Preprint typeset in JHEP style - HYPER VERSION

Classical Field Theory

Gleb Arutyunov[a*†]

[a] *Institute for Theoretical Physics and Spinoza Institute, Utrecht University, 3508 TD Utrecht, The Netherlands*

Abstract: The aim of the course is to introduce the basic methods of classical field theory and to apply them in a variety of physical models ranging from classical electrodynamics to macroscopic theory of ferromagnetism. In particular, the course will cover the Lorentz-covariant formulation of Maxwell's electromagnetic theory, advanced radiation problems, elements of soliton theory. The students will get acquainted with the Lagrangian and Hamiltonian description of infinite-dimensional dynamical systems, the concept of global and local symmetries, conservation laws. A special attention will be paid to mastering the basic computation tools which include the Green function method, residue theory, Laplace transform, elements of group theory, orthogonal polynomials and special functions.

Last Update 8.05.2011

*Email: G.Arutyunov@phys.uu.nl

†Correspondent fellow at Steklov Mathematical Institute, Moscow.

Contents

---

1. Classical Fields: General Principles

Classical field theory is a very vast subject which traditionally includes the Maxwell theory of electromagnetism describing electromagnetic properties of matter and the Einstein theory of General Relativity. The main scope of classical field theory is

to construct the mathematical description of dynamical systems with an infinite number of degrees of freedom. As such, this discipline also naturally incorporates the classics aspects of fluid dynamics. The basic mathematical tools involved are partial differential equations with given initial and boundary conditions, theory of special functions, elements of group and representation theory.

1.1 Lagrangian and Hamiltonian formalisms

We start with recalling the two ways the physical systems are described in classical mechanics. The first description is known as the Lagrangian formalism which is equivalent to the "principle of least action[1]" (Maupertuis's principle). Consider a point particle which moves in a n-dimensional space with coordinates $(q^1, \dots, q^n)$ and in the potential $U(q)$. The Newtons equations describing the corresponding motion (trajectory) are

$$m\ddot{q}^i = -\frac{\partial U}{\partial q^i}\,. \tag{1.1}$$

These equations can be obtained by extremizing the following functional

$$S = \int_{t_1}^{t_2} \mathrm{d}t\, L(q, \dot{q}, t) = \int_{t_1}^{t_2} \mathrm{d}t \left(\frac{m\dot{q}^2}{2} - U(q)\right). \tag{1.2}$$

Here S is the functional on the space of particle trajectories: to any trajectory which satisfies given initial $q^i(t_1) = q^i_{\text{in}}$ and final $q^i(t_2) = q^i_{\text{f}}$ conditions it puts in correspondence a number. This functional is called the *action*. The specific function L depending on particle coordinates and momenta is called *Lagrangian*. According to the principle of stationary action, the actual trajectories of a dynamical system (particle) are the ones which deliver the extremum of S.

Compute the variation of the action

$$\delta S = -\int_{t_1}^{t_2} \mathrm{d}t \left[\frac{d}{dt}(m\dot{q}^i) + \frac{\partial U}{\partial q^i}\right]\delta q^i + \text{total derivative}\,,$$

where we have integrated by parts. The total derivative term vanishes provided the end points of a trajectory are kept fixed under the variation. The quantity δS vanishes for any δq^i provided eq.(1.1) is satisfied. Note that in our particular example, the Lagrangian coincides with the difference of the kinetic and the potential energy $L = T - U$ and it does not explicitly depend on time.

In general, we simply regard L as an arbitrary function of q, $\dot{q}$ and time. The equations of motion are obtained by extremizing the corresponding action

$$\frac{\delta S}{\delta q^i} = \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big) - \frac{\partial L}{\partial q^i} = 0$$

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

and they are called the *Euler-Lagrange equations*. We assume that L does not depend on higher derivatives $\dddot{q}$, $\ddddot{q}$ and so on, which reflects the fact that the corresponding dynamical system is fully determined by specifying coordinates and velocities. Indeed, for a system with n degrees of freedom there are n Euler-Lagrange equations of the second order. Thus, an arbitrary solution will depend on $2n$ integration constants, which are determined by specifying, e.g. the initial coordinates and velocities.

Suppose L does not explicitly depend[2] on t, then

$$\frac{dL}{dt} = \frac{\partial L}{\partial \dot{q}^i}\ddot{q}^i + \frac{\partial L}{\partial q^i}\dot{q}^i \,.$$

Substituting here $\frac{\partial L}{\partial q^i}$ from the Euler-Lagrange equations, we get

$$\frac{dL}{dt} = \frac{\partial L}{\partial \dot{q}^i}\ddot{q}^i + \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big)\dot{q}^i = \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i\Big) \,.$$

Therefore, we find that

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L\Big) = 0 \tag{1.3}$$

as the consequence of the equations of motion. Thus, the quantity

$$H = \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L \,, \tag{1.4}$$

is conserved under the time evolution of our dynamical system. For our particular example,

$$H = m\dot{q}^2 - L = \frac{m\dot{q}^2}{2} + U(q) = T + U \equiv E \,.$$

Thus, H is nothing else but the energy of our system; energy is conserved due to equations of motion. *Dynamical quantities which are conserved during the time evolution of a dynamical system are called conservation laws or integrals of motion.* Energy is our first non-trivial example of a conservation law.

Introduce a quantity called the (canonical) momentum

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

For a point particle $p_i = m\dot{q}^i$. Suppose that $U = 0$. Then

$$\dot{p}_i = \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big) = 0$$

by the Euler-Lagrange equations. Thus, in the absence of the external potential, the momentum p is an integral of motion. This is our second example of a conservation law.

[2]This is homogenuity of time.

Now we remind the second description of dynamical systems which exploits the notion of the Hamiltonian. The conserved energy of a system expressed via canonical coordinates and momenta is called the *Hamiltonian*

$$H \equiv H(p,q) = \frac{1}{2m}p^2 + U(q)\,.$$

Let us again verify by direct calculation that it does not depend on time,

$$\frac{dH}{dt} = \frac{1}{m}p_i\dot{p}_i + \dot{q}^i\frac{\partial U}{\partial q^i} = \frac{1}{m}m^2\dot{q}_i\ddot{q}_i + \dot{q}^i\frac{\partial U}{\partial q^i} = 0$$

due to the Newton equations of motion.

Having the Hamiltonian, the Newton equations can be rewritten in the form

$$\dot{q}^j = \frac{\partial H}{\partial p_j}\,, \qquad \dot{p}_j = -\frac{\partial H}{\partial q^j}\,.$$

These are the fundamental Hamiltonian equations of motion. Their importance lies in the fact that they are valid for arbitrary dependence of $H \equiv H(p,q)$ on the dynamical variables p and q.

In the general setting the Hamiltonian equations are obtained as follows. We take the full differential of the Lagrangian

$$dL = \frac{\partial L}{\partial q^i}dq^i + \frac{\partial L}{\partial \dot{q}^i}d\dot{q}^i = \dot{p}_i dq^i + p_i d\dot{q}^i = \dot{p}_i dq^i + d(p_i\dot{q}^i) - \dot{q}^i dp_i\,,$$

where we have used the definition of the canonical momentum and the Euler-Lagrange equations. From here we find

$$d(\underbrace{p_i\dot{q}^i - L}_{H}) = \dot{q}^i dp_i - \dot{p}_i dq^i\,.$$

From the differential equality the Hamiltonian equations follow. Transformation

$$H(p,q) = p_i\dot{q}^i - L(q,\dot{q})|_{\dot{q}^i\to p_i}$$

is the Legendre transform.

The last two equations can be rewritten in terms of the single equation. Introduce two $2n$-dimensional vectors

$$x = \begin{pmatrix} p \\ q \end{pmatrix}\,, \qquad \nabla H = \begin{pmatrix} \frac{\partial H}{\partial p_j} \\ \frac{\partial H}{\partial q^j} \end{pmatrix}$$

and $2n \times 2n$ matrix J:

$$J = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\,.$$

Then the Hamiltonian equations can be written in the form

$$\dot{x} = J\cdot\nabla H\,, \qquad \text{or} \qquad J\cdot\dot{x} = -\nabla H\,.$$

In this form the Hamiltonian equations were written for the first time by Lagrange in 1808.

A point $x = (x^1, \ldots, x^{2n})$ defines a state of a system in classical mechanics. The set of all these points form a *phase space* $\mathscr{P} = \{x\}$ of the system which in the present case is just the $2n$-dimensional Euclidean space with the metric $(x, y) = \sum_{i=1}^{2n} x^i y^i$.

To get more familiar with the concept of a phase space, consider a one-dimensional example: the harmonic oscillator. The potential is $U(q) = \frac{q^2}{2}$. The Hamiltonian $H = \frac{p^2}{2} + \frac{q^2}{2}$, where we choose $m = 1$. The Hamiltonian equations of motion are given by *ordinary* differential equations:

$$\dot{q} = p\,, \quad \dot{p} = -q \qquad \Longrightarrow \qquad \ddot{q} = -q\,.$$

Solving these equations with given initial conditions (p_0, q_0) representing a point in the phase space[3], we obtain a phase space curve

$$p \equiv p(t; p_0, q_0)\,, \qquad q \equiv q(t; p_0, q_0)\,.$$

Through every phase space point there is one and only one phase space curve (uniqueness theorem for ordinary differential equations). 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 only one point. Such a point is called an *equilibrium position*. The Hamiltonian vector field at an equilibrium position vanishes.

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 harmonic oscillator

$$p^2 + q^2 = 2h$$

and the phase space curves are concentric circles and the origin.

The matrix J serves to define the so-called *Poisson brackets* on the space $\mathcal{F}(\mathscr{P})$ of differentiable functions on $\mathscr{P}$:

$$\{F, G\}(x) = (J\nabla F, \nabla G) = -J^{ij}\partial_i F \partial_j G = \sum_{j=1}^{n} \Big(\frac{\partial F}{\partial p_j}\frac{\partial G}{\partial q^j} - \frac{\partial F}{\partial q^j}\frac{\partial G}{\partial p_j} \Big)\,.$$

The Poisson bracket satisfies the following conditions

$$\begin{aligned} &\{F, G\} = -\{G, F\}\,, \\ &\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0 \end{aligned}$$

[3]The two-dimensional plane in the present case.

for arbitrary functions F, G, H.

Thus, the Poisson bracket introduces on $\mathcal{F}(\mathscr{P})$ the structure of an infinite-dimensional Lie algebra. The bracket also satisfies the Leibnitz rule

$$\{F, GH\} = \{F, G\}H + G\{F, H\}$$

and, therefore, it is completely determined by its values on the basis elements x^i:

$$\{x^j, x^k\} = -J^{jk}$$

which can be written as follows

$$\{q^i, q^j\} = 0\,, \quad \{p_i, p_j\} = 0\,, \quad \{p_i, q^j\} = \delta_i^j\,.$$

The Hamiltonian equations can be now rephrased in the form

$$\dot{x}^j = \{H, x^j\} \qquad \Leftrightarrow \qquad \dot{x} = \{H, x\} = X_H\,.$$

It follows from Jacobi identity that the Poisson bracket of two integrals of motion is again an integral of motion. The Leibnitz 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.

In the case under consideration the matrix J is non-degenerate so that there exists the inverse

$$J^{-1} = -J$$

which defines a skew-symmetric bilinear form ω on phase space

$$\omega(x, y) = (x, J^{-1}y)\,.$$

In the coordinates we consider it can be written in the form

$$\omega = \sum_j dp_j \wedge dq^j\,.$$

This form is closed, i.e. $d\omega = 0$.

A non-degenerate closed two-form is called symplectic and a manifold endowed with such a form is called a symplectic manifold. Thus, the phase space we consider is the symplectic manifold.

Imagine we make a change of variables $y^j = f^j(x^k)$. Then

$$\dot{y}^j = \underbrace{\frac{\partial y^j}{\partial x^k}}_{A^j_k}\dot{x}^k = A^j_k J^{km}\nabla^x_m H = A^j_k J^{km}\frac{\partial y^p}{\partial x^m}\nabla^y_p \widetilde{H}$$

or in the matrix form

$$\dot{y} = AJA^t \cdot \nabla_y \widetilde{H} \,.$$

The new equations for y are Hamiltonian with the new Hamiltonian is $\widetilde{H}(y) = H(f^{-1}(y)) = H(x)$ if and only if

$$AJA^t = J \,.$$

Hence, this construction motivates the following definition.

Transformations of the phase space which satisfy the condition

$$AJA^t = J$$

are called canonical[4].

Canonical transformations[5] do not change the symplectic form ω:

$$\omega(Ax, Ay) = -(Ax, JAy) = -(x, A^t JAy) = -(x, Jy) = \omega(x, y) \,.$$

In the case we considered the phase space was Euclidean: $\mathscr{P} = \mathbb{R}^{2n}$. This is not always so. The generic situation is that the phase space is a manifold. Consideration of systems with general phase spaces is very important for understanding the structure of the Hamiltonian dynamics.

Short summary

A Hamiltonian system is characterized by a triple $(\mathscr{P}, \{,\}, H)$: a phase space $\mathscr{P}$, a Poisson structure $\{,\}$ and by a Hamiltonian function H. The vector field X_H is called the *Hamiltonian vector field* corresponding to the Hamiltonian H. For any function $F = F(p, q)$ on phase space, the evolution equations take the form

$$\frac{dF}{dt} = \{H, F\} = X_H \cdot F \,.$$

Again we conclude from here that the Hamiltonian H is a time-conserved quantity

$$\frac{dH}{dt} = \{H, H\} = 0 \,.$$

Thus, the motion of the system takes place on the subvariety of phase space defined by $H = E$ constant.

[4]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*.

[5]Notice that $AJA^t = J$ implies that $A^t JA = J$. Indeed, multiplying by J both sides of the first equality from the right, we get $AJA^t J = J^2 = -\mathbb{1}$, which further implies $A^t J = -J^{-1}A^{-1} = JA^{-1}$. Finally, multiplying both sides of the last expression from the right by A, we obtain the desired formula.

1.2 Noether's theorem in classical mechanics

Noether's theorem is one of the most fundamental and general statements concerning the behavior of dynamical systems. It relates symmetries of a theory with its conservation laws.

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: $L \to L + \frac{d}{dt}G(q,t)$. Indeed, the change of the action under the variation will be

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

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.

Let now an infinitezimal transformation $q \to q + \delta q$ be such that the variation of the Lagrangian takes the form (*without usage of equations of motion*!)[6] of a total time derivative of some function F:

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

Transformation δq is called a symmetry of the action.

Now we are ready to discuss Noether's theorem. Suppose that $q' = q + \delta q$ is a symmetry of the action. Then

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

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

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

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

$$J = \frac{\partial L}{\partial \dot{q}^i}\delta q^i - F = p_i \delta q^i - F$$

is called Noether's current. Now we consider some important applications.

[6] As we have already seen, a variation of the Lagrangian *computed on the equations of motion* is always a total derivative!

- *Momentum conservation.* Momentum conservation is related to the freedom of arbitrary choosing the origin of the coordinate system. Consider a Lagrangian

$$L = \frac{m}{2}\dot{q}_i^2 \,.$$

Consider a displacement

$$\begin{aligned} q'^i = q^i + a^i \quad &\Rightarrow \quad \delta q^i = a^i, \\ \dot{q}'^i = \dot{q}^i \quad &\Rightarrow \quad \delta \dot{q}^i = 0 \,. \end{aligned}$$

Obviously, under this transformation the Lagrangian remains invariant and we can take $F = 0$ or $F =$ any constant. Thus,

$$J = p_i \delta q^i = p_i a^i \,,$$

Since a_i arbitrary, all the components p_i are conserved.

