{p}_i]$ is the angular momentum of the i'th particle and we have used the fact that for $v \ll c$ the expression $m\vec{v}$ coincides with the particle momentum $\vec{p}$. If for all the particles the ratio of charge to mass is the same, $e_i/m_i \equiv e/m$, then

$$\vec{M} = \frac{e}{2mc}\sum_i \vec{L}_i = \frac{e}{2mc}\vec{L}\,,$$

where $\vec{L}$ is the total angular momentum of a system of particles. The relation

$$\vec{M} = \frac{e}{2mc}\vec{L} \quad \Rightarrow \quad \frac{M}{L} = \frac{e}{2mc}$$

is an important classical relation between the magnetic and the angular momenta. This relation is *remarkable* – for a loop of current it expresses the ratio of two macroscopic quantities (the magnetic moment of the current loop and the total angular momentum of electrons) via a combination of microscopic quantities characterizing the charge carriers! The quantity

$$\gamma = \frac{M}{L} = \frac{e}{2mc}$$

is called a *gyromagnetic ratio*. In a conductor charge carriers are electrons, *i. e.*

$$\gamma = \frac{e}{2m_e c}\,.$$

Gyromagnetic ratio is often measured in units of $\gamma = \frac{e}{2m_e c}$, in particular, γ is taken for unity. Indeed, if the current in a conductor would be carried by ions rather than electrons, then the gyromagnetic ratio will be thousand times less. It is difficult to imagine that gyromagnetic ratio could be bigger than one – electrons the lightest particles carrying the charge!

Chapter 5

Relativistic Mechanics

"The special theory of relativity owes its origin to Maxwell's equations of the electromagnetic field."

Albert Einstein (1949)

5.1 Relativity principle

5.1.1 Newton's relativity principle

In order to describe a dynamical system one has to choose a reference frame. The reference frame is a system of coordinates and a clock which measures the time in this coordinate system, see Fig. 5.1. In mechanics one introduces the notion of an *intertial frame*. In such frames a free motion (*i.e.* the motion in the absence of forces) happens with a uniform velocity. Excluding trivial translations of coordinates, any two inertial frames are related by an orthogonal transformation, *i.e.* by a rotation with possible reflections of coordinate axes.

Experience shows that that *the relativity principle* is valid. According to this principle, all laws of Nature are the same in all inertial frames. In other words, the equations which encode the laws of Nature are invariant with respect to transformations from one inertial system of coordinates to another. This means that an equation encoding a physical law when expressed through spatial coordinates and time in different inertial frames must have the one and the same form.

In order to give a mathematical description of the relativity principle, one has to find formulas which relate special coordinates and time in different inertial frames. In Newtonian mechanics it was assumed for a long time that inertial frames are related by Galilean transformations

$$\begin{aligned} \vec{x}' &= R(\vec{x} - \vec{v}t) \\ t' &= t \end{aligned} \tag{II.5.1}$$

Here R is a matrix of orthogonal transformations of coordinates.

5.1.2 Einstein's relativity principle

In classical mechanics interaction of particles is described by means of potential energy, which is a function of coordinates of interacting particles. Such a description is based on an assumption of

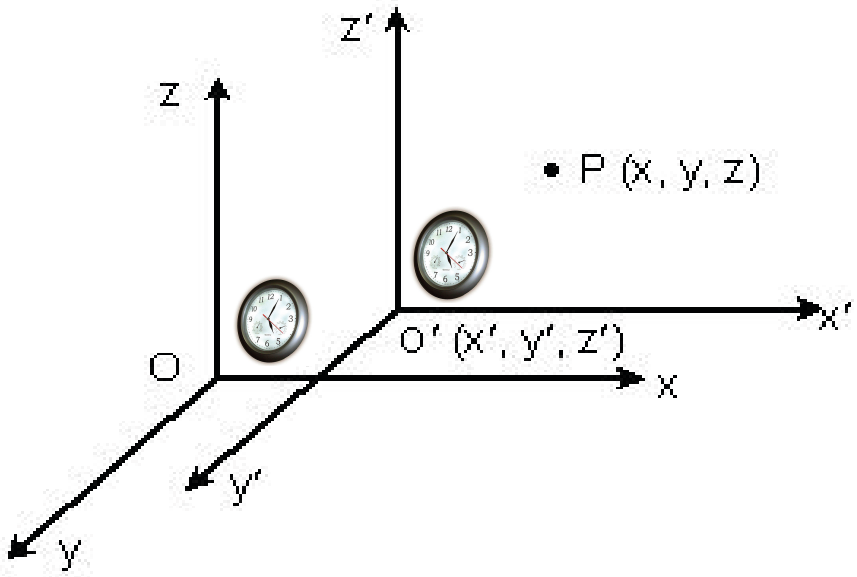

Figure 5.1: Reference frame – a coordinate system and a clock.

instantaneous interactions. Indeed, forces which act on particles depend only on the positions of particles in the same moment when these positions are measured. Any change in the motion of any of the particles immediately reflects on the others with no time delay. On the other hand, experience shows that instantaneous interactions are impossible in Nature. Therefore, any mechanics which is based on the principle of instantaneous interactions has certain limitations. If something happens to one body, the time is needed for the corresponding changes to reach another body. Therefore, there must exist a maximal velocity of propagating the interactions and it must be the same in all inertial frames. This universal velocity happens to coincide with the speed of light in vacuum and it is equal to

$$c = 2.99792458 \cdot 10^8 \quad \text{m/sec.}$$

This is a fundamental physical constant.[1].

Conjunction of the relativity principle with the finiteness of the speed of interaction propagation (speed of light) is called *Einstein's relativity principle* (Einstein, 1905). The mechanics which is based on Einstein's relativity principle is called *relativistic*. The mechanics which arises in the limiting case when formally $c \to \infty$ is called Newtonian or classical.

Three fundamental effects of Special Relativity are

1) Time delay measured by a moving clock;

2) Lorentz contraction of the length of a moving body;

3) Abberation of light (deviation of light when passing from one to another inertial system).

[1]Before the 17th century it was generally thought that the light is transmitted instantaneously. Galileo was the first one who doubted it. The first measurement has been done by Roemer in 1676 (observations of the moons of Jupiter). The first measurement on Earth has been done by Fizeau in 1849. He used a beam of light reflected from a mirror 8km away. He found the value 301300 km/sec.

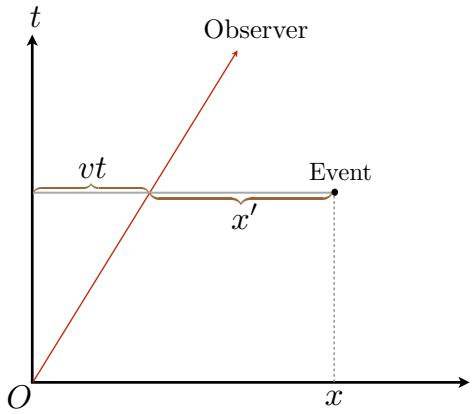

Figure 5.2: Galilean boost in two dimensions. The oblique (red) line represents the trajectory of the origin of the reference frame M' which moves with velocity v in the x-direction with respect to the reference frame M. An event which happens in M at the position x at time t occurs at x' at time $t' = t$ in the moving frame M'. Hence, $x' = x - vt$.

5.2 Lorentz group

Here we introduce Lorentz transformations as transformations between two inertial frames that preserve the length of the 4-interval between two events. We then show that this transformations form a Lie group known as the *Lorentz group*, and we discuss the structure of this group from various viewpoints.

5.2.1 Defining Lorentz transformations

We will use the notion of "event". Every event is characterized by the place (coordinates) where it happened and by the time when it happened. Define the so-called interval between two events

$$s_{12} = c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2 \,.$$

If two events are close to each other we have an infinitezimal interval:

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \,.$$

The fact that the speed of light is the one and the same constant in all inertial frames leads to the fact that the infinitezimal interval between two events is also the same in all inertial frames

$$ds^2 = ds'^2 \,.$$

From the equality of infinitezimal intervals, the equality of finite intervals follows $s = s'$.

The interval between two events is the same in all inertial frames, i.e. it is invariant under transformations from one inertial frame to another. This invariance encodes the constancy of the speed of light.

The intervals can be naturally classified as follows. Introduce

$$\ell_{12}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2.$$

Then, $s_{12}^2 = c^2 t_{12}^2 - \ell_{12}^2$ and the equality of the intervals in two different inertial frames is expressed as

$$s_{12}^2 = c^2 t_{12}^2 - \ell_{12}^2 = c^2 {t'}_{12}^2 - {\ell'}_{12}^2 \,.$$

- *Time-like interval.* This is an interval for which $s_{12}^2 > 0$, *i.e.* the interval is real. For such an interval there exists an inertial system for which the two events happen in the one and the same space point, *i.e.* ${\ell'}_{12}^2 = 0$. If two events happened to the one and the same body then the interval between them is always time-like. Indeed, the distance $\ell_{12} = v t_{12}$ which the body passes cannot be bigger than $c t_{12}$ as $v < c$.

  Remember: Real intervals are time-like. They describe events which happen to a (massive) body.

- *Space-like intervals.* For these intervals $s_{12}^2 < 0$, *i.e.* they are imaginary. For a space-like interval one can always find an inertial system in which the corresponding two events happened as the same moment of time, so that ${t'}_{12} = 0$. The distance between these events is ${\ell'}_{12} = i s_{12}$.

- *Light-like intervals (null intervals).* For these intervals $s_{12} = 0$.

It is convenient introduce the diagonal 4×4-matrix

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1) \,.$$

It is called the Minkowski metric and it defines a quadratic form

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \,,$$

which is an infinitezimal interval and we consider the index μ running from 0 to 3, so that $x^0 = ct$ and $x^1 \equiv x$, $x^2 \equiv y$, and $x^3 \equiv z$ stand for three spacial coordinates.

Thus, the set (ct, x, y, z) can be considered as components of a vector in a four-dimensional space. The square of the "length" of the vector is

$$x^2 \equiv (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = \eta_{\mu\nu} x^\mu x^\nu \,.$$

Geometry in which the length of a vector is given by the above formula is called *pseudo-euclidean.*

According to the discussion above, the transformations from one inertial frame to another must be such that they preserve the interval. In the four-dimensional space they can be only the global shifts of the coordinate system

$$x^\mu \to x^\mu + a^\mu$$

or rotations

$$x^\mu \to \Lambda^\mu_\nu x^\nu \,.$$

Under the rotations the quadratic form transforms as

$$x'^2 = \eta_{\mu\nu} \Lambda^\mu_\alpha x^\alpha \Lambda^\nu_\beta x^\beta = \eta_{\mu\nu} \Lambda^\mu_\alpha \Lambda^\nu_\beta x^\alpha x^\beta = x^2 \,,$$

so that the transformation matrices must satisfy the requirement

$$\eta_{\mu\nu} \Lambda^\mu_\alpha \Lambda^\nu_\beta = \eta_{\alpha\beta} \,.$$

The matrices satisfying this requirement are called *Lorentz transformations.*

5.2.2 Structure of the Lorentz group

We recall that a definition of a group was given in subsection 2.2.2. An important class of groups constitute Lie groups. A *Lie group* is a group which is also a smooth manifold.[2]

Let us show that Lorentz transformations form a group. In the matrix form the Lorentz transformations can be written as

$$\Lambda^t \eta \Lambda = \eta \,.$$

Any matrix Λ which satisfies this relation (*defining relation*) defines a Lorentz transformation.[3] Suppose we have two such matrices

$$\Lambda_1^t \eta \Lambda_1 = \eta \,, \qquad \Lambda_2^t \eta \Lambda_2 = \eta \,,$$

then their product is also satisfies the defining relation of the Lorentz group:

$$(\Lambda_1 \Lambda_2)^t \eta (\Lambda_1 \Lambda_2) = \Lambda_2^t (\Lambda_1^t \eta \Lambda_1) \Lambda_2 = \Lambda_2^t \eta \Lambda_2 = 1 \,.$$

Identity matrix is a (trivial) Lorentz transformation. Finally, any Λ has an inverse which also a Lorentz transformation. Indeed,

$$\det(\Lambda^t \eta \Lambda) = \det(\Lambda)^2 \det\eta = \det\eta \quad \Longrightarrow \quad \det\Lambda = \pm 1 \,.$$

This means that Λ is non-degenerate. Then, from the defining relation[4]

$$\Lambda^{-1} = \eta \Lambda^t \eta \,.$$

Thus,

$$(\Lambda^{-1})^t \eta \Lambda^{-1} = (\eta \Lambda^t \eta)^t \eta (\eta \Lambda^t \eta) = \Lambda \eta \Lambda^t = \eta \,,$$

that is Λ^{-1} is a Lorentz transformation. Thus, Lorentz transformations form a group. We have also shown that if Λ is a Lorentz transformation, then

$$\Lambda^{-1} \,, \quad \Lambda^t \,, \quad (\Lambda^t)^{-1}$$

are also Lorentz transformations.

Note that the defining relation of the Lorentz group implies that

$$\eta_{\mu\nu} \Lambda_0^\mu \Lambda_0^\nu = (\Lambda_0^0)^2 - (\Lambda_0^i)^2 = 1 \,,$$

that is $(\Lambda_0^0)^2 = 1 + (\Lambda_0^i)^2 \geq 1$. Thus, for any Lorentz transformation either $\Lambda_0^0 \geq 1$ or $\Lambda_0^0 \leq -1$.

Topological structure of the Lorentz group

The Lorentz group is a 6-dimensional non-compact Lie group $\mathrm{O}(1,3)$ which consists of four connected components (four topologically separated pieces), each of them is not simply connected, see Fig. 5.3. To understand this topological structure of the Lorentz group, let us notice that a Lorentz transformation may or may not

- reverse the direction of time (or more precisely, transform a future-pointing time-like vector into a past-pointing one),

[2] In other words, group elements of a Lie group can be continuously parametrised by a set of parameters.

[3] Would η be identity matrix, then the relation $\Lambda^t \Lambda = 1$ would define the group of orthogonal transformations.

[4] It follows from $\Lambda^{-1} = \eta \Lambda^t \eta$ by multiplying it from the right with Λ that the following relation is also true $\Lambda \eta \Lambda^t = \eta$. This shows that matrix Λ^t is also a Lorentz transformation.

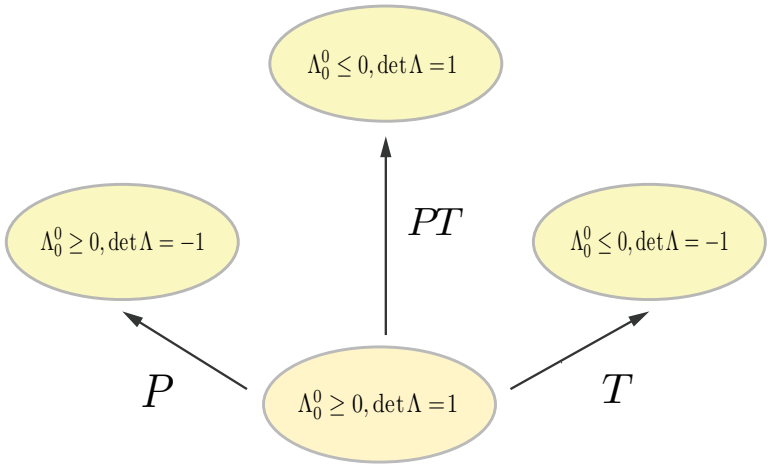

Figure 5.3: Four connected components of the Lorentz group. The component with $\Lambda_0^0 \leq 1$ and $\det\Lambda = 1$ is a subgroup of proper orthochronous transformations $\mathrm{SO}^+(1,3)$ (the restricted Lorentz group).

- reverse the orientation of a four-dimensional reference frame.

Lorentz transformations with $\Lambda_0^0 > 0$ preserve the direction of time and are called *orthochronous*. The product of two orthochronous transformations is also an orthochronous transformation. To see this, we notice that $(\Lambda_0^0)^2 = 1 + (\Lambda_0^i)^2 \geq 1$ implies that $|\Lambda_0^0| > ||\Lambda_0^i||$ and analogously, by changing $\Lambda \to \Lambda^t$, one gets $|\Lambda_0^0| > ||\Lambda_i^0||$, where Λ_0^i and Λ_i^0 are understood as vectors with components $i = 1, 2, 3$. For a product of two transformations Λ and Λ' one has

$$(\Lambda\Lambda')_0^0 = \Lambda_0^0\Lambda_0'^0 + \Lambda_0^i\Lambda_i'^0 \,.$$

By the Cauchy-Bunyakovsky-Schwarz inequality[5], one obtains that

$$|\Lambda_0^i\Lambda_i'^0| \leq ||\Lambda_0^i||\,||\Lambda_i'^0|| < |\Lambda_0^0||\Lambda_0'^0| = \Lambda_0^0\Lambda_0'^0 \,.$$

Hence, $(\Lambda\Lambda')_0^0 > 0$ if both Λ_0^0 and $\Lambda_0'^0$ are positive. The subgroup of orthochronous transformations is often denoted $\mathrm{O}^+(1,3)$.

Lorentz transformations which preserve orientation are called *proper*, and as linear transformations they have determinant +1. (The improper Lorentz transformations have determinant -1.) The subgroup of proper Lorentz transformations is denoted $\mathrm{SO}(1,3)$.

The identity component of the Lorentz group, *i.e.* the component containing the identity element, is the set of all Lorentz transformations preserving both orientation and the direction of time. It is the proper, orthochronous Lorentz group, which is sometimes also called the *restricted Lorentz group* $\mathrm{SO}^+(1,3)$.

Every element in $\mathrm{O}(1,3)$ can be written as the semidirect product of a proper, orthochronous transformation and an element of the discrete group

$$\{1, P, T, PT\}$$

[5]For any two vectors x and y: $|(x,y)| \leq ||x||||y||$.

where P and T are the space inversion and time reversal operators:

$$\begin{aligned} P &= \mathrm{diag}(1,-1,-1,-1) \\ T &= \mathrm{diag}(-1,1,1,1) \end{aligned}$$

The four elements of this isomorphic copy of the Klein four-group label the four connected components of the Lorentz group.

As stated above, the restricted Lorentz group is the identity component of the Lorentz group. This means that it consists of all Lorentz transformations which can be connected to the identity by a continuous curve lying in the group. The restricted Lorentz group is a connected normal subgroup[6] of the full Lorentz group with the same dimension (in this case, 6 dimensions).

Structure of Lorentz transformations: spatial rotations and boosts

Introduce two four-vectors in the original and a Lorentz-transformed coordinate systems, respectively,

$$x = \begin{pmatrix} x^0 \\ \vec{x} \end{pmatrix}, \qquad x' = \begin{pmatrix} x'^0 \\ \vec{x}' \end{pmatrix}.$$

The relation is $x' = \Lambda x$ and $x = \Lambda^{-1} x'$. In what follows it is convenient to parametrize

$$\Lambda = \begin{pmatrix} a & v_1^t \\ v_2 & S \end{pmatrix}, \qquad \Lambda^t = \begin{pmatrix} a & v_2^t \\ v_1 & S^t \end{pmatrix}, \qquad \Lambda^{-1} = \eta \Lambda^t \eta = \begin{pmatrix} a & -v_2^t \\ -v_1 & S^t \end{pmatrix}.$$

Here a is a scalar, v_1 and v_2 are vectors and S is a 3×3 matrix. We recall that a matrix Λ of Lorentz transformation satisfies the conditions $\Lambda^t \eta \Lambda = \eta$ and, as a consequence, $\Lambda \eta \Lambda^t = \eta$. In particular, $\Lambda^t \eta \Lambda = \eta$ implies

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & v_2^t \\ v_1 & S^t \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & v_1^t \\ v_2 & S \end{pmatrix} = \begin{pmatrix} a^2 - v_2^2 & a v_1^t - v_2^t S \\ a v_1 - S^t v_2 & v_1 \otimes v_1^t - S^t S \end{pmatrix}.$$

Thus, we find three conditions

$$a^2 - v_2^2 = 1\,, \quad a v_1^t - v_2^t S = 0\,, \quad v_1 \otimes v_1^t - S^t S = -1\,.$$

The change $\Lambda \to \Lambda^t$ gives

$$a^2 - v_1^2 = 1\,, \quad a v_2 - S^t v_1 = 0\,, \quad v_2 \otimes v_2^t - S^t S = -1\,.$$

Now we are going to clarify the meaning of the vectors v_1 and v_2 and the matrix S. To this end, consider the transformation $x = \Lambda^{-1} x'$. Explicitly, it is

$$\begin{aligned} x^0 &= a x'^0 - (\vec{v}_2 \vec{x}')\,, \\ \vec{x} &= -\vec{v}_1 x'^0 + S^t \vec{x}'. \end{aligned}$$

In the moving coordinate system M', it's origin O' has coordinates $\vec{x'} = 0$, therefore, the formulae before takes the form

$$\begin{aligned} x^0 &= a x'^0\,, \\ \vec{x} &= -\vec{v}_1 x'^0\,. \end{aligned}$$

[6] A subgroup $N \subset G$ is called normal, if $gNg^{-1} \subset N$ for any $g \in G$.

Dividing second formula by the first, we get $\frac{\vec{x}}{x^0} = \frac{\vec{v}}{c} = -\frac{\vec{v}_1}{a}$, where $\vec{v}$ is the velocity of O' with respect to the stationary coordinate system M. Thus, $\vec{v}_1 = -a\frac{\vec{v}}{c}$. Further, from the condition $a^2 - v_1^2 = 1$ it follows that

$$a = \pm \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \,.$$

We chose "+" sign here which corresponds considering orthochronous transformations $\Lambda_0^0 \geq 1$.

Now we turn our attention to the equation

$$S^t S = 1 + v_1 \otimes v_1^t \,.$$

Explicitly, the matrix $1 + v_1 \otimes v_1^t$ has the following matrix elements

$$(1 + v_1 \otimes v_1^t)_{ij} = \delta_{ij} + (v_1)_i (v_1)_j \,.$$

Consider for the moment another matrix

$$Q \equiv (1 + \alpha\, v_1 \otimes v_1^t)_{ij} = \delta_{ij} + \alpha\, (v_1)_i (v_1)_j \,, .$$

where α is an arbitrary number. Compute its squire

$$Q_{ij}^2 = (\delta_{ik} + \alpha\, (v_1)_i (v_1)_k)(\delta_{kj} + \alpha\, (v_1)_k (v_1)_j) = \delta_{ij} + \big(2\alpha + \alpha^2 v_1^2\big)(v_1)_i (v_1)_j \,.$$

Thus, we see that if we subject the coefficient α to the condition

$$\alpha^2 (a^2 - 1) + 2\alpha = 1 \,,$$

then the following property will be satisfied

$$1 + v_1 \otimes v_1^t = Q^2 \,.$$

Solving the quadratic equation for α, one finds

$$\alpha = \frac{1}{1 \pm a} \,.$$

We pick the solution with "+" and denote the corresponding Q by Q_+. Explicitly,

$$Q_{+ij} = \delta_{ij} + \frac{1}{1+a}\frac{a^2}{c^2} v_i v_j = \delta_{ij} + \frac{a-1}{v^2} v_i v_j \,.$$

The relation $S^t S = Q_+^2$ can be written as $Q_+^{-1} S^t S Q_+^{-1} = (S Q_+^{-1})^t (S Q_+^{-1}) = 1$, since Q_+ is a symmetric matrix. Hence $R = S Q_+^{-1}$ is an orthogonal matrix, as $R^t R = 1$. Furthermore, since now $S = R Q_+$, we get that

$$v_2^t = \frac{1}{a} v_1^t S^t = \frac{1}{a} v_1^t Q_+ R^t = \frac{1}{a} v_1^t (1 + \alpha\, v_1 \otimes v_1^t) R^t = \frac{1}{a}(1 + \alpha(a^2 - 1)) v_1^t R^t = v_1^t R^t \,,$$

that is $v_2 = R v_1$. To summarize, we have established that generic matrix Λ has the following structure

$$\Lambda = \begin{pmatrix} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} & -\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{v^t}{c} \\ -\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} R \frac{v}{c} & R Q_+ \end{pmatrix} .$$

We immediately see that this 4×4 matrix factorizes into a product of the following matrices

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{v^t}{c} \\ -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{v}{c} & \mathbb{1} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{v \otimes v^t}{v^2} \end{pmatrix},$$

where $\mathbb{1}$ signifies 3×3 identity matrix. The first matrix is just an orthogonal transformation of a three-dimensional vector of spacial coordinates, while the second matrix is the Lorentz boost. With this matrix Λ at hand, we find for $x' = \Lambda x$ the following explicit formulae, where we use that $x^0 = ct$ and $x'^0 = ct'$,

$$\begin{aligned} t' &= \frac{t - \frac{(\vec{x}\vec{v})}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}, \\ \vec{x}' &= R\left[\vec{x} - \frac{\vec{v}t}{\sqrt{1-\frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{\vec{v}(\vec{v}\vec{x})}{v^2}\right]. \end{aligned} \tag{II.5.2}$$

These are Lorentz transformations[7] which describe how coordinates $(\vec{x}, t)$ of an even in a stationary reference frame transform to coordinates $(\vec{x}', t')$ of a reference frame which moves with respect to the stationary frame with an arbitrary velocity $\vec{v}$. Note that for $c \to \infty$, *i.e.* when $v \ll c$, the factor $\sqrt{1-\frac{v^2}{c^2}} \to 1$ and the Lorentz transformations reduce to the Galilean ones:

$$\begin{aligned} t' &= t \\ \vec{x}' &= R(x - vt). \end{aligned}$$

Inverse Lorentz transformations are obtained from $x = \Lambda^{-1}x'$, but they can be alternatively obtained from direct transformations above by changing primed indices for unprimed and changing the sign of velocity $\vec{v}$. One obtains

$$\begin{aligned} t &= \frac{t' + \frac{(\vec{x}'\vec{v})}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}, \\ \vec{x} &= R\left[\vec{x}' + \frac{\vec{v}t'}{\sqrt{1-\frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{\vec{v}(\vec{v}\vec{x}')}{v^2}\right]. \end{aligned} \tag{II.5.3}$$

It is of interest to see what the second solution with $\alpha = \frac{1}{1-a}$ gives. Denoting the corresponding Q by Q_-, we get

$$S^t S = Q_-^2 = Q_+^2$$

or

$$Q_-^{-1} S^t S Q_-^{-1} = Q_+^{-1} S^t S Q_+^{-1} = 1$$

which gives rise to two orthogonal matrices $R_+ = SQ_+^{-1}$ and $R_- = SQ_-^{-1}$. Obviously, R_+ and R_- differ from each other by an orthogonal transformation $R_-^{-1}R_+$. The nature of this transformation

[7] Regrouping terms, the expression for x' can be also written in the following form

$$\vec{x}' = R\left[\frac{[\vec{v}, [\vec{x}, \vec{v}]]}{\sqrt{1-\frac{v^2}{c^2}}} - \vec{v}\frac{t - \frac{(\vec{v}\vec{x})}{v^2}}{\sqrt{1-\frac{v^2}{c^2}}}\right].$$

can be understood by computing its determinant

$$\det(R_-^{-1}R_+) = \det(Q_- S^{-1} S Q_+^{-1}) = \frac{\det Q_-}{\det Q_+}\,.$$

The direct computation shows that for $Q(\alpha)$, the corresponding determinant is $\det Q(\alpha) = 1+\alpha v^2 = 1+\alpha(a^2-1)$. Thus, $\det Q_+ = a$ and $\det Q_- = -a$, so that $\det(R_-^{-1}R_+) = -1$ contains a reflection of the coordinate axes. Hence, we see that the choice of $\pm a$ and $\alpha = \frac{1}{1\pm a}$ precisely give rise to four connected components of the Lorentz group.