- *Angular momentum conservation.* Consider again

$$L = \frac{m}{2}\dot{q}_i^2$$

and make a transformation

$$q'^i = q^i + \epsilon^{ij} q^j \quad \Rightarrow \quad \delta q^i = \epsilon^{ij} q^j \,.$$

Then,

$$\delta L = m \dot{q}^i \epsilon^{ij} \dot{q}^j \,.$$

Thus, if ϵ^{ij} is anti-symmetric, the variation of the Lagrangian vanishes. Again, we can take $F = 0$ or $F =$ any constant and obtain

$$J = p_i \delta q^i = p_i \epsilon^{ij} q^j \,,$$

Since ϵ^{ij} is arbitrary, we find the conservation of angular momentum components

$$J_i^j = p_i q^j - p_j q^i \,.$$

- *Particle in a constant gravitational field .* The Lagrangian

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

Shift $z \to z + a$, *i.e.* $\delta z = a$. We get $\delta L = -mga = \frac{d}{dt}(-mgat)$. Thus, the quantity

$$J = m\dot{z}\delta z - F = m\dot{z}a + mgat$$

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

- *Conservation of energy.* Energy conservation is related to the freedom of arbitrary choosing the origin of time (you can perform you experiment today or after a several years but the result will be the same provided you 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 $\delta t = \epsilon$. The Lagrangian response on it is

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

  On the other hand,

  $$\delta L = \frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i + \frac{\partial L}{\partial t}\delta t = \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big)\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i\,,$$

  where we have used the Euler-Lagrange equations and assumed that L does not explicitly depends on time. Obviously, $\delta q^i = \dot{q}^i\epsilon$ and $\delta \dot{q}^i = \ddot{q}^i\epsilon$, so that

  $$\delta L = \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big)\dot{q}^i\epsilon + \frac{\partial L}{\partial \dot{q}^i}\ddot{q}^i\epsilon = \frac{dL}{dt}\epsilon\,.$$

  Cancelling ϵ, we recover the conservation law for the energy

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

Finally, it remains to note that in 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 last example $\mathrm{d}t \to \mathrm{d}(t+\epsilon) = \mathrm{d}t$).

1.3 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 constant k. 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

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

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 $F_i = -\frac{\partial U}{\partial \phi_i}$:

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

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

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

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 $k\Delta a \to Y$, where Y is a finite quantity. Thus, we have

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

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

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

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

$$\phi(x,t) \to \phi(x,t) + \delta\phi(x,t)\,. \tag{1.5}$$

The derivatives change accordingly,

$$\frac{\partial}{\partial t}\phi(x,t) \to \frac{\partial}{\partial t}\phi(x,t) + \frac{\partial}{\partial t}\delta\phi(x,t)\,, \tag{1.6}$$

$$\frac{\partial}{\partial x}\phi(x,t) \to \frac{\partial}{\partial x}\phi(x,t) + \frac{\partial}{\partial x}\delta\phi(x,t)\,. \tag{1.7}$$

This gives

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

Integrating by parts, we find

$$\begin{aligned}\delta S[\phi] = &\int_{t_1}^{t_2} \mathrm{d}t \int_0^\ell \mathrm{d}x \left[\frac{\partial\mathscr{L}}{\partial\phi} - \partial_t\frac{\partial\mathscr{L}}{\partial\dot{\phi}} - \partial_x\frac{\partial\mathscr{L}}{\partial(\partial_x\phi)}\right]\delta\phi \\ &+ \int_0^\ell \mathrm{d}x \frac{\partial\mathscr{L}}{\partial(\partial_t\phi)}\delta\phi|_{t=t_1}^{t=t_2} + \int_{t_1}^{t_2} \mathrm{d}t \frac{\partial\mathscr{L}}{\partial(\partial_x\phi)}\delta\phi|_{x=0}^{x=\ell}\,.\end{aligned} \tag{1.8}$$

The action principle requires that the action principle be stationary w.r.t. infinitezimal 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 $\dot{\phi}(x,0)$.

Scalar and Vector Fields

The generalization 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.

In general, fields depending on the space-time variables are tensors, *i.e.* they transforms under general coordinate transformations in a definite way. Namely, a tensor field $\phi^{i_1\ldots i_p}_{j_1\ldots j_q}$ of rank (p,q) under general coordinate transformations of the coordinates x^i: $x^i \to x'^i(x^j)$ transforms as follows[7]

$$\phi'^{k_1\ldots k_p}_{l_1\ldots l_q}(x') = \frac{\partial x'^{k_1}}{\partial x^{i_1}}\cdots\frac{\partial x'^{k_p}}{\partial x^{i_p}}\,\frac{\partial x^{j_1}}{\partial x'^{l_1}}\cdots\frac{\partial x^{j_q}}{\partial x'^{l_q}}\;\phi^{i_1\ldots i_p}_{j_1\ldots j_q}(x)\,.$$

Here tensor indices are acted with the matrices $\frac{\partial x'^i}{\partial x^j}$ which form a group $\mathrm{GL}(d,\mathbb{R})$. This is a group of all invertible real $d\times d$ matrices. A 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.

1.4 Noether's theorem in field theory

In order to fully describe a dynamical system, it is not enough to only know the equations of motion. It is also important to be able to express the basic physical characteristics, in particular, the dynamical invariants, of the systems via solutions of these equations.

Noether's theorem: *To any finite-parametric, i.e. dependent on s constant parameters, continuous transformation of the fields and the space-time coordinates which leaves the action invariant corresponds s dynamical invariants, i.e. the conserved functions of the fields and their derivatives.*

To prove the theorem, consider an infinitezimal transformation

$$\begin{aligned} &x^i \to x'^i = x^i + \delta x^i\,, \qquad i = 1,\ldots,d,\\ &\phi_I(x) \to \phi'_I(x') = \phi_I(x) + \delta\phi_I(x)\,. \end{aligned}$$

As in the finite-dimensional case, the variations δx^i and $\delta\phi_I$ are expressed via infinitezimal linearly independent parameters $\delta\omega_n$:

$$\delta x^i = \sum_{1\le n\le s} X^i_n\delta\omega_n\,, \qquad \delta\phi_I(x) = \sum_{1\le n\le s}\Phi_{I,n}\delta\omega_n\,. \tag{1.9}$$

Here all $\delta\omega_n$ are independent of the coordinates x. Such transformations are called *global*. The coefficients X^i_n and $\Phi_{I,n}$ may depend on x and the fields, and they

[7]There is a simple rule to remember 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", then they must label "primed" coordinates in the right hand side of the formula.

describe a response of coordinates and fields on the infinitezimal transformation with a parameter $\delta\omega_n$.

Obviously, particular cases of the transformations above arise, when $X_n^k = 0$ or $\Phi_{I,n} = 0$. In the first case the coordinates x^i do not change under symmetry transformations at all, while the fields are transformed according to

$$\phi_I(x) \to \phi'_I(x) = \phi_I(x) + \delta\phi_I(x)\,.$$

In the second case the symmetry acts on the space-time coordinates only and the condition $\Phi_{I,n} = 0$ implies that $\phi'_I(x') = \phi_I(x)$, *i.e.* the fields under considerations are scalars. We point out that in the case when ϕ_I is not a scalar but rather a tensor, $\Phi_{I,n}$ is not zero even if the symmetry acts on the space-time coordinates only! To illustrate this point, consider a vector field $\phi^i(x)$. Under coordinate transformation $x^i \to x'^i = x^i + \delta x^i$ one gets

$$\phi'^i(x') = \frac{\partial x'^i}{\partial x^j}\phi^j(x) = \frac{\partial(x^i + \delta x^i)}{\partial x^j}\phi^j(x) = \phi^i(x) + \underbrace{\frac{\partial \delta x^i}{\partial x^j}\phi^j(x)}_{\delta\phi^i}\,,$$

which implies that the corresponding quantity Φ_I is non-trivial; the trivial case occurs only when δx^i does not depend on coordinates, *i.e.* it is a constant.

In the general case symmetry transformations act on both the space-time coordinates and the fields, cf. eq.(1.9). Consider

$$\phi'_I(x') = \phi'_I(x+\delta x) = \phi'_I(x) + \partial_k\phi'_I(x)\delta x^k + \ldots = \phi'_I(x) + \partial_k\phi_I(x)\, X_n^k\delta\omega_n + \ldots$$

It is important to realize that the operations δ and $\partial/\partial x$ do not commute. This is because δ is the variation of the fields due to both the change of their form and their arguments x^i. We therefore introduce the notion of the variation of the form of the field function

$$\bar\delta\phi_I(x) = \phi'_I(x) - \phi_I(x) = (\Phi_{I,n} - \partial_k\phi_I\, X_n^k)\delta\omega_n\,.$$

Variation of the form does commute with the derivative $\partial/\partial x$. For the variation of the Lagrangian density we, therefore, have

$$\mathscr{L}'(x') = \mathscr{L}'(x) + \frac{d\mathscr{L}}{dx^k}\delta x^k = \mathscr{L}(x) + \underbrace{\mathscr{L}'(x) - \mathscr{L}(x)}_{\bar\delta\mathscr{L}(x)} + \frac{d\mathscr{L}}{dx^k}\delta x^k\,.$$

The change of the action is[8]

$$\delta S = \int \mathrm{d}x'\,\mathscr{L}'(x') - \int \mathrm{d}x\,\mathscr{L}(x) = \int \mathrm{d}x'\,[\mathscr{L}(x) + \bar\delta\mathscr{L}(x) + \frac{d\mathscr{L}}{dx^k}\delta x^k] - \int \mathrm{d}x\,\mathscr{L}(x)\,.$$

[8]We consider a field theory in d-dimensions, so that the integration measure $\mathrm{d}x$ must be understood as $\mathrm{d}x = \mathrm{d}x_1\mathrm{d}x_2\ldots\mathrm{d}x_d \equiv \mathrm{d}^d x$.

Transformation of the integration measure is

$$\mathrm{d}x' = J \cdot \mathrm{d}x \equiv \underbrace{\det \begin{pmatrix} \frac{\partial x'^1}{\partial x^1} & \cdots & \frac{\partial x'^d}{\partial x^1} \\ \vdots & & \vdots \\ \frac{\partial x'^1}{\partial x^d} & \cdots & \frac{\partial x'^d}{\partial x^d} \end{pmatrix}}_{\text{Jacobian}} \mathrm{d}x = \det \begin{pmatrix} 1+\frac{\partial \delta x^1}{\partial x^1} & \cdots & \frac{\partial \delta x^d}{\partial x^1} \\ \vdots & & \vdots \\ \frac{\partial \delta x^1}{\partial x^d} & \cdots & 1+\frac{\partial \delta x^d}{\partial x^d} \end{pmatrix} \mathrm{d}x\,.$$

Thus, at leading order in $\delta\omega_n$ we have

$$\mathrm{d}x' = \mathrm{d}x(1 + \partial_k \delta x^k + \ldots).$$

Plugging this into the variation of the action, we find

$$\delta S = \int \mathrm{d}x \left[\bar{\delta}\mathscr{L}(x) + \frac{d\mathscr{L}}{dx^k}\delta x^k + \partial_k \delta x^k\, \mathscr{L}\right] = \int \mathrm{d}x \left[\bar{\delta}\mathscr{L}(x) + \frac{d}{dx^k}(\mathscr{L}\delta x^k)\right].$$

We further note that

$$\bar{\delta}\mathscr{L}(x) = \frac{\partial \mathscr{L}}{\partial \phi_I}\bar{\delta}\phi_I + \frac{\partial \mathscr{L}}{\partial(\partial_k \phi_I)}\partial_k \bar{\delta}\phi_I = \partial_k\Big(\frac{\partial \mathscr{L}}{\partial(\partial_k\phi_I)}\Big)\bar{\delta}\phi_I + \frac{\partial \mathscr{L}}{\partial(\partial_k\phi_I)}\partial_k\bar{\delta}\phi_I = \\ = \partial_k\Big(\frac{\partial \mathscr{L}}{\partial(\partial_k\phi_I)}\bar{\delta}\phi_I\Big)\,,$$

where we have used the Euler-Lagrange equations. Thus, we arrive at the following formula for the variation of the action

$$\delta S = \int \mathrm{d}x\, \frac{d}{dx^k}\Big[\frac{\partial \mathscr{L}}{\partial(\partial_k\phi_I)}\bar{\delta}\phi_I + \mathscr{L}\delta x^k\Big] = \int \mathrm{d}x\, \frac{d}{dx^k}\Big[\frac{\partial \mathscr{L}}{\partial(\partial_k\phi_I)}(\Phi_{I,n} - \partial_m\phi_I\, X_n^m) + \mathscr{L}X_n^k\Big]\delta\omega_n\,.$$

Since the integration volume is arbitrary we conclude that

$$\frac{dJ_n^k}{dx^k} = 0 \qquad \Longleftrightarrow \qquad \mathrm{div} J_n = 0\,,$$

where

$$J_n^k = -\frac{\partial \mathscr{L}}{\partial(\partial_k\phi_I)}(\Phi_{I,n} - \partial_m\phi_I\, X_n^m) - \mathscr{L}X_n^k$$

and $n = 1, \ldots s$. Thus, we have shown that the invariance of the action under the s-parametric symmetry transformations implies the existence of s conserved currents.

An important remark is in order. The quantities J_n^k are not uniquely defined. One can add

$$J_n^k \to J_n^k + \partial_m \chi_n^{km}\,,$$

where $\chi_n^{km} = -\chi_n^{mk}$. Adding such anti-symmetric functions does not influence the conservation law $\partial_k J_n^k = 0$.

Now we are ready to investigate concrete examples of symmetry transformations and derive the corresponding conserved currents.

- *Energy-momentum tensor.* Consider the infinitezimal space-time translations

$$x'^k = x^k + \delta x^k = x^k + \delta^k_n \delta\omega_n \qquad \Longrightarrow \qquad X^k_n = \delta^k_n$$

and $\Phi_{I,n} = 0$. Thus, the conserved current J^k_n becomes in this case a second rank tensor T^k_n

$$T^k_n = \frac{\partial \mathscr{L}}{\partial(\partial_k \phi_I)} \partial_n \phi_I - \delta^k_n \mathscr{L} \,.$$

Here, as usual, the sum over the index I is assumed. The quantity T^k_n is the so-called *stress-energy* or *energy-momentum* tensor. If all the fields vanish at spacial infinity then the integral[9]

$$P_n = \int \mathrm{d}^{n-1}x T^0_n$$

is a conserved quantity. Here 0 signifies the time direction and the integral is taken over the whole $(n-1)$-dimensional space. Indeed,

$$\frac{dP_n}{dt} = \int \mathrm{d}x \frac{dT^0_n}{dt} = -\int \mathrm{d}^{n-1}x \, \frac{dT^i_n}{dx^i} = -\int_{\Omega\to\infty} \mathrm{d}\Omega \, (\vec{T}_n \cdot \vec{n}) \,,$$

where Ω is a $(n-2)$-dimensional sphere which surrounds a $n-1$-dimensional volume; its radius tends to infinity. The vector $\vec{n}$ is a unit vector orthogonal to Ω.

- *Angular momentum.* Consider infinitezimal rotations $x'^n \to x^n + x_m \delta\Omega^{nm}$, where $\delta\Omega^{nm} = -\delta\Omega^{mn}$. Because of anti-symmetry, we can choose $\delta\Omega^{nm} = \delta\omega^{nm}$ with $n < m$ as linearly independent transformation parameters. We find

$$\begin{aligned} \delta x^k &= X^k_j \delta\omega^j = \sum_{n<m} X^k_{nm} \delta\omega^{nm} = x_l \delta\omega^{kl} = x_l \delta^k_m \delta\omega^{ml} \\ &= \sum_{m<l} x_l \delta^k_m \delta\omega^{ml} + \sum_{m>l} x_l \delta^k_m \delta\omega^{ml} = \sum_{m<l} (x_l \delta^k_m - x_m \delta^k_l) \delta\omega^{ml} \,. \end{aligned} \tag{1.10}$$

From here we deduce that

$$X^k_{nm} = x_m \delta^k_n - x_n \delta^k_m, \qquad n < m \,.$$

If we consider a scalar field then $\phi'(x') = \phi(x)$ and $\delta\phi = 0$. As a result, $\Phi_{I,n} = 0$. Using the general formula

$$J^k_n = -\frac{\partial \mathscr{L}}{\partial(\partial_k \phi_I)} (\Phi_{I,n} - \partial_m \phi_I \, X^m_n) - \mathscr{L} X^k_n \,,$$

[9]Here we explicitly distinguished a time direction t and write the integration measure in the action as $\mathrm{d}x = \mathrm{d}t \, \mathrm{d}^{n-1}x$.

we therefore find the following angular momentum tensor

$$M^k_{lm} = \frac{\partial \mathscr{L}}{\partial(\partial_k \phi)}(\partial_l \phi\, x_m - \partial_m \phi\, x_l) + \mathscr{L}(x_l \delta^k_m - x_m \delta^k_l)\,.$$

Notice that the last formula can be written in the form

$$M^k_{lm} = x_m\Big(\frac{\partial \mathscr{L}}{\partial(\partial_k \phi)}\partial_l \phi\, - \mathscr{L}\delta^k_l\Big) - x_l\Big(\frac{\partial \mathscr{L}}{\partial(\partial_k \phi)}\partial_m \phi\, - \mathscr{L}\delta^k_m\Big) = x_m T^k_l - x_l T^k_m\,,$$

where T^k_l is the stress-energy tensor.