Note that the simplest example of the Lorentz transformation is a rotation in the tx-plane. This rotation must leave the interval $(ct)^2 - x^2$ invariant. The relation between the old and the new coordinates is described by the formulas

$$x = x'\cosh\psi + ct'\sinh\psi\,, \qquad ct = x'\sinh\psi + ct'\cosh\psi\,.$$

Indeed,

$$(ct)^2 - x^2 = (x'\sinh\psi + ct'\cosh\psi)^2 - (x'\cosh\psi + ct'\sinh\psi)^2 = (ct')^2 - x'^2\,.$$

Substituting here the coordinate $x' = 0$ of the center of the moving system, we get

$$x = ct'\sinh\psi\,, \qquad ct = ct'\cosh\psi \quad\Longrightarrow\quad \tanh\psi = \frac{x}{ct} = \frac{v}{c}\,.$$

From here we find

$$\sinh\psi = \frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}}\,, \qquad \cosh\psi = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\,.$$

and, therefore,

$$x = \frac{x'+vt'}{\sqrt{1-\frac{v^2}{c^2}}}\,, \qquad y = y'\,, \qquad z = z'\,, \qquad t = \frac{t'+\frac{v}{c^2}x'}{\sqrt{1-\frac{v^2}{c^2}}}\,,$$

This transformation is called the *Lorentz boost* as it describes the change of coordinates and time due to boosting one coordinate system with respect to the other. The reader can verify that this particular example fits our general discussion of arbitrary Lorentz transformations.

Addition of velocities

Suppose in the moving frame M' a particle is moving with velocity $\vec{u}$, that is $\vec{u} = \frac{d\vec{x}'}{dt'}$. We want to find its velocity in the stationary frame M. To this end, we consider the differentials of the inverse Lorentz transformations

$$dt = \frac{dt' + \frac{(d\vec{x}'\vec{v})}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} = dt'\frac{1+\frac{(\vec{u}\vec{v})}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}\,,$$

$$d\vec{x} = R\left[d\vec{x}' + \frac{\vec{v}dt'}{\sqrt{1-\frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{\vec{v}(\vec{v}d\vec{x}')}{v^2}\right]\,.$$

Dividing the differential dx by dt, we find the velocity $\vec{w}$ in the stationary frame ($R = 1$)

$$\vec{w} = \frac{d\vec{x}}{dt} = \frac{\vec{u} + \frac{\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{\vec{v}(\vec{v}\vec{u})}{v^2}}{\frac{1+\frac{(\vec{u}\vec{v})}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}}\,.$$

This is the law for addition of velocities in the relativistic case. In the non-relativistic limit $c\to\infty$, it reduces to the Galilean law: $\vec{w} = \vec{u} + \vec{v}$.

Lie algebra of the Lorentz group

First we recall the basic facts about the rotation group in three dimensions and then concentrate our attention on certain aspects of the Lorentz group.

Any rotation has the form

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad \text{or} \qquad r' = R\, r\,.$$

Under rotations the distance to the origin remains unchanged, that is

$$x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2\,, \qquad \text{or} \qquad r'^t r' = r^t r\,.$$

This means that

$$r^t R^t R r = r^t r \qquad i.e. \qquad R^t R = 1\,.$$

This means that R is an orthogonal 3×3 matrix. Orthogonal matrices form a group called O(3).

Rotation of a vector on a finite angle θ around z-axis is

$$\begin{pmatrix} V'_x \\ V'_y \\ V'_z \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix}$$

so that

$$R_z(\theta) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Analogously, the rotation matrices around the axes x and y have the form

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix}, \qquad R_y(\psi) = \begin{pmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{pmatrix}.$$

These matrices do not commute between themselves:

$$R_z(\theta) R_x(\phi) \neq R_x(\phi) R_z(\theta)\,.$$

This means that the rotation group is a *non-abelian* group. That is also a Lie group, *i.e.* a continuous group with infinite number of elements, because the values of the group parameters (angles) form a continuum. Any rotation is determined by three parameters: the matrix R has 9 elements and the relation $R^t R = 1$ imposes on them 6 conditions. These three parameters can be chosen to be the Euler angles. Three parameters give rise to three generators defined as

$$J_z = \frac{1}{i}\frac{dR_z(\theta)}{d\theta}|_{\theta=0} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$J_x = \frac{1}{i}\frac{dR_x(\phi)}{d\phi}|_{\phi=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix},$$

$$J_y = \frac{1}{i}\frac{dR_y(\psi)}{d\psi}|_{\psi=0} = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}.$$

These generators are hermitian. The infinitezimal rotations are given by

$$R_z(\delta\theta) = 1 + iJ_z\delta\theta\,, \qquad R_x(\delta\phi) = 1 + iJ_x\delta\phi\,, \qquad R_y(\delta\psi) = 1 + iJ_y\delta\psi\,.$$

Commutators of two generators

$$[J_x, J_y] = iJ_z + \text{cyclic permutations}$$

coincide with the commutation relations of angular momentum. Rotation on a finite angle around z-axis is

$$R_z(\theta) = e^{iJ_z\theta} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If one considers a rotation around an arbitrary axis $\vec{n}$, then

$$R_{\vec{n}}(\theta) = e^{i(\vec{J}\cdot\vec{n})\theta}\,.$$

Now we turn our attention to the Lorentz group. Mathematically, the Lorentz group may be described as the generalized orthogonal group $\mathrm{O}(1,3)$, the matrix Lie group which preserves the quadratic form

$$(ct, x, y, z) \to (ct)^2 - x^2 - y^2 - z^2\,.$$

This quadratic form is the *metric* tensor of Minkowski spacetime, so this definition is simply a restatement of the fact that Lorentz transformations are precisely the linear transformations which are also isometries of Minkowski spacetime.[8]

The restricted Lorentz group is generated by ordinary spatial rotations and Lorentz boosts (which can be thought of as hyperbolic rotations in a plane that includes a time-like direction). The set of all boosts, however, does not form a subgroup, since composing two boosts does not, in general, result in another boost. Indeed, introducing the identification

$$x^0 = ct\,, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z$$

we can write the Lorentz boost as

$$\begin{pmatrix} x^{0'} \\ x^{1'} \\ x^{2'} \\ x^{3'} \end{pmatrix} = \begin{pmatrix} \cosh\varphi & \sinh\varphi & 0 & 0 \\ \sinh\varphi & \cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

The generator corresponding to the infinitezimal boost is defined as

$$K_x = \frac{1}{i}\frac{dB_x(\varphi)}{d\varphi}|_{\varphi=0} = \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The other boost generators are

$$K_y = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad K_z = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}.$$

[8]The Lorentz group is a subgroup of the Poincaré group, the group of all isometries of Minkowski spacetime. The Lorentz transformations are precisely the isometries which leave the origin fixed. Thus, the Lorentz group is an isotropy subgroup of the isometry group of Minkowski spacetime. For this reason, the Lorentz group is sometimes called the homogeneous Lorentz group while the Poincaré group is sometimes called the inhomogeneous Lorentz group.

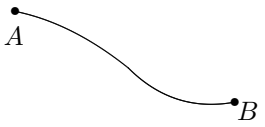

Figure 5.4: The simplest form of action is given by the length of the space-time interval between points A and B.

The set of all rotations forms a Lie subgroup isomorphic to the ordinary rotation group SO(3). The usual rotation generators now look like

$$J_x = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \qquad J_y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \qquad J_z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

One can compute the commutators

$$\begin{aligned} &[K_x, K_y] = -iJ_z + \text{cyclic permutations} \\ &[J_x, K_x] = [J_y, K_y] = [J_z, K_z] = 0 \\ &[J_x, K_y] = iK_z + \text{cyclic permutations} \end{aligned} \tag{II.5.4}$$

Boosts do not form a group; commutator of two boosts is a rotation.

A boost in some direction, or a rotation about some axis, each generate a one-parameter subgroup. An arbitrary rotation is specified by 3 real parameters, as is an arbitrary boost. Since every proper, orthochronous Lorentz transformation can be written as a product of a rotation and a boost, it takes 6 real numbers (parameters) to specify an arbitrary proper orthochronous Lorentz transformation.

The 6 generators K and J can be combined into one skew-symmetric matrix M_{ab} with the following commutation relations

$$[M_{\mu\nu}, M_{\rho\lambda}] = i(\eta_{\mu\rho}M_{\nu\lambda} - \eta_{\nu\rho}M_{\mu\lambda} - \eta_{\mu\lambda}M_{\nu\rho} + \eta_{\nu\lambda}M_{\mu\rho})$$

representing the Lie algebra relations of the Lorentz group.

5.3 Relativistic particle

Let us first revisit some of the basics of special relativity written using tensor notation. The Minkowski metric $\eta_{\mu\nu}$ that we will use has the signature $(+,-,-,-)$ and we will use the convention that the Latin indices run only over the space coordinates (i.e. $i, j, k... = 1, 2, 3$), whereas the Greek indices will include both time and space coordinates (*i.e.* $\mu, \nu, \sigma, \rho... = 0, 1, 2, 3$). Additionally, in special relativity we will have to distinguish between 3-vectors (those with only space components) and 4-vectors (having both space and time components). The convention that we will use is that $\vec{A}$ will denote a 3-vector, whereas A^μ will denote a 4-vector.

Using these definitions, we can define the Lorentz invariant relativistic interval given by the expression

$$ds^2 = dx_\mu dx^\mu = c^2dt^2 - \left(dx^i\right)^2 . \tag{II.5.5}$$

The action for a relativistic particle has the following form

$$S = -\alpha \int_a^b \sqrt{ds^2} = -\alpha \int_a^b ds\,.$$

Reparametrisation invariance. Suppose we parametrise the trajectory of a particle with a parameter τ: $x^\mu = x^\mu(\tau)$. Then the action can be written as

$$S = -\alpha \int_{\tau_1}^{\tau_2} \sqrt{\dot{x}_\mu \dot{x}^\mu}\, d\tau\,, \tag{II.5.6}$$

where $\dot{x}^\mu = \frac{dx^\mu}{d\tau}$. The new feature of this action in comparison to the case of non-relativistic mechanics is that this action is invariant under reparametrizations of τ:

$$\delta x^\mu = \xi(\tau)\partial_\tau x^\mu \quad \text{ as long as } \quad \xi(\tau_1) = \xi(\tau_2) = 0\,.$$

Let us show this

$$\begin{aligned}
\delta(\sqrt{\dot{x}_\mu \dot{x}^\mu}) &= \frac{1}{2\sqrt{\dot{x}_\mu \dot{x}^\mu}}(2\dot{x}^\nu \delta\dot{x}_\nu) = \frac{1}{\sqrt{\dot{x}_\mu \dot{x}^\mu}}\dot{x}^\nu \partial_\tau(\xi \dot{x}_\nu) = \\
&= \frac{1}{\sqrt{\dot{x}_\mu \dot{x}^\mu}}\Big[\dot{x}^\nu \dot{x}_\nu \dot{\xi} + \xi \dot{x}^\nu \ddot{x}_\nu\Big] = \frac{1}{\sqrt{\dot{x}_\mu \dot{x}^\mu}}\dot{x}^\nu \dot{x}_\nu \dot{\xi} + \xi\partial_\tau(\sqrt{\dot{x}_\mu \dot{x}^\mu}) \\
&= \sqrt{\dot{x}_\mu \dot{x}^\mu}\dot{\xi} + \xi\partial_\tau(\sqrt{\dot{x}_\mu \dot{x}^\mu}) = \partial_\tau(\xi\sqrt{\dot{x}_\mu \dot{x}^\mu})\,.
\end{aligned}$$

Therefore, we arrive at

$$\delta S = -\alpha \int_{\tau_1}^{\tau_2} \mathrm{d}\tau\, \partial_\tau(\xi\sqrt{\dot{x}_\mu \dot{x}^\mu}) = -\alpha\Big[\xi\sqrt{\dot{x}_\mu \dot{x}^\mu}\Big]|_{\tau=\tau_1}^{\tau=\tau_2} = 0\,,$$

i.e. the action is indeed invariant w.r.t. the *local* reparametrisation transformations. Physically, the reparametrisation invariance means that the value of the action does not depend on with which velocity one runs over the trajectory connecting the initial and final points. This independence is due to the fact that the action S is a true geometric object, namely, the length of the trajectory between the initial and final points.

Non-relativistic limit. Since the action does not depend on the choice of a local parameter, we can pick up one. The convenient choice is the so-called static choice, also called *static gauge*, where the variable $t \equiv \tau$ is chosen to parametrise a trajectory as

$$x^0 = ct\,, \qquad x^i = x^i(t)\,.$$

With this choice we can now establish the physical meaning of the parameter α. First, we have

$$\frac{dx^\mu}{dt} = (c, \vec{v})\ , \qquad ds = \sqrt{c^2 - \vec{v}^2} = c\sqrt{1 - \frac{\vec{v}^2}{c^2}}\,.$$

Therefore, the action becomes

$$S = -\alpha c \int_{t_1}^{t_2} \sqrt{1 - \frac{\vec{v}^2}{c^2}}\, dt\,,$$

which corresponds to the following lagrangian

$$L = -\alpha c\sqrt{1 - \frac{\vec{v}^2}{c^2}}\,. \tag{II.5.7}$$

The non-relativistic limit corresponds to taking $\frac{\vec{v}^2}{c^2} \ll 1$. Expanding the lagrangian in this limit, we find

$$L \approx -\alpha c\left(1 - \frac{\vec{v}^2}{2c^2} + \cdots\right) \approx -\alpha c + \alpha\frac{\vec{v}^2}{2c}\,. \tag{II.5.8}$$

If we want to recover the usual form of the lagrangian $L = T - U$ where for of a free non-relativistic particle $T = \frac{m\vec{v}^2}{2}$ is the kinetic energy and $U = 0$, we need to set $\alpha = mc$. When we do so, equation (II.5.8) turns into

$$L = -mc^2 + \frac{1}{2}m\vec{v}^2 \,.$$

Mass-shell condition. Returning back to the action in generic parametrisation

$$S = -mc \int_{\tau_1}^{\tau_2} \sqrt{\dot{x}_\mu \dot{x}^\mu}\, d\tau \,, \tag{II.5.9}$$

so that L is

$$L = -mc\sqrt{\dot{x}_\mu \dot{x}^\mu} \,.$$

As we argue below by comparison to the description in the static gauge, the canonical *4-momentum* p^μ should be defined as the derivative of L with respect to $-\dot{x}_\mu$. We get

$$p^\mu = -\frac{\partial L}{\partial \dot{x}^\mu} = mc\frac{\dot{x}^\mu}{\sqrt{\dot{x}_\nu \dot{x}^\nu}} \,.$$

Now when we take

$$p^2 \equiv p_\mu p^\mu = m^2c^2 \frac{\dot{x}_\mu \dot{x}^\mu}{\left(\sqrt{\dot{x}_\nu \dot{x}^\nu}\right)^2} = m^2c^2 \,.$$

Hence, the particle trajectories which minimize the action must satisfy the constraint $p^2 - m^2c^2 = 0$, which is referred to as *the mass-shell condition*.

In the static gauge $t = \tau$, we have $x^0 = ct$ and, therefore, the components of the 4-momentum are

$$p^\mu = \left(\frac{mc}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}}, \frac{m\vec{v}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \right) \,. \tag{II.5.10}$$

Here

$$p^0 = \frac{E}{c} = \frac{mc}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \,, \tag{II.5.11}$$

where

$$E = \frac{mc^2}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \tag{II.5.12}$$

is the energy of the relativistic particle and

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \tag{II.5.13}$$

is its 3-momentum. The mass-shell condition takes the form

$$p_0 p^0 - \vec{p}^2 = m^2c^2 \quad \Longrightarrow \quad \frac{E^2}{c^2} - \vec{p}^2 = m^2c^2 \,.$$

Lagrangian and hamiltonian dynamics in the static gauge. Justification of the definitions above comes from consideration of the action in the static gauge, which is

$$S = -mc^2 \int dt \sqrt{1 - \frac{\vec{v}^2}{c^2}} \,. \tag{II.5.14}$$

This action involves the usual 3-velocity $\vec{v} = \frac{d\vec{x}}{dt}$. Thus, according to the canonical formalism, we have

$$\vec{p} = \frac{\partial L}{\partial \vec{v}} = \frac{m\vec{v}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \,. \tag{II.5.15}$$

The Euler-Lagrange equations are

$$\frac{d}{dt}\left(\frac{m\vec{v}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \right) = 0 \,. \tag{II.5.16}$$

Let us solve this equation. We have

$$\frac{d}{dt}\left(\frac{m\vec{v}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \right) = \frac{m\dot{\vec{v}}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} + \frac{m\vec{v}(\vec{v}\cdot\dot{\vec{v}})/c^2}{(1 - \frac{\vec{v}^2}{c^2})^{3/2}} = 0 \,. \tag{II.5.17}$$

Multiplying this equation with $\vec{v}$, we get

$$\frac{m(v\cdot\dot{\vec{v}})}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} + \frac{m\vec{v}^2/c^2(\vec{v}\cdot\dot{\vec{v}})}{(1 - \frac{\vec{v}^2}{c^2})^{3/2}} = \frac{m(\vec{v}\cdot\dot{\vec{v}})}{(1 - \frac{\vec{v}^2}{c^2})^{3/2}} = 0 \,,$$

from where we conclude that $(\vec{v}\cdot\dot{\vec{v}}) = 0$. Substituting this result into (II.5.17), we get

$$\frac{m\dot{\vec{v}}}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} = 0 \,,$$

which yields $\dot{\vec{v}} = 0$. Thus, a free relativistic particle moves with a constant velocity. The hamiltonian is

$$H = \vec{p}\vec{v} - L = \frac{m\vec{v}^2}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} + mc^2\sqrt{1 - \frac{\vec{v}^2}{c^2}} = \frac{mc^2}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \,. \tag{II.5.18}$$

To write the Hamiltonian as the function of $\vec{p}$, we take

$$\vec{p}^2 = \frac{m^2\vec{v}^2}{1 - \frac{\vec{v}^2}{c^2}} \qquad \longrightarrow \qquad \vec{v}^2 = \frac{\vec{p}^2c^2}{\vec{p}^2 + m^2c^2}$$

and substituting this expression for $\vec{v}^2$ into the expression for H, we find

$$H = c\sqrt{\vec{p}^2 + m^2c^2} \,. \tag{II.5.19}$$

This is the hamiltonian of the free relativistic particle.

5.4 Relativistic particle in electromagnetic field

Let us now define the vector potential, which is an underlying field (a Lorentz invariant 4-vector) in electrodynamics that we will base our further derivations on. It reads

$$A^\mu = \left(\varphi(x), \vec{A}(x) \right) \,.$$

Notice that

$$A^\mu \to A_\mu = \eta_{\mu\nu} A^\nu = \left(\varphi\left(x\right), -\vec{A}\left(x\right)\right) . \tag{II.5.20}$$

The properties of a charged particle with respect to its interaction with electromagnetic field are characterized by a single parameter: the electric charge e. The properties of the electromagnetic field itself are determined by the vector A^μ, the electromagnetic potential introduced above. Using these quantities, one can introduce the action of a charged particle in electromagnetic field as follows

$$S = -mc \int_a^b ds - \frac{e}{c} \int A_\mu dx^\mu .$$

Using Hamilton's principle, stating that a particle follows the path that extremises the action ($\delta S = 0$), we can derive the equations of motion in which we neglect the back reaction of the charge on the electromagnetic field

$$0 = \delta S \quad = \quad -mc \int \frac{\mathrm{d}x_\mu d\delta x^\mu}{\sqrt{\mathrm{d}x_\nu \mathrm{d}x^\nu}} - \frac{e}{c} \int \left[(\delta A_\mu) dx^\mu + A_\mu d(\delta x^\mu)\right] , \tag{II.5.21}$$

where we used that $ds = \sqrt{\mathrm{d}x_\nu \mathrm{d}x^\nu}$. We can further take the length s itself as a parameter along the trajectory and, therefore, define the so-called *4-velocity* $U^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}s}$. The explicit form of U^μ is

$$U^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}s} = \frac{\mathrm{d}x^\mu}{c\sqrt{1 - \frac{\vec{v}^2}{c^2}}\mathrm{d}t} = \left(\frac{1}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}}, \frac{\vec{v}}{c\sqrt{1 - \frac{\vec{v}^2}{c^2}}}\right) . \tag{II.5.22}$$

and it has an important property that

$$U_\mu U^\mu = \frac{\mathrm{d}x_\mu}{\mathrm{d}s} \frac{\mathrm{d}x^\mu}{\mathrm{d}s} = 1 .$$

Using the fact that $\delta A_\mu = A_\mu(x_\nu + \delta x_\nu) - A_\mu(x_\nu) = \partial_\nu A_\mu \delta x^\nu + \cdots$ and differentiating by parts, we can rewrite equation (II.5.21) as follows

$$\delta S = mc \int \mathrm{d}U_\mu \delta x^\mu + \frac{e}{c} \int \left(\partial_\nu A_\mu \mathrm{d}x^\nu \delta x^\mu - \partial_\nu A_\mu \delta x^\nu \mathrm{d}x^\mu\right) = 0 .$$

This imposes the following condition for the extremum

$$mc \frac{\mathrm{d}U_\mu}{\mathrm{d}s} + \frac{e}{c} \left(\partial_\nu A_\mu - \partial_\mu A_\nu\right) U^\nu = 0 .$$

Introducing the tensor $F_{\mu\nu}$ of the electromagnetic field

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = -F_{\nu\mu} , \tag{II.5.23}$$

we can write the equation of motion of the charge in the electromagnetic field as follows

$$mc \frac{\mathrm{d}U^\mu}{\mathrm{d}s} = \frac{e}{c} F^{\mu\nu} U_\nu . \tag{II.5.24}$$

This expression can also be written in a more suggestive form if we define the momentum[9] $p^\mu = mcU^\mu$, so that one can express the acceleration term $\frac{\mathrm{d}U^\mu}{\mathrm{d}s} = \frac{d^2 x^\mu}{\mathrm{d}s^2}$ as

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}s} = \frac{\mathrm{d}p^\mu}{\mathrm{d}t} \frac{\mathrm{d}t}{\mathrm{d}s} = \frac{e}{c} F^{\mu\nu} U_\nu , \tag{II.5.25}$$

[9]This is consistent with the requirement $p^2 = m^2 c^2$ since $U^2 = 1$.

Figure 5.5: In the presence of the vector potential $A^\mu = \left(\varphi, \vec{A}\right)$ the action of a charged particle contains an additional term describing an interaction with the vector potential.

where the right hand side of the equation is referred to as *the Lorentz force*, whereas the left hand side is simply the rate of change of momentum with respect to the relativistic interval. This equation is comparable with the Newtonian statement: force is the rate of change of momentum. Note that this derivation has assumed that the electromagnetic field is given (fixed) and that we vary the trajectory of the particle only (the endpoints remain fixed).