If we consider now a vector field ϕ^i, then according to the discussion above, we will have

$$\delta\phi^i = \sum_{m<l} \Phi^i_{ml}\delta w^{ml} = \frac{\partial \delta x^i}{\partial x^j}\phi^j(x) = \frac{\partial}{\partial x^j}\Big(\sum_{m<l}(x_l\delta^i_m - x_m\delta^i_l)\delta\omega^{ml}\Big)$$

so that

$$\Phi^i_{ml} = (g_{jl}\delta^i_m - g_{jm}\delta^i_l)\phi^j = \phi_l\delta^i_m - \phi_m\delta^i_l\,,$$

where g_{ij} is a space-time metric. According to our general formula, the set of corresponding Noether currents will have the form

$$J^k_{mn} = -\frac{\partial \mathscr{L}}{\partial(\partial_k \phi^i)}(\Phi^i_{mn} - \partial_l\phi^i X^l_{mn}) - \mathscr{L}X^k_{mn}\,.$$

Substitution of all the quantities gives

$$J^k_{mn} = -\frac{\partial \mathscr{L}}{\partial(\partial_k \phi^i)}\big[\phi_n\delta^i_m - \phi_m\delta^i_n - \partial_l\phi^i(x_n\delta^l_m - x_m\delta^l_n)\big] - \mathscr{L}(x_n\delta^k_m - x_m\delta^k_n)\,.$$

We, therefore, see that for the vector field, the angular-momentum tensor takes the form

$$J^k_{mn} = x_n T^k_m - x_m T^k_n - \Big(\frac{\partial \mathscr{L}}{\partial(\partial_k \phi^n)}\phi_m - \frac{\partial \mathscr{L}}{\partial(\partial_k \phi^m)}\phi_n\Big)\,.$$

The first piece here, which depends on the stress-energy tensor is called *the orbital momentum* and the second piece characterizes polarization properties of the field and is related with a notion of *spin*.

The final remark concern continuous s-parametric transformations which leave the action invariant up to a total derivative term (in the original formulation of the Noether's an exact invariance of the action was assumed!)

$$\delta S = \delta\omega_n \int \mathrm{d}x\, \partial_k F^k_n\,.$$

These transformations also lead to conservation laws. It obtain them, it is enough to subtract from the canonical current J^k_n the term F^k_n:

$$\mathscr{J}^k_n = J^k_n - F^k_n\,.$$

One can verify that this new current is conserved $\partial_k \mathscr{J}^k_n$ as the consequence of the equations of motion.

1.5 Hamiltonian formalism in field theory

As was discussed above, in the Lagrangian formalism the dynamics of classical fields ϕ^i is described by the action functional

$$S = \int L\mathrm{d}t = \int \mathrm{d}t\mathrm{d}\vec{x}\,\mathscr{L}(\phi^i, \partial_\mu\phi^i)\,,$$

where $\mathscr{L}$ is the Lagrangian density being a function of ϕ^i and $\partial_\mu\phi^i$ taken at the same point x. The transition to the Hamiltonian formalism is performed by introducing the canonical momenta conjugate to the "coordinates" ϕ^i:

$$p_i(x) = \frac{\delta L}{\delta\dot{\phi}^i(x)} = \frac{\partial\mathscr{L}}{\partial\dot{\phi}^i(x)}\,.$$

The Hamiltonian has the form

$$H = \int \mathrm{d}\vec{x}\,\mathscr{H}\,, \qquad \mathscr{H} = \frac{\partial\mathscr{L}}{\partial\dot{\phi}^i(x)}\dot{\phi}^i(x) - \mathscr{L}\,,$$

where in the right hand side of the last formula one has to substitute the expression for $\dot{\phi}^i(x)$ via $p_i(x)$ and $\phi^i(x)$.

The definition of the Poisson brackets is also generalized to the field-theoretic case. For any two local in time functionals F and G of fields and their momenta we define their Poisson bracket as the following functional

$$\{F, G\} = \int \mathrm{d}\vec{x}\Big[\frac{\delta F}{\delta p_i(x)}\frac{\delta G}{\delta\phi^i(x)} - \frac{\delta G}{\delta p_i(x)}\frac{\delta F}{\delta\phi^i(x)}\Big]\,,$$

where F and G are taken at the same moment of time. The Hamiltonian equations are then

$$\dot{\phi}^i = \{H, \phi^i\}\,, \qquad\qquad \dot{p}_i = \{H, p_i\}\,.$$

The canonical Poisson brackets are

$$\begin{aligned}
&\{\phi^i(t,\vec{x}), \phi^j(t,\vec{y})\} = 0\,,\\
&\{p_i(t,\vec{x}), p_j(t,\vec{y})\} = 0\,,\\
&\{p_i(t,\vec{x}), \phi^j(t,\vec{y})\} = \delta_i^j\delta(\vec{x}-\vec{y})\,.
\end{aligned}$$

Note that all the fields for which the brackets are computed are taken at the one and the same moment of time!

Consider the simplest example of a real massive scalar field ϕ described by the Lagrangian density

$$\mathscr{L} = \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2)\,.$$

The momentum is

$$p(x) = \frac{\partial\mathscr{L}}{\partial\dot{\phi}(x)} = \dot{\phi}(x)$$

and, therefore, the Hamiltonian density is

$$\mathscr{H} = \frac{1}{2}\big(p^2 - \partial_i\phi\,\partial^i\phi + m^2\phi^2\big)\,.$$

2. Electrostatics

Classical electrodynamics is a theory of electric and magnetic fields caused by *macroscopic* distributions of electric charges and currents. 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.

2.1 Laws of electrostatics

Electrostatics is the study of electric fields produced by static charges. It is based entirely on Coulomb's law (1785). This law defines the force that two electrically charged bodies (point charges) exert on each other

$$\vec{F}\left(\vec{x}\right) = k\, q_1 q_2 \frac{\vec{x}_1 - \vec{x}_2}{|\vec{x}_1 - \vec{x}_2|^3}\,, \tag{2.1}$$

where k is Coulomb's constant (depends on the system of units used[10]), q_1 and q_2 are the magnitudes of the two charges, and $\vec{x}_1$ and $\vec{x}_2$ are their position vectors (as presented in Figure 1).

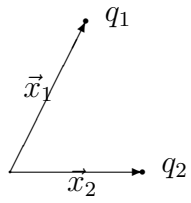


Figure 1: Two charges q_1 and q_2 and their respective position vectors $\vec{x}_1$ and $\vec{x}_2$. The charges exert an electric force on one another.

One can introduce the concept of an *electric field* $\vec{E}$ as the force experienced by a point-like charge q in the limit of vanishing q

$$\vec{E}\left(\vec{x}\right) = \lim_{q\to 0} \frac{\vec{F}\left(\vec{x}\right)}{q}\,.$$

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

[10] In SI units (SI – the international system of units), the Coulomb's constant is $k = \frac{1}{4\pi\epsilon_0}$, while force is measured in newtons, charge in coulombs, length in meters, and the vacuum permittivity ϵ_0 is given by $\epsilon_0 = \frac{10^7}{4\pi c^2} = 8.8542 \cdot 10^{-12} F/m$. Here, F indicates farad, a unit of capacitance being equal to one coulomb per volt. One can also use the Gauss system of units (CGS – centimetre-gram-second). In CGS units, force is expressed in dynes, charge in statcoulombs, length in centimeters, and the vacuum permittivity then reduces to $\epsilon_0 = \frac{1}{4\pi}$.

Coulomb's law, we obtain an expression for the electric field of a point charge

$$\vec{E}\,(\vec{x}) = kq\frac{\vec{x}-\vec{x}'}{|\vec{x}-\vec{x}'|^3}\,.$$

Since $\vec{E}$ is a vector quantity, for multiple charges we can apply the principle of linear superposition. Consequently, the field strength will simply be a sum of all of the contributions, which we can write as

$$\vec{E}\,(\vec{x}) = k\sum_i q_i\frac{\vec{x}-\vec{x}_i}{|\vec{x}-\vec{x}_i|^3}\,. \tag{2.2}$$

Introducing the electric charge density $\rho\,(\vec{x})$, the electric field for a continuous distribution of charge is given by

$$\vec{E}\,(\vec{x}) = k\int \rho\,(\vec{x}')\,\frac{\vec{x}-\vec{x}'}{|\vec{x}-\vec{x}'|^3}\mathrm{d}^3x'\,. \tag{2.3}$$

The Dirac delta-function (distribution) allows one to write down the electric charge density which corresponds to local charges

$$\rho\,(\vec{x}) = \sum_{i=1}^{N} q_i\delta\,(\vec{x}-\vec{x}_i)\ . \tag{2.4}$$

Substituting this formula into eq.(2.3), one recovers eq.(2.2).

However, eq.(2.3) 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 Figure 2,

$$\begin{aligned}\mathrm{d}N = \big(\vec{E}\cdot\vec{n}\big)\mathrm{d}S &= \frac{q}{r^3}\,(\vec{r}\cdot\vec{n})\,\mathrm{d}S\\ &= \frac{q}{r^2}\cos\,(\vec{r},\vec{n})\,\mathrm{d}S = \frac{q}{r^2}\mathrm{d}S'\,,\end{aligned}$$

where on the first step we have used that $\vec{E} = q\frac{\vec{r}}{r^3}$. 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. We introduce the solid angle $\mathrm{d}\Omega'$

$$\mathrm{d}\Omega' = \frac{\mathrm{d}S'}{r^2}\,. \tag{2.5}$$

Plugging this relation into eq.(2.5) leaves us with the following expression for the flux

$$\mathrm{d}N = q\cdot\mathrm{d}\Omega'\,. \tag{2.6}$$

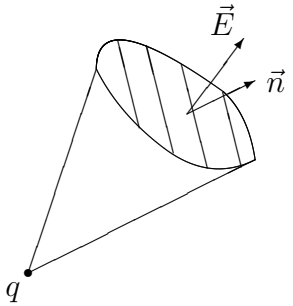


Figure 2: The electric flux through a surface, which is proportional to the charge within the surface.

By integrating eq.(2.6), we obtain the following equation for the flux N

$$\oint_S \left(\vec{E}\cdot\vec{n}\right)\mathrm{d}S = \begin{cases} 4\pi q & \text{if } q \text{ is inside the surface} \\ 0 & \text{otherwise} \end{cases}$$

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

$$N = \oint_S \left(\vec{E}\cdot\vec{n}\right)\mathrm{d}S = 4\pi \int \rho(x)\,\mathrm{d}^3x\,.$$

By making use of the Gauss-Ostrogradsky theorem, one may rewrite the above integral in terms of the volume integral of the divergence of the vector field $\vec{E}$

$$\oint_S \left(\vec{E}\cdot\vec{n}\right)\mathrm{d}S = \int_V \mathrm{div}\,\vec{E}\,(\vec{x})\,\mathrm{d}^3x\,.$$

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.

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. Hence, some information is necessary about the curl of electric field. This is in fact given by the second equation of electrostatics

$$\mathrm{rot}\,\vec{E} = 0\,. \tag{2.7}$$

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 derive this equation in the following way. Starting from the definition of the electric field (Coulomb's law) given by equation (2.3), we rewrite it in terms of a gradient and pull the differential operator outside of the integral

$$\begin{aligned}\vec{E}\,(\vec{x}) &= \int \rho\,(\vec{x}')\,\frac{\vec{x}-\vec{x}'}{|\vec{x}-\vec{x}'|^3}\mathrm{d}^3x' = -\int \rho\,(\vec{x}')\,\vec{\nabla}_x\frac{1}{|\vec{x}-\vec{x}'|}\mathrm{d}^3x' \\ &= -\vec{\nabla}_x\int\frac{\rho\,(\vec{x}')}{|\vec{x}-\vec{x}'|}\mathrm{d}^3x' = -\mathrm{grad}\int\frac{\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}\mathrm{d}^3x'\,.\end{aligned} \tag{2.8}$$

From vector calculus we know that the curl of gradient is always equal to zero, such that

$$\mathrm{rot}\,(\mathrm{grad}\,f) = 0 \qquad \Rightarrow \qquad \mathrm{rot}\,\vec{E} = 0\,.$$

This derivation shows that the vanishing of $\mathrm{rot}\,\vec{E}$ is not related to the inverse square law. It also shows that the electric field is the minus gradient of some scalar potential

$$\vec{E}(\vec{x}) = -\mathrm{grad}\,\varphi\,.$$

From the above, it then follows that this scalar potential is given by

$$\varphi(x) = \int\frac{\rho(x')}{|x-x'|}\mathrm{d}^3x'\,, \tag{2.9}$$

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 $\varphi(x)$ does not change the corresponding electric field $\vec{E}$.

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 $\vec{E}$

$$W = -\int_A^B \vec{F}\cdot\mathrm{d}\vec{l} = -q\int_A^B \vec{E}\cdot\mathrm{d}\vec{l}.$$

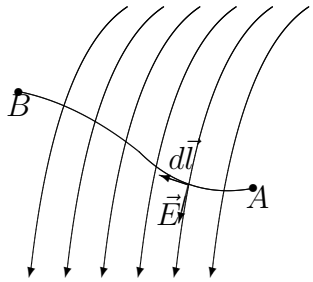


Figure 3: The work that has to be done over a charged particle to move it along the path from A to B through an electric field $\vec{E}$.

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

$$\begin{aligned} W &= q \int_A^B \mathrm{grad}\varphi \cdot \mathrm{d}\vec{l} = q \int_A^B \frac{\partial \varphi}{\partial x}\mathrm{d}x + \frac{\partial \varphi}{\partial y}\mathrm{d}y + \frac{\partial \varphi}{\partial z}\mathrm{d}z \\ &= \int_{t_A}^{t_B} \Big(\frac{\partial \varphi}{\partial x}\frac{dx}{dt} + \frac{\partial \varphi}{\partial y}\frac{dy}{dt} + \frac{\partial \varphi}{\partial z}\frac{dz}{dt} \Big) dt = q\left(\varphi_B - \varphi_A\right) , \end{aligned}$$

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

$$\oint \left(\vec{E} \cdot \mathrm{d}\vec{l}\right) = \oint_S \mathrm{rot}\,\vec{E} \cdot \mathrm{d}\vec{S} = 0\,,$$

where we have used the fact that $\mathrm{rot}\,\vec{E} = 0$.

To summarize, we have derived two laws of electrostatics in the differential form

$$\vec{\nabla} \cdot \vec{E}\,(\vec{x}) = \mathrm{div}\,\vec{E}\,(\vec{x}) = 4\pi\rho\,(\vec{x})\ , \tag{2.10}$$

$$\vec{\nabla} \times \vec{E}\,(\vec{x}) = \mathrm{rot}\,\vec{E}\,(\vec{x}) = 0\,. \tag{2.11}$$

2.2 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

$$\text{rot}\,\vec{E}\,(\vec{x}) = 0 \quad \Rightarrow \quad \vec{E}\,(\vec{x}) = -\vec{\nabla}\varphi\,(\vec{x})\ .$$

Substituting the right hand side of this expression into equation (2.10), we obtain

$$\text{div}\,\vec{\nabla}\varphi\,(\vec{x}) = -4\pi\rho\,(\vec{x})\ .$$

This gives

$$\nabla^2\varphi\,(\vec{x}) \equiv \Delta\varphi\,(\vec{x}) = -4\pi\rho\,(\vec{x})\ . \tag{2.12}$$

Equation (2.12) is known as *the Poisson equation*. In case $\rho\,(\vec{x}) = 0$, i.e. in a region of no charge, the left hand side of (2.12) is zero, which is known as the Laplace equation. Substituting into (2.12) the form scalar potential φ, given by (2.9) , we get

$$\nabla^2\varphi\,(\vec{x}) = \nabla^2\int\frac{\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}\mathrm{d}^3x' = \int\mathrm{d}^3x'\,\rho(\vec{x}')\nabla^2\left(\frac{1}{|\vec{x}-\vec{x}'|}\right)\ .$$

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

$$\nabla^2\frac{1}{|\vec{x}-\vec{x}'|} = \nabla^2\frac{1}{r} = \frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}\left(r\frac{1}{r}\right) = 0\,.$$

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 law and the fact that $\vec{\nabla}r = \vec{n}$, one obtains

$$\begin{aligned}\int\limits_V\nabla^2\left(\frac{1}{r}\right)\mathrm{d}^3x &= \int\limits_V\text{div}\,\vec{\nabla}\left(\frac{1}{r}\right)\mathrm{d}^3x = \oint\limits_S\vec{n}\cdot\vec{\nabla}\frac{1}{r}\,\mathrm{d}S\\ &= \oint\limits_S\vec{n}\cdot\frac{\partial}{\partial r}\left(\frac{1}{r}\right)\vec{n}\,\mathrm{d}S = \oint\limits_S\frac{\partial}{\partial r}\left(\frac{1}{r}\right)\underbrace{r^2\mathrm{d}\Omega}_{\mathrm{d}S} = -4\pi\,.\end{aligned}$$

Therefore,

$$\nabla^2\left(\frac{1}{r}\right) = -4\pi\delta(\vec{x})\ ,$$

or

$$\nabla_x^2\frac{1}{|\vec{x}-\vec{x}'|} = -4\pi\delta\,(\vec{x}-\vec{x}')\ .$$

Thus, we find

$$\nabla^2\varphi = \int\rho(x')\,(-4\pi\delta(x-x'))\,\mathrm{d}^3x' = -4\pi\rho(x)\,.$$

Hence, we have proved that $\frac{1}{r}$ solves the Poisson equation with the point charge source. In general, the functions satisfying $\vec{\nabla}^2\varphi = 0$ are called *harmonic functions*.