Electromagnetic tensor $F_{\mu\nu}$. Before we proceed with the discussion of the Lorentz force, let us understand a relation between $F_{\mu\nu}$ and electromagnetic fields $\vec{E}, \vec{H}$. First, we have

$$F_{0i} = \partial_0 A_i - \partial_i A_0 = \frac{\partial A_i}{\partial x^0} - \frac{\partial A_0}{\partial x^i} \,. \tag{II.5.26}$$

Taking into account (II.5.20) and $x^0 = ct$, we find that

$$F_{0i} = \left[-\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \right]_i = (\vec{E})_i \,,$$

where $\vec{E} = (E_x, E_y, E_z)$ is the electric field. Second, we consider

$$F_{ij} = \partial_i A_j - \partial_j A_i = -(\partial_i A^j - \partial_j A^i) \,.$$

Here we have to recall that $\vec{H} = \vec{\nabla} \times \vec{A}$ or in components $(\vec{H})_i = \epsilon_{imn} \partial_m \vec{A}_n$, where $\vec{A}_i$ are components of the 3-vector $\vec{A}$. Multiplying this relation with ϵ_{ijk} and summing over i, we get

$$(\vec{H})_i \epsilon_{ijk} = \epsilon_{ijk}\epsilon_{imn}\partial_m \vec{A}_n = \partial_j \vec{A}_k - \partial_k \vec{A}_j \,,$$

where we have used the summation formula

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km} \,.$$

Thus, we obtain

$$F_{ij} = \partial_i A_j - \partial_j A_i = -(\partial_i A^j - \partial_j A^i) = -\epsilon_{ijk}(\vec{H})_k \,, \qquad \vec{H} = (H_x, H_y, H_z) \,.$$

These formulae allow us to obtain an explicit formula for the tensor of the electromagnetic field, the latter being 4×4 matrix, in terms of components of electric and magnetic fields

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -H_z & H_y \\ -E_y & H_z & 0 & -H_x \\ -E_z & -H_y & H_x & 0 \end{pmatrix} \tag{II.5.27}$$

and, as a consequence,

$$F^{\mu\nu} = \eta^{\mu\sigma}\eta^{\nu\rho}F_{\sigma\rho} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -H_z & H_y \\ E_y & H_z & 0 & -H_x \\ E_z & -H_y & H_x & 0 \end{pmatrix} . \tag{II.5.28}$$

We have also established the following relations between electric and magnetic fields and the corresponding components of the 4-potential

$$\vec{E} = -\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \qquad \text{and} \qquad \vec{H} = \vec{\nabla} \times \vec{A}\,. \tag{II.5.29}$$

Lorentz force. Now we come back at the expression for the Lorentz force and write it in terms of electric and magnetic fields. Rearranging (II.5.25), we get

$$\begin{aligned} \frac{\mathrm{d}p^i}{\mathrm{d}t} &= \left(\frac{e}{c}F^{i0}U_0 + \frac{e}{c}F^{ij}U_j\right)\frac{\mathrm{d}s}{\mathrm{d}t} = \\ &= \left(\frac{e}{c}E_i\frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}} - \frac{e}{c}F_{ij}\frac{v^j}{c\sqrt{1-\frac{\vec{v}^2}{c^2}}}\right)c\sqrt{1-\frac{\vec{v}^2}{c^2}} = eE_i + \frac{e}{c}\epsilon_{ijk}v^jH_k\,. \end{aligned} \tag{II.5.30}$$

Here we used the fact that $F^{i0} = \vec{E}_i$, $F_{ij} = F^{ij}$ and $U_j = -\frac{v^j}{c\sqrt{1-\frac{v^2}{c^2}}}$. We can thus rewrite the equations of motion in the form

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = e\vec{E} + \frac{e}{c}\vec{v}\times\vec{H}\,. \tag{II.5.31}$$

This equation is of the fundamental nature and can be experimentally verified. It is an equation of motion for a changed particle in electromagnetic field. In the non-relativistic limit $\vec{p} = m\vec{v}$ the above equation turns into Newton's equation

$$m\frac{d\vec{v}}{dt} = e\vec{E} + \frac{e}{c}\vec{v}\times\vec{H}\,.$$

Concerning the equation for p^0, we have

$$\frac{dp^0}{dt} = \frac{e}{c}F^{0i}U_i\frac{ds}{dt} = -\frac{e}{c}E_i\left(-\frac{v^i}{c\sqrt{1-\frac{v^2}{c^2}}}\right)c\sqrt{1-\frac{\vec{v}^2}{c^2}} = \frac{e}{c}(\vec{E}\cdot\vec{v})\,.$$

This result is not independent but follows from (II.5.31). Indeed, since

$$p^0 = \frac{mc}{\sqrt{1-\frac{\vec{v}^2}{c^2}}}\,, \qquad E = E_{\mathrm{kin}} = cp^0\,, \tag{II.5.32}$$

we find[10]

$$\frac{dE_{\mathrm{kin}}}{\mathrm{d}t} = \frac{d}{dt}\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} = \vec{v}\cdot\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = e\big(\vec{E}\cdot\vec{v}\big)\,.$$

Further we note that the last formula represents the work of the electromagnetic field on the charge. Hence, the magnetic field does not change the kinetic energy, but rather only affects the direction of the particle trajectory!

Hamiltonian of particle in a static gauge. In a static gauge $t = \tau$, where t is a time measured by a non-moving (static) observer. In this gauge the action takes the form

$$S = \int L\,dt = \int dt\left[-mc^2\sqrt{1-\frac{v^2}{c^2}}\,\mathrm{d}t - \frac{e}{c}A_0\mathrm{d}x^0 - \frac{e}{c}A_i\mathrm{d}x^i\right]\,,$$

[10] We have

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \frac{m\dot{\vec{v}}}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{m\vec{v}}{\left(1-\frac{v^2}{c^2}\right)^{3/2}}\frac{(\vec{v}\cdot\dot{\vec{v}})}{c^2} \quad \longrightarrow \quad \vec{v}\cdot\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \frac{m(\vec{v}\cdot\dot{\vec{v}})}{\left(1-\frac{v^2}{c^2}\right)^{3/2}} = \frac{\mathrm{d}E_{\mathrm{kin}}}{\mathrm{d}t}\,.$$

i.e. the Lagrangian is

$$L = -mc^2\sqrt{1-\frac{v^2}{c^2}} + \frac{e}{c}\vec{A}\cdot\vec{v} - e\,\varphi$$

The momentum in the presence of the electromagnetic field is

$$\vec{P} = \frac{\partial L}{\partial \vec{v}} = \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{e}{c}\vec{A} = \vec{p} + \frac{e}{c}\vec{A}, \tag{II.5.33}$$

where $\vec{p}$ is the momentum of the free particle, and the hamiltonian

$$H = \frac{\partial L}{\partial \vec{v}}\vec{v} - L = \underbrace{\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}}_{\text{kinetic energy}} + \underbrace{e\,\varphi}_{\text{potential energy}}.$$

Expressing from eq.(II.5.33) the velocity $\vec{v}$ in terms of the canonical momentum $\vec{P}$, we find that

$$H = \sqrt{m^2c^4 + c^2\Big(\vec{P} - \frac{e}{c}\vec{A}\Big)^2} + e\,\varphi\,. \tag{II.5.34}$$

We stress that such an expression for the hamiltonian arises only due to our choice of the static gauge. In the absence of electromagnetic field (II.5.34) reduces to (II.5.19).

Gauge invariance. All the physical properties of the electromagnetic field as well as the properties of charge in the electromagnetic field are determined not by A_μ, but rather by $F_{\mu\nu}$. The underlying reason for this is that electrodynamics exhibits an important new type of symmetry[11]. To understand this issue, we may decide to change the vector potential in the following way

$$A_\mu \to A_\mu - \partial_\mu\chi\,, \tag{II.5.35}$$

which can be rewritten in a less abstract form of space and time components separately:

$$\vec{A} \to \vec{A} + \vec{\nabla}\chi \quad \text{and} \quad \varphi \to \varphi - \frac{1}{c}\frac{\partial\chi}{\partial t}\,. \tag{II.5.36}$$

These transformations are referred to as *the gauge transformations*. Let us see what effect they have on the tensor of the electromagnetic field:

$$\begin{aligned}\delta F_{\mu\nu} &= \partial_\mu\left(A_\nu + \partial_\nu\chi\right) - \partial_\nu\left(A_\mu + \partial_\mu\chi\right) - F_{\mu\nu} \\ &= \partial_\mu\partial_\nu\chi - \partial_\nu\partial_\mu\chi = 0\,.\end{aligned} \tag{II.5.37}$$

Thus, the transformation (II.5.35) does not change the form of the electromagnetic field tensor. For this reason electromagnetism is a gauge invariant theory!

The same conclusion on the gauge invariance can be achieved by inspecting the gauge invariance of the electric and magnetic fields $\vec{E}$ and $\vec{H}$ as expressed in terms of the scalar and vector potential

$$\vec{E} = -\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t} \qquad \text{and} \qquad \vec{H} = \vec{\nabla}\times\vec{A}\,. \tag{II.5.38}$$

One can easily see that in the first case an extra φ term cancels with an extra $\vec{A}$ term and in the second case we have the gauge transformation contribution vanishing due to the fact that $\vec{\nabla}\times\vec{\nabla}\chi = 0$.

[11] This symmetry extends to many other physical theories besides electrodynamics.

Chapter 6

Maxwell equations

"The velocity of light is one of the most important of the fundamental constants of Nature. Its measurement by Foucault and Fizeau gave as the result a speed greater in air than in water, thus deciding in favor of the undulatory and against the corpuscular theory. Again, the comparison of the electrostatic and the electromagnetic units gives as an experimental result a value remarkably close to the velocity of light – a result which justified Maxwell in concluding that light is the propagation of an electromagnetic disturbance. Finally, the principle of relativity gives the velocity of light a still greater importance, since one of its fundamental postulates is the constancy of this velocity under all possible conditions."

Albert Abraham Michelson, Studies in Optics

"From a long view of the history of mankind the most significant event of the nineteenth century will be judged as Maxwell's discovery of the laws of electrodynamics."

Richard Feynman

6.1 Fields produced by moving charges

Let us now consider the case where the moving particles produce the fields themselves. The new action will be then

$$S = S_{\rm p} + S_{\rm int} + S_{\rm f}\,,$$

where we have added a new term $S_{\rm f}$, which describes the action for electromagnetic field itself. The action of electrodynamics is

$$S = -mc\int ds - \frac{e}{c}\int A_\mu dx^\mu - \frac{1}{16\pi c}\int F_{\mu\nu}F^{\mu\nu}\,d^4x\,.$$

We recall that it is written in the Gauss system of units, where $\mu_0 = 4\pi$ and $\varepsilon_0 = \frac{1}{4\pi}$. Note that we can rewrite the second term as

$$\frac{e}{c}\int A_\mu dx^\mu \quad = \quad \frac{1}{c}\int \rho A_\mu dx^\mu dV = \frac{1}{c}\int \rho A_\mu \frac{dx^\mu}{dt} dV dt$$

$$= \frac{1}{c}\int j^\mu A_\mu dV dt = \frac{1}{c^2}\int j^\mu A_\mu \, d^4x \,, \tag{II.6.1}$$

where in the second line we have introduced, the current $j^\mu = \rho\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = (\rho c, \rho\vec{v})$. Including this, we can now write the action of the moving test charge as

$$S = -mc\int ds - \frac{1}{c^2}\int j^\mu A_\mu \, d^4x - \frac{1}{16\pi c}\int F_{\mu\nu}F^{\mu\nu} d^4x \,.$$

Keeping sources constant and the path unchanged (i.e. $\delta j^\mu = 0$ and $\delta s = 0$), we can write the deviation from the action as follows

$$\begin{aligned} \delta S &= -\frac{1}{c^2}\int j^\mu \delta A_\mu \, d^4x - \frac{1}{8\pi c}\int F^{\mu\nu}\delta F_{\mu\nu} d^4x \\ &= -\frac{1}{c}\left[\frac{1}{c}\int j^\mu \delta A_\mu \, d^4x + \frac{1}{4\pi}\int \frac{\partial F^{\mu\nu}}{\partial x^\nu}\delta A_\mu \mathrm{d}^4x\right], \end{aligned} \tag{II.6.2}$$

where in the last term in the first line, we have used that

$$\delta F_{\mu\nu} = \partial_\mu \delta A_\nu - \partial_\nu \delta A_\mu \,.$$

To find the extremum, we need to satisfy $\delta S = 0$, which due to eq.(II.6.2), is equivalent to the second pair of Maxwell's equations

$$\frac{\partial F^{\mu\nu}}{\partial x^\nu} = -\frac{4\pi}{c} j^\mu \,.$$

Identifying the respective components of the electromagnetic tensor we can rewrite the second pair of Maxwell's equations in a more familiar form

$$\vec{\nabla}\times\vec{H} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial\vec{E}}{\partial t} \qquad \text{and} \qquad \vec{\nabla}\cdot\vec{E} = 4\pi\rho \,, \tag{II.6.3}$$

where $\frac{4\pi}{c}\vec{j}$ and $4\pi\rho$ are the sources and $\frac{1}{c}\frac{\partial\vec{E}}{\partial t}$ is the so-called displacement current. The first expression is Ampére's law (also known as the Biot-Savart law), whereas the second one is the Gauss law. Finally, we notice that the covariant conservation of the current $\frac{\partial j^\mu}{\partial x^\mu} = 0$ is equivalent to the continuity equation

$$\frac{\partial\rho}{\partial t} + \mathrm{div}\vec{j} = 0 \,.$$

Electromagnetic duality. Below we include here a short digression on the tensor of the electromagnetic field. It is easy to check that, using the definition of the tensor, the following is true:

$$\begin{aligned} \mathrm{d}F &= \partial_\sigma F_{\mu\nu} + \partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} \\ &= \partial_\sigma(\partial_\mu A_\nu - \partial_\nu A_\mu) + \partial_\mu(\partial_\nu A_\sigma - \partial_\sigma A_\nu) + \partial_\nu(\partial_\sigma A_\mu - \partial_\mu A_\sigma) = 0 \,. \end{aligned}$$

With a change of indices, this takes the form

$$\varepsilon^{\mu\nu\sigma\rho}\frac{\partial F_{\nu\sigma}}{\partial x^\rho} = 0 \,, \tag{II.6.4}$$

which are four equations in disguise, since we are free to pick any value of the index μ. Let us introduce the so-called *dual electromagnetic tensor*

$$F^{*\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \,. \tag{II.6.5}$$

Then we can rewrite equation (II.6.4) as

$$\frac{\partial F^{*\mu\nu}}{\partial x^\nu} = 0\,. \tag{II.6.6}$$

These equations are nothing else but the 1st pair of Maxwell's equations.

Omitting the currents in the second pair, on can now see that the first and second pair of Maxwell's equations are similar. Indeed, we have

$$\begin{aligned} \text{1st pair}: \quad \frac{\partial F^{*\mu\nu}}{\partial x^\mu} &= 0\,, \\ \text{2nd pair}: \quad \frac{\partial F^{\mu\nu}}{\partial x^\mu} &= 0\,. \end{aligned}$$

The main difference between them is that the first pair never involves any currents:

- first pair of Maxwell's equations does not involve any density or current: $\rho, \vec{j}$;
- second pair of Maxwell's equations does involve the density and current: $\rho, \vec{j}$.

This distinction has a deeper meaning. The magnetic field, as opposed to the electric field, is an axial vector, i.e. one that does not change sign under reflection of all coordinate axes. Thus, if there would be sources for the first pair of Maxwell equations, they should be an axial vector and a pseudoscalar[1]. The classical description of particles does not allow to construct such quantities from dynamical variables associated to particle.

6.2 Electromagnetic waves

Maxwell equations are partial differential equations on components of electric $\vec{E}$ and magnetic $\vec{H}$ fields. In absence of sources these equations read as

$$\vec{\nabla}\cdot\vec{H} = 0\,, \qquad \frac{\partial \vec{H}}{\partial t} = -c\,\vec{\nabla}\times\vec{E}\,, \tag{II.6.7}$$

$$\vec{\nabla}\cdot\vec{E} = 0\,, \qquad \frac{\partial \vec{E}}{\partial t} = c\,\vec{\nabla}\times\vec{H}\,. \tag{II.6.8}$$

Equations depend on a parameter c which appears to coincide with the speed of light. Usually, one refers to two equations in (II.6.7) as the first pair and to (II.6.8) as the second pair, respectively. The first pair has $\vec{H}$ on the left hand side, while the second one has $\vec{E}$.

Equations (II.6.7), (II.6.8) admit non-zero solutions meaning that electromagnetic fields can exist without any charges or currents. *Electromagnetic fields which exist in the absence of any sources are called electromagnetic waves.*

A progress with solving Maxwell's equations is based on the introduction of the 4-vector electromagnetic potential A^μ, which combines a scalar φ and a vector $\vec{A}$ components

$$A^\mu = (\varphi, \vec{A})\,, \qquad A_\mu = \eta_{\mu\nu}A^\nu = (\varphi, -\vec{A})\,. \tag{II.6.9}$$

The relationship between electric and magnetic fields and the corresponding components of the 4-potential is

$$\vec{E} = -\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \qquad \text{and} \qquad \vec{H} = \operatorname{rot}\vec{A}\,. \tag{II.6.10}$$

[1] A physical quantity that behaves like a scalar, only it changes sign under parity inversion e.g. an improper rotation.

Such a parametrisation of $\vec{E}$ and $\vec{H}$ in terms of arbitrary 4-potential *automatically* satisfies the 1-st pair of Maxwell equations.

Now we consider the equation

$$\vec{\nabla}\cdot\vec{E}=\vec{\nabla}(-\vec{\nabla}A^0-\frac{1}{c}\frac{\partial\vec{A}}{\partial t})=0$$

also known as the Gauss law. This can be written as

$$\vec{\nabla}^2A_0+\frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla}\cdot\vec{A})=0$$

This equation shows that the scalar potential $A_0\equiv\varphi$ is not dynamical (occurs without its time derivative). In fact A_0 can be unambiguously found from this equation by inverting the 3-dimensional Laplace operator $\vec{\nabla}^2$

$$A_0=-\vec{\nabla}^{-2}\frac{1}{c}\frac{\partial}{\partial t}(\vec{\nabla}\cdot\vec{A})=\int \mathrm{d}\vec{x}'\,\frac{\frac{\partial(\vec{\nabla}\cdot\vec{A})}{\partial t}(\vec{x}')}{4\pi c|\vec{x}-\vec{x}'|}\,. \tag{II.6.11}$$

Thus, A_0 is not independent – we do not need to specify it at $t=t_0$ (initial time slice). Thus, the number of independent degrees of freedom cannot be more than three.

We therefore end up with the last equation we have to solve, namely,

$$\frac{\partial\vec{E}}{\partial t}=c\,\vec{\nabla}\times\vec{H}\,.$$

Substituting here the electromagnetic potential we will get

$$\frac{\partial}{\partial t}\Big(\vec{\nabla}A_0+\frac{1}{c}\frac{\partial\vec{A}}{\partial t}\Big)=-c\,\vec{\nabla}\times\vec{\nabla}\times\vec{A}\,,$$

which with the use of the known formula

$$\vec{\nabla}\times(\vec{\nabla}\times\vec{A})=\vec{\nabla}(\vec{\nabla}\cdot\vec{A})-\vec{\nabla}^2\vec{A}$$

boils down to the following very complicated equation for $\vec{A}$

$$\frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2}-\vec{\nabla}^2\vec{A}=\vec{\nabla}\Big(\vec{\nabla}^{-2}\frac{\partial^2}{c^2\partial t^2}-1\Big)(\vec{\nabla}\cdot\vec{A})\,. \tag{II.6.12}$$

At this point it is unclear how to solve such an equation.

The progress how can be made by invoking the gauge freedom (II.5.36). According to (II.5.36), two configurations of A_μ which differ from each other by a gauge transformation are, thus, physically indistinguishable. We can use this freedom in the redefinition of the electromagnetic potential to bring it by a gauge transformation to a convenient form and in this way to solve equation (II.6.12).

Coulomb gauge. Looking at (II.6.12), it is obvious that one has to use gauge transformations to subject A_μ a condition

$$\vec{\nabla}\cdot\vec{A}\equiv\mathrm{div}\vec{A}\equiv\partial_iA^i=0\,. \tag{II.6.13}$$

Subjecting A_μ to an extra condition is called *gauge fixing* and the condition itself is known as the *gauge*. The gauge (II.6.13) is known as the *Coulomb (or radiation) gauge*. For any given $\tilde{A}_i$ one can always find a representative A_i in its gauge orbit which satisfies the Coulomb gauge condition. Indeed,

$$A_i=\tilde{A}_i+\partial_i\alpha\,,\qquad \partial_iA^i=0\,.$$

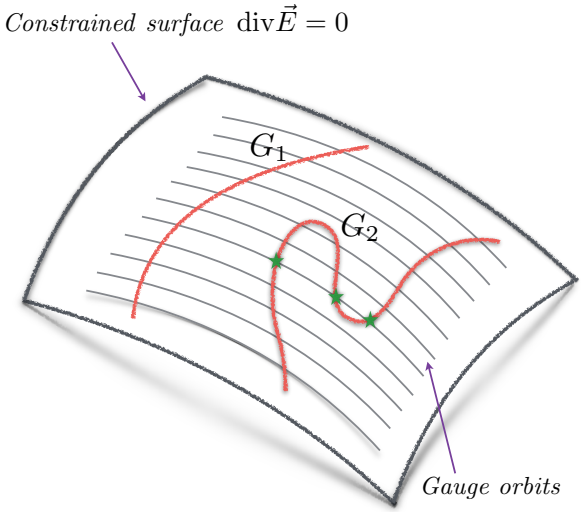

Figure 6.1: Constrained surface $\mathrm{div}\,\vec{E} = 0$ in the space of field configurations A_μ. This surface is foliated by the action of the gauge group into a set of disjoint orbits. The red curves exemplify two possible gauge fixings – the first one G_1 is complete and the second G_2 is not.

We have

$$\partial_i A^i = \partial_i \tilde{A}^i - \vec{\nabla}^2 \alpha = 0\,,$$

that is

$$\alpha = \vec{\nabla}^{-2}\partial_i \tilde{A}^i = -\int \mathrm{d}\vec{x}' \frac{\vec{\nabla}\cdot\vec{A}(\vec{x}')}{4\pi|\vec{x}-\vec{x}'|}\,.$$

We see that in the Coulomb gauge equation (II.6.12) turns into the *wave equation* (*d'Alembert's equation*) for each component of $\vec{A}$

$$\frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} - \vec{\nabla}^2\vec{A} = 0\,. \tag{II.6.14}$$

In fact from (II.6.11) it follows that in the Coulomb gauge $A_0 = 0$ which together with the gauge fixing condition $\vec{\nabla}\cdot\vec{A} = 0$ means that electromagnetic potential (and therefore electromagnetic field) has only two degrees of freedom, whose time evolution is described by (II.6.14). These two degrees of freedom correspond to two possible polarisations of a photon. In this way we, in fact, solved Maxwell's equations.

We further point out that one of the convenient gauge choices involves setting $\partial_\mu A^\mu = 0$, which is the covariant gauge choice known as the Lorenz gauge[2]. This however is not a complete gauge choice, because, as will be shown later, there are still the gauge transformations that leave the electromagnetic field tensor unchanged. A further specification of the Lorenz gauge known as the Coulomb gauge sets the divergence of the vector or the scalar potential equal to zero, i.e. $\mathrm{div}\vec{A} = 0$ and $\varphi = 0$.

[2]Often erroneously referred to as the Lorentz gauge, due to the similarity with the name Lorentz as in Lorentz transformations, developed by Dutch physicist Hendrik Lorentz. However it was a Danish physicist, Ludvig Lorenz, who actually introduced the Lorenz gauge.

Solving the wave equation. Consider the wave equation

$$\Delta\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = 0\,.$$

When we only consider the plane-wave solutions (*i.e.* the ones that depend on the coordinate x only), then the equation reduces to

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = 0\,.$$

It can be further written in the factorized form

$$\left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right) f = 0\,.$$

With a change of variables $\xi = t - \frac{x}{c}$ and $\eta = t + \frac{x}{c}$, one gets

$$\frac{\partial}{\partial \xi} = \frac{1}{2}\left(\frac{\partial}{\partial t} - c\frac{\partial}{\partial x}\right)\,, \qquad \frac{\partial}{\partial \eta} = \frac{1}{2}\left(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x}\right)\,,$$

so that the wave equation in these coordinates takes the form

$$\frac{\partial^2 f}{\partial \xi \partial \eta} = 0\,. \tag{II.6.15}$$

The general solution of this equation reads as

$$f = f_1(\xi) + f_2(\eta)\,,$$

where f_1 and f_2 are two arbitrary functions of their arguments. Changing our variables back to x and t, we find that the general solution for f is given by

$$f = f_1\left(t - \frac{x}{c}\right) + f_2\left(t + \frac{x}{c}\right)\,.$$

Notice that this solution simply represents the sum of right- and left-moving plane waves of any arbitrary profile, respectively.

Let us now consider plane wave solution to the d'Alambert equation. In this case the derivatives of the y and z component of the vector potential with respect to y and z components respectively should vanish as we will only look at oscillations in the x direction. For the Coulomb gauge condition this implies that

$$\text{div}\,\vec{A} = 0 = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \Rightarrow \frac{\partial A_x}{\partial x} = 0\,.$$

If $\frac{\partial A_x}{\partial x} = 0$ everywhere, then $\frac{\partial^2 A_x}{\partial x^2} = 0$, which rendes the wave equation for the component A_x in the form

$$\begin{aligned} \frac{\partial^2 A_x}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 A_x}{\partial t^2} &= 0 \\ -\frac{1}{c^2}\frac{\partial^2 A_x}{\partial t^2} &= 0 \Rightarrow \frac{\partial^2 A_x}{\partial t^2} = 0 \Rightarrow \frac{\partial A_x}{\partial t} = \text{const.} \end{aligned}$$

Since we are not interested in a constant electric field E_x, we put $A_x = 0$. Since $\vec{E} = -\frac{1}{c}\frac{\partial\vec{A}}{\partial t}$ and $\vec{H} = \text{rot}\,\vec{A}$, then

$$\vec{H} = \vec{\nabla}\times\vec{A} = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ \partial_x & \partial_y & \partial_z \\ 0 & A_y(t-x/c) & A_z(t-x/c) \end{vmatrix}$$

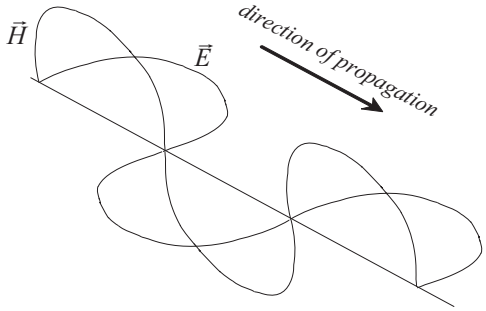

Figure 6.2: Oscillations of the electric and magnetic fields in electromagnetic wave.

$$
\begin{aligned}
&= -\vec{e}_y \partial_x A_z(t-x/c) + \vec{e}_z \partial_x A_y(t-x/c) = \vec{e}_y \frac{1}{c}\frac{\partial A_z}{\partial t} + \vec{e}_z \left(-\frac{1}{c}\frac{\partial A_y}{\partial t}\right) \\
&= \vec{e}_y(-E_z) + \vec{e}_z E_y = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ 1 & 0 & 0 \\ 0 & E_y & E_z \end{vmatrix} = \vec{n} \times \vec{E}\,,
\end{aligned}
$$

where $\vec{n} = (1,0,0)$ is the unit vector in the x-direction. Thus, the electric field $\vec{E}$ and the magnetic field $\vec{H}$ are perpendicular to each other. Waves with this property are referred to as *transversal waves*.

Energy flux and energy density. Electromagnetic waves are known to carry energy; we can define the *energy flux* also known as the *Poynting vector*

$$\vec{S} = \frac{c}{4\pi}\vec{E} \times \vec{H} = \frac{c}{4\pi}\vec{E} \times (\vec{n} \times \vec{E})\,.$$

Since $\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}\big(\vec{a},\vec{c}\big) - \vec{c}\big(\vec{a},\vec{b}\big)$, where$\big(\vec{a},\vec{b}\big)$ denotes the scalar product between vectors $\vec{a}$ and $\vec{b}$, we find the following result

$$\vec{S} = \frac{c}{4\pi}\vec{n}\vec{E}^2\,,$$

where due to orthogonality of $\vec{n}$ and $\vec{E}$ the contribution of the second term vanishes. The energy density is given by

$$W = \frac{1}{8\pi}\big(\vec{E}^2 + \vec{H}^2\big)\,.$$

For electromagnetic waves $\big|\vec{E}\big| = \big|\vec{H}\big|$, so that $W = \frac{1}{4\pi}\vec{E}^2$. Hence, there exists a simple relationship

$$\vec{S} = cW\vec{n}\,.$$

We define the momentum associated to the electromagnetic wave to be

$$\vec{p} = \frac{\vec{S}}{c^2} = \frac{W}{c}\vec{n}\,.$$

For a particle moving along $\vec{n}$, we have $p = \frac{W}{c}$. Consider a particle moving with velocity $\vec{v}$. We then have $p = \frac{vE}{c^2}$ which for $v \to c$ becomes $p = \frac{E}{c}$; the dispersion relation for a relativistic particle moving at the speed of light (photon).

6.3 Maxwell's equations with sources

Continuing, we are now interested in the case of fields created by moving charges. So far we have discussed

- Motion of a charged particle in an external electromagnetic field (the Lorentz force);
- Time-dependent fields but without charges (electromagnetic waves).

We will now study time-dependent fields in the presence of arbitrary moving charges[3]. Consider the second pair of Maxwell's equations

$$\begin{aligned} \frac{\partial F^{\mu\nu}}{\partial x^\nu} &= -\frac{4\pi}{c} j^\mu \,, \\ \frac{\partial}{\partial x^\nu}\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) &= \frac{\partial^2}{\partial x^\nu \partial x_\mu} A^\nu - \frac{\partial^2}{\partial x^\nu \partial x_\nu} A^\mu = -\frac{4\pi}{c} j^\mu \,. \end{aligned}$$

Imposing the Lorenz condition

$$\frac{\partial A^\mu}{\partial x^\mu} = 0 \,,$$

we obtain from the previous equation

$$\Box A^\mu \equiv \frac{\partial^2}{\partial x^\nu \partial x_\nu} A^\mu = \frac{4\pi}{c} j^\mu \,.$$

The last equation can be split into two

$$\begin{aligned} \Delta \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} &= -\frac{4\pi}{c}\vec{j} \,, \\ \Delta \varphi - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} &= -4\pi\rho \,. \end{aligned}$$

Here we deal with an equation of the type

$$\Delta\Phi - \frac{1}{c^2}\frac{\partial^2 \Phi}{\partial t^2} = -4\pi f\left(\vec{x}, t\right) \,. \tag{II.6.16}$$

To solve this problem, as in electrostatics, it is useful to first find the Green's function $G\left(\vec{x}, t; \vec{x}', t'\right)$, defined as a solution of the following equation

$$\left(\Delta_x - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) G\left(\vec{x}, t; \vec{x}', t'\right) = -4\pi\delta\left(\vec{x} - \vec{x}'\right)\delta\left(t - t'\right) \,. \tag{II.6.17}$$

Note that $G\left(\vec{x}, t; \vec{x}', t'\right)$ is not unique and it has to be specified in a number of ways. Additionally, it is referred to as the propagator (especially in the field of quantum electrodynamics). The solution to equation (II.6.16) reads

$$\Phi\left(\vec{x}, t\right) = \int G\left(\vec{x}, t; \vec{x}', t'\right) f\left(\vec{x}', t'\right) \mathrm{d}^3 x' \mathrm{d}t \,.$$

To check that this is actually the solution, one can apply the operator $\Delta_x - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$ and move it under the integral - two delta functions will emerge by virtue of (II.6.17), which upon integration will turn $f\left(\vec{x}', t'\right)$ into $f\left(\vec{x}, t\right)$.

[3]The motion of charges has to be set up, *i.e.* even though the charges produce an electromagnetic field, their motion will not be influenced by the presence of these fields. This excludes the back-reaction of fields on charges.

To proceed with solving (II.6.16), we will need the Fourier transform[4] of the function entering equation (II.6.17)

$$\delta\left(\vec{x}-\vec{x}'\right)\delta\left(t-t'\right)=\frac{1}{\left(2\pi\right)^4}\int_{-\infty}^{\infty}d\vec{k}\int_{-\infty}^{\infty}d\omega\, e^{i\vec{k}\cdot\left(\vec{x}-\vec{x}'\right)}e^{-i\omega\left(t-t'\right)}\,,$$
$$G\left(\vec{x},t;\vec{x}',t'\right)=\int_{-\infty}^{\infty}d\vec{k}\int_{-\infty}^{\infty}d\omega\, g\big(\vec{k},\omega\big)e^{i\vec{k}\cdot\left(\vec{x}-\vec{x}'\right)-i\omega\left(t-t'\right)}\,.$$

Plugging these into the equation, we obtain

$$g\big(\vec{k},\omega\big)\left(-\vec{k}^2+\frac{\omega^2}{c^2}\right)=-4\pi\frac{1}{\left(2\pi\right)^4}=-\frac{1}{4\pi^3}\,,$$

which amounts to

$$g\big(\vec{k},\omega\big)=\frac{1}{4\pi^3}\frac{1}{\vec{k}^2-\frac{\omega^2}{c^2}}\,.$$

From this one can find an integral expression for $G\left(\vec{x},t;\vec{x}',t'\right)$

$$G\left(\vec{x},t;\vec{x}',t'\right)=\frac{1}{4\pi^3}\int_{-\infty}^{\infty}d\vec{k}\int_{-\infty}^{\infty}d\omega\frac{e^{i\vec{k}\cdot\left(\vec{x}-\vec{x}'\right)-i\omega\left(t-t'\right)}}{\vec{k}^2-\frac{\omega^2}{c^2}}\,.$$

The complex function inside the integral is singular at $\vec{k}^2=\frac{\omega^2}{c^2}$ and thus has two first order poles at $\omega=\pm c\big|\vec{k}\big|$. We have to find the proper way to treat this singularity. This is done by using the following physical reasoning. The Green function is a wave-type perturbation produced by a point source sitting at x' and emanating during an infinitesimal time at $t=t'$. We can expect that this wave propagates with the speed of light as a spherical wave. Thus, we should require that

a) $G=0$ in the whole space for $t<t'$

b) G is a diverging wave for $t>t'$

We shall see that the above only represents one of the possible Green's functions, since a different treatment of the poles produces different Green's functions - an advanced or a retarded one:

Retarded Green function states $G_{\rm ret}=0$ if $t<t'$

Advanced Green function states $G_{\rm adv}=0$ if $t>t'$

Notice that the difference of the two $G_{\rm adv}-G_{\rm ret}$, called the Pauli function $G_{\rm Pauli}$, satisfies the homogenous equation.

Consider the retarded Green's function. For $t>t'$, it should give a wave propagating from a point-like source. Let us define $\tau=t-t'$, $\vec{R}=\vec{x}-\vec{x}'$ and $R=\big|\vec{R}\big|$. Then we have

$$e^{-i\omega\left(t-t'\right)}\sim e^{\Im\omega\tau}\,,$$

since $\tau>0$. Thus we need to require that $\Im\omega<0$ in order to have a decaying function at large ω, hence we have to integrate over the lower complex plane. In opposite, for $t<t'$, the contour over which we integrate in the upper half of the complex plane should give zero contribution due to the aforementioned physical reasons. As a result, one could infinitesimally shift the poles into the lower

[4]The role of the Fourier transform is to convert a linear differential equation for the function into an algebraic one for its Fourier image.

half plane when performing the analytic continuation. According to this prescription, the Green's function is specified as follows

$$G_{\rm ret}(\vec{x},t;\vec{x}',t') = \frac{1}{4\pi^3}\int d\vec{k}\int d\omega \frac{e^{i\vec{k}R-i\omega\tau}}{k^2-\frac{1}{c^2}(\omega+i\varepsilon)^2}\,.$$

We can conveniently rewrite the previous statement, by making use of partial fractions

$$\begin{aligned} G_{\rm ret}\left(\vec{x},t;\vec{x}',t'\right) &= \\ &= -\frac{1}{4\pi^3}\int_{-\infty}^{\infty} d\vec{k}\int_{-\infty}^{\infty} d\omega\, e^{i\vec{k}R}\frac{c}{2k}\left[\frac{1}{\omega-ck+i\varepsilon}-\frac{1}{\omega+ck+i\varepsilon}\right]e^{-i\omega\tau}\,. \end{aligned} \tag{II.6.18}$$

Applying Cauchy's theorem[5] and taking the limit $\varepsilon\to 0$, we find

$$\begin{aligned} G_{\rm ret}\left(\vec{x},t;\vec{x}',t'\right) &= \frac{1}{4\pi^3}\int_{-\infty}^{\infty} d\vec{k}\, e^{i\vec{k}\cdot\vec{R}}\, 2\pi i\,\frac{c}{2k}\left[e^{-ick\tau}-e^{ick\tau}\right] \\ &= \frac{c}{2\pi^2}\int_{-\infty}^{\infty} d\vec{k}\,\frac{e^{i\vec{k}\cdot\vec{R}}}{k}\sin(ck\tau)\,. \end{aligned} \tag{II.6.19}$$

To compute this integral, we pass to spherical coordinates.

$$\begin{aligned} G_{\rm ret}\left(\vec{x},t;\vec{x}',t'\right) &= \frac{c}{2\pi^2}\int_0^{\infty} dk\,k\,\sin(ck\tau)\int_0^{\pi}\sin\theta d\theta\int_0^{2\pi} d\varphi\, e^{ikR\cos\theta} \\ &= \frac{c}{\pi}\int_0^{\infty} dk\,k\,\sin(ck\tau)\int_{-1}^{1} dx\, e^{ikRx} \\ &= \frac{2c}{\pi R}\int_0^{\infty} dk\sin(kR)\sin(ck\tau) && \text{(II.6.20)} \\ &= \frac{1}{\pi R}\int_{-\infty}^{\infty} d\,(ck)\sin\left(\frac{(ck)\,R}{c}\right)\sin\left((ck)\,\tau\right) && \text{(II.6.21)} \\ &= -\frac{1}{4\pi R}\int_{-\infty}^{\infty} dx\left(e^{ix\frac{R}{c}}-e^{-ix\frac{R}{c}}\right)\left(e^{ix\tau}-e^{-ix\tau}\right) && \text{(II.6.22)} \\ &= \frac{1}{2\pi R}\int_{-\infty}^{\infty} dx\left(e^{ix\left(\tau-\frac{R}{c}\right)}-e^{ix\left(\tau+\frac{R}{c}\right)}\right) \\ &= \frac{1}{R}\delta\left(\tau-\frac{R}{c}\right)-\frac{1}{R}\delta\left(\tau+\frac{R}{c}\right) && \text{(II.6.23)} \\ &= \frac{1}{R}\delta\left(\tau-\frac{R}{c}\right) && \text{(II.6.24)} \end{aligned}$$

To summarise, in this computation we have used: partial fractions (II.6.18), the Cauchy theorem in (II.6.18-II.6.19), switched to spherical coordinates and integrated over the angles (II.6.20), substituted $ck=x$ (II.6.21), expanded the trigonometric functions in terms of their complex exponentials (II.6.22), and identified Fourier transforms of delta functions (II.6.23). On the last step we have rejected $\delta\left(\tau+\frac{R}{c}\right)$, because for $\tau,R,c>0$, the result will always be zero. Substituting back our original variables, we get

$$G_{\rm ret}\left(\vec{x},t;\vec{x}',t'\right) = \frac{\delta\left(t'+\frac{|\vec{x}-\vec{x}'|}{c}-t\right)}{|\vec{x}-\vec{x}'|}\,.$$

[5] Cauchy integral formula reads

$$f(a) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}\,dz\,,$$

where a function $f(z)$ is holomorphic inside the region surrounded by a contour C and integration is performed in counter-clockwize direction.

The result can be understood as the signal propagating at the speed of light, which was emitted at t' and will travel for $\frac{|\vec{x}-\vec{x}'|}{c}$ and will be observed at time t. Thus, this Green function reflects a natural causal sequence of events. The time t is then expressed in terms of the *retarded time* t'

$$t = t' + \frac{|\vec{x}-\vec{x}'|}{c}\,.$$

Substituting this solution and integrating over t', we obtain the *retarded potentials*

$$\begin{aligned}\varphi(\vec{x},t) &= \int \frac{\delta\left(t' + \frac{|\vec{x}-\vec{x}'|}{c} - t\right)}{|\vec{x}-\vec{x}'|}\rho(\vec{x}',t')\,d\vec{x}'dt' + \varphi_0 \\ &= \int \frac{\rho\left(\vec{x}', t - \frac{|\vec{x}-\vec{x}'|}{c}\right)}{|\vec{x}-\vec{x}'|}d\vec{x}' + \varphi_0\,, \end{aligned} \tag{II.6.25}$$

$$\begin{aligned}\vec{A}(\vec{x},t) &= \frac{1}{c}\int \frac{\delta\left(t' + \frac{|\vec{x}-\vec{x}'|}{c} - t\right)}{|\vec{x}-\vec{x}'|}\vec{j}(\vec{x}',t')\,d\vec{x}'dt' + \vec{A}_0 \\ &= \frac{1}{c}\int \frac{\vec{j}\left(\vec{x}', t - \frac{|\vec{x}-\vec{x}'|}{c}\right)}{|\vec{x}-\vec{x}'|}d\vec{x}' + \vec{A}_0\,, \end{aligned} \tag{II.6.26}$$

where φ_0 and $\vec{A}_0$ are the solutions of the homogeneous d'Alambert equations (those corresponding to the free electromagnetic field) $\Box A_0^\mu = 0$ and the Lorentz gauge condition $\partial_\mu A_0^\mu = 0$.

Note that for φ in the case of time-independent ρ and $\vec{j}$ we have

$$\varphi = \int \frac{\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'\,.$$

This is just the electrostatic formula for the scalar potential. Moreover, if the current $\vec{j}$ is time-independent, we obtain

$$\vec{A}(\vec{x}) = \frac{1}{c}\int \frac{\vec{j}(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'\,.$$

Lorentz invariance of Green's function. Let us now show that G_{ret} is Lorentz invariant. We write

$$G_{\rm ret}(\vec{x},t;\vec{x}',t') = \Theta(t-t')\,\frac{\delta\left(t' + \frac{|\vec{x}-\vec{x}'|}{c} - t\right)}{|\vec{x}-\vec{x}'|}\,.$$

Here the extra term $\Theta(t-t')$ ensures that $G_{\rm ret}(\vec{x},t;\vec{x}',t') = 0$ for $t < t'$, because

$$\Theta(t-t') = \begin{cases} 0, & t < t' \\ 1, & t \geq t' \end{cases}$$

To rewrite the retarded Green's function in a Lorentz-invariant way, we have to use

$$\delta(f(x)) = \sum_i \frac{\delta(x-x_i)}{|f'(x_i)|}\,.$$

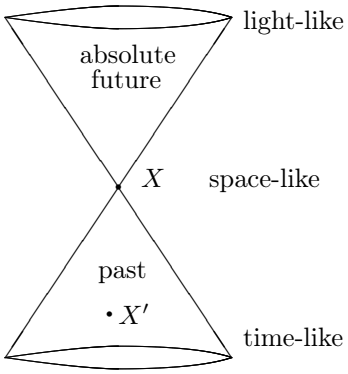

Figure 6.3: At every point in time every observer has his past light cone, which is a set of all events that could have influenced his presence, and a future light cone, the set of events which the observer can influence. The boundaries of the light cones also define the split between different kinds of space-time intervals. On the light cone itself the intervals are all light-like, time-like on the inside and space-like on the outside.

In the last formula the derivative is evaluated at the set of points x_i, such that $f(x_i) = 0$. Hence, introduce $u = |\vec{x} - \vec{x}'| - c(t - t')$. Then

$$\delta\Big(|\vec{x} - \vec{x}'|^2 - c^2 (t - t')^2 \Big) = \delta\Big(u(u + 2c(t - t'))\Big) = \delta\Big(u^2 + 2uc(t - t'))\Big) .$$

Now define $f(u) = u^2 + 2uc(t-t')$ with $f'(u) = 2u + 2c(t-t')$. Equation $f(u) = 0$ has two solutions: $u = 0$ and $u = -2c(t-t')$. The second one will not contribute into the formula describing the change of variables in the delta-function because of $\Theta(t - t')$. Thus,

$$\delta\left(|\vec{x} - \vec{x}'|^2 - c^2 (t - t')^2\right) = \frac{\delta\left(|\vec{x} - \vec{x}'| - c(t - t')\right)}{(2u + 2c(t - t'))|_{u=0}} = \frac{\delta\left(|\vec{x} - \vec{x}'| - c(t - t')\right)}{2|\vec{x} - \vec{x}'|} .$$

In this way, we arrive at the following formula

$$\begin{aligned} G_{\rm ret}(\vec{x}, t; \vec{x}', t') &= 2c\,\Theta(t - t')\,\frac{\delta\left(|\vec{x} - \vec{x}'| - c(t - t')\right)}{2\,|\vec{x} - \vec{x}'|} \\ &= 2c\,\Theta(t - t')\,\delta\left(|\vec{x} - \vec{x}'|^2 - c^2 (t - t')^2\right) = 2c\,\Theta(t - t')\,\delta(s^2) , \end{aligned}$$

where the argument of the delta function is the square of the 4-interval s between two events $(\vec{x}, t)$ and $(\vec{x}', t')$, which is a Lorentz invariant object. From this we can conclude that the Green's function is invariant under proper orthochronous Lorentz transformations. Note that for orthochronous Lorentz transformations the $\Theta(t-t')$-function is not invariant only for space-like intervals, but these intervals are discarded by the presence of the δ-function $\delta(s^2)$.

Green's function and causality principle. Let us recall a classification of 4-intervals

$$\mathrm{d}s^2 = c^2 \mathrm{d}t^2 - \mathrm{d}x_i^2 \qquad \text{(II.6.27)}$$

We refer to them differently depending on the sign of $\mathrm{d}s^2$:

1) time-like intervals if $\mathrm{d}s^2 > 0$,

2) space-like intervals if $\mathrm{d}s^2 < 0$,

3) light-like intervals (also called null intervals) if $\mathrm{d}s^2 = 0$.

Consider Fig. 6.3 representing the light-cone built over a point X. Signals in X can come only from points X', which are in the past light-cone of X. We say $X > X'$ (X is later than X'). The influence of a current j in X' on potential A at X is a signal from X' to X. Thus, the causality principle is reflected in the fact that $A(X)$ can depend on 4-currents $j(X')$ only for those X' for which $X > X'$. Thus,

$$\frac{\delta A(X)}{\delta j(X')} \sim G(X - X') = 0 \tag{II.6.28}$$

for $X < X'$ or points X' that are space-like to X. Hence, the causality principle for the Green function is

$$G(X' - X) = 0\,, \tag{II.6.29}$$

in terms of the conditions described above. The retarded Green's function is the only relativistic Green's function which has this property.

Chapter 7

Radiation

The last part of these lectures will treat two classical radiation problems: Liénard-Wiechert potentials and the dipole radiation.

7.1 Fields of a uniformly moving charge

Before studying the radiation problems, we consider the field produced by an electric charge which moves uniformly with velocity $\vec{v}$. To this end we need fist to understand how electromagnetic fields transform under Lorentz transformations.

7.1.1 Lorentz transformations of electromagnetic field

First we consider the 4-potential A^μ. Under Lorentz transformations of space-time coordinates, A^μ transforms as a vector:

$$A'^\mu(x') = \Lambda^\mu_\nu A^\nu(x)\,.$$

Recall that the matrix Λ of a Lorentz transformation from a stationary to a moving with velocity $\vec{v}$ frame is of the form

$$\Lambda = \begin{pmatrix} a & -\frac{a}{c}v^t \\ -\frac{a}{c}v & \Lambda_{ij} \end{pmatrix}, \tag{II.7.1}$$

where $\Lambda_{ij} = \delta_{ij} + \frac{a-1}{v^2}v_i v_j$ and $a = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$. Thus, the scalar and vector potentials in the moving frame are

$$\varphi' = a\varphi - \frac{a}{c}(v\cdot A) = \frac{\varphi - \frac{(A\cdot v)}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}\,,$$

$$\vec{A}' = -\frac{a}{c}\varphi v + \vec{A} + d\vec{v}(\vec{v}\cdot\vec{A}) = \vec{A} - \frac{\varphi\frac{\vec{v}}{c}}{\sqrt{1-\frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{\vec{v}(\vec{v}\cdot\vec{A})}{\vec{v}^2}\,,$$

where we adopted a concise notation $d = \frac{a-1}{v^2}$. Now we come to the electromagnetic field $(\vec{E},\vec{H})$. It is important to realize that components of the electromagnetic field transform as components of the *second rank tensor*! Namely, one has

$$F^{\mu\nu'}(x') = \Lambda^\mu_\rho\Lambda^\nu_\tau F^{\rho\tau}(x)\,.$$

For E_i one therefore gets

$$\begin{aligned} E_i' &= F^{i0'} = \Lambda^i_\mu \Lambda^0_\nu F^{\mu\nu} = \Lambda^i_0\Lambda^0_k F^{0k} + \Lambda^i_k\Lambda^0_0 F^{k0} + \Lambda^i_k\Lambda^0_j F^{kj} \\ &= (\Lambda^i_k\Lambda^0_0 - \Lambda^i_0\Lambda^0_k)E_k + \Lambda^i_k\Lambda^0_j(-\epsilon_{kjm}H_m) \\ &= a(\delta_{ik} + dv_iv_k)E_k - \frac{a^2}{c^2}v_iv_kE_k - (\delta_{ik}+dv_iv_k)\frac{a}{c}v_j(-\epsilon_{kjm}H_m) \\ &= aE_i + \Big(ad - \frac{a^2}{c^2}\Big)v_i(\vec{v}\cdot\vec{E}) + \frac{a}{c}\epsilon_{ijm}v_jH_m\,. \end{aligned}$$

The final formula reads as

$$E_i' = aE_i - \frac{a-1}{v^2}v_i(\vec{v}\cdot\vec{E}) + \frac{a}{c}(\vec{v}\times\vec{H})_i\,.$$

Now we come to the magnetic field. We have

$$\begin{aligned} H_i' &= -\frac{1}{2}\epsilon_{ijk}F'^{jk} = -\frac{1}{2}\epsilon_{ijk}(\Lambda^j_0\Lambda^k_nF^{0n} + \Lambda^j_n\Lambda^k_0F^{n0} + \Lambda^j_m\Lambda^k_nF^{mn}) \\ &= -\frac{1}{2}\epsilon_{ijk}(\Lambda^j_0\Lambda^k_n - \Lambda^j_n\Lambda^k_0)F^{0n} - \frac{1}{2}\epsilon_{ijk}\Lambda^j_m\Lambda^k_nF^{mn}\,. \end{aligned}$$

We proceed by substituting the matrix elements of Λ:

$$\begin{aligned} H_i' &= \frac{1}{2}\epsilon_{ijk}\Big(-\frac{a}{c}v_j(\delta_{nk}+dv_nv_k) + \frac{a}{c}v_k(\delta_{nj}+dv_nv_j)\Big)E_n \\ &- \frac{1}{2}\epsilon_{ijk}\Big((\delta_{mj}+dv_mv_j)(\delta_{nk}+v_nv_k)\Big)F^{mn}\,. \end{aligned}$$

Making use of the formula for the pairing of two ϵ-tensors in the second line of the last formula, we arrive at

$$\begin{aligned} H_i' &= -\frac{a}{c}\epsilon_{ijn}v_jE_n + H_i - \frac{d}{2}\Big((\delta_{in}\delta_{ks}-\delta_{is}\delta_{nk})v_kv_nH_s + (\delta_{im}\delta_{js}-\delta_{is}\delta_{jm})v_jv_mH_s\Big) \\ &= H_i - \frac{a}{c}\epsilon_{ijn}v_jE_n + d(H_iv^2 - v_i(\vec{v}\cdot\vec{H})) \end{aligned} \tag{II.7.2}$$

The final expression is

$$H_i' = aH_i - \frac{a-1}{v^2}v_i(\vec{v}\cdot\vec{H}) - \frac{a}{c}(\vec{v}\times\vec{E})_i\,.$$

We summarize the transformation formulae

$$\begin{aligned} \varphi' &= a\varphi - \frac{a}{c}(\vec{v}\cdot\vec{A})\,, \\ \vec{A}' &= \vec{A} - \frac{a}{c}\varphi\,\vec{v} + \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{A}) \end{aligned} \tag{II.7.3}$$

and

$$\begin{aligned} \vec{E}' &= a\vec{E} - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{E}) + \frac{a}{c}(\vec{v}\times\vec{H})\,, \\ \vec{H}' &= a\vec{H} - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{H}) - \frac{a}{c}(\vec{v}\times\vec{E})\,. \end{aligned} \tag{II.7.4}$$

The inverse transformations are

$$\begin{aligned} \varphi &= a\varphi' + \frac{a}{c}(\vec{v}\cdot\vec{A}')\,, \\ \vec{A} &= \vec{A}' + \frac{a}{c}\varphi'\,\vec{v} + \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{A}') \end{aligned} \tag{II.7.5}$$

and

$$\begin{aligned}\vec{E} &= a\vec{E}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{E}') - \frac{a}{c}(\vec{v}\times\vec{H}')\,,\\ \vec{H} &= a\vec{H}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{H}') + \frac{a}{c}(\vec{v}\times\vec{E}')\,.\end{aligned} \tag{II.7.6}$$

This completes our discussion of the transformation properties of the 4-potential and the electromagnetic field under Lorentz transformations.