2.3 The Green theorems

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)

$$\varphi(x) = \int \rho(x') \frac{\mathrm{d}^3 x'}{|x - x'|} \tag{2.13}$$

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 (2.13) cannot be applied with the exception of some particular cases (as in the method of images). To understand boundary problems, one has to invoke the Green theorems.

Consider an arbitrary vector field[11] $\vec{A}$. We have

$$\int\limits_V \operatorname{div} \vec{A}\, \mathrm{d}^3 x = \oint\limits_S \left(\vec{A} \cdot \vec{n}\right) \mathrm{d}S\,. \tag{2.14}$$

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

$$\begin{aligned} \operatorname{div} \vec{A} &= \operatorname{div}\left(\varphi \cdot \vec{\nabla}\psi\right) = \operatorname{div}\left(\varphi \frac{\partial \psi}{\partial x_i}\right) = \frac{\partial}{\partial x^i}\left(\varphi \frac{\partial \psi}{\partial x_i}\right) \\ &= \vec{\nabla}\varphi \cdot \vec{\nabla}\psi + \varphi \nabla^2 \psi\,. \end{aligned}$$

Substituting this back into eq.(2.14), we get

$$\int\limits_V \left(\vec{\nabla}\varphi \cdot \vec{\nabla}\psi + \varphi\nabla^2\psi\right) \mathrm{d}^3 x = \oint\limits_S \varphi \cdot \left(\vec{\nabla}\psi \cdot \vec{n}\right) \mathrm{d}S = \oint\limits_S \varphi \left(\frac{\mathrm{d}\psi}{\mathrm{d}n}\right) \mathrm{d}S\,.$$

which 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*

$$\int\limits_V \left(\varphi\nabla^2\psi - \psi\nabla^2\varphi\right) \mathrm{d}^3 x = \oint\limits_S \left(\varphi \frac{\mathrm{d}\psi}{\mathrm{d}n} - \psi \frac{\mathrm{d}\varphi}{\mathrm{d}n}\right) \mathrm{d}S\,. \tag{2.15}$$

[11] Now introduced for mathematical convenience, but it will later prove to be of greater importance.

By using this formula, the differential Poisson equation can be reduced to an integral equation. Indeed, consider a function ψ such that

$$\psi \equiv \frac{1}{R} = \frac{1}{|\vec{x} - \vec{x}'|} \quad \Rightarrow \quad \nabla^2 \psi = -4\pi\delta(\vec{x}) \; . \tag{2.16}$$

Substituting it into the second Green formula (2.15) and assuming x is inside the space V integrated over, one gets

$$\int\limits_V \left(-4\pi\varphi(\vec{x}')\delta(\vec{x} - \vec{x}') + \frac{4\pi\rho(\vec{x}')}{|\vec{x} - \vec{x}'|} \right) \mathrm{d}^3x' = \oint\limits_{S'} \left[\varphi \frac{\mathrm{d}}{\mathrm{d}n'} \left(\frac{1}{R} \right) - \frac{1}{R} \frac{\mathrm{d}\varphi}{\mathrm{d}n'} \right] \mathrm{d}S' \, .$$

Here we have chosen $\varphi(\vec{x}')$ to satisfy the Poisson equation $\Delta\varphi(\vec{x}') = -4\pi\rho(\vec{x}')$. By using the sampling property of the delta function, i.e. $\int_V \varphi(\vec{x}')\,\delta(\vec{x} - \vec{x}') = \varphi(\vec{x})$, the expression above allows one to express $\varphi(\vec{x})$ as

$$\varphi(\vec{x}) = \int_V \frac{\rho(\vec{x}')}{R} \mathrm{d}^3x' + \frac{1}{4\pi} \oint_S \left[\frac{1}{R} \frac{\partial\varphi}{\partial n'} - \varphi \frac{\partial}{\partial n'} \left(\frac{1}{R} \right) \right] \mathrm{d}S' \, , \tag{2.17}$$

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:

- 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.(2.13).
- 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 the 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*[12].

[12]Note that both Dirichlet as well as Neumann boundary conditions are not only limited to electrodynamics, but are more general and appear throughout the field of ordinary or partial differential equations.

One can prove, with the help of the first Green formula, that the Poisson equation

$$\vec{\nabla}^2\varphi = -4\pi\rho\,,$$

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 $\nabla^2 U = 0$ 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

$$\int\limits_V \left(\left|\vec{\nabla} U\right|^2 + U\nabla^2 U \right) \mathrm{d}^3x = \oint\limits_S U \left(\frac{\partial U}{\partial n} \right) \mathrm{d}S\,. \tag{2.18}$$

Here the second term in the integral vanishes as $\vec{\nabla}^2 U = 0$ 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. This equation is true *iff*[13]

$$\begin{aligned} \int_V |\vec{\nabla} U|^2 = 0 &\longrightarrow |\vec{\nabla} U| = 0 \\ &\longrightarrow \vec{\nabla} U = 0 \end{aligned} \tag{2.19}$$

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.

2.4 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 problems. The method is based on an impulse from a source, which is later integrated with the source function over entire space. Recall

$$\nabla^2 \frac{1}{|\vec{x} - \vec{x}'|} = -4\pi\delta\left(\vec{x} - \vec{x}'\right)\,. \tag{2.20}$$

However, the function $\frac{1}{|\vec{x}-\vec{x}'|}$ is just one of many functions which obeys $\nabla^2\psi = -4\pi\delta\left(\vec{x} - \vec{x}'\right)$. The functions that are solutions of this second order differential equation are known as *Green's functions*. In general,

$$\vec{\nabla}^2 G\left(\vec{x}, \vec{x}'\right) = -4\pi\delta\left(\vec{x} - \vec{x}'\right)\,, \tag{2.21}$$

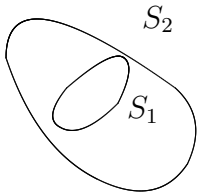


Figure 4: Choosing arbitrarily the surfaces S_1 and S_2, where S is the area between them, we let them expand so that the average value of the scalar potential tends to zero.

where $G(\vec{x},\vec{x}') = \frac{1}{|\vec{x}-\vec{x}'|} + F(\vec{x},\vec{x}')$, so that $\vec{\nabla}^2 F(\vec{x},\vec{x}') = 0$, i.e. it obeys the Laplace equation inside V.

The point is now to find such $F(\vec{x},\vec{x}')$, that gets rid of one of the terms in the integral equation (2.17) we had for $\varphi(\vec{x})$. Letting $\varphi = \varphi(\vec{x})$ and $\psi = G(\vec{x},\vec{x}')$, we then get

$$\varphi(\vec{x}) = \int_V \rho(\vec{x}')\, G(\vec{x},\vec{x}')\, \mathrm{d}^3x' + \frac{1}{4\pi}\oint_S \left[G(\vec{x},\vec{x}')\, \frac{\partial\varphi(\vec{x}')}{\partial n'} - \varphi(\vec{x}')\, \frac{\partial G(\vec{x},\vec{x}')}{\partial n'} \right] \mathrm{d}S' \,.$$

By using the arbitrariness in the definition of the Green function we can leave in the surface integral the desired boundary conditions. For the Dirichlet case we can choose $G_{boundary}(\vec{x},\vec{x}') = 0$ when $\vec{x}' \in S$, then $\varphi(\vec{x})$ simplifies to

$$\varphi(\vec{x}) = \int_V \rho(\vec{x}')\, G(\vec{x},\vec{x}')\, \mathrm{d}^3x' - \frac{1}{4\pi}\oint_S \varphi(\vec{x}')\, \frac{\partial G(\vec{x},\vec{x}')}{\partial n'} \mathrm{d}S' \,,$$

where $G(\vec{x},\vec{x}')$ is referred to as the bulk-to-bulk propagator and $\frac{\partial G(\vec{x},\vec{x}')}{\partial n'}$ is the bulk-to-boundary propagator.
For the Neumann case we could try to choose $\frac{\partial G(\vec{x},\vec{x}')}{\partial n'} = 0$ when $\vec{x}' \in S$. However, one has

$$\begin{aligned} \oint \frac{\partial G(\vec{x},\vec{x}')}{\partial n'} \mathrm{d}S' = \oint_S \left(\vec{\nabla} G \cdot \vec{n} \right) \mathrm{d}S' = \int \mathrm{div}\vec{\nabla} G\, \mathrm{d}^3x' = \int \nabla^2 G\, \mathrm{d}^3x' \\ = -4\pi \int \delta(\vec{x}-\vec{x}')\, \mathrm{d}^3x' = -4\pi \,. \end{aligned} \tag{2.22}$$

For this reason we can not demand $\frac{\partial G(\vec{x},\vec{x}')}{\partial n'} = 0$. Instead, one chooses another simple condition $\frac{\partial G(\vec{x},\vec{x}')}{\partial n'} = -\frac{4\pi}{S}$, where S is the total surface area, and the left hand side of

[13] "If and only if".

the equation is referred to as the Neumann Green function. Using this condition:

$$\varphi(\vec{x}) = \int_V \rho(\vec{x}')\, G_N(x, x')\, \mathrm{d}^3x' + \frac{1}{4\pi} \oint_S G_N(\vec{x}, \vec{x}')\, \frac{\partial \varphi(\vec{x}')}{\partial n'} \mathrm{d}S' + \frac{1}{S} \oint_S \varphi(\vec{x}')\, \mathrm{d}S' \tag{2.23}$$

The last term represents $\langle\varphi\rangle$, the averaged value of the potential on S. If one takes the limit $S = S_1 + S_2 \to \infty$, where S_1 and S_2 are two surfaces enclosing the volume V and such that S_2 tends to infinity, this average disappears. In any case, the extra term $\frac{1}{S} \oint_S \varphi(\vec{x}')\, \mathrm{d}S'$ is just a constant (does not depend on x) and, therefore, does not contribute to the electric field $\vec{E} = -\vec{\nabla}\varphi$.

2.5 Electrostatic problems with spherical symmetry

Frequently, when dealing with electrostatics, one encounters the problems exhibiting spherical symmetry. As an example, take the Coulomb law (2.1), 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.

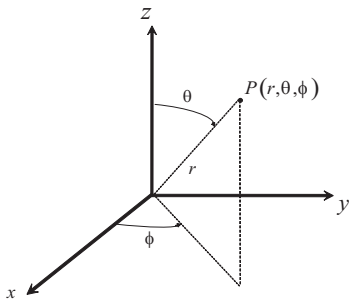


Figure 5: Spherical coordinate system.

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

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}\,, \\ \theta &= \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)\,, \\ \phi &= \arctan\left(\frac{y}{x}\right)\,, \end{aligned} \tag{2.24}$$

To obtain the Cartesian coordinates from the spherical ones, we use

$$\begin{aligned} x &= r \sin\theta \cos\phi\,, \\ y &= r \sin\theta \sin\phi\,, \\ z &= r \cos\theta\,. \end{aligned} \tag{2.25}$$

In terms of spherical coordinates our differential operators look different. The one we will be most interested in, the Laplace operator, becomes

$$\vec{\nabla}^2 = \frac{1}{r^2}\left(\frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\left(\frac{\partial}{\partial \theta}\sin\theta \frac{\partial}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial \phi^2}\,.$$

Hence, in these coordinates the Laplace equation reads as

$$\vec{\nabla}^2 \varphi = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\varphi) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \varphi}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 \varphi}{\partial \phi^2} = 0\,.$$

We use the *ansatz* that $\varphi(r,\theta,\phi) = \frac{U(r)}{r} P(\theta) Q(\phi)$. 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 $Q = e^{\pm i m \phi}$, where m is an integer such that Q 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

$$\frac{\mathrm{d}^2 U}{\mathrm{d}r^2} - \frac{l(l+1)}{r^2} U = 0\,,$$

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

$$U(r) = A r^{l+1} + B r^{-l}\,,$$

where l is assumed to be positive and A and B are arbitrary constants. 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

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(1-x^2)\frac{\mathrm{d}P}{\mathrm{d}x}\right] + l(l+1)P = 0\,. \tag{2.26}$$

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 $-1 \leq x \leq 1$. We try as a solution the following power series (the Frobenius method)

$$P\left(x\right) = x^{\alpha}\sum_{j=0}^{\infty} a_j x^j\,, \tag{2.27}$$

where α is unknown. Substituting our trial solution (2.27) into the Legendre equation (2.26), we obtain

$$\sum_{j=0}^{\infty}\Big(\left(\alpha+j\right)\left(\alpha+j-1\right)a_j x^{\alpha+j-2} - \left[\left(\alpha+j\right)\left(\alpha+j+1\right) - l\left(l+1\right)\right]a_j x^{\alpha+j}\Big) = 0\,.$$

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 equation equal to zero 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

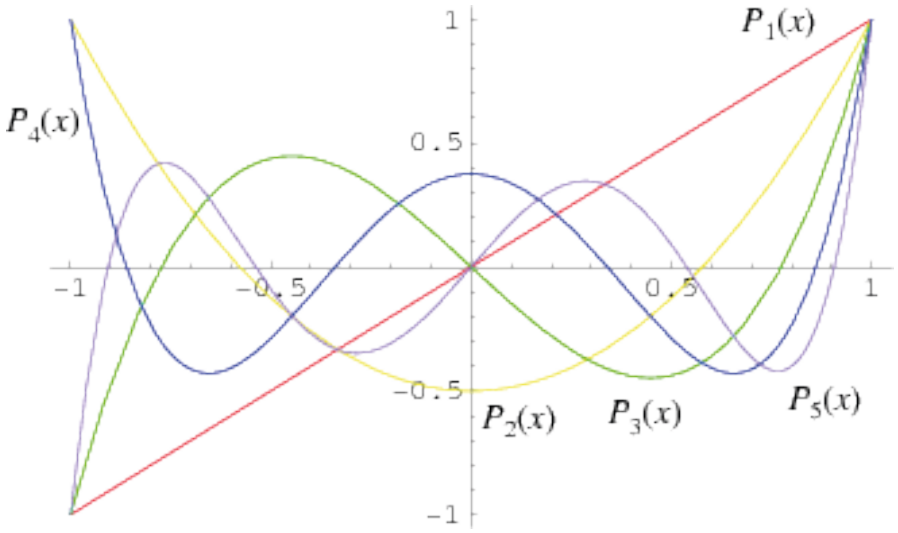


Figure 6: Profiles of a few Legendre polynomials.

such a way that they are all equal to identity at $x = 1$. These are the Legendre polynomials $P_l(x)$:

$$\begin{aligned} P_0(x) &= 1\,; \\ P_1(x) &= x\,; \\ P_2(x) &= \frac{1}{2}\left(3x^2-1\right)\,; \\ P_3(x) &= \frac{1}{3}\left(5x^3-2x\right)\,; \\ &\cdots \\ P_l(x) &= \frac{1}{2^l l!}\frac{d^l}{dx^l}\left(x^2-1\right)^l\,. \end{aligned}$$

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

$$\int_{-1}^{1}\left[(x^2-1)\frac{\mathrm{d}P_{l'}}{\mathrm{d}x}\frac{\mathrm{d}P_l}{\mathrm{d}x} + l(l+1)P_{l'}P_l\right)\right]\mathrm{d}x = 0\,.$$

Now subtract the same equation, but with the interchange of l and l', such that the following expression is left

$$[(l'(l'+1) - l(l+1)] \int_{-1}^{1} P_{l'} P_l = 0\,.$$

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 an identity

$$\int_{-1}^{1} P_{l'}(x) P_l(x) \mathrm{d}x = \frac{2}{2l+1} \delta_{l',l}\,.$$

For any function defined on $-1 \leq x \leq 1$

$$\begin{aligned} f(x) &= \sum_{l=0}^{\infty} A_l P_l(x)\,, \\ A_l &= \frac{2l+1}{2} \int_{-1}^{1} f(x) P_l(x) \mathrm{d}x\,. \end{aligned}$$

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 $B_l = 0$. Take a sphere of radius a with the potential $V(\theta)$. Then

$$V(\theta) = \sum_{l=0}^{\infty} A_l a^l P_l(\cos\theta)$$

so that

$$A_l = \frac{2l+1}{2a^l} \int_0^{\pi} V(\theta) P_l(\cos\theta) \sin\theta \mathrm{d}\theta\,.$$

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

$$\begin{aligned} \frac{d}{dx}(1-x^2)P' + l(l+1)P &= 0 \\ \frac{d}{dx}(1-x^2)Q' + l(l+1)Q &= 0\,. \end{aligned}$$

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 $y=\pm1$. $Q_n(x)$ are called as *Legendre functions of the second kind.*

Example: find the potential of an empty sphere of radius $r=a$ 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

$$\begin{aligned}\varphi(r,\theta)&=\frac{V}{\sqrt{\pi}}\sum_{j=1}^{\infty}(-1)^{j-1}\frac{(2j-\frac{1}{2})\Gamma(j-\frac{1}{2})}{j!}\left(\frac{r}{a}\right)^{2j}P_{2j-1}(\cos\theta)\\&=V\left[\frac{3}{2}\left(\frac{r}{a}\right)P_1(\cos\theta)-\frac{7}{8}\left(\frac{r}{a}\right)^3P_3(\cos\theta)+\frac{11}{16}\left(\frac{r}{a}\right)^5P_5(\cos\theta)-\ldots\right]\,.\end{aligned}$$

Here $\Gamma(z)$ for $\Re(z)>0$ is defined as

$$\Gamma(z)=\int_0^\infty t^{z-1}e^{-t}\mathrm{d}t\,.$$

Finally, we would like to comment on the solutions of the Laplace equation $\triangle\varphi=0$. It is not difficult to show that one cannot have an absolute minimum or maximum in the region (in both directions, x and y) because for an extremum to exist one requires $\frac{\partial\varphi}{\partial x_i}=0$ which results in $\frac{\partial^2\varphi}{\partial x_i^2}>0$ or $\frac{\partial^2\varphi}{\partial x_i^2}<0$ implying that in the other direction the second derivative must have an opposite sign.