From the electric and magnetic fields one can make invariants, i.e. objects that remain unchanged under Lorentz transformations. In terms of the tensor of the electromagnetic field two such invariants are

$$F_{\mu\nu}F^{\mu\nu} = \text{inv}\,; \tag{II.7.7}$$

$$\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = \text{inv}\,. \tag{II.7.8}$$

7.1.2 Uniformly moving charge – no radiation

The Lorentz transformation of $\vec{x}$ is

$$\vec{x}' = \vec{x} - a\vec{v}t + \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{x})\,.$$

In what follows we need to know x'^2. We compute

$$\begin{aligned}x'^2 &= x^2 + a^2v^2t^2 + \frac{(a-1)^2}{v^2}(\vec{v}\vec{x})^2 - 2a(\vec{v}\vec{x})t + 2\frac{a-1}{v^2}(\vec{v}\vec{x})^2 - 2a(a-1)(\vec{v}\vec{x})t\\ &= x^2 + a^2v^2t^2 + \frac{(\vec{v}\vec{x})^2}{v^2}\left(\underbrace{(a-1)^2 + 2(a-1) + 1}_{}-1\right) - 2a^2(\vec{v}\vec{x})t\\ &= x^2 + a^2v^2t^2 + \frac{a^2-1}{v^2}(\vec{v}\vec{x})^2 - 2a^2(\vec{v}\vec{x})t = x^2 + a^2v^2t^2 + \frac{a^2}{c^2}(\vec{v}\vec{x})^2 - 2a^2(\vec{v}\vec{x})t\,,\end{aligned}$$

since $\frac{a^2-1}{v^2} = \frac{a^2}{c^2}$. Then, we proceed as follows

$$\begin{aligned}x'^2 &= \underbrace{x^2 - 2(\vec{v}\vec{x})t + v^2t^2}_{} + (a^2-1)(v^2t^2 - 2(\vec{v}\vec{x})t + x^2 - x^2) + \frac{a^2}{c^2}(\vec{v}\vec{x})^2\\ &= (\vec{x}-\vec{v}t)^2 + a^2\frac{v^2}{c^2}(\vec{x}-\vec{v}t)^2 - \frac{a^2}{c^2}(v^2x^2 - (\vec{v}\vec{x})^2)\\ &= (\vec{x}-\vec{v}t)^2\underbrace{\left(1 + a^2\frac{v^2}{c^2}\right)}_{a^2} - \frac{a^2}{c^2}(v^2x^2 - (\vec{v}\vec{x})^2)\end{aligned}$$

Thus,

$$x'^2 = a^2(\vec{x}-\vec{v}t)^2 - \frac{a^2}{c^2}(v^2x^2 - (\vec{v}\vec{x})^2) = a^2(\vec{x}-\vec{v}t)^2 - \frac{a^2}{c^2}(\vec{v}\times\vec{x})^2\,.$$

We further note that in the vector product $\vec{v}\times\vec{x}$ one can replace $\vec{x}$ for $\vec{x}-\vec{v}t$ without changing the result. The final answer we need reads as

$$x'^2 = a^2(\vec{x}-\vec{v}t)^2 - \frac{a^2}{c^2}(v^2x^2 - (\vec{v}\vec{x})^2) = a^2\left((\vec{x}-\vec{v}t)^2 - \frac{1}{c^2}(\vec{v}\times\vec{x}-\vec{v}t)^2\right).$$

Since in the moving frame $\vec{H}' = 0$, the electric and magnetic fields in the stationary frame are

$$\begin{aligned}\vec{E} &= a\vec{E}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{E}') - \frac{a}{c}(\vec{v}\times\vec{H}') = a\vec{E}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{E}')\,,\\ \vec{H} &= a\vec{H}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{H}') + \frac{a}{c}(\vec{v}\times\vec{E}') = \frac{a}{c}(\vec{v}\times\vec{E}')\,,\end{aligned} \tag{II.7.9}$$

The electric field is $\vec{E}' = e\frac{\vec{x}'}{x'^3}$. Thus, we compute

$$\vec{E} \;=\; \frac{e}{x'^3}\left(a\vec{x} - a^2\vec{v}t + \frac{a(a-1)}{v^2}\vec{v}(\vec{v}\vec{x}) - \frac{a-1}{v^2}\vec{v}\Big((\vec{v}\vec{x}) - av^2t + (a-1)(\vec{v}\vec{x})\Big)\right) = \frac{ae}{x'^3}(\vec{x} - \vec{v}t)\,,$$

which upon substituting x' results into a very simple formula

$$\vec{E}(\vec{x},t) = \frac{e(\vec{x} - \vec{v}t)}{a^2\Big((\vec{x} - \vec{v}t)^2 - \frac{1}{c^2}(\vec{v}\times\vec{x} - \vec{v}t)^2\Big)^{3/2}}\,.$$

We recall that in the last formula $(\vec{x},t)$ is a (observation) point in a stationary frame where the field $\vec{E}(\vec{x},t)$ is measured and $\vec{R} = \vec{x} - \vec{v}t$ is vector from the charge to the observation point. Note that $\vec{E}$ is collinear to $\vec{R}$. Introducing an angle θ between velocity $\vec{v}$ (the direction of motion) and $\vec{R}$, the last formula can be written as

$$\vec{E}(\vec{x},t) = \frac{e\vec{R}}{R^3}\frac{\Big(1 - \frac{v^2}{c^2}\Big)}{\Big(1 - \frac{v^2}{c^2}\sin^2\theta\Big)^{3/2}}\,.$$

As to the magnetic field, one gets

$$\vec{H} = \frac{a}{c}\vec{v}\times\frac{e}{x'^3}\Big(\vec{x} - a\vec{v}t + \frac{a-1}{v^2}\vec{v}(\vec{v}\vec{x})\Big) = \frac{1}{c}\Big[\vec{v}, \frac{ae}{x'^3}\vec{x}\Big]\,.$$

Obviously, the last expression can be written as

$$\vec{H}(\vec{x},t) = \frac{1}{c}\vec{v}\times\vec{E}\,.$$

The corresponding energy flux is

$$\vec{S}(\vec{x},t) = \frac{c}{4\pi}\vec{E}\times\vec{H} = \frac{1}{4\pi}\vec{E}\times(\vec{E}\times\vec{v}) = \frac{1}{4\pi}(\vec{v}E^2 - \vec{E}(\vec{E}\cdot\vec{v}))\,.$$

A charge moving with a uniform velocity is not radiating energy. It is not radiating energy in the rest frame, and, therefore, the same must hold in any other inertial frame.

7.2 Fields of an arbitrary moving charge

The charge distribution in space and time of a single point-like charge is given by

$$\begin{aligned}\rho\,(\vec{x},t) &= e\delta\,(\vec{x} - \vec{r}\,(t))\,,\\ \vec{j}\,(\vec{x},t) &= e\vec{v}\delta\,(\vec{x} - \vec{r}\,(t))\;.\end{aligned}$$

Here $\vec{x}$ is the position of the observer, $\vec{r}\,(t)$ is the trajectory of the charge and $\vec{v} = \dot{r}\,(t)$, its velocity. The potential then reads

$$\varphi\,(\vec{x},t) = \int \frac{\delta\left(t' + \frac{|\vec{x}-\vec{x}'|}{c} - t\right)}{|\vec{x} - \vec{x}'|}e\delta\,(\vec{x}' - \vec{r}\,(t'))\,d\vec{x}'dt' \qquad \text{(II.7.10)}$$

Let us take $\vec{x}' = \vec{r}\,(t')$, because only then the integrand is non-zero. Then eq.(II.7.10) can be integrated over $\vec{x}'$ and we get

$$\varphi\,(\vec{x},t) = e\int \frac{\delta\left(t' + \frac{|\vec{x}-\vec{r}(t')|}{c} - t\right)}{|\vec{x} - \vec{r}\,(t')|}dt'\,. \qquad \text{(II.7.11)}$$

Take $f(t') = t' + \frac{|\vec{x}-\vec{r}(t')|}{c} - t$ and use $\delta(f(x)) = \frac{\delta(x)}{|f'(x)|}$, where $f'(x)$ is evaluated at the point were $f(x) = 0$, i.e. at t' which solves $t' + \frac{|\vec{x}-\vec{r}(t')|}{c} - t = 0$

$$\frac{\mathrm{d}f(t')}{\mathrm{d}t'} = 1 - \frac{1}{c}\frac{(\vec{x}-\vec{r}(t'))\cdot\dot{\vec{r}}(t')}{|\vec{x}-\vec{r}(t')|} = 1 - \frac{1}{c}\frac{\vec{R}\cdot\vec{v}}{R}\,.$$

In the last equation we have used the fact that $\vec{R} = \vec{x} - \vec{r}(t')$ and $\vec{v} = \dot{\vec{r}}(t)$. The potential then becomes

$$\varphi(\vec{x},t) = \frac{e}{R}\frac{1}{1-\frac{1}{c}\frac{\vec{R}\cdot\vec{v}}{R}} = \frac{e}{R-\frac{\vec{R}\cdot\vec{v}}{c}}\,. \tag{II.7.12}$$

We can use the same line of reasoning to show

$$\vec{A}(\vec{x},t) = \frac{e}{c}\frac{\vec{v}}{(R-\frac{\vec{R}\cdot\vec{v}}{c})}\,. \tag{II.7.13}$$

The formulae (II.7.12) and (II.7.13) are *the Liénard-Wiechert potentials*. Let us compute the corresponding electric and magnetic fields.

We have

$$\begin{aligned}\vec{E} &= -\frac{1}{c}\frac{\partial\vec{A}}{\partial t} - \vec{\nabla}\varphi\,;\\ \vec{H} &= \mathrm{rot}\vec{A}\,.\end{aligned}$$

Moreover, $R(t')$ is given by the difference in the times t and t' with an overall factor of c

$$R(t') = c(t-t')\,.$$

Therefore,

$$\frac{\partial R(t')}{\partial t} = \frac{\partial R(t')}{\partial t'}\frac{\partial t'}{\partial t} = -\frac{\vec{R}\cdot\vec{v}}{R}\frac{\partial t'}{\partial t} = c\left(1-\frac{\partial t'}{\partial t}\right)\,. \tag{II.7.14}$$

From this relation, it follows that

$$\frac{\partial t'}{\partial t} = \frac{1}{1-\frac{\vec{R}\cdot\vec{v}}{Rc}}\,.$$

Analogously, one can also start from the expressions $R(t') = c(t-t')$ and $t' = t'(t,\vec{x})$, such that

$$\begin{aligned}\vec{\nabla}R(t') = -c\vec{\nabla}t' \Rightarrow \vec{\nabla}t' = -\frac{1}{c}\vec{\nabla}R(t') &= -\frac{1}{c}\vec{\nabla}_x\left|\vec{x}-\vec{r}\left(t'(\vec{x},t)\right)\right|\\ &= -\frac{1}{c}\left(\frac{\vec{R}}{R} + \frac{\partial R}{\partial t'}\vec{\nabla}t'\right)\,,\end{aligned}$$

where one can again identify $\frac{\partial R}{\partial t'}$ with the previous result from (II.7.14) and finally obtain

$$\vec{\nabla}t' = -\frac{\vec{R}}{c\left(R-\frac{\vec{R}\cdot\vec{v}}{c}\right)} \quad \text{and} \quad \vec{\nabla}R = \frac{\vec{R}}{R-\frac{\vec{R}\cdot\vec{v}}{c}}\,.$$

Now we have all the necessary ingredients, which we can use to find $\vec{E}$ and $\vec{H}$, i.e. to obtain the Liénard-Wiechert fields.

First let's calculate the quantity $\nabla\varphi$,

$$\nabla\varphi = \frac{-e}{(R-\frac{\vec{R}\cdot\vec{v}}{c})^2}\nabla(R-\frac{\vec{R}\cdot\vec{v}}{c}).$$

The first term is

$$\nabla R = -c\nabla t'$$

and we can rewrite the second term by using of the vector identities

$$\nabla(\vec{R}\cdot\vec{v}) = (\vec{R}\cdot\nabla)\vec{v} + (\vec{v}\cdot\nabla)\vec{R} + \vec{R}\times(\nabla\times\vec{v}) + \vec{v}\times(\nabla\times\vec{R}).$$

Now we have to calculate these quantities one at a time. A difficult quantity is

$$(\vec{v}\cdot\nabla)\vec{R} \quad = \quad (\vec{v}\cdot\nabla)\vec{x} - (\vec{v}\cdot\nabla)\vec{r}(t').$$

Switching to index notation hugely simplifies this

$$\begin{aligned} v_m\partial_m R_i &= v_m\partial_m x_i - v_m\partial_m r_i \\ &= v_m\delta_{mi} - v_m v_i \partial_m t' \\ &= v_i - v_i v_m \partial_m t' v_i. \end{aligned}$$

Here I have used that $\partial_m r_i = \frac{dr_i}{dx_m} = \frac{dr_i}{dt'}\frac{dt'}{dx_m} = v_i\partial_m t'$. Going back to vector notation

$$(\vec{v}\cdot\nabla)\vec{R} \quad = \quad \vec{v} - (\vec{v}\cdot\nabla t')\vec{v}.$$

Similarly

$$(\vec{R}\cdot\nabla)\vec{v} = (\vec{R}\cdot\nabla t')\dot{\vec{v}}.$$

Now we calculate

$$\begin{aligned} (\nabla\times\vec{v})_i &= \epsilon_{ijk}\partial_j v_k \\ &= \epsilon_{ijk}\partial_j t'\dot{v}_k \\ &= ((\nabla t')\times\dot{\vec{v}})_i, \end{aligned}$$

and similarly

$$\nabla\times\vec{R} = \nabla\times\vec{x} - \nabla\times\vec{r} = -(\nabla t')\times\vec{v}.$$

Now use an identity $\vec{A}\times(\vec{B}\times\vec{C}) = \vec{B}(\vec{A}\cdot\vec{C}) - \vec{C}(\vec{A}\cdot\vec{B})$, and we finally get

$$\nabla(\vec{R}\cdot\vec{v}) = \vec{v} + \nabla t'(\vec{R}\cdot\dot{\vec{v}} - v^2).$$

Substituting all the quantities finally gives

$$\nabla\varphi = \frac{e}{c^2(R - \frac{\vec{R}\cdot\vec{v}}{c})^3}\left(-\vec{R}(c^2 - v^2 + \vec{R}\cdot\dot{\vec{v}}) + c\vec{v}(R - \frac{\vec{R}\cdot\vec{v}}{c})\right).$$

A similar (but a little bit easier) exercise for $\frac{d\vec{A}}{dt}$ gives

$$\frac{d\vec{A}}{dt} = \frac{e}{c(R - \frac{\vec{R}\cdot\vec{v}}{c})^3}\left((R - \frac{\vec{R}\cdot\vec{v}}{c})(\dot{\vec{v}}R - c\vec{v}) + \frac{\vec{v}R}{c}(c^2 - v^2 + \vec{R}\cdot\dot{\vec{v}})\right).$$

Putting these together we obtain

$$\begin{aligned} \vec{E} \quad = \quad & \frac{e}{(R - \frac{\vec{R}\cdot\vec{v}}{c})^3}\Big((\vec{R} - \frac{\vec{v}R}{c})(1 - \frac{v^2}{c^2}) + \frac{1}{c^2}(\vec{R}(\vec{R}\cdot\dot{\vec{v}}) - R^2\dot{\vec{v}}) \\ & - \frac{R}{c^3}(\vec{v}(\vec{R}\cdot\dot{\vec{v}}) - \dot{\vec{v}}(\vec{R}\cdot\vec{v}))\Big). \end{aligned}$$

By using $R^2 = \vec{R}\cdot\vec{R}$ and again the relation $\vec{A}\times(\vec{B}\times\vec{C}) = \vec{B}(\vec{A}\cdot\vec{C}) - \vec{C}(\vec{A}\cdot\vec{B})$ we now find

$$\vec{E} \quad = \quad \frac{e}{(R - \frac{\vec{R}\cdot\vec{v}}{c})^3}\Big((\vec{R} - \frac{\vec{v}R}{c})(1 - \frac{v^2}{c^2}) + \frac{1}{c^2}(\vec{R}\times((\vec{R} - \frac{R\vec{v}}{c})\times\dot{\vec{v}}))\Big).$$

For the magnetic field we use

$$\vec{H} = \nabla\times\vec{A} = \frac{1}{c}\nabla\times(\varphi\vec{v}) = \frac{1}{c}\left(\varphi(\nabla\times\vec{v}) + (\nabla\varphi)\times\vec{v}\right).$$

Substituting the quantities gives

$$\begin{aligned} \vec{H} \quad = \quad & \frac{\vec{R}}{R}\times\frac{e}{(R - \frac{\vec{R}\cdot\vec{v}}{c})^3}\Big((-\frac{\vec{v}R}{c})(1 - \frac{v^2}{c^2}) + \frac{1}{c^2}(-R^2\dot{\vec{v}}) \\ & - \frac{R}{c^3}(\vec{v}(\vec{R}\cdot\dot{\vec{v}}) - \dot{\vec{v}}(\vec{R}\cdot\vec{v}))\Big). \end{aligned}$$

We see that we almost have the electric field (from the equation just above the final result for $\vec{E}$), but we are missing the quantities $\vec{R}(1 - \frac{v^2}{c^2})$ and $\frac{1}{c^2}\vec{R}(\vec{R}\cdot\dot{\vec{v}})$. However, the cross product with these quantities will vanish, since $\vec{R}\times\vec{R} = 0$, and therefore we can simply add these quantities. We finally have

$$\vec{H} = \frac{\vec{R}}{R}\times\vec{E}.$$

To summarize, the Liénard-Wiechert fields are given by the following expressions

$$\begin{aligned}\vec{H} &= \frac{1}{R}\vec{R}\times\vec{E}\,,\\ \vec{E} &= e\frac{\left(1-\frac{v^2}{c^2}\right)\left(\vec{R}-\frac{\vec{v}}{c}R\right)}{\left(R-\frac{\vec{R}\cdot\vec{v}}{c}\right)^3}+\frac{e\vec{R}\times((\vec{R}-\frac{\vec{v}}{c}R)\times\dot{\vec{v}})}{c^2\left(R-\frac{\vec{R}\cdot\vec{v}}{c}\right)^3}\,.\end{aligned}$$

Notice that in the last equation the first term only depends on the velocity of the moving particle and is proportional to $\frac{1}{R^2}$ (short distance), whereas the second term depends on acceleration and is proportional to $\frac{1}{R}$ providing, therefore, the long-distance dominating contribution, the so-called wave-zone. Note also that flux is proportional to $\vec{E}^2$ hence is also proportional to $\frac{1}{R^2}$. Therefore,

$$\int \vec{E}^2 \mathrm{d}V \sim \int \frac{1}{R^2}R^2\mathrm{d}\Omega = 4\pi\,,$$

which is a constant flux of $\vec{E}$ at large distances. It is worth stressing that there is no energy (radiation) coming from a charge moving at a constant velocity, because we can always choose a frame where it is stationary, hence $\vec{H}=0\Rightarrow\vec{E}\cdot\vec{H}=0$, consequently it cannot emit energy.

7.3 Dipole radiation

Field of a neutral system is expressed with the help of the so-called electric moment given in its discretized form as

$$\vec{d}=\sum_{i=1}^{N}e_i\vec{R}_i\,, \tag{II.7.15}$$

where e_i is the magnitude of a charge at some distance R_i taken from an arbitrary point, in this case chosen to be the origin. For a neutral system we require that

$$\sum_{i=1}^{N}e_i=0\,.$$

Note that for such a system, electric moment does not depend on the choice of the origin of the reference frame, i.e. shifting all $\vec{R}_i\to\vec{R}_i-\vec{a}$ gives

$$\vec{d}_{\vec{a}}=\sum_{i=1}^{N}e_i\left(\vec{R}_i-\vec{a}\right)=\sum_{i=1}^{N}e_i\vec{R}_i-\vec{a}\sum_{i=1}^{N}e_i=\sum_{i=1}^{N}e_i\vec{R}_i=\vec{d}\,.$$

Let us now consider a neutral system of moving charges. From diagram 7.1 using Pythagorean theorem and assuming that $\vec{l}\ll R_0$, l being the characteristic size, we get[1]

$$\begin{aligned}R &= \sqrt{\left(\vec{R}_0-\vec{R}'\right)^2}=\sqrt{\vec{R}_0^2-2\vec{R}_0\cdot\vec{R}'+\vec{R}'^2}\approx\\ &\approx \sqrt{\vec{R}_0^2\left(1-2\frac{\vec{R}_0\cdot\vec{R}'}{\vec{R}_0^2}\right)}\approx R_0\left(1-\frac{\vec{R}_0\cdot\vec{R}'}{\vec{R}_0^2}\right)=R_0-\frac{\vec{R}_0\cdot\vec{R}'}{R_0}\,.\end{aligned}$$

[1] Here $\vec{R}'\equiv(x',y',z')$.

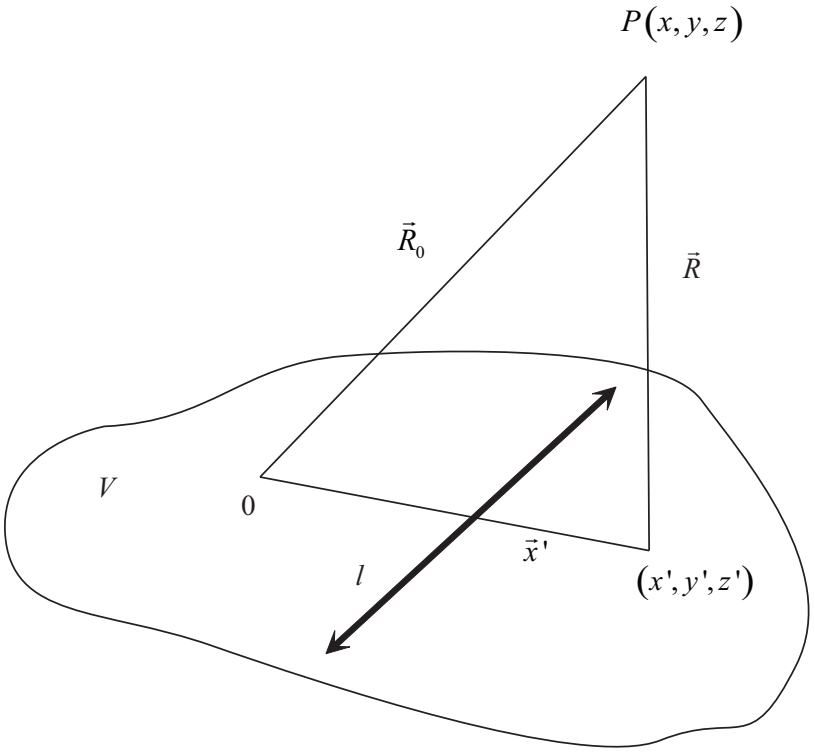

Figure 7.1: A diagrammatic representation of a dipole

By using (II.6.25), we then find the retarded scalar potential

$$\begin{aligned}\varphi &= \int \frac{\rho\left(x', t-\frac{R}{c}\right)}{R}\mathrm{d}^3x' = \\ &= \int \mathrm{d}^3x' \frac{\rho\left(x', t-\frac{R_0}{c}\right)}{R_0} - \frac{\vec{R}_0\cdot\vec{R}'}{R_0}\frac{\partial}{\partial R_0}\frac{\rho\left(x', t-\frac{R_0}{c}\right)}{R_0} + \cdots = \\ &= -\frac{\vec{R}_0}{R_0}\cdot\frac{\partial}{\partial R_0}\frac{1}{R_0}\int \mathrm{d}^3x'\vec{R}'\rho\left(x', t-\frac{R_0}{c}\right),\end{aligned}$$

where the first term vanishes because it is proportional the complete charge of the system, which we have set to zero, by defining the system to be neutral. In the remaining term we will write the integral as $\vec{d}\left(t-\frac{R_0}{c}\right)$, the electric moment at time $t-\frac{R_0}{c}$, which is just a continuous version of (II.7.15)

$$\vec{d}\left(t-\frac{R_0}{c}\right) = \int \mathrm{d}^3x'\vec{R}'\rho\left(x', t-\frac{R_0}{c}\right). \tag{II.7.16}$$

Therefore[2],

$$\varphi = -\frac{\vec{R}}{R}\cdot\frac{\partial}{\partial R}\frac{\vec{d}\left(t-\frac{R}{c}\right)}{R}.$$

Further, we find

$$\mathrm{div}\frac{\vec{d}\left(t-\frac{R}{c}\right)}{R} = \vec{d}\cdot\vec{\nabla}\frac{1}{R} + \frac{1}{R}\mathrm{div}\,\vec{d} = -\frac{\vec{d}\cdot\vec{R}}{R^3} + \frac{1}{R}\mathrm{div}\,\vec{d},$$

[2]To simplify our further treatment, the have changed the notation $R_0 \to R$.

$$\mathrm{div}\,\vec{d} \;=\; \frac{\partial d_i}{\partial x^i} = \frac{\partial d_i}{\partial R}\frac{\partial R}{\partial x^i} = \left(\frac{\vec{R}}{R}\cdot\frac{\partial \vec{d}}{\partial R}\right),$$

so that

$$\mathrm{div}\frac{\vec{d}\left(t-\frac{R}{c}\right)}{R} \;=\; -\frac{\vec{d}\cdot\vec{R}}{R^3} + \frac{\vec{R}}{R^2}\frac{\partial \vec{d}}{\partial R}.$$

On the other hand,

$$\varphi = \frac{\vec{d}\cdot\vec{R}}{R^3} - \frac{\vec{R}}{R^2}\frac{\partial \vec{d}}{\partial R}.$$

Thus,

$$\varphi = -\mathrm{div}\frac{\vec{d}\left(t-\frac{R}{c}\right)}{R}.$$

Here divergence is taken over coordinates of the point $P\,(x,y,z)$ where the observer is located. Using expression (II.6.26), the vector potential becomes

$$\begin{aligned}\vec{A} &= \frac{1}{c}\int \frac{\vec{j}\left(x',t-\frac{R}{c}\right)}{R}\mathrm{d}^3x' = \\ &= \frac{1}{c}\int \mathrm{d}^3x'\Big[\frac{\vec{j}\left(x',t-\frac{R_0}{c}\right)}{R_0} - \frac{\vec{R}_0\cdot\vec{R}'}{\vec{R}_0}\frac{\partial}{\partial R_0}\frac{\vec{j}\left(x',t-\frac{R_0}{c}\right)}{R_0}+\cdots\Big].\end{aligned}$$

First integral can also be expressed via electric moment, which can be achieved by using the continuity equation

$$\frac{\partial}{\partial t}\rho\left(x',t-\frac{R_0}{c}\right) = -\mathrm{div}'\,\vec{j}\left(x',t-\frac{R_0}{c}\right).$$

Multiplying both sides of this equation by time independent $\vec{R}'$, integrating over entire space and using the definition (II.7.16), we can then state that

$$\frac{\partial}{\partial t}\vec{p}\left(t-\frac{R_0}{c}\right) = -\int \mathrm{d}^3x'\vec{R}'\mathrm{div}'\,\vec{j}\left(x',t-\frac{R_0}{c}\right).$$

To proceed, let us sidetrack and consider an arbitrary unit vector $\vec{a}$, i.e. $|\vec{a}|=1$. Then

$$\begin{aligned}(\vec{a}\vec{R}')\mathrm{div}\vec{j} &= \mathrm{div}\left(\vec{j}(\vec{a}\vec{R}')\right) - \vec{j}\cdot\vec{\nabla}'(\vec{a}\vec{R}') \\ &= \mathrm{div}\left(\vec{j}(\vec{a}\vec{R}')\right) - \vec{j}\cdot\vec{a},\end{aligned}$$

where the last step follows from $\vec{a}$ being a constant and $\nabla'\vec{R}'=1$. Based on that we can write

$$\vec{a}\cdot\frac{\partial}{\partial t}\vec{d}\left(t-\frac{R_0}{c}\right) = -\int \mathrm{d}^3x'\mathrm{div}'\left(\vec{j}(\vec{a}\vec{R}')\right) + \vec{a}\cdot\int \mathrm{d}^3x'\vec{j}\left(x',t-\frac{R_0}{c}\right).$$

Since currents do not leave the volume V, we find that

$$\int \mathrm{d}^3x'\mathrm{div}'\left[\vec{j}\left(\vec{a}\vec{R}'\right)\right] = \oint (aR')\,j_n\mathrm{d}S = 0$$

as the normal component j_n of the current vanishes (all currents never leave the integration volume V). This gives

$$\vec{a}\cdot\frac{\partial}{\partial t}\vec{d}\left(t-\frac{R_0}{c}\right) = \vec{a}\cdot\int \mathrm{d}^3x'\vec{j}\left(x',t-\frac{R_0}{c}\right).$$

Since the last relation is valid for any unit vector $\vec{a}$, we obtain that

$$\frac{\partial}{\partial t}\vec{d}\left(t-\frac{R_0}{c}\right)=\int \mathrm{d}^3x'\vec{j}\left(x',t-\frac{R_0}{c}\right).$$

Therefore, we arrive at[3]

$$\vec{A}=\frac{1}{cR}\cdot\frac{\partial}{\partial t}\vec{d}\left(t-\frac{R}{c}\right).$$

We see that both the scalar and the vector potential of any arbitrary neutral system on large distances are defined via the electric moment of this system.

The simplest system of this type is a dipole, *i.e.* two opposite electric charges separated by a certain distance from each other. A dipole whose moment $\vec{d}$ changes in time is called an oscillator (or a vibrator).

Radiation of an oscillator plays an important role in the electromagnetic theory (radiotelegraphic antennae, radiating bodies, proton-electron systems, etc.). To advance our investigation of a dipole, let us introduce the Hertz vector

$$\vec{P}\,(t,R)=\frac{\vec{d}\left(t-\frac{R}{c}\right)}{R}. \qquad \text{(II.7.17)}$$

It is interesting to see that

$$\Delta\vec{P}\,(t,R) \;=\; \vec{\nabla}^2\vec{P}\,(t,R)=\frac{1}{c^2}\frac{\partial^2\vec{P}}{\partial t^2}.$$

This can be derived as follows. First, we notice that

$$\frac{\partial}{\partial x}\vec{P} \;=\; -\frac{1}{R^2}\frac{\partial R}{\partial x}\vec{d}-\frac{1}{cR}\frac{\partial\vec{d}}{\partial t}\frac{\partial R}{\partial x}=-\frac{x}{R^3}\vec{d}-\frac{x}{cR^2}\frac{\partial\vec{d}}{\partial t},$$

since $\frac{\partial R}{\partial x}=\frac{x}{R}$. Differentiating once again, we get

$$\frac{\partial^2}{\partial x^2}\vec{P} \;=\; -\frac{1}{R^3}\vec{d}+3\frac{x^2}{R^5}\vec{d}+\frac{3}{c}\frac{x^2}{R^4}\frac{\partial\vec{d}}{\partial t}-\frac{1}{cR^2}\frac{\partial\vec{d}}{\partial t}+\frac{1}{c^2}\frac{x^2}{R^3}\frac{\partial^2\vec{d}}{\partial t^2},$$

so that

$$\sum_{i=1}^{3}\frac{\partial^2}{\partial x_i^2}\vec{P} \;=\; \frac{1}{c^2R}\frac{\partial^2\vec{d}}{\partial t^2},$$

which represents the spherically symmetric solution of the wave equation.

Consider the retarded potentials

$$\varphi(\vec{R},t) \;=\; -\mathrm{div}\vec{P}\,(t,R)\;,\qquad \vec{A}(\vec{R},t)=\frac{1}{c}\frac{\partial\vec{P}\,(t,R)}{\partial t};$$

The potentials are spherically symmetric, *i.e.* they depend on the distance R only. For the electromagnetic fields we have

$$\begin{aligned}\vec{H} &= \mathrm{rot}\vec{A}\,(t)=\frac{1}{c}\frac{\partial}{\partial t}\mathrm{rot}\vec{P}\,(t,R)\;;\\ \vec{E} &= -\frac{1}{c}\frac{\partial\vec{A}\,(t)}{\partial t}-\vec{\nabla}\varphi=-\frac{1}{c^2}\frac{\partial^2\vec{P}\,(t,R)}{\partial t^2}-\vec{\nabla}\mathrm{div}\vec{P}\,(t,R)\end{aligned}$$

[3]Here we again changed the notation $R_0\to R$.

$$= -\frac{1}{c^2}\frac{\partial^2 \vec{P}\,(t,R)}{\partial t^2} + \vec{\nabla}^2 \vec{P}\,(t,R) + \text{rot rot}\,\vec{P}\,(t,R)\ .$$

On the last line the sum of the first two terms is equal to zero by virtue of the wave equation. This results in

$$\vec{E} = \text{rot rot}\,\vec{P}\,(t,R)\ . \tag{II.7.18}$$

Assume that the electric moment changes only its magnitude, but not its direction, *i.e.*

$$\vec{d}\,(t) = \vec{d}_0 f\,(t)\ .$$

This is not a restriction because moment $\vec{d}$ of an arbitrary oscillator can be decomposed into three mutually orthogonal directions and a field in each direction can be studied separately. Based on this we have

$$\begin{aligned}
\vec{P}\,(t,R) &= \vec{d}_0 \frac{f\left(t-\frac{R}{c}\right)}{R}\,, \\
\text{rot}\,\vec{P} &= \frac{f}{R}\text{rot}\,\vec{d}_0 + \left[\vec{\nabla}\frac{f}{R}, \vec{d}_0\right] = \frac{\partial}{\partial R}\left(\frac{f\left(t-\frac{R}{c}\right)}{R}\right)\left[\frac{\vec{R}}{R}, \vec{d}_0\right] = \\
&= \frac{1}{R}\frac{\partial}{\partial R}\left(\frac{f\left(t-\frac{R}{c}\right)}{R}\right)\left[\vec{R}, \vec{d}_0\right]
\end{aligned}$$

as $\text{rot}\,\vec{d}_0 = 0$. In the spherical coordinate system we compute the corresponding components

$$\begin{aligned}
\left|\left[\vec{R}, \vec{d}_0\right]\right| &= R d_0 \sin\theta\,, \\
\left[\vec{R}, \vec{d}_0\right]_R &= \left[\vec{R}, \vec{d}_0\right]_\theta = 0\,, \\
\left[\vec{R}, \vec{d}_0\right]_\phi &= -R d_0 \sin\theta\,.
\end{aligned}$$

and get[4]

$$\begin{aligned}
\left(\text{rot}\,\vec{P}\right)_R &= \left(\text{rot}\,\vec{P}\right)_\theta = 0\,, \\
\left(\text{rot}\,\vec{P}\right)_\phi &= -d_0 \sin\theta \frac{\partial}{\partial R}\left(\frac{f\left(t-\frac{R}{c}\right)}{R}\right) = -\sin\theta\frac{\partial}{\partial R}P\,(t,R)\ .
\end{aligned}$$

Since the magnetic field components are the components of the curl of the vector potential, the latter is written in terms of the Hertz vector (II.7.17), where we find

$$\begin{aligned}
H_R &= H_\theta = 0 \\
H_\phi &= -\sin\theta\frac{1}{c}\frac{\partial^2 P\,(t,R)}{\partial t\,\partial R}\,.
\end{aligned}$$

The components of curl of any vector field $\vec{a}$ in spherical coordinates are given by

$$\begin{aligned}
(\text{rot}\ \vec{a})_R &= \frac{1}{R\sin\theta}\left(\frac{\partial}{\partial\theta}\left(\sin\theta a_\phi\right) - \frac{\partial a_\theta}{\partial R}\right)\,; \\
(\text{rot}\ \vec{a})_\theta &= \frac{1}{R\sin\theta}\left(\frac{\partial a_R}{\partial\phi} - \frac{\partial}{\partial R}\left(R\sin\theta a_\phi\right)\right)\,; \\
(\text{rot}\ \vec{a})_\phi &= \frac{1}{R}\left(\frac{\partial}{\partial R}\left(R a_\theta\right) - \frac{\partial a_R}{\partial\theta}\right)\,.
\end{aligned}$$

[4]Note that P here is the numerical value of the Herz vector $\vec{P}$.

Using these formulae together with equation (II.7.18), we also find the components of the electric field

$$\begin{aligned} E_R &= \frac{1}{R\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\left(-\sin\theta\right)\frac{\partial}{\partial R}P\left(t,R\right)\right] \\ &= -\frac{1}{R\sin\theta}\frac{\partial}{\partial\theta}\left[\sin^2\theta\frac{\partial P}{\partial R}\right] = -\frac{2\cos\theta}{R}\frac{\partial P}{\partial R}\,; \\ E_\theta &= -\frac{1}{R\sin\theta}\sin\theta\frac{\partial}{\partial R}\left[R\left(-\sin\theta\right)\frac{\partial}{\partial R}P\left(t,R\right)\right] = \\ &= \frac{\sin\theta}{R}\frac{\partial}{\partial R}\left(R\frac{\partial P}{\partial R}\right)\,; \\ E_\phi &= 0\,. \end{aligned}$$

From the above expressions we can see that electric and magnetic fields are always perpendicular; magnetic lines coincide with circles parallel to the equator, while electric field lines are in the meridian planes. Now let us further assume that

$$f\left(t\right)=\cos\omega t \qquad \Rightarrow \qquad \vec{d}\left(t-\frac{R}{c}\right)=\vec{d}_0\cos\omega\left(t-\frac{R}{c}\right)$$

or in a complex form

$$\vec{d}\left(t-\frac{R}{c}\right) = \vec{d}_0 e^{i\omega\left(t-\frac{R}{c}\right)}\,. \tag{II.7.19}$$

Then

$$\begin{aligned} \frac{\partial P}{\partial R} &= \frac{\partial}{\partial R}\left(\frac{d_0 e^{i\omega\left(t-\frac{R}{c}\right)}}{R}\right) = -\frac{1}{R^2}d_0 e^{i\omega\left(t-\frac{R}{c}\right)} - \frac{i\omega}{c}\frac{1}{R}\,d_0 e^{i\omega\left(t-\frac{R}{c}\right)} = \\ &= -\left(\frac{1}{R}+\frac{i\omega}{c}\right)P\left(R,t\right)\,, \end{aligned}$$

and

$$\frac{\partial}{\partial R}\left(R\frac{\partial P}{\partial R}\right) = -\frac{\partial}{\partial R}\left[\left(1+\frac{i\omega R}{c}\right)P\right] = \left(\frac{1}{R}+\frac{i\omega}{c}-\frac{\omega^2 R}{c}\right)P\,.$$

Thus, for this particular case we get the following result

$$\begin{aligned} H_\phi &= \frac{i\omega}{c}\sin\theta\left(\frac{1}{R}+\frac{i\omega}{c}\right)P\left(R,t\right)\,; \\ E_R &= 2\cos\theta\left(\frac{1}{R^2}+\frac{i\omega}{cR}\right)P\left(R,t\right)\,; \\ E_\theta &= \sin\theta\left(\frac{1}{R^2}+\frac{i\omega}{cR}-\frac{\omega^2}{c^2}\right)P\left(R,t\right)\,. \end{aligned}$$

These are the exact expressions for electromagnetic fields of a harmonic oscillator. They are complicated and we will look more closely only on what happens close and far away from the oscillator. To do that we will aid ourselves with the concept of a characteristic scale, which is determined by the competition between

$$\frac{1}{R} \quad \text{and} \quad \frac{\omega}{c}=\frac{2\pi}{Tc}=\frac{2\pi}{\lambda}\,,$$

where T and λ are the period and the wavelength of the electromagnetic wave, respectively.

Close to the oscillator

By "close to the oscillator" we mean:

$$R \ll \frac{\lambda}{2\pi} \quad \text{or} \quad \frac{1}{R} \gg \frac{\omega}{c} = \frac{2\pi}{\lambda}\,,$$

i.e. distances from oscillator are smaller than the wavelength. Thus we can simplify

$$\omega\left(t - \frac{R}{c}\right) = \omega t - R\frac{\omega}{c} = \omega t - \frac{2\pi R}{\lambda} \approx \omega t\,,$$

so that

$$P\left(t, R\right) = \frac{d\left(t - \frac{R}{c}\right)}{R} \approx \frac{d\left(t\right)}{R}\,.$$

Using the "close to oscillator condition", fields are determined by the electric moment $d\left(t\right)$ and its derivative $\frac{\partial d}{\partial t}$ without retarding

$$H_\phi \approx \frac{i\omega}{c}\sin\theta\frac{P}{R} \approx \frac{i\omega}{c}\sin\theta\frac{d\left(t\right)}{R^2} = \frac{1}{c}\frac{\sin\theta}{R^2}\frac{\partial d\left(t\right)}{\partial t}\,,$$

because $i\omega d\left(t\right) = \frac{\partial d(t)}{\partial t}$, which follows from the particular choice of the time dependence of the oscillator that we have made in (II.7.19). Similarly in this limit the electric field components become

$$\begin{aligned} E_R &= \frac{2\cos\theta}{R^2}P = \frac{2\cos\theta}{R^3}d\left(t\right)\,; \\ E_\theta &= \frac{\sin\theta}{R^2}P = \frac{\sin\theta}{R^3}d\left(t\right)\,. \end{aligned}$$

At any given moment t, this is a field of a static dipole. For the magnetic field we find

$$\vec{H} = \frac{1}{cR^3}\frac{\partial\vec{d}\left(t\right)}{\partial t} \times \vec{R} = \frac{J}{cR^3}\,\vec{\ell}\times\vec{R}\,.$$

Given that this introduced current J obeys $J\vec{\ell} = \frac{\partial\vec{d}(t)}{\partial t}$, this expression gives the magnetic field of a current element of length ℓ. This is known as *the Biot-Savart law*[5].

Far away from the oscillator

Let us now consider the region far away from the oscillator, i.e. the region where

$$R \gg \frac{\lambda}{2\pi} \quad \text{or} \quad \frac{1}{R} \ll \frac{\omega}{c} = \frac{2\pi}{\lambda}\,.$$

Distances greater than the wavelength are called wave-zone. In this particular limit our field components become

$$\begin{aligned} H_\phi &= -\frac{\omega^2}{c^2}\sin\theta P = -\frac{\omega^2}{c^2}\sin\theta\frac{d\left(t - \frac{R}{c}\right)}{R}\,; \\ E_R &= 0\,; \\ E_\theta &= -\frac{\omega^2}{c^2}\sin\theta\frac{d\left(t - \frac{R}{c}\right)}{R} = H_\phi\,. \end{aligned}$$

Thus summarizing we get

$$E_R = E_\phi = H_R = H_\theta = 0\,,$$

[5]Note that $E \sim \frac{1}{R^3}$ and $H \sim \frac{1}{R^2}$.

and

$$E_\theta = H_\phi = -\frac{\omega^2 \sin\theta}{c^2 R} d_0 \cos\omega\left(t - \frac{R}{c}\right),$$

or

$$E_\theta = H_\phi = \frac{\sin\theta}{c^2 R} \frac{\partial^2 d\left(t - \frac{R}{c}\right)}{\partial t^2}.$$

This last result is valid for any arbitrary $d(t)$, not necessarily $d_0 f(t)$, because we can always perform a harmonic Fourier decomposition of any function. Thus in the wave zone the electric and magnetic fields are equal to each other and vanish as $\frac{1}{R}$. Additionally, vectors $\vec{E}$, $\vec{H}$, and $\vec{R}$ are perpendicular[6]. Note that the phase of $\vec{E}$ and $\vec{H}$, i.e. $\omega\left(t - \frac{R}{c}\right)$ moves with the speed of light.

Thus, in the wave zone of the oscillator an electromagnetic wave is propagating!

$$\lambda = cT = \frac{2\pi c}{\omega}.$$

This wave propagates in the radial direction, *i.e.* its phase depends on the distance to the center.

Let us now look at the Poynting vector

$$S = \frac{c}{4\pi}\left|\left[\vec{E}, \vec{H}\right]\right| = \frac{c}{4\pi} EH = \frac{1}{4\pi}\frac{\sin^2\theta}{c^3 R^2}\left(\frac{\partial^2 d\left(t - \frac{R}{c}\right)}{\partial t^2}\right)^2,$$

where on the first step we have used the fact that the electric and the magnetic fields are perpendicular. Additionally note that the second derivative with respect to time inside the square is an acceleration. Energy flux through the sphere of radius R is

$$\begin{aligned}\Sigma &= \int_0^{2\pi}\int_0^{\pi} SR^2 \sin\theta \mathrm{d}\phi \mathrm{d}\theta = \\ &= \int_0^{2\pi}\int_0^{\pi} \frac{1}{4\pi}\frac{\sin^2\theta}{c^3 R^2}\left(\frac{\partial^2 d\left(t - \frac{R}{c}\right)}{\partial t^2}\right)^2 R^2 \sin\theta \mathrm{d}\phi \mathrm{d}\theta = \frac{2}{3c^3}\left[\frac{\partial^2 d\left(t - \frac{R}{c}\right)}{\partial t^2}\right]^2 = \frac{2}{3c^3}\ddot{d}^2.\end{aligned}$$

For $d\left(t - \frac{R}{c}\right) = d_0 \cos\omega\left(t - \frac{R}{c}\right)$ the flux for one period is

$$\begin{aligned}\int_0^T \Sigma \,\mathrm{d}t &= \frac{2}{3c^3} d_0^2 \omega^4 \int_0^T \cos^2\omega\left(t - \frac{R}{c}\right)\mathrm{d}t = \\ &= \frac{d_0^2 \omega^4 T}{3c^3} = \frac{2\pi d_0^2 \omega^3}{3c^3} = \frac{2\pi d_0^2}{3}\left(\frac{2\pi}{\lambda}\right)^3.\end{aligned}$$

The averaged radiation in a unit time is then

$$\langle\, \Sigma \,\rangle = \frac{1}{T}\int_0^T \Sigma \mathrm{d}t = \frac{c d_0^2}{3}\left(\frac{2\pi}{\lambda}\right)^4. \tag{II.7.20}$$

Thus, the oscillator continuously radiates energy into surrounding space with average rate $\langle\, \Sigma \,\rangle \sim d_0^2 \frac{1}{\lambda^4}$. In particular this explains that when transmitting radio signals by telegraphing one should

[6]Note that $\vec{E}$, $\vec{H}$ and $\vec{R}$ have completely mismatching components i.e. if one vector has a particular non-zero component, for the other two this component is zero.

use waves of relatively short wavelengths[7] (or equivalently high frequencies ω). On the other hand, radiation of low frequency currents is highly suppressed, which explains the effect of the sky appearing in blue, which is to the high frequency end of the visible light[8] spectrum.

Lastly, let us finally focus on the concept of resistance to radiation, which is given by R_λ such that

$$\langle\, \Sigma \,\rangle = R_\lambda \langle\, J^2 \,\rangle\,.$$

Recall that we have previously defined J such that it obeys $J\vec{\ell} = \frac{\partial \vec{d}\left(t-\frac{R}{c}\right)}{\partial t}$. Using this definition, we get

$$\begin{aligned}\langle\, J^2 \,\rangle &= \frac{1}{T}\int\limits_0^T J^2 \mathrm{d}t = \frac{1}{T\ell^2}\int\limits_0^T \left(\frac{\partial \vec{p}\left(t-\frac{R}{c}\right)}{\partial t}\right)^2 \mathrm{d}t = \\ &= \frac{1}{T\ell^2}\int\limits_0^T d_0^2\omega^2 \sin^2\omega\left(t-\frac{R}{c}\right)\mathrm{d}t = \frac{d_0^2\omega^2}{T\ell^2}\frac{\pi}{\omega} = \frac{\pi d_0^2\omega^2}{\ell^2\frac{2\pi}{\omega}\omega} = \frac{d_0^2\omega^2}{2\ell^2}\,.\end{aligned}$$

Using the result (II.7.20), it is now easy to find R_λ

$$R_\lambda = \frac{cd_0^2}{3}\left(\frac{2\pi}{\lambda}\right)^4 \frac{2\ell^2}{d_0^2\omega^2} = \frac{2c}{3\ell^2}\left(\frac{2\pi}{\lambda}\right)^4 \frac{1}{\left(\frac{2\pi}{\lambda}c\right)^2} = \frac{2}{3c}\left(\frac{2\pi\ell}{\lambda}\right)^2\,.$$

7.4 Applicability of classical electrodynamics

We conclude this section by pointing out the range of applicability of classical electrodynamics.

The energy of the charge distribution in electrodynamics is given by

$$U = \int \mathrm{d}V \rho(x)\varphi(x)\,.$$

Putting electron at rest, one can assume that the entire energy of the electron coincides with its electromagnetic energy (electric charge is assumed to be homogeneously distributed over a ball of the radius r_e)

$$mc^2 \sim \frac{e^2}{r_e}\,,$$

where m and e are the mass and the charge of electron. Thus, we can define *the classical radius of electron*

$$r_e = \frac{e^2}{mc^2}\,.$$

In SI units it reads as $r_e = \frac{1}{4\pi\epsilon_0}\frac{e^2}{mc^2} \sim 2.818\cdot 10^{-15}\,\mathrm{m}$. At distances less than r_e, the classical electrodynamics is not applicable.

In reality, due to quantum effects the classical electrodynamics fails even at larger distances. The characteristic scale is the *Compton wavelength* λ_e, which is the fundamental limitation on measuring

[7] Generally these range from tens of meters to tens of kilometers.

[8] In this case charge polarized chemical bonds between the atoms in the particles in the atmosphere act as little oscillators.

the position of a particle taking both quantum mechanics and special relativity into account. Its theoretical value is given by

$$\lambda_e = \frac{\hbar}{mc} \sim 137\, r_e \sim 10^{-13}\,\mathrm{m}\,,$$

where $\alpha = \frac{1}{137} = \frac{e^2}{\hbar c}$ is the fine structure constant for electromagnetism. The most recent experimental measurement of Compton wavelength (CODATA 2002) is one order of magnitude larger and is approximately equal to $2.426 \cdot 10^{-12}\,\mathrm{m}$.

7.5 Darvin's Lagrangian

In classical mechanics a system of interacting particles can be described by a proper Lagrangian which depends on coordinates and velocities of all particles taken at the one and the same moment. This is possible because in mechanics the speed of propagation of signals is assumed to be *infinite*.

On the other hand, in electrodynamics field should be considered as an independent entity having its own degrees of freedom. Therefore, if one has a system of interacting charges (particles) for its description one should consider a system comprising both these particles and the field. Thus, taking into account that the propagation speed of interactions is finite, we arrive at the conclusion that the rigorous description of a system of interacting particles with the help of the Lagrangian depending on their coordinates and velocities but do not containing degrees of freedom related to the field is *impossible*.

However, if velocities v of all the particles are small with respect to the speed of light, then such a system can be approximately described by some Lagrangian. The introduction of the Lagrangian function is possible up to the terms of order $\frac{v^2}{c^2}$. This is related to the fact that radiation of electromagnetic waves by moving charges (that is an appearance of independent field) arises in the third order of $\frac{v}{c}$ only.

At zero approximation, *i.e.* by completely neglecting retarding of the potentials, the Lagrangian for a system of charges has the form

$$L^{(0)} = \sum_i \frac{m_i v_i^2}{2} - \sum_{i>j} \frac{e_i e_j}{r_{ij}}\,.$$

The second term is the potential energy of non-moving charges.

In order to find higher approximation, we first write the Lagrangian for a charge e_i in an external electromagnetic field $(\varphi, \vec{A})$:

$$L_i = -mc^2\sqrt{1-\frac{v_i^2}{c^2}} - e_i\varphi + \frac{e_i}{c}(\vec{A}\cdot\vec{v}_i)\,.$$

Picking up one of the charges, we determine electromagnetic potentials created by all the other charges in a point where this charge sits and express them via coordinates and velocities of the corresponding charges (this can be done only approximately: φ can be determined up to the order $\frac{v^2}{c^2}$ and $\vec{A}$ up to $\frac{v}{c}$). Substituting the found expressions for the potentials in the previous formula, we will find the Lagrangian for the whole system.