Now we come back to the general case when azimuthal symmetry is absent. In this case we have an equation

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\left(1-x^2\right)\frac{\mathrm{d}P}{\mathrm{d}x}\right]+\left[l(l+1)-\frac{m^2}{1-x^2}\right]P=0\,,$$

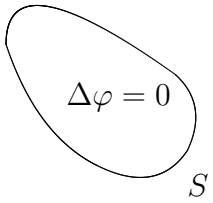


Figure 7: The field $\varphi(\vec{x})$, which obeys the Laplace equation, has no maximum or minimum inside a region S.

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 $-l, -(l-1), \ldots, 0, \ldots, l-1, l$.

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 are called *spherical harmonics*. Such functions are chosen to be

$$Y_{lm}(\theta,\phi) = \left(\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\right)^{\frac{1}{2}} P_l^m(\cos\theta)e^{im\phi}\,.$$

They are normalized as

$$\int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \mathrm{d}\theta\, \sin\theta\, Y^*_{lm}(\theta,\phi)Y_{l'm'}(\theta,\phi) = \delta_{ll'}\delta_{mm'}\,.$$

An arbitrary function $f(\theta,\phi)$ on a sphere can be expanded in a series over spherical harmonics

$$f(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{m=1} A_{lm}Y_{lm}(\theta,\phi)\,.$$

Coefficients A_{lm} are found by using orthogonality condition for spherical harmonics. This completes our discussion of solving the Laplace equation in spherical coordinates.[14]

[14] Analogously, one can treat the case of cylindrical, elliptical or other orthogonal coordinate systems.

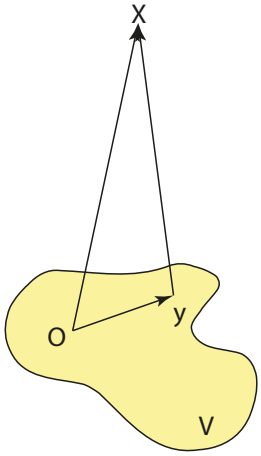


Figure 8: Multipole expansion is an expansion of the exact expression for the scalar potential on distances that are large in comparison with a region of charge localization.

2.6 Multipole expansion for scalar potential

Let us assumed that electric charge is localized with the local charge density $\rho(x)$ inside a bounded region V. We chose an origin of a coordinate system somewhere inside V. Let us call $\max|y| = L$, where y is an arbitrary point in V, "the size" of our system of charges.

It is interesting to know the scalar potential $\varphi(x)$ outside V, that is in the region $r \equiv |x| \geq L$. 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

$$\varphi(x) = \int_V \mathrm{d}y \frac{\rho(y)}{|x-y|}$$

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 $|x-y|^{-1}$ into Taylor series in variable y:

$$\frac{1}{|x-y|} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} y_{i_1} \cdots y_{i_n} \, \partial_{i_1} \cdots \partial_{i_n} \frac{1}{r} \,,$$

where $|y| < |x| = r$. Substituting this expansion into the expression for the potential, we get

$$\varphi(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \, T_{i_1 \ldots i_n} \, \partial_{i_1} \cdots \partial_{i_n} \frac{1}{r} \,,$$

where

$$T_{i_1 \ldots i_n} = \int \mathrm{d}y \, \rho(y) y_{i_1} \cdots y_{i_n} \,.$$

This is a multipole expansion and $T_{i_1\ldots i_n}$ are called the multipole momenta. The first ones are

$$\begin{aligned} Q &= \int \mathrm{d}y\, \rho(y) \quad - \quad \text{total electric charge} \\ d_i &= \int \mathrm{d}y\, \rho(y) y_i \quad - \quad \text{dipole moment} \\ T_{ik} &= \int \mathrm{d}y\, \rho(y) y_i y_k \quad - \quad \text{quadrupole moment} \end{aligned} \tag{2.28}$$

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 $dy' = dy$, one gets

  $$T'_{i_1\ldots i_n} = \int \mathrm{d}y\, \rho(y)(y_{i_1} + a_i) \cdots (y_{i_n} + a_{i_n})$$

  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.[15]

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

[15]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

$$\vec{d}_{\vec{a}} = \sum_{i=1}^{N} e_i\,(\vec{x}_i - \vec{a}) = \sum_{i=1}^{N} e_i \vec{x}_i - \vec{a} \sum_{i=1}^{N} e_i = \sum_{i=1}^{N} e_i \vec{x}_i = \vec{d}.$$

cannot construct by contacting indices a simpler object – a tenor 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

$$\mathscr{T}_{ij} = \int \mathrm{d}y\, \rho(y) \big(y_i y_j - \frac{y^2}{3} \delta_{ij} \big)\,.$$

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 x. Thus, the multipole expansion can be written as

$$\varphi(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\, \mathscr{T}_{i_1 \ldots i_n}\, \partial_{i_1} \cdots \partial_{i_n} \frac{1}{r}\,.$$

We further notice that

$$\begin{aligned} \partial_i \frac{1}{r} &= -\frac{x_i}{r^3}\,, \\ \partial_i \partial_j \frac{1}{r} &= -\frac{\delta_{ij}}{r^3} + (-1)(-3) \frac{x^i x^j}{r^3} \end{aligned} \tag{2.29}$$

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

$$\varphi(x) = \sum_{n=0}^{\infty} \frac{(2n-1)!!}{n!}\, \mathscr{T}_{i_1 \ldots i_n}\, \frac{x_{i_1} \cdots x_{i_n}}{r^{2n+1}}\,.$$

Explicitly,

$$\varphi(x) = \frac{q}{r} + \frac{d_i x_i}{r^3} + \frac{3 \mathscr{T}_{ij} x_i x_j}{r^5} + \ldots$$

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 $\varphi(x)$ 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

$$\varphi = \frac{q}{r} + \frac{(\vec{x}\cdot\vec{d})}{r^3} + \ldots\,,$$

where we have used neutrality of the system of charges. Thus, the electric field is

$$\vec{E} = -\vec{\nabla}\frac{(\vec{x}\cdot\vec{d})}{r^3} = \frac{3\vec{n}(\vec{n}\cdot\vec{d}) - \vec{d}}{r^3}\,.$$

Thus, for a neutral system the electric field at large distances from this system behaves itself as $1/r^3$!

3. Magnetostatics

3.1 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 take 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

$$\vec{A}(x) = \frac{1}{c}\int \mathrm{d}^3x' \, \frac{\vec{j}(x')}{|x-x'|}\,.$$

Now we can determine the corresponding magnetic field

$$\vec{H} = \mathrm{rot}\,\vec{A} = \frac{1}{c}\int \mathrm{d}^3x' \left[\vec{\nabla}\frac{1}{|x-x'|}, \vec{j}(x')\right] = \frac{1}{c}\int \mathrm{d}^3x' \, \frac{[\vec{j}(x'), \vec{R}]}{R^3}\,,$$

where the bracket means the vector product[16]. 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 Stocks theorem gives

$$\oint_C \vec{H}\cdot\mathrm{d}\vec{\ell} = \frac{4\pi}{c}\int_S (\vec{j}\cdot\vec{n})\mathrm{d}S = \frac{4\pi}{c}I\,,$$

where $I = \int_S (\vec{j}\cdot\vec{n})\mathrm{d}S$ is the full current through the surface S. This is the Ampère law.

3.2 Magnetic (dipole) moment

Free magnetic charges do not exist. The really existing object which plays the basic role[17] 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 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 $x' << x$ and we can expand

$$\frac{1}{|x-x'|} = \frac{1}{|x|} + \frac{(\vec{x}\cdot\vec{x}')}{|x|^3} + \dots$$

Therefore, for the vector potential we get

$$A_i(x) = \frac{1}{c|x|}\int j_i(x')\mathrm{d}^3x' + \frac{1}{c|x|^3}\cdot\int j_i(x')(\vec{x}\cdot\vec{x}')\mathrm{d}^3x' + \cdots$$

[16]Here we have used the formula $\mathrm{rot}\, f\vec{A} = f\mathrm{rot}\,\vec{A} + [\mathrm{grad}\, f, \vec{A}]$.

[17]The same role as elementary electric charge in electrostatics.

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

$$0 = \int f(x') \operatorname{div} \vec{j} \, \mathrm{d}^3 x' = - \int (\vec{\nabla} f \cdot \vec{j}) \, \mathrm{d}^3 x' \,,$$

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

$$\int j_i(x') \mathrm{d}^3 x' = 0 \quad \text{for any } i \,.$$

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

$$\vec{A}(x) = \frac{1}{c|x|^3} \cdot \int \vec{j}(x')(\vec{x} \cdot \vec{x}') \, \mathrm{d}^3 x'$$

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

$$\int \mathrm{d}^3 x' \, j_k x'_i = \int \mathrm{d}^3 x' \operatorname{div}(x'_k \vec{j}) x'_i \overset{\text{by parts}}{=} - \int \mathrm{d}^3 x' \, x'_k (\vec{j} \cdot \operatorname{grad}') x'_i = - \int \mathrm{d}^3 x' \, x'_k j_i \,.$$

From here we deduce that

$$\int \mathrm{d}^3 x' \, (\vec{x} \cdot \vec{j}) \, x'_i = - \int \mathrm{d}^3 x' \, (\vec{x} \cdot \vec{x'}) \, j_i \,,$$

or, in the vector form,

$$\int \mathrm{d}^3 x' \, (\vec{x} \cdot \vec{j}) \, \vec{x}' = - \int \mathrm{d}^3 x' \, (\vec{x} \cdot \vec{x'}) \, \vec{j} \,.$$

Therefore, we arrive at

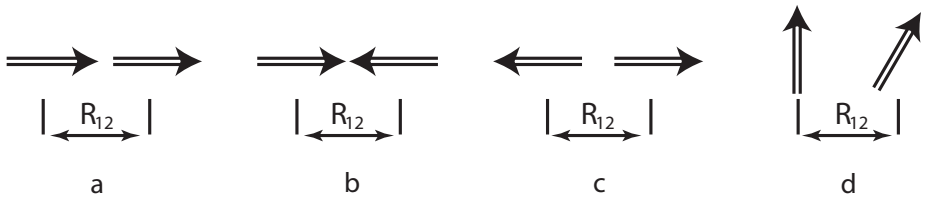
$$\vec{A} = - \frac{\vec{x}}{|x|^3} \times \frac{1}{2c} \int \mathrm{d}^3 x' \, \vec{x}' \times \vec{j}(x') \,.$$

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

$$\vec{M} = \int \mathrm{d}^3x' \, \vec{\mathcal{M}}(x') = \frac{1}{2c} \int \mathrm{d}^3x' \, \vec{x}' \times \vec{j}(x') \, .$$



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 $U_M = \frac{(\vec{M}_1 \cdot \vec{M}_2) - 3(\vec{M}_1 \cdot \vec{n}_{12})(\vec{M}_2 \cdot \vec{n}_{12})}{R_{12}^3}$, $\quad \vec{n}_{12} = \frac{\vec{R}_{12}}{R_{12}}$.

We, therefore, find

$$\vec{A}(x) = \frac{\vec{M} \times \vec{x}}{|x|^3} \, .$$

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

$$\vec{H} = \mathrm{rot} \, \vec{A} = \frac{3\vec{n}(\vec{n} \cdot \vec{M}) - \vec{M}}{|x|^3} \, ,$$

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.

3.3 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,

$$|M| = \frac{IS}{c}$$

independently of the shape of the contour. Here $|M|$ 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 $\vec{v}_i << c$, 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 $v << c$ 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

$$\vec{M} = \frac{e}{2c\,m}\sum_i \vec{L}_i = \frac{e}{2c\,m}\vec{L}\,,$$

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!

4. Relativistic Mechanics

4.1 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 Figure 9. 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

$$\begin{aligned} \vec{x}' &= R(\vec{x} - \vec{v}t) \\ t' &= t \end{aligned} \tag{4.1}$$

Here R is a matrix of orthogonal transformations of coordinates.

4.2 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 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

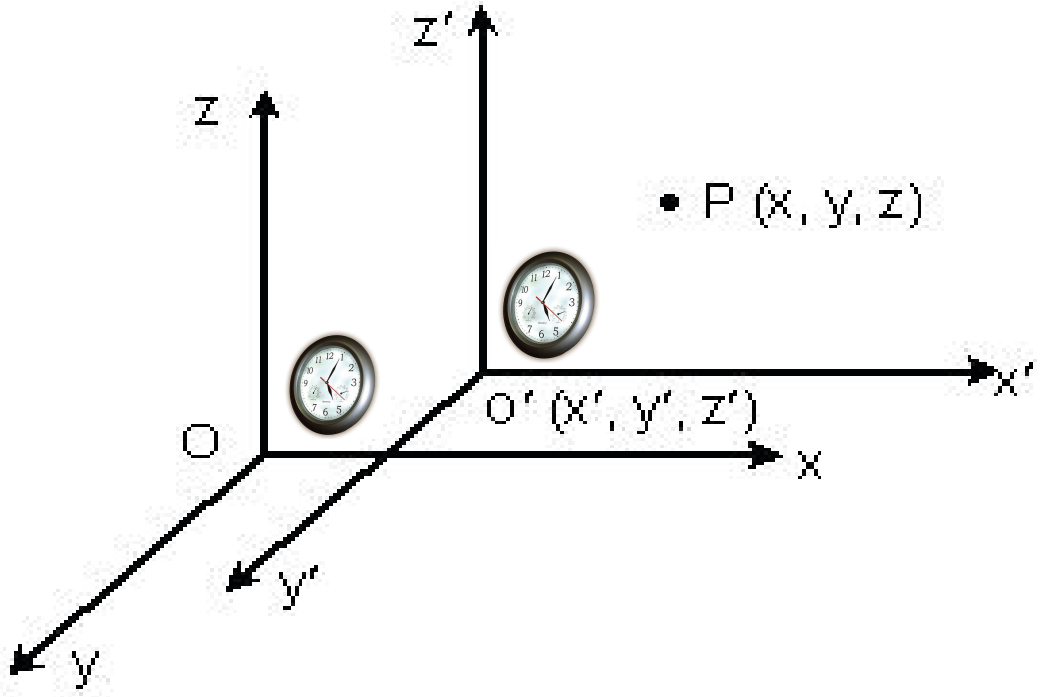


Figure 9: Reference frame – a coordinate system and a clock.

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

$$c = 2.99792458 \cdot 10^8 \ \text{ m/sec.}$$

This is a fundamental physical constant. Since this speed is so high, in our everyday life the classical mechanics is a good approximation.