Consider the retarded potentials

$$\varphi(x,t) \;=\; \int \mathrm{d}^3x'\mathrm{d}t'\,\frac{\delta\left(t'+\frac{|x-x'|}{c}-t\right)}{|x-x'|}\rho(x',t')\,,$$

$$\vec{A}(x,t) \;=\; \frac{1}{c}\int \mathrm{d}^3x'\mathrm{d}t'\,\frac{\delta\left(t'+\frac{|x-x'|}{c}-t\right)}{|x-x'|}\vec{j}(x',t')\,.$$

As before, integrating over t' we get

$$\varphi(x,t)=\int \mathrm{d}^3x'\,\frac{\rho\left(t-\frac{|x-x'|}{c}\right)}{|x-x'|}\,,\qquad \vec{A}(x,t)=\frac{1}{c}\int \mathrm{d}^3x'\,\frac{\vec{j}\left(t-\frac{|x-x'|}{c}\right)}{|x-x'|}\,.$$

If velocities of all the charges are small in comparison to the speed of light, then the distribution of charges does not change much for the time $\frac{|x-x'|}{c}$. Thus, the sources can be expanded in series in $\frac{|x-x'|}{c}$. we have

$$\varphi(x,t)=\int \mathrm{d}^3x'\,\frac{\rho(t)}{R}-\frac{1}{c}\frac{\partial}{\partial t}\int \mathrm{d}^3x'\,\rho(t)+\frac{1}{2c^2}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\,R\rho(t)+\ldots$$

where $R=|x-x'|$. Since $\int \mathrm{d}^3x'\,\rho(t)$ is a constant charge of the system, we have at leading and subleading orders the following expression for the scalar potential

$$\varphi(x,t)=\int \mathrm{d}^3x'\,\frac{\rho(t)}{R}+\frac{1}{2c^2}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\,R\rho(t)\,.$$

Analogous expansion takes place for the vector potential. Since expression for the vector potential via the current already contains $1/c$ and after the substitution in the Lagrangian is multiplied by another power $1/c$, it is enough to keep in the expansion of $\vec{A}$ the leading term only, *i.e.*

$$\vec{A}=\frac{1}{c}\int \mathrm{d}x'\,\frac{\rho\vec{v}}{R}\,.$$

If the field is created by a single charge, we have

$$\varphi=\frac{e}{R}+\frac{e}{2c^2}\frac{\partial^2 R}{\partial t^2}\,,\qquad \vec{A}=\frac{e\vec{v}}{cR}\,.$$

To simplify further treatment, we will make the gauge transformation

$$\varphi'=\varphi-\frac{1}{c}\frac{\partial\chi}{\partial t}\,,\qquad \vec{A}'=\vec{A}+\vec{\nabla}\chi\,,$$

where

$$\chi=\frac{e}{2c}\frac{\partial R}{\partial t}\,.$$

This gives

$$\varphi'=\frac{e}{R}\,,\qquad \vec{A}'=\frac{e\vec{v}}{cR}+\frac{e}{2c}\vec{\nabla}\frac{\partial R}{\partial t}\,.$$

Here $\vec{\nabla}\frac{\partial R}{\partial t}=\frac{\partial}{\partial t}\vec{\nabla}_x R$ and $\vec{\nabla}_x R=\frac{\vec{R}}{R}=\vec{n}$, where $\vec{n}$ is the unit vector directed from the charge to the observation point. Thus,

$$\vec{A}'=\frac{e\vec{v}}{cR}+\frac{e}{2c}\frac{\partial}{\partial t}\left(\frac{\vec{R}}{R}\right)=\frac{e\vec{v}}{cR}+\frac{e}{2c}\left(\frac{\dot{\vec{R}}}{R}-\frac{\vec{R}\dot{R}}{R^2}\right)=\frac{e\vec{v}}{cR}+\frac{e}{2c}\left(\frac{-\vec{v}}{R}-\frac{\vec{R}\dot{R}}{R^2}\right)\,.$$

Finally, since $R^2=\vec{R}^2$, we find $R\dot{R}=\vec{R}\cdot\dot{\vec{R}}=-\vec{R}\cdot\vec{v}$. In this way we find

$$\varphi'=\frac{e}{R}\,,\qquad \vec{A}'=\frac{e\left[\vec{v}+(\vec{v}\cdot\vec{n})\,\vec{n}\right]}{2cR}\,.$$

If the field is created by several charges then this expression must be summed for all the charges.

Now substituting the potentials created by all the other charges into the Lagrangian for a given charge e_i we obtain

$$L_i = \frac{m_i v_i^2}{2} + \frac{1}{8}\frac{m_i v_i^4}{c^2} - e_i \sum_{j\neq i} \frac{e_j}{r_{ij}} + \frac{e_i}{2c^2}\sum_{j\neq i}\frac{e_j}{r_{ij}}\left[(\vec{v}_i\cdot\vec{v}_j) + (\vec{v}_i\cdot\vec{n}_{ij})(\vec{v}_j\cdot\vec{n}_{ij})\right].$$

Here we have also expanded the relativistic Lagrangian for the point particle up to the order $\frac{v^2}{c^2}$. From this expression we can find the total Lagrangian

$$L = \sum_i \frac{m_i v_i^2}{2} + \sum_i \frac{m_i v_i^4}{8c^2} - \sum_{i>j}\frac{e_i e_j}{r_{ij}} + \sum_{i>j}\frac{e_i e_j}{2c^2 r_{ij}}\left[(\vec{v}_i\cdot\vec{v}_j) + (\vec{v}_i\cdot\vec{n}_{ij})(\vec{v}_j\cdot\vec{n}_{ij})\right].$$

This Lagrangian was obtained by Darvin in 1922 and it expresses an effect of electromagnetic interaction between charges up to the second order in $\frac{v}{c}$.

It is interesting to find out what happens if we expand the potential further. For the scalar potential at third order in $1/c$ and for the vector potential at second order in $1/c$ one finds

$$\varphi^{(3)} = -\frac{1}{6c^3}\frac{\partial^3}{\partial t^3}\int \mathrm{d}^3x'\ R^2\rho\,, \qquad \vec{A}^{(2)} = -\frac{1}{c^2}\frac{\partial}{\partial t}\int \mathrm{d}^3x'\vec{j}\,.$$

Performing a gauge transformation

$$\varphi' = \varphi - \frac{1}{c}\frac{\partial\chi}{\partial t}\,, \qquad \vec{A}' = \vec{A} + \vec{\nabla}\chi$$

with

$$\chi = -\frac{1}{6c^2}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\ R^2\rho\,,$$

we transform $\varphi^{(3)}$ into zero. The new vector potential will take the form

$$\begin{aligned}
\vec{A}'^{(2)} &= -\frac{1}{c^2}\frac{\partial}{\partial t}\int \mathrm{d}^3x'\vec{j} - \frac{1}{6c^2}\frac{\partial^2}{\partial t^2}\vec{\nabla}\int \mathrm{d}^3x'\ R^2\rho \\
&= -\frac{1}{c^2}\frac{\partial}{\partial t}\int \mathrm{d}^3x'\vec{j} - \frac{1}{3c^2}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\ \vec{R}\rho\ = \\
&= -\frac{1}{c^2}\sum e\dot{\vec{v}} - \frac{1}{3c^2}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\ (\vec{R}_0 - \vec{r})\rho\ = -\frac{2}{3c^2}\sum e\dot{\vec{v}}. \qquad \text{(II.7.21)}
\end{aligned}$$

In the last formula we pass to the discrete distribution of charges. This potential leads to a vanishing magnetic field $\vec{H} = \mathrm{rot}_x\vec{A}'^{(2)}$, as curl is taken with respect to the coordinates x of observation point which $\vec{A}'^{(2)}$ does not depend on. For the electric field one finds $\vec{E} = -\dot{\vec{A}}'^{(2)}/c$, so that

$$\vec{E} = \frac{2}{3c^3}\dddot{\vec{d}}\,,$$

where $\vec{d}$ is the dipole moment of the system. Thus, additional terms of the third order in the expansion of fields lead to the appearance of additional forces which are not contained in Darvin's Lagrangian; these forces do depend on time derivatives of charge accelerations.

Compute the averaged work performed by fields for one unit of time. Each charge experienced a force $\vec{F} = e\vec{E}$ so that

$$\vec{F} = \frac{2e}{3c^3}\dddot{\vec{d}}\,.$$

The work produced is

$$\sum(\vec{F}\cdot\vec{v}) = \frac{2e}{3c^3}(\dddot{\vec{d}}\cdot\sum e\vec{v}) = \frac{2}{3c^2}(\dddot{\vec{d}}\cdot\dot{\vec{d}}) = \frac{2}{3c^3}\frac{d}{dt}(\dot{d}\cdot\ddot{\vec{d}}) - \frac{2}{3c^3}\ddot{\vec{d}}^2\,.$$

Performing time average we arrive at

$$\sum(\vec{F}\cdot\vec{v}) = -\frac{2}{3c^3}\ddot{\vec{d}}^2\,.$$

Now one can recognize that the expression of the right hand side of the last formula is nothing else but the average radiation of the system for one unit of time. Thus, the forces arising at third order describe the backreaction which radiation causes on charges. These forces are known as *bracing by radiation or Lorentz friction forces.*

Part III

Appendices

7.6 Trigonometric formulae

Some important trigonometric formulae

$$\begin{aligned}
\sin(x \pm y) &= \sin x \cos y \pm \sin y \cos x \\
\cos(x \pm y) &= \cos x \cos y \mp \sin x \sin y \\
\sin x \pm \sin y &= 2 \sin \frac{x \pm y}{2} \cos \frac{x \mp y}{2} \\
\cos x + \cos y &= 2 \cos \frac{x + y}{2} \cos \frac{x - y}{2} \\
\cos x - \cos y &= -2 \sin \frac{x + y}{2} \sin \frac{x - y}{2}
\end{aligned}$$

7.7 Helmholtz theorem

A vector field can be reconstructed from its curl and divergence. More precisely, the Helmholtz theorem takes place.

The Helmholtz theorem. *An arbitrary vector field $\vec{A}(\vec{x})$ can always be decomposed into the sum of two vector fields: one with zero divergence and one with zero curl,*

$$\vec{A} = \vec{A}_\perp + \vec{A}_{||}\,,$$

where

$$\vec{\nabla} \cdot \vec{A}_\perp = 0\,, \qquad \vec{\nabla} \times \vec{A}_{||} = 0\,.$$

An explicit representation of special interest is

$$\vec{A}(\vec{x}) = \vec{\nabla} \times \frac{1}{4\pi} \int d^3x' \frac{\vec{\nabla}' \times \vec{A}(\vec{x}')}{|\vec{x} - \vec{x}'|} - \vec{\nabla} \frac{1}{4\pi} \int d^3x' \frac{\vec{\nabla}' \cdot \vec{A}(\vec{x}')}{|\vec{x} - \vec{x}'|}\,.$$

This representation expresses the vector field via its curl and divergence.

Here is the proof. Suppose $\vec{A}(\vec{x})$ is a vector for which its divergence and its curl are known, that is

$$\vec{\nabla} \cdot \vec{A} = f\,, \qquad \vec{\nabla} \times \vec{A} = \vec{K}\,, \tag{III.7.1}$$

where a function f and a vector $\vec{K}$ are given. The question is how to reconstruct from this data the field $\vec{A}$. This can be done as follows. We take the curl of the second equation in (III.7.1) and obtain

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla} \times \vec{K}\,.$$

Now for the left hand side we apply the known formula $\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - \Delta \vec{A}$, so that

$$\vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - \Delta \vec{A} = \vec{\nabla} f - \Delta \vec{A} = \vec{\nabla} \times \vec{K}\,.$$

From here we get the Poisson equation for each component of $\vec{A}$

$$\Delta \vec{A} = \vec{\nabla} f - \vec{\nabla} \times \vec{K} = -4\pi \left[\frac{1}{4\pi} \vec{\nabla} \times \vec{K} - \frac{1}{4\pi} \vec{\nabla} f \right].$$

This equation has a unique solution given by

$$\vec{A}(\vec{x}) = \frac{1}{4\pi} \int d^3x' \frac{\vec{\nabla}' \times \vec{K}(\vec{x}')}{|\vec{x} - \vec{x}'|} - \frac{1}{4\pi} \int d^3x' \frac{\vec{\nabla}' f(\vec{x}')}{|\vec{x} - \vec{x}'|}\,.$$

Further, we integrate the derivatives by parts, omit the boundary terms[9] and then replace the gradient $\vec{\nabla}'$ taken with respect to $\vec{x}'$ by $\vec{\nabla}$ taken with respect to $\vec{x}$

$$\begin{aligned}\vec{A}(\vec{x}) &= -\frac{1}{4\pi}\int d^3x'\vec{\nabla}'\left(\frac{1}{|\vec{x}-\vec{x}'|}\right)\times\vec{K}(\vec{x}')+\frac{1}{4\pi}\int d^3x'\vec{\nabla}'\left(\frac{1}{|\vec{x}-\vec{x}'|}\right)f(\vec{x}')\\ &= \frac{1}{4\pi}\vec{\nabla}\times\int d^3x'\frac{\vec{K}(\vec{x}')}{|\vec{x}-\vec{x}'|}-\frac{1}{4\pi}\vec{\nabla}\int d^3x'\frac{f(\vec{x}')}{|\vec{x}-\vec{x}'|}\,.\end{aligned}$$

Now we can substitute here the expressions (III.7.1) and obtain the statement of the Helmholtz theorem

$$\vec{A}(\vec{x})=\frac{1}{4\pi}\vec{\nabla}\times\int d^3x'\frac{\vec{\nabla}'\times A(\vec{x}')}{|\vec{x}-\vec{x}'|}-\frac{1}{4\pi}\vec{\nabla}\int d^3x'\frac{\vec{\nabla}'\cdot\vec{A}(\vec{x}')}{|\vec{x}-\vec{x}'|}\,. \tag{III.7.2}$$

7.8 Tensors

Many geometric and physical quantities can be described only as a set of functions depending on a chosen coordinate system $(x^1,\ldots,x^n)$. The representation of these quantities may drastically change if another coordinate system is chosen $(z^1,\ldots,z^n)$:

$$x^i=x^i(z^1,\ldots,z^n)\,,\qquad i=1,\ldots,n.$$

Vectors

Consider, for instance, a velocity vector along a given curve $z^j=z^j(t)$. In z-coordinates the components of the velocity vector are

$$\left(\frac{dz^1}{dt},\ldots,\frac{dz^n}{dt}\right)=(\eta^1,\ldots,\eta^n)\,.$$

In the other coordinate system we will have

$$\left(\frac{dx^1}{dt},\ldots,\frac{dx^n}{dt}\right)=(\xi^1,\ldots,\xi^n)\,.$$

Obviously,

$$\frac{dx^i}{dt}=\sum_{j=1}^{n}\frac{\partial x^i}{\partial z^j}\frac{dz^j}{dt}\,.$$

Therefore, for the components of the velocity vector one finds

$$\xi^i=\sum_{j=1}^{n}\eta^j\,\frac{\partial x^i}{\partial z^j}\,.$$

Here ξ^i are components of the vector in coordinates $(x^1,\ldots,x^n)$ at a given point, while η^i are components of the vector in coordinates $(z^1,\ldots,z^n)$ at the same point.

Co-vectors

Consider the gradient of a function $f(x^1,\ldots,x^n)$:

$$\nabla f=\left(\frac{\partial f}{\partial x^1},\ldots,\frac{\partial f}{\partial x^n}\right)=(\xi_1,\ldots,\xi_n)\,.$$

[9] We assume that they vanish as the boundary tends to infinity.

In z-coordinates one has

$$\nabla f = \left(\frac{\partial f}{\partial z^1}, \ldots, \frac{\partial f}{\partial z^n} \right) = (\eta_1, \ldots, \eta_n) \, .$$

Obviously,

$$\frac{\partial f}{\partial z^i} = \sum_{j=1}^{n} \frac{\partial f}{\partial x^j} \frac{\partial x^j}{\partial z^i} \implies \quad \eta_i = \frac{\partial x^j}{\partial z^i} \xi_j \, .$$

To compare vector and co-vector transformation laws, let us introduce the Jacobi matrix A with elements $A^i_j = \frac{\partial x^i}{\partial z^j}$. It is convenient to think about a vector as being a column and about a co-vector as being a row, *i.e.* transposed column. Then we have

$$\begin{aligned} \text{Velocity vector} \qquad & \xi \quad = A\eta \, , \\ \text{Gradient} \qquad & \eta^t \quad = \xi^t A \, . \end{aligned}$$

After taking transposition of the second line, we get

$$\begin{aligned} \text{Velocity vector} \qquad & \xi \quad = A\eta \, , \\ \text{Gradient} \qquad & \eta \quad = A^t \xi \, . \end{aligned}$$

This clearly shows that vectors and co-vectors have different transformation laws.

Metric

Recall that the length of a curve is the length of the velocity vector integrated over time. Therefore, in order for the length to be an invariant quantity, that is not to depend on a choice of the coordinate system, the square of the length of the velocity vector

$$|v|^2 = g_{ij} \xi^i \xi^j$$

should be independent of the coordinates chosen. This requirement together with the transformation law for vectors leads to the following transformation law for the metric under general coordinate transformation

$$g'_{ij}(z) = g_{kl}(x) \frac{\partial x^k}{\partial z^i} \frac{\partial x^l}{\partial z^j} \, , \qquad x^i = x^i(z) \, .$$

Metric constitutes an example of a second rank tensor (it has two indices) with two lower indices, both of them transforming in the co-vector fashion.

These examples of tensorial objects can be continued. For instance, a linear operator A^j_i represents an example of a tensor with one index up and another index down signifying that under general coordinate transformations the index j transforms in the same way as the index of a vector, while i transforms in the same way as the index of a co-vector.

Tensor fields

Let us associate to each point x of space-time a collection of numbers encoded into an object (tensor) $\phi^{j_1 \ldots j_q}_{i_1 \ldots i_p}(x)$ with p upper indices transforming in the vector fashion and q lower indices transforming in the co-vector one. This object is called a tensor field of (p, q)-type if under a transformation of coordinates x^μ: $x^\mu \to x'^\mu(x^\nu)$, it transforms as follows[10]

$$\phi'^{\mu_1 \ldots \mu_p}{}_{\nu_1 \ldots \nu_q}(x') = \frac{\partial x'^{\mu_1}}{\partial x^{\lambda_1}} \cdots \frac{\partial x'^{\mu_p}}{\partial x^{\lambda_p}} \frac{\partial x^{\rho_1}}{\partial x'^{\nu_1}} \cdots \frac{\partial x^{\rho_q}}{\partial x'^{\nu_q}} \, \phi^{\lambda_1 \ldots \lambda_p}{}_{\rho_1 \ldots \rho_q}(x) \, .$$

[10]There is a simple rule to memorise the appearance of primed and unprimed indices in the tensor transformation rule. Assuming that all indices on the left hand side of the tensor transformation formula are 'primed', the 'primed' coordinates corresponding to these primed indices must appear on the right hand side and in the same position. Notationally we keep sets of upper and lower indices shifted with respect to each other to prevent possible confusion on their position under lowering or raising, respectively.

Here tensor indices are acted with the matrices $\frac{\partial x'^\mu}{\partial x^\nu}$ which form a group $\mathrm{GL}(d,\mathbb{R})$. This is a group of all invertible real $d\times d$ matrices. Further restrictions on possible transformations of coordinates can be imposed by physical requirements. For instance, Galilean covariance restricts general coordinate transformations to that of the rotation group. Analogously, Einstein relativity principle allows only for Lorentz transformations. In this case one speaks about tensors on the rotation group, Lorentz group, etc.

The simplest example is a scalar field that does not carry any indices. Its transformation law under coordinate transformations is $\phi'(x') = \phi(x)$. We stress that a point with coordinates x in the original frame and a point with coordinates x' in the transformed frame is the one and the same geometric point.

Properties of the Levi-Civita tensor

$$\epsilon_{kil}\epsilon_{mpq} = \begin{vmatrix} \delta_{km} & \delta_{im} & \delta_{lm} \\ \delta_{kp} & \delta_{ip} & \delta_{lp} \\ \delta_{kq} & \delta_{iq} & \delta_{lq} \end{vmatrix} . \tag{III.7.3}$$

Single sum over repeated index i

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km} . \tag{III.7.4}$$

7.9 Functional derivative

Let $F[f]$ be a functional and η is a differentiable function. The functional derivative $\delta F \equiv \frac{\delta F}{\delta f(x)}$ is a distribution defined for a test function η as

$$\langle \delta F, \eta \rangle = \lim_{\epsilon\to 0} \frac{d}{d\epsilon} F[f+\epsilon\eta] .$$

Consider for instance the following functional

$$F[x(t)] = \frac{1}{2}\int \mathrm{d}t\, g_{ij}(x(t))\dot{x}^i\dot{x}^j .$$

Here $g_{ij}(x)$ is a metric on a smooth n-dimensional manifold M^n which has local coordinates $x^k(t)$. Then

$$\begin{aligned} \langle \delta F, \eta \rangle &= \lim_{\epsilon\to 0}\frac{d}{d\epsilon}\frac{1}{2}\int \mathrm{d}t\, g_{ij}(x(t)+\epsilon\eta)(\dot{x}^i+\epsilon\dot{\eta}^i)(\dot{x}^j+\epsilon\dot{\eta}^j) = \\ &= \lim_{\epsilon\to 0}\frac{d}{d\epsilon}\frac{1}{2}\int \mathrm{d}t\Big[g_{ij}(x)+\epsilon\frac{\partial g_{ij}}{\partial x^k}\eta^k+\dots\Big]\Big[\dot{x}^i\dot{x}^j+2\epsilon\dot{x}^i\dot{\eta}^j+\dots\Big] \\ &= \int \mathrm{d}t\Big[-\frac{d}{dt}(g_{ik}\dot{x}^i)+\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j\Big]\eta^k . \end{aligned}$$

Thus, for the corresponding variational derivative we find

$$\frac{\delta F}{\delta x^k(t)} = -\frac{d}{dt}(g_{ik}\dot{x}^i)+\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j = -\frac{d}{dt}(g_{ik})\dot{x}^i - g_{ik}\ddot{x}^i+\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j .$$

Vanishing of this functional derivative gives an extremality condition for the corresponding functional, which is nothing else but the geodesic equation

$$\ddot{x}^i + \Gamma^i_{kl}\dot{x}^k\dot{x}^l = 0 ,$$

where

$$\Gamma^i_{kl} = \frac{1}{2} g^{ij} \Big(\frac{\partial g_{jk}}{\partial x^l} + \frac{\partial g_{jl}}{\partial x^k} - \frac{\partial g_{kl}}{\partial x^j} \Big)$$

is the Christoffel symbol.

Note that a function itself, *i.e.* $u(x)$, can be considered as the functional

$$u(x) = \int \mathrm{d}x\, u(y) \delta(x-y)\,.$$

From this one can deduce the functional derivative

$$\frac{\delta u(x)}{\delta u(y)} = \delta(x-y)\,.$$

7.10 Introduction to Lie groups and Lie algebras

To introduce a concept of a Lie group we need two notions: the notion of a group and the notion of a smooth manifold.

Definition of a group. A set of elements G is called a group if it is endowed with two operations: for any pair g and h from G there is a third element from G which is called the product gh, for any element $g \in G$ there is the inverse element $g^{-1} \in G$. The following properties must be satisfied

- $(fg)h = f(gh)$
- there exists an identity element $\mathbb{I} \in G$ such that $\mathbb{I}g = g\mathbb{I} = g$
- $gg^{-1} = \mathbb{I}$

Definition of a smooth manifold. Now we introduce the notion of a differentiable manifold. Any set of points is called a differentiable manifold if it is supplied with the following structure

- M is a union: $M = \cup_q U_q$, where U_q is homeomorphic (i.e. a continuous one-to-one map) to the n-dimensional Euclidean space
- Any U_q is supplied with coordinates x^α_q called the *local coordinates*. The regions U_q are called coordinate *charts*.
- any intersection $U_q \cap U_p$, if it is not empty, is also a region of the Euclidean space where two coordinate systems x^α_q and x^α_p are defined. It is required that any of these two coordinate systems is expressible via the other by a differentiable map:

$$\begin{aligned} x^\alpha_p &= x^\alpha_p(x^1_q, \cdots x^n_q)\,, \qquad \alpha = 1, \cdots n \\ x^\alpha_q &= x^\alpha_q(x^1_p, \cdots x^n_p)\,, \qquad \alpha = 1, \cdots n \end{aligned} \tag{III.7.5}$$

  Then the Jacobian $\det\Big(\frac{\partial x^\alpha_p}{\partial x^\beta_q}\Big)$ is different from zero. The functions (III.7.5) are called *transition functions* from coordinates x^α_q to x^α_p and vice versa. If all the transition functions are infinitely differentiable (i.e. have all partial derivatives) the corresponding manifold is called *smooth*.

Definition of a Lie group: A smooth manifold G of dimension n is called a Lie group if G is supplied with the structure of a group (multiplication and inversion) which is compatible with the

structure of a smooth manifold, i.e., the group operations are smooth. In other words, a Lie group is a group which is simultaneously a smooth manifold and the group operations are smooth.

The list of basic matrix Lie groups

- The group of $n \times n$ invertible matrices with complex or real matrix elements:

$$A = a_i^j\,, \qquad \det A \neq 0$$

It is called *the general linear group* $GL(n,\mathbb{C})$ or $GL(n,\mathbb{R})$. Consider for instance $GL(n,\mathbb{R})$. Product of two invertible matrices is an invertible matrix is invertible; an invertible matrix has its inverse. Thus, $GL(n,\mathbb{R})$ is a group. Condition $\det A \neq 0$ defines a domain in the space of all matrices $M(n,\mathbb{R})$ which is a linear space of dimension n^2. Thus, the general linear group is a domain in the linear space $\mathbb{R}^{n^2}$. Coordinates in $M(n,\mathbb{R})$ are the matrix elements a_i^j. If A and B are two matrices then their product $C = AB$ has the form

$$c_i^j = a_i^k b_k^j$$

It follows from this formula that the coordinates of the product of two matrices is expressible via their individual coordinates with the help of smooth functions (polynomials). In other words, the group operation which is the map

$$GL(n,\mathbb{R}) \times GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$$

is smooth. Matrix elements of the inverse matrix are expressible via the matrix elements of the original matrix as no-where singular rational functions (since $\det A \neq 0$) which also defines a smooth mapping. Thus, the general Lie group is a Lie group.

- *Special linear group* $SL(n,\mathbb{R})$ or $SL(n,\mathbb{C})$ is a group of real or complex matrices satisfying the condition

$$\det A = 1\,.$$

- *Special orthogonal group* $SO(n,\mathbb{R})$ or $SO(n,\mathbb{C})$ is a group or real or complex matrices satisfying the conditions

$$AA^t = \mathbb{I}\,, \qquad \det A = 1\,.$$

- *Pseudo-orthogonal groups* $SO(p,q)$. Let g will be pseudo-Euclidean metric in the space $\mathbb{R}^n_{p,q}$ with $p+q=n$. The group $SO(p,q)$ is the group of real matrices which preserve the form g:

$$AgA^t = g\,, \qquad \det A = 1\,.$$

- *Unitary group* $U(n)$ – the group of unitary $n \times n$ matrices:

$$UU^\dagger = \mathbb{I}\,.$$

- *Special unitary group* $SU(n)$ – the group of unitary $n \times n$ matrices with the unit determinant

$$UU^\dagger = \mathbb{I}\,, \qquad \det U = 1\,.$$

- *Pseudo-unitary group* $U(p,q)$:

$$AgA^\dagger = g\,,$$

where g is the pseudo-Euclidean metric. Special pseudo-unitary group requires in addition the unit determinant $\det A = 1$.

- *Symplectic group* $Sp(2n,\mathbb{R})$ or $Sp(2n,\mathbb{C})$ is a group or real or complex matrices satisfying the condition

$$AJA^t = J$$

where J is $2n \times 2n$ matrix

$$J = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}$$

and $\mathbb{I}$ is $n \times n$ unit matrix.

Question to the class: What are the eigenvalues of J? Answer:

$$J = \mathrm{diag}(\mathrm{i}, \cdots \mathrm{i}; -\mathrm{i}, \cdots, -\mathrm{i}).$$

Thus, the group $Sp(2n)$ is really different from $SO(2n)$!

The powerful tool in the theory of Lie groups are the Lie algebras. Let us see how they arise by using as an example $SO(3)$. Let A be "close" to the identity matrix

$$A = \mathbb{I} + \epsilon a$$

is an orthogonal matrix $A^t = A^{-1}$. Therefore,

$$\mathbb{I} + \epsilon a^t = (\mathbb{I} + \epsilon a)^{-1} = \mathbb{I} - \epsilon a + \epsilon^2 a^2 + \cdots$$

From here $a^t = -a$. The space of matrices a such that $a^t = -a$ is denoted as $so(3)$ and called the Lie algebra of the Lie group $SO(3)$. The properties of this Lie algebra: $so(3)$ is a linear space, in $so(3)$ the commutator is defined: if $a, b \in so(3)$ then $[a,b]$ also belongs to $so(3)$. A linear space of matrices is called a Lie algebra if the commutator does not lead out of this space. Commutator of matrices naturally arises from the commutator in the group:

$$\begin{aligned} ABA^{-1}B^{-1} &= (\mathbb{I} + \epsilon a)(\mathbb{I} + \epsilon b)(\mathbb{I} + \epsilon a)^{-1}(\mathbb{I} + \epsilon b)^{-1} \\ &= (\mathbb{I} + \epsilon a)(\mathbb{I} + \epsilon b)(\mathbb{I} - \epsilon a + \epsilon^2 a^2 + \cdots)(\mathbb{I} - \epsilon b + \epsilon^2 b^2 + \cdots) = \\ &= \mathbb{I} + \epsilon(a + b - a - b) + \epsilon^2(ab - a^2 - ab - ba - b^2 + ab + a^2 + b^2) + \cdots = \\ &= \mathbb{I} + \epsilon^2 [a,b] + \cdots \end{aligned}$$

The algebra and the Lie group in our example are related as

$$\exp a = \sum_{n=0}^{\infty} \frac{a^n}{n!} = A \in SO(3)$$

Exponential of matrix. The exponent $\exp a$ of the matrix a is the sum of the following series

$$\exp a = \sum_{m=0}^{\infty} \frac{a^m}{m!}\,.$$

This series shares the properties of the usual exponential function, in particular it is convergent for any matrix A. The following obvious properties are

- If matrices X and Y commute then

$$\exp(X + Y) = \exp(X)\exp(Y)$$

- The matrix $A = \exp X$ is invertible and $A^{-1} = \exp(-X)$.
- $\exp(X^t) = (\exp X)^t$.

Definition of a Lie algebra: A linear vector space $\mathcal{J}$ (over a field $\mathbb{R}$ or $\mathbb{C}$) supplied with the multiplication operation (this operation is called *the commutator*) $[\xi, \eta]$ for $\xi, \eta \in \mathcal{J}$ is called a Lie algebra if the following properties are satisfied

1. The commutator $[\xi, \eta]$ is a bilinear operation, i.e.
$$[\alpha_1\xi_1 + \alpha_2\xi_2, \beta_1\eta_1 + \beta_2\eta_2] = \alpha_1\beta_1[\xi_1, \eta_1] + \alpha_2\beta_1[\xi_2, \eta_1] + \alpha_1\beta_2[\xi_1, \eta_2] + \alpha_2\beta_2[\xi_2, \eta_2]$$
2. The commutator is skew-symmetric: $[\xi, \eta] = -[\eta, \xi]$
3. The Jacobi identity
$$[[\xi, \eta], \zeta] + [[\eta, \zeta], \xi] + [[\zeta, \xi], \eta] = 0$$

Let $\mathcal{J}$ be a Lie algebra of dimension n. Choose a basis $e_1, \cdots, e_n \in \mathcal{J}$. We have

$$[e_i, e_j] = C_{ij}^k e_k$$

The numbers C_{ij}^k are called *structure constants* of the Lie algebra. Upon changing the basis these structure constants change as the tensor quantity. Let $e'_i = A_i^j e_i$ and $[e'_i, e'_j] = C'^k_{ij} e'_k$ then

$$C'^k_{ij} A_k^m e_m = A_i^r A_j^s [e_r, e_s] = A_i^r A_j^s C_{rs}^m e_m$$

Thus, the structure constants in the new basis are related to the constants in the original basis as

$$C'^k_{ij} = A_i^r A_j^s C_{rs}^m (A^{-1})_m^k \,. \tag{III.7.6}$$

Skew-symmetry and the Jacobi identity for the commutator imply that the tensor C_{ij}^k defines the Lie algebra if and only if

$$C_{ij}^k = -C_{ij}^k \,, \qquad C_{p[i}^m C_{jk]}^p = 0 \,.$$

Classify all Lie algebras means in fact to find all solutions of these equations modulo the equivalence relation (III.7.6).

Example. The Lie algebra $so(3, \mathbb{R})$ of the Lie group $SO(3, \mathbb{R})$. It consists of 3×3 skew-symmetric matrices. We can introduce a basis in the space of these matrices

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} , \qquad X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} , \qquad X_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} .$$

In this basis the Lie algebra relations take the form

$$[X_1, X_2] = X_3 \,, \qquad [X_2, X_3] = X_1 \,, \qquad [X_3, X_1] = X_2 \,.$$

These three relation can be encoded into one

$$[X_i, X_j] = \epsilon_{ijk} X_k \,.$$

Example. The Lie algebra $su(2)$ of the Lie group $SU(2)$. It consists of 2×2 skew-symmetric matrices. The basis can be constructed with the help of the so-called Pauli matrices σ_i

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

These matrices satisfy the relations

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k\,, \qquad \{\sigma_i, \sigma_j\} = 2\delta_{ij}\,.$$

If we introduce $X_i = -\frac{i}{2}\sigma_i$ which are three linearly independent anti-hermitian matrices then the $su(2)$ Lie algebra relations read

$$[X_i, X_j] = \epsilon_{ijk}X_k$$

Note that the structure constants are real! Comparing with the previous example we see that the Lie algebra $su(2)$ is isomorphic to that of $so(3,\mathbb{R})$:

$$su(2) \approx so(3,\mathbb{R})\,.$$

With every matrix group we considered above one can associate the corresponding matrix Lie algebra. The vector space of this Lie algebra is the tangent space at the identity element of the group. For this case the operation "commutator" is the usual matrix commutator. The tangent space to a Lie group at the identity element naturally appears in this discussion. To understand why let us return to the case of the Lie group $GL(n,\mathbb{R})$. Consider a one-parameter curve $A(t) \in GL(n,\mathbb{R})$, i.e, a family of matrices $A(t)$ from $GL(n,\mathbb{R})$ which depend on the parameter t. Let this curve to pass though the identity at $t = 0$, i.e., $A(0) = \mathbb{I}$. Then the tangent vector (the velocity vector!) at $t = 0$ is the matrix $\dot{A}(t)|_{t=0}$. Other way around, let X be an arbitrary matrix. Then the curve $A(t) = \mathbb{I} + tX$ for t sufficiently closed to zero lies in $GL(n,\mathbb{R})$. It is clear that

$$A(0) = \mathbb{I}\,, \qquad \dot{A}(0) = X\,.$$

In this way we demonstrated that the space of vectors which are tangent to the group $GL(n,\mathbb{R})$ at the identity coincide with the space of all $n \times n$ matrices. This example of $GL(n,\mathbb{R})$ demonstrates a universal connection between Lie group G and its Lie algebra: *The tangent space to G at the identity element is the Lie algebra w.r.t. to the commutator. This Lie algebra is called the Lie algebra of the group G.*

Exercise to do in the class: making infinitesimal expansion of a group element close to the identity compute the Lie algebras for the classical matrix groups discussed above. The answer is the following list:

The list of basic matrix Lie algebras

- The general Lie group $GL(n,\mathbb{R})$ or $GL(n,\mathbb{C})$ has the matrix Lie algebra which is $M(n,\mathbb{R})$ or $M(n,\mathbb{C})$, where $M(n)$ is the space of all real or complex matrices.

- Special linear group $SL(n,\mathbb{R})$ or $SL(n,\mathbb{C})$ has the Lie algebra $sl(n,\mathbb{R})$ or $sl(n,\mathbb{C})$ which coincides with the space of all real or complex matrices with zero trace.

- Special orthogonal group $SO(n,\mathbb{R})$ or $SO(n,\mathbb{C})$ has the Lie algebra $so(n,\mathbb{R})$ or $so(n,\mathbb{C})$ which are real or complex matrices satisfying the condition

$$X^t = -X\,.$$

- Pseudo-orthogonal group $SO(p,q)$ has the Lie algebra which is the algebra of matrices X satisfying the condition
$$Xg + gX^t = 0\,.$$
We see that if we introduce the matrix $u = Xg$ then the relation defining the Lie algebra reads
$$u + u^t = 0\,.$$
Thus, the matrix u is skew-symmetric $u^t + u = 0$. This map establishes the isomorphism between $so(p,q)$ and the space of all skew-symmetric matrices.

- Unitary group $U(n)$ has the Lie algebra which is the space of all anti-hermitian matrices
$$X^\dagger = -X\,.$$

- Special unitary group $SU(n)$ has the Lie algebra which is the space of all anti-hermitian matrices with zero trace
$$X^\dagger = -X\,, \qquad \mathrm{tr}X = 0\,.$$

- Pseudo-unitary group $U(p,q)$ has the Lie algebra which is the space of all matrices obeying the relation
$$Xg + gX^\dagger = 0\,.$$
The space $u(p,q)$ is isomorphic to the space of anti-hermitian matrices. The isomorphism is established by the formula $u = Xg$. Finally the Lie algebra of the special pseudo-unitary group is defined by further requirement of vanishing trace for X.

- The symplectic group $Sp(2n,\mathbb{R})$ or $Sp(2n,\mathbb{C})$ has the Lie algebra which comprises all is the is a group or real or complex matrices satisfying the condition
$$XJ + JX^t = 0$$
where J is $2n \times 2n$ matrix
$$J = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix}$$
and $\mathbb{I}$ is $n \times n$ unit matrix.

Linear representations of Lie groups Consider an action of a Lie group a n-dimensional vector space $\mathbb{R}^n$. This action is called *a linear representation* of Lie group G on $\mathbb{R}^n$ if for any $g \in G$ the map
$$\rho: \quad g \to \rho(g)$$
is a linear operator on $\mathbb{R}^n$. In other words, by a linear representation of G on $\mathbb{R}^n$ we call the homomorphism ρ which maps G into $GL(n,\mathbb{R})$, the group of linear transformations of $\mathbb{R}^n$. The homomorphism means that under this map the group structure is preserved, i.e.
$$\rho(g_1 g_2) = \rho(g_1)\rho(g_2)\,.$$

Any Lie group G has a distinguished element – $g_0 = \mathbb{I}$ and the tangent space T at this point. Transformation
$$G \to G: \qquad g \to hgh^{-1}$$

is called *internal automorphism* corresponding to an element $h \in G$. This transformation leaves unity invariant: $h\mathbb{I}h^{-1} = \mathbb{I}$ and it transforms the tangent space T into itself:

$$\mathrm{Ad}(h): \quad T \to T\,.$$

This map has the following properties:

$$\mathrm{Ad}(h^{-1}) = (\mathrm{Ad}h)^{-1}\,, \qquad \mathrm{Ad}(h_1h_2) = \mathrm{Ad}h_1\mathrm{Ad}h_2\,.$$

In other words, the map $h \to \mathrm{Ad}h$ is a *linear representation* of G:

$$\mathrm{Ad}: \quad G \to GL(n,\mathbb{R})\,,$$

where n is the dimension of the group.

Generally, one-parameter subgroups of a Lie group G are defined as parameterized curves $F(t) \subset G$ such that $F(0) = \mathbb{I}$ and $F(t_1 + t_2) = F(t_1)F(t_2)$ and $F(-t) = F(t)^{-1}$. As we have already discussed for matrix groups they have the form

$$F(t) = \exp(At)$$

where A is an element of the corresponding Lie algebra. In an abstract Lie group G for a curve $F(t)$ one defines the t-dependent vector

$$F^{-1}\dot{F} \in T\,.$$

If this curve $F(t)$ is one-parameter subgroup then this vector does not depend on t! Indeed,

$$\dot{F} = \frac{dF(t+\epsilon)}{d\epsilon}|_{\epsilon=0} = F(t)\Big(\frac{dF(\epsilon)}{d\epsilon}\Big)_{\epsilon=0}\,,$$

i.e. $\dot{F} = F(t)\dot{F}(0)$ and $F^{-1}(t)\dot{F}(t) = \dot{F}(0) = \mathrm{const}$. Oppositely, for any non-zero $a \in T$ there exists a unique one-parameter subgroup with

$$F^{-1}\dot{F} = a\,.$$

This follows from the theorem about the existence and uniqueness of solutions of usual differential equations.

It is important to realize that even for the case of matrix Lie groups there are matrices which are not images of any one-parameter subgroup. The exercise to do in the class: Consider the following matrix:

$$g = \begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix} \in GL^+(2,\mathbb{R})\,,$$

where $GL^+(2,\mathbb{R})$ is a subgroup of $GL(2,\mathbb{R})$ with positive determinant. Show that there does not exist any real matrix ξ such that

$$e^{\xi} = g\,.$$

The answer: it is impossible because since the matrix ξ is real the eigenvalues $\lambda_{1,2}$ of ξ must be either real of complex conjugate. The eigenvalues of e^{ξ} are e^{λ_1} and e^{λ_2}. If λ_i are real then $e^{\lambda_i} > 0$. If λ_i are complex conjugate then e^{λ_i} are also complex conjugate.

It is also important to realize that different vectors ξ under the exponential map can be mapped on the one and the same group element. As an example, consider the matrices of the form

$$\xi = \alpha\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

where $\alpha, \beta \in \mathbb{R}$. Exponent e^{ξ} can be computed by noting that

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then we have

$$e^{\xi} = e^{\alpha}\Big[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\cos\beta + \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\sin\beta\Big].$$

It is clear that

$$\alpha\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \alpha\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (\beta + 2\pi k)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

has the the same image under the exponential map. In the sufficiently small neighbourhood of 0 in $M(n, \mathbb{R})$ the map $\exp A$ is a diffeomorphism. The inverse map is constructed by means of series

$$\ln x = (x - \mathbb{I}) - \frac{1}{2}(x - \mathbb{I})^2 + \frac{1}{3}(x - \mathbb{I})^3 - \cdots$$

for x sufficiently close to the identity.

Linear representation of a Lie algebra. Adjoint representation. Let $\mathcal{J}$ be a Lie algebra. We say that a map

$$\rho: \quad \mathcal{J} \to M(n, \mathbb{R})$$

defines a representation of the Lie algebra $\mathcal{J}$ is the following equality is satisfied

$$\rho[\zeta, \eta] = [\rho(\eta), \rho(\zeta)]$$

for any two vectors $\zeta, \eta \in \mathcal{J}$.

Let $F(t)$ be a one-parameter subgroup in G. Then $g \to FgF^{-1}$ generates a one-parameter group of transformations in the Lie algebra

$$\mathrm{Ad}F(t): \quad T \to T\,.$$

The vector $\frac{d}{dt}\mathrm{Ad}F(t)|_{t=0}$ lies in the Lie algebra. Let $a \in T$ and let $F(t) = \exp(bt)$ then

$$\frac{d}{dt}\mathrm{Ad}F(t)|_{t=0}\ a = \frac{d}{dt}\Big(\exp(bt)a\exp(-bt)\Big)|_{t=0} = [b, a]$$

Thus to any element $b \in \mathcal{J}$ we associate an operator ad_b which acts on the Lie algebra:

$$\mathrm{ad}_b: \quad \mathcal{J} \to \mathcal{J}, \qquad \mathrm{ad}_b a = [b, a]\,.$$

This action defines a representation of the Lie algebra on itself. This representation is called *adjoint*. To see that this is indeed representation we have to show that it preserves the commutation relations, i.e. that from $[x, y] = z$ it follows that

$$[\mathrm{ad}x, \mathrm{ad}y] = \mathrm{ad}z\,.$$

We compute

$$\begin{aligned} [\mathrm{ad}x, \mathrm{ad}y]w &= \mathrm{ad}x\ \mathrm{ad}yw - \mathrm{ad}y\ \mathrm{ad}xw = [x,[y,w]] - [y,[x,w]] = [x,[y,w]] + [y,[w,x]] = \\ &- [w,[x,y]] = [[x,y],w] = [z,w] = \mathrm{ad}zw\,. \end{aligned}$$

Here the Jacobi identity has been used.

Semi-simple and simple Lie algebras. General classification of Lie algebras is a very complicated problem. To make a progress simplifying assumptions about the structure of the algebra are needed. The class of the so-called simple and semi-simple Lie algebras admits a complete classification.

A Lie subalgebra $\mathcal{H}$ of a Lie algebra $\mathcal{J}$ is a linear subspace $\mathcal{H} \subset \mathcal{J}$ which is closed w.r.t. to the commutation operation. An *ideal* $\mathcal{H} \subset \mathcal{J}$ is a subspace in $\mathcal{J}$ such that for any $x \in \mathcal{J}$ the following relation holds

$$[x, \mathcal{H}] \subset \mathcal{H}\,.$$

A Lie algebra $\mathcal{J}$ which does not have any ideals except the trivial one and the one coincident with $\mathcal{J}$ is called *simple*. A Lie algebra which have no commutative (i.e. abelian) ideals is called semi-simple. One can show that any semi-simple Lie algebra is a sum of simple Lie algebras. Consider for instance the Lie algebra $u(n)$ which is the algebra of anti-hermitian matrices

$$u + u^\dagger = 0\,.$$

The Lie algebra $su(n)$ is further distinguished by imposing the condition of vanishing trace: $\mathrm{tr}u = 0$. The difference between $u(n)$ and $su(n)$ constitute all the matrices which are proportional to the identity matrix $i\mathbb{I}$. Since

$$[\lambda i\mathbb{I}, u] = 0$$

the matrices proportional to $i\mathbb{I}$ form an ideal in $u(n)$ which is abelian. Thus, $u(n)$ has the abelian ideal and, therefore, $u(n)$ is not semi-simple. In opposite, $su(n)$ has no non-trivial ideals and therefore it is the simple Lie algebra.

A powerful tool in the Lie theory is the so-called Cartan-Killing from on a Lie algebra. Consider the adjoint representation of $\mathcal{J}$. The Cartan-Killing form on $\mathcal{J}$ is defined as

$$(a, b) = -\mathrm{tr}(\mathrm{ad}_a \mathrm{ad}_b)$$

for any two $a, b \in \mathcal{J}$. The following central theorem in the Lie algebra theory can be proven: *A Lie algebra is semi-simple if and only if its Cartan-Killing form is non-degenerate.*

For a simple Lie algebra $\mathcal{J}$ of a group G the internal automorphisms $\mathrm{Ad}g$ constitute the linear irreducible representation (i.e. a representation which does not have invariant subspaces) of G in $\mathcal{J}$. Indeed, if $\mathrm{Ad}(g)$ has an invariant subspace $\mathcal{H} \subset \mathcal{J}$, i.e. $g\mathcal{H}g^{-1} \subset \mathcal{H}$ for any g then sending g to the identity we will get

$$[\mathcal{J}, \mathcal{H}] \subset \mathcal{H}$$

i.e. $\mathcal{H}$ is an ideal which contradicts to the assumption that $\mathcal{J}$ is the semi-simple Lie algebra.

Cartan subalgebra. To demonstrate the construction of the adjoint representation and introduce the notion of the Cartan subalgebra of the Lie algebra we use the concrete example of $su(3)$. The Lie algebra $su(3)$ comprises the matrices of the form iM, where M is traceless 3×3 hermitian matrix. The basis consists of eight matrices which we chose to be the Gell-Mann matrices:

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
\lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned}$$

There are two diagonal matrices among these: λ_3 and λ_8 which we replace by $T_z = \frac{1}{2}\lambda_3$ and $Y = \frac{1}{\sqrt{3}}\lambda_8$. We introduce the following linear combinations of the generators

$$t_\pm = \frac{1}{2}(\lambda_1 \pm i\lambda_2)\,, \qquad v_\pm = \frac{1}{2}(\lambda_4 \pm i\lambda_5)\,, \qquad u_\pm = \frac{1}{2}(\lambda_6 \pm i\lambda_y)\,.$$

One can easily compute, e.g.,

$$\begin{aligned}&[t_+, t_+] = 0\,, \quad [t_+, t_-] = 2t_z, \quad [t_+, t_z] = -t_+\,, \quad [t_+, u_+] = v_+\,, \quad [t_+, u_-] = 0\,,\\ &[t_+, v_+] = 0\,, \quad [t_+, v_-] = -u_-\,, \quad [t_+, y] = 0\,.\end{aligned}$$

Since the Lie algebra of $su(3)$ is eight-dimensional the adjoint representation is eight-dimensional too. Picking up $(t_+, t_-, t_z, u_+, u_-, v_+, v_-, y)$ as the basis we can realize the adjoint action by 8×8 matrices. For instance,

$$\mathrm{ad}_{t_+}\begin{pmatrix} t_+\\ t_-\\ t_z\\ u_+\\ u_-\\ v_+\\ v_-\\ y\end{pmatrix} = \underbrace{\begin{pmatrix} 0&0&0&0&0&0&0&0\\ 0&0&2&0&0&0&0&0\\ -1&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&-1&0&0&0\\ 0&0&0&0&0&0&0&0\end{pmatrix}}_{\text{matrix realization of } t_+}\begin{pmatrix} t_+\\ t_-\\ t_z\\ u_+\\ u_-\\ v_+\\ v_-\\ y\end{pmatrix}$$

Note that both ad_{t_z} and ad_y are diagonal. Thus, if $x = at_z + by$ then ad_x is also diagonal. Explicitly we find

$$\mathrm{ad}_x = \begin{pmatrix} a&0&0&0&0&0&0&0\\ 0&-a&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&-\frac{1}{2}a+b&0&0&0&0\\ 0&0&0&0&\frac{1}{2}a-b&0&0&0\\ 0&0&0&0&0&\frac{1}{2}a+b&0&0\\ 0&0&0&0&0&0&-\frac{1}{2}a-b&0\\ 0&0&0&0&0&0&0&0\end{pmatrix}.$$

In other words, the basis elements $(t_+, t_-, t_z, u_+, u_-, v_+, v_-, y)$ are all eigenvectors of ad_x with eigenvalues $a, -a, 0, -\frac{1}{2}a + b, \frac{1}{2}a - b, -\frac{1}{2}a - b$ and 0 respectively. The procedure we followed in crucial for analysis of other (larger) Lie algebras. We found a two-dimensional subalgebra generated by t_z and y which is abelian. Further, we have chosen a basis for the rest of the Lie algebra such that each element of the basis is an eigenvector of ad_x if x is from this abelian subalgebra. This abelian subalgebra is called *the Cartan subalgebra*.

In general the Cartan subalgebra H is determined in the following way. An element $h \in H$ is called *regular* if ad_h has as simple as possible number of zero eigenvalues (i.e. multiplicity of zero eigenvalue is minimal). For instance, for $su(3)$ the element ad_{t_z} has two zero eigenvalues, while ad_y has for zero eigenvalues. Thus, the element ad_{t_z} is regular, while ad_y is not. *A Cartan subalgebra is a maximal commutative subalgebra which contains a regular element.* In our example the subalgebra generated by t_z and y is commutative and its maximal since there is no other element we can add to it which would not destroy the commutativity.

Roots. It is very important fact proved in the theory of Lie algebras that any simple Lie algebra has a Cartan subalgebra and it admits a basis where each basis vector is an eigenstate of all Cartan

generators; the corresponding eigenvalues depend of course on a Cartan generator. In our example of $su(3)$ for an element $x = at_z + by$ we have

$$\begin{aligned}
\mathrm{ad}_x t_+ &= at_+ \\
\mathrm{ad}_x t_- &= at_- \\
\mathrm{ad}_x t_z &= 0t_z \\
\mathrm{ad}_x u_+ &= (-\frac{1}{2}a + b)u_+ \\
\mathrm{ad}_x u_- &= (\frac{1}{2}a - b)u_- \\
\mathrm{ad}_x v_+ &= (\frac{1}{2}a + b)v_+ \\
\mathrm{ad}_x v_- &= (-\frac{1}{2}a - b)v_- \\
\mathrm{ad}_x y &= 0y\,.
\end{aligned}$$

We see that all eigenvalues are *linear* functions of the Cartan element x, in other words, if we denote by e_α the six elements $t_\pm, v_\pm, u_\pm$ and by h_i the two Cartan elements t_z, y we can write all the relations above as

$$[h_i, h_j] = 0$$

$$[h_i, e_\alpha] = \alpha(h_i)e_\alpha\,,$$

where $\alpha(h_i)$ is a linear function of h_i. The generators e_α, which are eigenstates of the Cartan subalgebra, are called *root vectors*, while the corresponding linear functions $\alpha(h)$ are called *roots*. To every root vector e_α we associate the root α which is a linear function on the Cartan sualgebra H. Linear functions on H, by definition, form the dual space H^* to the Cartan subalgebra H.

The Cartan-Weyl basis. Now we can also investigate what is the commutator of the root vectors. By using the Jacobi identity we find

$$[h, [e_\alpha, e_\beta]] = -[e_\alpha, [e_\beta, h]] - [e_\beta, [h, e_\alpha]] = (\alpha(h) + \beta(h))[e_\alpha, e_\beta]\,.$$

This clearly means that there are three distinct possibilities

- $[e_\alpha, e_\beta]$ is zero
- $[e_\alpha, e_\beta]$ is a root vector with the root $\alpha + \beta$
- $\alpha + \beta = 0$ in which case $[e_\alpha, e_\beta]$ commutes with every $h \in H$ and, therefore, is an element of the Cartan subalgebra.

  Thus,
  $$[e_\alpha, e_\beta] = N_{\alpha\beta} e_{\alpha+\beta}$$
  if $\alpha + \beta$ is a root,
  $$[e_\alpha, e_{-\alpha}] \sim h_\alpha$$
  and $[e_\alpha, e_\beta] = 0$ if $\alpha + \beta$ is not a root. The numbers $N_{\alpha\beta}$ depend on the normalization of the root vectors. The basis (h_i, e_α) of a Lie algebra with the properties described above is called *the Cartan-Weyl basis.*