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

- Time delay measured by a moving clock
- Lorentz contraction of the length of a moving body
- Abberation of light

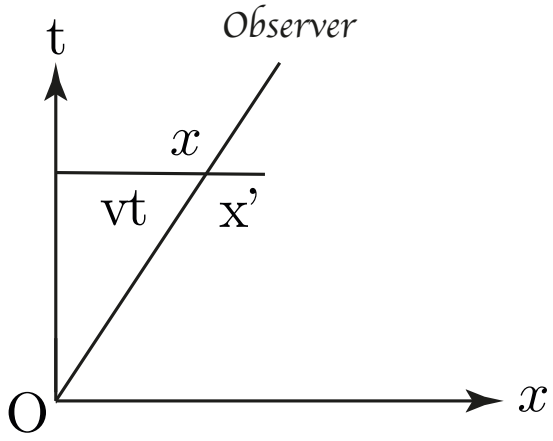


Figure 10: Galilean boost. The inclined line represents the trajectory of the origin of the reference frame M' which moves with velocity v in the x-direction with respect to the reference frame M. An event which happens in M at the position x at time t occurs at x' at time $t' = t$ in the moving frame M'. Hence, $x' = x - vt$.

4.3 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:

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 \,.$$

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

$$s_{12}^2 = c^2 t_{12}^2 - \ell_{12}^2 = c^2 t'^2_{12} - \ell'^2_{12} \, .$$

- *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.* $\ell'^2_{12} = 0$. 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.*

4.4 Lorentz group and its connected components

Lorentz transformations for a *group.* Before showing this, we give a general definition of a group.

A group G is a set of any nature which satisfies the following set of axioms:

1. For any two elements $g_1, g_2 \in G$ one can define their product $g_1 g_2$ which is also an element of G. The product is associative

$$(g_1 g_2) g_3 = g_1 (g_2 g_3) \,.$$

2. There exists a unit element $e \in G$ such that for any $g \in G$

$$ge = eg = g \,.$$

3. For any $g \in G$ there exists it's inverse $g^{-1} \in G$. that is

$$g g^{-1} = g^{-1} g = e \,.$$

In other words, all elements in a group are invertible.

An important class of groups constitute Lie groups. A *Lie group* is a group which is also a smooth manifold.[18] The Lorentz group is a Lie group.

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.[19] Suppose we have two such matrices

$$\Lambda_1^t \eta \Lambda_1 = \eta \,, \qquad \Lambda_2^t \eta \Lambda_2 = \eta \,,$$

[18] In other words, group elements of a Lie group can be continuously parametrized by a set of parameters.

[19] Would η be identity matrix, then the relation $\Lambda^t \Lambda = 1$ would define the group of orthogonal transformations.

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[20]

$$\Lambda^{-1} = \eta\Lambda^t\eta\,.$$

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.

Notice 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 $\Lambda_0^0 \leq -1$.

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 Figure 11. 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),
- 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

[20] 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.

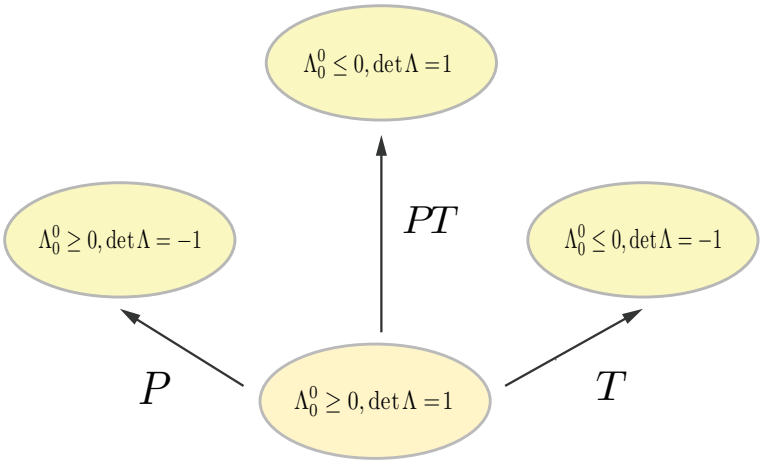


Figure 11: Four connected components of the Lorentz group. The component with $\Lambda_0^0 \leq 1$ and $\det\Lambda = 1$ is a subgroup of proper orthochronous transformations $\mathrm{SO}^+(1,3)$ (the restricted Lorentz group).

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[21], 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

$$\{1, P, T, PT\}$$

[21] For any two vectors x and y: $|(x,y)| \leq ||x||||y||$.

where P and T are the space inversion and time reversal operators:

$$P = \mathrm{diag}(1, -1, -1, -1)$$
$$T = \mathrm{diag}(-1, 1, 1, 1)$$

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[22] of the full Lorentz group with the same dimension (in this case, 6 dimensions).

4.5 Structure of Lorentz transformations

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

$$a^2 - v_2^2 = 1\,, \qquad a v_1^t - v_2^t S = 0\,, \qquad v_1 \otimes v_1^t - S^t S = -1\,.$$

The change $\Lambda \to \Lambda^t$ gives

$$a^2 - v_1^2 = 1\,, \qquad a v_2 - S^t v_1 = 0\,, \qquad v_2 \otimes v_2^t - S^t S = -1\,.$$

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

$$x^0 = a x'^0 - (\vec{v}_2 \vec{x}')\,,$$
$$\vec{x} = -\vec{v}_1 x'^0 + S^t \vec{x}'.$$

[22] A subgroup $N \subset G$ is called normal, if $gNg^{-1} \subset N$ for any $g \in G$.

In the moving coordinate system M', it's origin O' has coordinates $\vec{x'} = 0$, therefore, the formulae before takes the form

$$x^0 = ax'^0\,,$$
$$\vec{x} = -\vec{v}_1 x'^0\,.$$

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

$$(1 + v_1 \otimes v_1^t)_{ij} = \delta_{ij} + v_i v_j\,.$$

Consider for the moment another matrix

$$Q \equiv (1 + \alpha\, v_1 \otimes v_1^t)_{ij} = \delta_{ij} + \alpha\, v_i v_j\,,.$$

where α is an arbitrary number. Compute its squire

$$Q_{ij}^2 = (\delta_{ik} + \alpha\, v_i v_k)(\delta_{kj} + \alpha\, v_k v_j) = \delta_{ij} + \big(2\alpha + \alpha^2(a^2-1)\big)v_i v_j\,.$$

Thus, we see that if we subject the coefficient α to the condition

$$2\alpha + \alpha^2(a^2-1) = 1\,,$$

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_+. The relation $S^t S = Q_+^2$ can be written as $Q_+^{-1} S^t S Q_+^{-1} = (SQ_+^{-1})^t (SQ_+^{-1}) = 1$, since Q_+ is a symmetric matrix. Hence $R = SQ_+^{-1}$ is an orthogonal matrix, as $R^t R = 1$. Furthermore, since now $S = RQ_+$, 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 $v_2 = Rv_1$. To summarize, we have established that generic matrix Λ has the following structure

$$\Lambda = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{v^t}{c} \\ -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}R\frac{v}{c} & RQ_+ \end{pmatrix} .$$

We immediately see that this matrix factorizes into a product of the following matrices

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{v^t}{c} \\ -\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\frac{v}{c} & 1+\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}-1\right)\frac{v\otimes v^t}{v^2} \end{pmatrix} .$$

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 $x'^0 = ct'$,

$$\begin{aligned} t' &= \frac{t - \frac{(\vec{x}\vec{v})}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} , \\ \vec{x}' &= R\left[\vec{x} - \frac{\vec{v}t}{\sqrt{1-\frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{\vec{v}(\vec{v}\vec{x})}{v^2}\right] . \end{aligned} \tag{4.2}$$

These are Lorentz transformations[23] 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 $v << c$, 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

[23]Regrouping terms, the expression for x' can be also written in the following form

$$\vec{x}' = R\left[\frac{[\vec{v},[\vec{x},\vec{v}]]}{\sqrt{1-\frac{v^2}{c^2}}} - \vec{v}\frac{t - \frac{(\vec{v}\vec{x})}{v^2}}{\sqrt{1-\frac{v^2}{c^2}}}\right] .$$

unprimed and changing the sign of velocity $\vec{v}$ (*Check it!*). One obtains

$$t = \frac{t' + \frac{(\vec{x}'\vec{v})}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}\,,$$
$$\vec{x} = R\left[\vec{x}' + \frac{\vec{v}t'}{\sqrt{1-\frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{\vec{v}(\vec{v}\vec{x}')}{v^2}\right]. \qquad (4.3)$$

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

$$S^tS = Q_-^2 = Q_+^2$$

or

$$Q_-^{-1}S^tSQ_-^{-1} = Q_+^{-1}S^tSQ_+^{-1} = 1$$

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 can be understood by computing its determinant

$$\det(R_-^{-1}R_+) = \det(Q_-S^{-1}SQ_+^{-1}) = \frac{\det Q_-}{\det Q_+}\,.$$

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.

4.6 Addition of velocities

Suppose in the moving frame M' a particle is moving with velocity $\vec{u}$, that is $\vec{x}' = \vec{u}t$. We want to find its velocity in the stationary frame M. To this end, we take the inverse Lorentz transformations and substitute there $\vec{x}' = \vec{u}t$. We get

$$t = \frac{1 + \frac{(\vec{u}\vec{v})}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} t' \,,$$
$$\vec{x} = \left[\vec{u} + \frac{\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right) \frac{\vec{v}(\vec{v}\vec{u})}{v^2} \right] t' \,. \tag{4.4}$$

In the stationary frame the particle moves according to $\vec{x} = \vec{w}t$, where $\vec{w}$ is the velocity we are looking for. Thus,

$$\vec{w} = \frac{\vec{x}}{t} = \frac{\vec{u} + \frac{\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right) \frac{\vec{v}(\vec{v}\vec{u})}{v^2}}{\frac{1 + \frac{(\vec{u}\vec{v})}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}} \,.$$

This is a low for addition of velocities in the relativistic case. In the non-relativistic limit $c \to \infty$, it reduces to the Galilean law: $\vec{w} = \vec{u} + \vec{v}$.

4.7 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.[24]

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

$$K_y = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\,, \qquad K_z = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}\,.$$

[24]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.

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

$$\begin{aligned} &[K_x, K_y] = -iJ_z + \text{cyclic permutations} \\ &[J_x, K_x] = [J_y, K_y] = [J_z, K_z] = 0 \\ &[J_x, K_y] = iK_z + \text{cyclic permutations} \end{aligned} \tag{4.5}$$

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.

4.8 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

$$\mathrm{d}s^2 = \mathrm{d}x_\mu \mathrm{d}x^\mu = c^2\mathrm{d}t^2 - \left(\mathrm{d}x^i\right)^2 . \tag{4.6}$$

The action for a relativistic particle has the following form

$$S = -\alpha \int_a^b \sqrt{\mathrm{d}s^2} = -\alpha \int_a^b \mathrm{d}s \,.$$

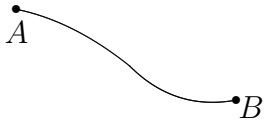


Figure 12: The simplest form of action is given by the length of the space-time interval between points A and B.

Rewriting (4.6), we get

$$\mathrm{d}s = \sqrt{\frac{\mathrm{d}x_\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\,\mathrm{d}t^2} = \sqrt{\frac{\mathrm{d}x_\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}}\;\mathrm{d}t\,. \tag{4.7}$$

Here we have used the convention $V_\mu V^\mu = \eta_{\mu\nu}V^\mu V^\nu$, where $\eta_{\mu\nu}$ is the Minkowski metric.

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = (c, \vec{v})\ , \qquad \mathrm{d}s = \sqrt{c^2 - \vec{v}^2} = c\sqrt{1 - \frac{\vec{v}^2}{c^2}}\,.$$

Therefore,

$$S = -\alpha c \int_{t_0}^{t_1} \sqrt{1 - \frac{\vec{v}^2}{c^2}}\mathrm{d}t\,,$$

where in non-relativistic physics we assume $\frac{\vec{v}^2}{c^2} \ll 1$. In general, $S = \int_{t_0}^{t_1} \mathcal{L}\,\mathrm{d}t$ where $\mathcal{L}$ is the so-called Lagrangian of the system, which in the non-relativistic limit is given by:

$$\mathcal{L} = -\alpha c\sqrt{1 - \frac{\vec{v}^2}{c^2}} \approx -\alpha c\left(1 - \frac{\vec{v}^2}{2c^2} + \cdots\right) \approx -\alpha c + \alpha\frac{\vec{v}^2}{2c}\,. \tag{4.8}$$

If we want to recover the usual form of the Lagrangian $\mathcal{L} = \text{Kin Energy} - V^{ext}$ for a free particle $V^{ext} = 0$ (hence $\mathcal{L} = \frac{1}{2}m\vec{v}^2$), we need to set $\alpha = mc$. When we do so, equation (4.8) turns into

$$\mathcal{L} = -mc^2 + \frac{1}{2}m\vec{v}^2\,.$$

Thus, one can rewrite $\mathcal{L}$ as

$$\mathcal{L} = -mc\sqrt{\dot{x}_\mu\dot{x}^\mu}\,.$$

When we use the canonical momentum p^μ defined as the derivative of $\mathcal{L}$ with respect to $\dot{x}^\mu$, we get

$$p^\mu = \frac{\partial\mathcal{L}}{\partial\dot{x}^\mu} = -mc\frac{\dot{x}_\mu}{\sqrt{\dot{x}_\nu\dot{x}^\nu}}\,.$$

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*.

The action is invariant under reparametrizations of τ:

$$\delta x^\mu = \xi(\tau)\partial_\tau x^\mu \quad \text{as long as} \quad \xi(\tau_0) = \xi(\tau_1) = 0\,.$$

Let us show this

$$\begin{aligned}\delta(\sqrt{\dot{x}_\mu \dot{x}^\mu}) &= \frac{1}{2\sqrt{\dot{x}_\mu \dot{x}^\mu}}(2\dot{x}^\nu \delta\dot{x}_\nu) = \frac{1}{\sqrt{\dot{x}_\mu \dot{x}^\mu}}\dot{x}^\nu \partial_\tau(\xi\dot{x}_\nu) = \\ &= \frac{1}{\sqrt{\dot{x}_\mu \dot{x}^\mu}}\left[\dot{x}^\nu \dot{x}_\nu \dot{\xi} + \xi\dot{x}^\nu \ddot{x}_\nu\right] = \frac{1}{\sqrt{\dot{x}_\mu \dot{x}^\mu}}\dot{x}^\nu \dot{x}_\nu \dot{\xi} + \xi\partial_\tau(\sqrt{\dot{x}_\mu \dot{x}^\mu}) \\ &= \sqrt{\dot{x}_\mu \dot{x}^\mu}\dot{\xi} + \xi\partial_\tau(\sqrt{\dot{x}_\mu \dot{x}^\mu}) = \partial_\tau(\xi\sqrt{\dot{x}_\mu \dot{x}^\mu})\,.\end{aligned}$$

Therefore, we arrive at

$$\delta S = -m\int_{\tau_0}^{\tau_1} \mathrm{d}\tau\, \partial_\tau(\xi\sqrt{\dot{x}_\mu \dot{x}^\mu}) = -m\left[\xi\sqrt{\dot{x}_\mu \dot{x}^\mu}\right]\Big|_{\tau=\tau_0}^{\tau=\tau_1} = 0\,,$$

i.e., the action is indeed invariant w.r.t. the *local* reparametrization transformations.

In the static (temporal) gauge $t = \tau$, the mass-shell condition takes the form

$$p_0 p^0 - \vec{p}^2 = m^2c^2 \quad \Longrightarrow \quad \frac{E^2}{c^2} - \vec{p}^2 = m^2c^2\,.$$

5. Classical Electrodynamics

5.1 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

$$A^\mu \to A_\mu = \eta_{\mu\nu}A^\nu = \left(\varphi(x)\,, -\vec{A}(x)\right)\,.$$

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, which has the form

$$S = -mc\int_a^b \mathrm{d}s - \frac{e}{c}\int A_\mu \mathrm{d}x^\mu\,.$$

Figure 13: 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.

Using Hamilton's principle, stating that particles follow paths that minimize their 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

$$0 = \delta S = -mc \int \frac{\mathrm{d}x_\mu}{\mathrm{d}s} d(\delta x^\mu) - \frac{e}{c} \int \left[(\delta A_\mu) \mathrm{d}x^\mu + A_\mu d(\delta x^\mu) \right] . \tag{5.1}$$

Using (4.7), the term δs in the first integral becomes $\delta \mathrm{d}s = \frac{\mathrm{d}x_\mu d\delta x^\mu}{\sqrt{\mathrm{d}x_\nu \mathrm{d}x^\nu}}$, whereas in the second integral we have simply used the product rule of differentiation. Let us consider for a moment the $U^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}s}$ term, which we will refer to as 4-velocity. The explicit form of U^μ is

$$U^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}s} = \frac{\mathrm{d}x^\mu}{c\sqrt{1 - \frac{\vec{v}^2}{c^2}}\mathrm{d}t} = \left(\frac{1}{\sqrt{1 - \frac{\vec{v}^2}{c^2}}}, \frac{\vec{v}}{c\sqrt{1 - \frac{\vec{v}^2}{c^2}}} \right) . \tag{5.2}$$

and it has an interesting property that

$$U_\mu U^\mu = \frac{\mathrm{d}x_\mu}{\mathrm{d}s} \frac{\mathrm{d}x^\mu}{\mathrm{d}s} = 1 .$$

Note that this result is only valid for the signature of the metric that we chose. If we were to invert the signature, the result would be -1 instead. Using the fact that $\delta A_\mu = A_\mu(x_\nu + \delta x_\nu) - A_\mu(x_\nu) = \partial_\nu A_\mu \delta x^\nu + \cdots$, we can rewrite equation (5.1) 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

$$mc \frac{\mathrm{d}U_\mu}{\mathrm{d}s} + \frac{e}{c} \left(\partial_\nu A_\mu - \partial_\mu A_\nu \right) U^\nu = 0 .$$

Identifying the *tensor $F_{\mu\nu}$ of the electromagnetic field*

$$\partial_\nu A_\mu - \partial_\mu A_\nu = F_{\nu\mu} = -F_{\mu\nu} , \tag{5.3}$$

we can write the equation of motion of the charge in the electromagnetic field as follows

$$mc\frac{\mathrm{d}U^\mu}{\mathrm{d}s} = \frac{e}{c}F^{\mu\nu}U_\nu\,. \tag{5.4}$$

This expression can also be written in a more suggestive form if we define the momentum $p^\mu = mcU^\mu$ (which is consistent with the requirement $p^2 = m^2c^2$ since $U^2 = 1$), so that one can express the acceleration term $\frac{\mathrm{d}U^\mu}{\mathrm{d}s} = \frac{d^2x^\mu}{\mathrm{d}s^2}$ as

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}s} = \frac{\mathrm{d}p^\mu}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}s} = \frac{e}{c}F^{\mu\nu}U_\nu\,, \tag{5.5}$$

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).*

The tensor of the electromagnetic field can be then written as

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -H_z & H_y \\ -E_y & H_z & 0 & -H_x \\ -E_z & -H_y & H_x & 0 \end{pmatrix} \tag{5.6}$$

and, therefore,

$$F^{\mu\nu} = \eta^{\mu\sigma}\eta^{\nu\rho}F_{\sigma\rho} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -H_z & H_y \\ E_y & H_z & 0 & -H_x \\ E_z & -H_y & H_x & 0 \end{pmatrix}\,, \tag{5.7}$$

where we have defined the F_{0i} components to be the electric fields and the F_{ij} components to be magnetic fields. Equation (5.3) leads to the following relation between electric and magnetic fields and the corresponding components of the 4-potential

$$\vec{E} = -\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t} \qquad \text{and} \qquad \vec{H} = \mathrm{rot}\,\vec{A}\,. \tag{5.8}$$

For reader's convenience we also present the relationship between the electromagnetic tensor and its components via indices

$$\varphi = A^0, \quad \vec{A}_i = A^i, \quad E_i = F^{i0} = -F^{0i} = F_{0i}, \quad F^{ik} = -\epsilon_{ikl}H_l\,, \quad H_i = -\frac{1}{2}\epsilon_{ikl}F^{kl}\,.$$

5.2 Lorentz transformations of the 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

$$\Lambda = \begin{pmatrix} a & -\frac{a}{c}v^t \\ -\frac{a}{c}v & \Lambda_{ij} \end{pmatrix}\,, \tag{5.9}$$

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

$$\varphi' = a\varphi - \frac{a}{c}(v\cdot A) = \frac{\varphi - \frac{(A\cdot v)}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}\,,$$

$$\vec{A}' = -\frac{a}{c}\varphi v + \vec{A} + d\vec{v}(\vec{v}\cdot\vec{A}) = \vec{A} - \frac{\varphi\frac{\vec{v}}{c}}{\sqrt{1-\frac{v^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right)\frac{\vec{v}(\vec{v}\cdot\vec{A})}{v^2}\,.$$

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

$$E'_i = aE_i - \frac{a-1}{v^2}v_i(\vec{v}\cdot\vec{E}) + \frac{a}{c}[\vec{v},\vec{H}]_i\,.$$

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:

$$\begin{aligned} H'_i = \frac{1}{2}\epsilon_{ijk}\Big(-\frac{a}{c}v_j(\delta_{nk} + dv_n v_k) + \frac{a}{c}v_k(\delta_{nj} + dv_n v_j)\Big)E_n \\ - \frac{1}{2}\epsilon_{ijk}\Big((\delta_{mj} + dv_m v_j)(\delta_{nk} + v_n v_k)\Big)F^{mn}\,. \end{aligned}$$

Making use of the formula for the pairing of two ϵ-tensors in the second line of the last formula, we arrive at

$$\begin{aligned} H'_i &= -\frac{a}{c}\epsilon_{ijn}v_j E_n + H_i - \frac{d}{2}\Big((\delta_{in}\delta_{ks} - \delta_{is}\delta_{nk})v_k v_n H_s + (\delta_{im}\delta_{js} - \delta_{is}\delta_{jm})v_j v_m H_s\Big) \\ &= H_i - \frac{a}{c}\epsilon_{ijn}v_j E_n + d(H_i v^2 - v_i(\vec{v}\cdot\vec{H})) \end{aligned} \tag{5.10}$$

The final expression is

$$H'_i = aH_i - \frac{a-1}{v^2}v_i(\vec{v}\cdot\vec{H}) - \frac{a}{c}[\vec{v},\vec{E}]_i\,.$$

We summarize the transformation formulae

$$\begin{aligned} \varphi' &= a\varphi - \frac{a}{c}(\vec{v}\cdot\vec{A})\,, \\ \vec{A}' &= \vec{A} - \frac{a}{c}\varphi\,\vec{v} + \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{A}) \end{aligned} \tag{5.11}$$

and

$$\begin{aligned} \vec{E}' &= a\vec{E} - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{E}) + \frac{a}{c}[\vec{v},\vec{H}]\,, \\ \vec{H}' &= a\vec{H} - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{H}) - \frac{a}{c}[\vec{v},\vec{E}]\,. \end{aligned} \tag{5.12}$$

The inverse transformations are

$$\begin{aligned} \varphi &= a\varphi' + \frac{a}{c}(\vec{v}\cdot\vec{A}')\,, \\ \vec{A} &= \vec{A}' + \frac{a}{c}\varphi'\,\vec{v} + \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{A}') \end{aligned} \tag{5.13}$$

and

$$\begin{aligned} \vec{E} &= a\vec{E}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{E}') - \frac{a}{c}[\vec{v},\vec{H}']\,, \\ \vec{H} &= a\vec{H}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{H}') + \frac{a}{c}[\vec{v},\vec{E}']\,. \end{aligned} \tag{5.14}$$

This completes our discussion of the transformation properties of the 4-potential and the electromagnetic field under Lorentz transformations.

5.3 Momentum and energy of a 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

$$S = \int L\,\mathrm{d}t = \int \Big[-mc^2\sqrt{1-\frac{v^2}{c^2}}\,\mathrm{d}t - \frac{e}{c}A_0\mathrm{d}x^0 - \frac{e}{c}A_i\mathrm{d}x^i\Big]\,,$$

i.e. the Lagrangian is

$$L = -mc^2\sqrt{1-\frac{v^2}{c^2}} + \frac{e}{c}\vec{A}\vec{v} - e\,\varphi$$

The momentum is

$$\vec{P} = \frac{\partial L}{\partial \vec{v}} = \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{e}{c}\vec{A} \tag{5.15}$$

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.(5.15) the velocity $\vec{v}$ in terms of the canonical momentum $\vec{P}$, we find that

$$H = \sqrt{m^2c^4 + c^2\Big(\vec{P} - \frac{e}{c}\vec{A}\Big)^2} + e\,\varphi\,.$$

We stress that such an expression for the hamiltonian arises only due to our choice of the static gauge.

5.4 Maxwell's equations and 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[25]. To understand this issue, we may decide to change the vector potential in the following way

$$A_\mu \to A_\mu - \partial_\mu\chi\,, \tag{5.16}$$

which can be rewritten in a less abstract form of space and time components separately:

$$\vec{A} \to \vec{A} + \vec{\nabla}\chi \quad \text{and} \quad \varphi \to \varphi - \frac{1}{c}\frac{\partial\chi}{\partial t}\,. \tag{5.17}$$

[25]This symmetry extends to many other physical theories besides electrodynamics.

These transformations are referred to as *the gauge transformations*. Let us see what effect they have on the tensor of the electromagnetic field:

$$\begin{aligned}\delta F_{\mu\nu} &= \partial_\mu \left(A_\nu + \partial_\nu \chi\right) - \partial_\nu \left(A_\mu + \partial_\mu \chi\right) - F_{\mu\nu} \\ &= \partial_\mu \partial_\nu \chi - \partial_\nu \partial_\mu \chi = 0\,. \end{aligned} \tag{5.18}$$

Thus, the transformation (5.16) does not change the form of the electromagnetic field tensor. For this reason electromagnetism is a gauge invariant theory!

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

$$F_{\mu\nu} F^{\mu\nu} = \text{inv}\,; \tag{5.19}$$

$$\varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = \text{inv}\,. \tag{5.20}$$

Let us inspect the gauge invariance of the electric and magnetic fields $\vec{E}$ and $\vec{H}$, which from the form and their in terms of the electromagnetic field tensor components can be expressed in terms of the vector potential as

$$\vec{E} = -\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \qquad \text{and} \qquad \vec{H} = \text{rot}\,\vec{A}\,. \tag{5.21}$$

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 $\text{rot}\,\text{grad}\chi = 0$. We look back at the expression for the Lorentz force and try to write it in terms of electric and magnetic fields. Rearranging (5.5), we get

$$\begin{aligned}\frac{\mathrm{d}p^i}{\mathrm{d}t} &= \left(\frac{e}{c}F^{i0}U_0 + \frac{e}{c}F^{ij}U_j\right)\frac{\mathrm{d}s}{\mathrm{d}t} = \\ &= \left(\frac{e}{c}E^i \frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}} - \frac{e}{c}F^{ij}\frac{v^j}{c\sqrt{1-\frac{\vec{v}^2}{c^2}}}\right) c\sqrt{1-\frac{\vec{v}^2}{c^2}}\,. \end{aligned} \tag{5.22}$$

Here we used the fact that $F^{i0} = -F^{0i} = -(-E^i) = E^i$ and $U_j = -\frac{v_j}{c\sqrt{1-\frac{v^2}{c^2}}}$. We can thus rewrite the expression for the Lorentz force as

$$\frac{\mathrm{d}p^i}{\mathrm{d}t} = eE^i + \frac{e}{c}\left[\vec{v}, \vec{H}\right]^i\,. \tag{5.23}$$

Concerning this result, it is interesting to point out that[26]

$$\frac{\mathrm{d}E_{\mathrm{kin}}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} = \vec{v}\cdot\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = e\big(\vec{E}\cdot\vec{v}\big)\,.$$

This is the work of the electromagnetic field on the charge. Hence, the magnetic field does not play any role is kinetic energy changes, but rather only affects the direction of the movement of the particle! Using basic vector calculus and the definitions of the electric and magnetic fields (5.21), the first two Maxwell's equations are attained

$$\mathrm{div}\,\vec{H} = \mathrm{div}\,\mathrm{rot}\,\vec{A} = 0 \Rightarrow \mathrm{div}\,\vec{H} = 0\,; \tag{5.24}$$

$$\mathrm{rot}\,\vec{E} = -\frac{1}{c}\mathrm{rot}\,\mathrm{grad}\,\varphi - \frac{1}{c}\frac{\partial}{\partial t}\mathrm{rot}\,\vec{A} \Rightarrow \mathrm{rot}\,\vec{E} = -\frac{1}{c}\frac{\partial\vec{H}}{\partial t}\,. \tag{5.25}$$

Equation (5.24) is known as the *no magnetic monopole rule* and (5.25) is referred to as *Faraday's law*, which we have already encountered in the previous section, but then the right hand side was suppressed due to time independence requirement. Together these two equations constitute the first pair of Maxwell's equations. Notice that these are 4 equations in total, as Faraday's law represents three equations - one for every space direction. Additionally, notice that Faraday's law is consistent with electrostatics; if the magnetic field is time independent then the right hand side of the equation is equal 0, which is exactly equation (2.11). These equations also have an integral form. Integrating (5.25) over a surface S with the boundary ∂S and using Stokes' theorem, we arrive at

$$\oint\limits_S \mathrm{rot}\,\vec{E}\cdot\mathrm{d}\vec{S} = \oint\limits_{\partial S}\vec{E}\cdot\mathrm{d}\vec{l} = -\frac{1}{c}\frac{\partial}{\partial t}\oint\limits_S \vec{H}\mathrm{d}\vec{S}\,. \tag{5.26}$$

For eq.(5.24) one integrates both sides over the volume and uses the Gauss-Ostrogradsky theorem to arrive at

$$\int\limits_V \mathrm{div}\,\vec{H}\mathrm{d}V = \oint\limits_{\partial V}\vec{H}\cdot\mathrm{d}\vec{S} = 0\,. \tag{5.27}$$

5.5 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

$$S = S_{particles} + S_{int} + S_{field}\,,$$

[26] We have

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \frac{m\dot{\vec{v}}}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{m\vec{v}}{\left(1-\frac{v^2}{c^2}\right)^{3/2}}\frac{(\vec{v}\cdot\dot{\vec{v}})}{c^2}$$

so that

$$\vec{v}\cdot\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = \frac{m(\vec{v}\cdot\dot{\vec{v}})}{\left(1-\frac{v^2}{c^2}\right)^{3/2}} = \frac{\mathrm{d}E_{\mathrm{kin}}}{\mathrm{d}t}\,.$$

where we have added a new term S_{field}, which represents the interaction between the particles and the field that they have produced themselves. We will write it as

$$S_{field} \sim \int F_{\mu\nu}F^{\mu\nu}\,\mathrm{d}^4x = \int F_{\mu\nu}F^{\mu\nu}c\mathrm{d}t\,\mathrm{d}^3x\,.$$

Then adding the proportionality constants the total action is written as

$$S = -mc\int \mathrm{d}s - \frac{e}{c}\int A_\mu \mathrm{d}x^\mu - \frac{1}{16\pi c}\int F_{\mu\nu}F^{\mu\nu}\,c\mathrm{d}t\,\mathrm{d}^3x\,,$$

where we have adopted the Gauss system of units, i.e. $\mu_0 = 4\pi$ and $\varepsilon_0 = \frac{1}{4\pi}$. Note that we can rewrite the second term as

$$\begin{aligned}\frac{e}{c}\int A_\mu \mathrm{d}x^\mu &= \frac{1}{c}\int \rho A_\mu \mathrm{d}x^\mu \mathrm{d}V = \frac{1}{c}\int \rho A_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\mathrm{d}V\mathrm{d}t \\ &= \frac{1}{c}\int j^\mu A_\mu \mathrm{d}V\mathrm{d}t = \frac{1}{c^2}\int j^\mu A_\mu \mathrm{d}^4x\,,\end{aligned} \tag{5.28}$$

where in the second line we have introduced, the current $j^i = \rho\frac{\mathrm{d}x^i}{\mathrm{d}t} = (\rho c, \rho\vec{v})$. Including this, we can now write the action of the moving test charge as

$$S = -mc\int \mathrm{d}s - \frac{1}{c^2}\int j^\alpha A_\alpha\,\mathrm{d}^4x - \frac{1}{16\pi c}\int F_{\mu\nu}F^{\mu\nu}c\mathrm{d}t\mathrm{d}^3x\,.$$

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

$$\begin{aligned}\delta S &= -\frac{1}{c^2}\int j^\alpha \delta A_\alpha\,\mathrm{d}^4x - \frac{1}{8\pi c}\int F_{\mu\nu}\delta F^{\mu\nu}c\mathrm{d}t\mathrm{d}^3x \\ &= -\frac{1}{c}\left[\frac{1}{c}\int j^\alpha \delta A_\alpha\,\mathrm{d}^4x + \frac{1}{4\pi}\int \frac{\partial F^{\mu\nu}}{\partial x^\nu}\delta A_\mu c\mathrm{d}t\mathrm{d}^3x\right]\,,\end{aligned} \tag{5.29}$$

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.(5.29), is equivalent

$$-\frac{1}{c^2}j^\mu - \frac{1}{4\pi c}\partial_\nu F^{\mu\nu} = 0\,.$$

Upon rearrangement, this gives us the second pair of Maxwell's equations

$$\frac{\partial F^{\mu\nu}}{\partial x^\nu} = -\frac{4\pi}{c}j^\mu\,.$$

Notice that for vanishing currents, these equation *resemble* the first pair of Maxwell's equations, when currents are to vanish (i.e. $j^\mu = 0$). Below we dwell more on this point.

Identifying the respective components of the electromagnetic tensor we can rewrite the second pair of Maxwell's equations in a more familiar form

$$\text{rot}\,\vec{H} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial \vec{E}}{\partial t} \qquad \text{and} \qquad \text{div}\,\vec{E} = 4\pi\rho\,, \tag{5.30}$$

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 Coulomb's law, which we have already found before, but using a different principle. Finally, we notice that the covariant conservation of the current $\frac{\partial j^\mu}{\partial x^\mu} = 0$ is equivalent to the continuity equation

$$\frac{\partial \rho}{\partial t} + \text{div}\vec{j} = 0\,.$$

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:

$$\mathrm{d}F = \frac{\partial F_{\mu\nu}}{\partial x^\sigma} + \frac{\partial F_{\nu\sigma}}{\partial x^\mu} + \frac{\partial F_{\sigma\mu}}{\partial x^\nu} = 0\,. \tag{5.31}$$

With a change of indices, this takes the form

$$\varepsilon^{\mu\nu\sigma\rho}\frac{\partial F_{\nu\sigma}}{\partial x^\rho} = 0\,, \tag{5.32}$$

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 tensor*

$$F^{*\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}\,. \tag{5.33}$$

Then we can rewrite equation (5.32) as

$$\frac{\partial F^{*\mu\nu}}{\partial x^\nu} = 0\,. \tag{5.34}$$

Omitting the currents in the second pair, the first and second pair of Maxwell's equations are similar. Indeed, we have

$$\frac{\partial F^{*\mu\nu}}{\partial x^\mu} = 0\,,$$
$$\frac{\partial F^{\mu\nu}}{\partial x^\mu} = 0\,.$$

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 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[27]. The classical description of particles does not allow to construct such quantities from dynamical variables associated to particle.

5.6 Electromagnetic waves

When the electric charge source and current terms are absent, we obtain the electromagnetic wave solutions. In this case the Maxwell equations reduce to

$$\begin{aligned} \mathrm{rot}\vec{E} &= -\frac{1}{c}\frac{\partial \vec{H}}{\partial t}\,, \qquad \mathrm{div}\vec{E} = 0\,, \\ \mathrm{rot}\vec{H} &= \frac{1}{c}\frac{\partial \vec{E}}{\partial t}\,, \qquad \mathrm{div}\vec{H} = 0\,. \end{aligned}$$

These equations can have non-zero solutions meaning that the electromagnetic fields can exist without any charges or currents. *Electromagnetic fields, which exist in the absence of any charges, are called electromagnetic waves.* Starting with the definitions of the electric and magnetic fields given in terms of the vector potential in equation (5.21), one can choose a gauge, i.e. fix A^μ, which will simplify the mathematical expressions as well as the calculations, we will be dealing with. The reason why we are allowed to make this choice is that gauge symmetry transforms one solution into another, both solutions being physically equivalent[28]. By making a gauge choice one breaks the gauge symmetry. This removes the excessive, unphysical degrees of freedom, which make two physically equivalent solutions to the equations of motion appear different. Obviously the simplicity of these equations and their solutions drastically depends on the gauge choice.

One of the convenient gauge choices involves setting $\partial_\mu A^\mu = 0$, which is the covariant gauge choice known as the Lorenz gauge[29]. 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 $\varphi = 0$. We will return back to the comparison of these gauge choices later.

[27] A physical quantity that behaves like a scalar, only it changes sign under parity inversion e.g. an improper rotation.

[28] Both solutions belong the same *gauge orbit*.

[29] 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.

To see the process of gauge fixing and how we can use it to simplify the equations of motion, consider the gauge transformations

$$\begin{aligned}\vec{A} &\to \vec{A} + \vec{\nabla} f\,,\\ \varphi &\to \varphi - \frac{1}{c}\frac{\partial f}{\partial t}\,.\end{aligned}$$

If f does not depend on t, φ will not change, however $\vec{A}$ will. On the other hand, $\mathrm{div}\vec{A}$ does not depend on t by the Maxwell equations[30]. Thus, in this gauge, equations (5.21) become

$$\begin{aligned}\vec{E} &= -\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t}\,,\\ \vec{H} &= \mathrm{rot}\vec{A}\,.\end{aligned}$$

Plugging this into (5.30), our Maxwell's equation describing the curl of the magnetic field, we obtain

$$\begin{aligned}\mathrm{rot}\vec{H} \;&=\; \mathrm{rot}\,\mathrm{rot}\vec{A} = \frac{1}{c}\frac{\partial}{\partial t}\left(\frac{-1}{c}\frac{\partial \vec{A}}{\partial t}\right) = -\frac{1}{c}\frac{\partial^2 \vec{A}}{\partial t^2}\,,\\ \Rightarrow\; &-\Delta\vec{A} + \mathrm{grad}\,\mathrm{div}\vec{A} = \frac{-1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2}\,.\end{aligned}$$

In this gauge we can choose f, such that the term involving the divergence of $\vec{A}$ disappears. The equation that remains is known as *d'Alembert's equation (or 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. only x-dependence), 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 $\eta = t + \frac{x}{c} \Rightarrow \frac{\partial^2 f}{\partial\xi\partial\eta} = 0$. Hence, the solution to the equation is

$$f = f(\xi) + f(\eta)\,.$$

[30] Under the gauge transformation with the time-independent function f we have $\mathrm{div}\vec{A} \to \mathrm{div}\vec{A} + \mathrm{div}\nabla f$, therefore, the function f should be determined from the Poisson equation $\Delta f = -\mathrm{div}\vec{A}$.

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 return to the issue of the Coulomb versus the Lorenz gauge choice, and first consider the latter. The Lorenz gauge condition reads as follows

$$0 = \frac{\partial A^\mu}{\partial x^\mu} = \mathrm{div}\vec{A} + \frac{1}{c}\frac{\partial \varphi}{\partial t} .$$

We see that under gauge transformations the Lorenz gauge condition transforms as

$$\partial_\mu \left(A^\mu + \partial^\mu \chi\right) = \frac{\partial A^\mu}{\partial x^\mu} + \partial_\mu \partial^\mu \chi$$

and it remains unchanged provided $\partial_\mu \partial^\mu \chi = 0$. Thus, the Lorenz gauge does not kill the gauge freedom completely. We still have a possibility to perform gauge transformations of the special type $\partial_\mu \partial^\mu \chi = 0$. Hence there will be still an excessive number of solutions that are physically equivalent and transform into each other under gauge transformations involving harmonic functions.

This problem is fixed with the introduction of the complete gauge choice. Starting over, one can always fix $\varphi = 0$ by choosing a suitable function $\chi(\vec{x}, t)$, i.e. a function such that $\varphi = \frac{1}{c}\frac{\partial \chi}{\partial t}$. Under the gauge transformations we have

$$\varphi \to \varphi - \frac{1}{c}\frac{\partial \chi}{\partial t} \quad \Rightarrow \quad \varphi = 0 .$$

Transforming the new $\varphi = 0$ with a new, only space-dependent function $\tilde{\chi}(x, y, z)$, we obtain[31]

$$0 = \varphi \to \varphi - \frac{1}{c}\frac{\partial \tilde{\chi}}{\partial t} = 0 \quad \text{and} \quad \vec{A} \to \vec{A} + \nabla \tilde{\chi} .$$

Since $\vec{E} = -\vec{\nabla}\varphi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}$ and $\varphi = 0$, we find

$$\mathrm{div}\ \vec{E} = -\frac{1}{c}\frac{\partial}{\partial t}\mathrm{div}\vec{A} \quad \text{and} \quad \mathrm{div}\vec{E} = 0 ,$$

where the right hand side has to be equal to zero from our original assumption - lack of sources of electromagnetic fields. From the above equation we can infer that $\frac{\partial}{\partial t}\,\mathrm{div}\,\vec{A} = 0$. We can use yet another gauge freedom to set the space-dependent and time-independent $\tilde{\chi}$, such that $\mathrm{div}\,\vec{A} = -\mathrm{div}\,\vec{\nabla}\tilde{\chi}$, which means that we have reached the Coulomb gauge

$$\mathrm{div}\,\vec{A} \to \mathrm{div}\,\vec{A} + \mathrm{div}\,\vec{\nabla}\tilde{\chi} = 0 .$$

[31] Note that $\frac{\partial \tilde{\chi}}{\partial t} = 0$.

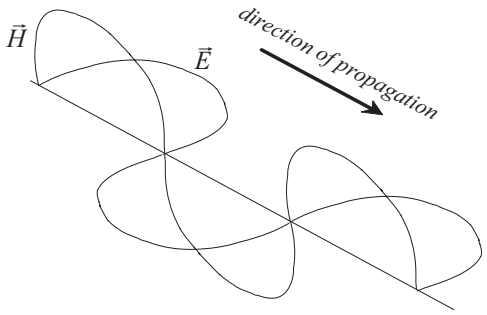


Figure 14: Oscillations of the electric and magnetic fields in electromagnetic wave.

Having fixed the gauge, 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. 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 leaves the wave equation in the form

$$\begin{aligned}\frac{\partial^2 A_x}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 A_x}{\partial t^2} &= 0\\ -\frac{1}{c^2}\frac{\partial^2 A_x}{\partial t^2} = 0 \Rightarrow \frac{\partial^2 A_x}{\partial t^2} = 0 &\Rightarrow \frac{\partial A_x}{\partial t} = const.\end{aligned}$$

Since we are not interested in a constant electric field E_x, we need to fix $A_x = 0$. Since $\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t}$ and $\vec{H} = \text{rot}\,\vec{A}$, then

$$\vec{H} = \left[\vec{\nabla}_{\left(t-\frac{x}{c}\right)}, \vec{A}\right] = -\frac{1}{c}\left[\vec{n}, \frac{\partial}{\partial t}\vec{A}\right] = \left[\vec{n}, \vec{E}\right],$$

where $\left[\vec{A}, \vec{B}\right]$ denotes the cross-product of two vectors. From the definition of the cross product one can see that 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.*

Electromagnetic waves are known to carry energy; we can define the *energy flux* to be

$$\vec{S} = \frac{c}{4\pi}\left[\vec{E}, \vec{H}\right] = \frac{c}{4\pi}\left[\vec{E}, \left[\vec{n}, \vec{E}\right]\right].$$

Since $\big[\vec{a},\big[\vec{b},\vec{c}\big]\big] = \vec{b}\big(\vec{a},\vec{c}\big) - \vec{c}\big(\vec{a},\vec{b}\big)$, 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).

5.7 Hamiltonian formulation of electrodynamics

To obtain the Hamiltonian formulation of classical electrodynamics (without sources), we start for the action for electromagnetic field (we put $c = 1$)

$$S = -\frac{1}{16\pi}\int \mathrm{d}^4x F_{\mu\nu}F^{\mu\nu}$$

and rewrite it in the first order formalism. To do so, we first compute the canonical momentum conjugate to A^μ. We have

$$p_\mu(x) = \frac{\delta L}{\delta \dot{A}^\mu(x)} = -\frac{1}{4\pi}\int \mathrm{d}^3y\, F^\rho{}_\nu(y)\frac{\delta(\partial_\rho A^\nu(y))}{\delta(\partial_t A^\mu(x))} = -\frac{1}{4\pi}F^0{}_\mu(x) = -\frac{1}{4\pi}F_{0\mu}(x)\,.$$

We see that we have a primary constraint[32]

$$p_0 = 0\,,$$

i.e. the momentum conjugate to A_0 vanishes. This is a straightforward consequence of the fact that the Lagrangian does not contain the time derivative of A_0. In other

[32]Thus, we are dealing with a singular Lagrangian system.

words, the velocity for A_0 is absent so that A_0 is not a dynamical field! As to the components of the canonical momentum, they simply coincide with the electric field:

$$p_i(x) = -\frac{1}{4\pi}F_{0i}(x) = -\frac{1}{4\pi}(\partial_0 A_i - \partial_i A_0) = -\frac{1}{4\pi}E_i\,.$$

This relation allows us to find the velocities $\dot{A}_i$ via the electric field

$$\dot{A}_i = E_i + \partial_i A_0\,.$$

Now we write the Lagrangian in the Hamiltonian form

$$L = \underbrace{\int \mathrm{d}^3x\, p_i(x)\dot{A}^i(x)}_{\text{symplectic structure}} - \text{rest}$$

or

$$\text{rest} = \int \mathrm{d}^3x\, p_i(x)\dot{A}^i(x) - L = \int \mathrm{d}^3x\, p_i(x)\dot{A}^i(x) + \frac{1}{16\pi}\int \mathrm{d}^3x\,(-2F_{0i}F_{0i} + F_{ij}F_{ij})\,.$$

The rest must be reexpressed via canonical coordinates and momenta (electric field), *i.e.* all the velocities must be excluded in favor of the canonical momenta. We have

$$\text{rest} = \frac{1}{4\pi}\int \mathrm{d}^3x\, E_i(E_i + \partial_i A_0) + \frac{1}{16\pi}\int \mathrm{d}^3x\,(-2E_i^2 + F_{ij}F_{ij})\,.$$

We also notice that $\vec{H} = \mathrm{rot}\vec{A}$ which can be also written as

$$H_i = -\frac{1}{2}\epsilon_{ijk}F_{jk}\,.$$

Since we have

$$\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}\,,$$

we see that

$$H_i^2 = \frac{1}{4}\epsilon_{ijk}\epsilon_{imn}F_{jk}F_{mn} = \frac{1}{2}F_{ij}F_{ij}\,.$$

Thus, we arrive at

$$\text{rest} = \frac{1}{8\pi}\int \mathrm{d}^3x\left(E_i^2 + H_i^2 - 2A_0\partial_i E_i\right).$$

Thus, the original Lagrangian takes the following form

$$L = \underbrace{\frac{1}{4\pi}\int \mathrm{d}^3x\, E_i\dot{A}_i}_{\text{symplectic structure}} - \underbrace{\frac{1}{8\pi}\int \mathrm{d}^3x\left(E_i^2 + H_i^2\right)}_{\text{Hamiltonian}} - \underbrace{\frac{1}{4\pi}\int \mathrm{d}^3x\, A_0\partial_i E_i}_{\text{Constraint}}\,.$$

Here

$$H = \frac{1}{8\pi}\int \mathrm{d}^3x\left(E_i^2 + H_i^2\right)$$

is the Hamiltonian of the electromagnetic field. This is nothing else as the energy of the electromagnetic field! The first term defines the Poisson bracket

$$\{E_i(\vec{x}), A_j(\vec{y})\} = \frac{1}{4\pi}\delta_{ij}\delta(\vec{x}-\vec{y})\,.$$

The last term in the Lagrangian is a constraint. Indeed, varying the Lagrangian with respect to A_0 we find the following constraint:

$$C(x) \equiv \partial_i E_i(x) = 0 \quad \Longrightarrow \quad \mathrm{div}\vec{E} = 0\,,$$

which is nothing else as the Gauss law. As an exercise, check that

$$\dot{C} = \{H, C(x)\} = 0\,,$$

that is the constraint is preserved in time. Also, one can easily see that

$$\{C(x), C(y)\} = 0.$$

We can also verify that the Lagrangian (written in the Hamiltonian form) is invariant with respect to gauge transformations

$$\begin{aligned} A_i &\to A_i + \partial_i\chi \\ A_0 = \varphi &\to A_0 - \dot{\chi}. \end{aligned}$$

Under the gauge transformations we find

$$\delta L = \frac{1}{4\pi}\int \mathrm{d}^3x\, E_i\partial_i\dot{\chi} + \frac{1}{4\pi}\int \mathrm{d}^3x\,\dot{\chi}\partial_i E_i\,.$$

After integrating by parts ∂_i we obtain $\delta L = 0$.

Concluding this chapter, we will list the gauge conditions usually used in the literature

$$\begin{aligned} \partial_\mu A^\mu &= 0 \qquad \text{Lorenz gauge} \\ A_0 &= 0 \qquad \text{Hamilton gauge} \\ \partial_i A^i &= 0 \qquad \text{Coulomb gauge} \\ A_3 &= 0 \qquad \text{Axial gauge} \\ x_\mu A^\mu &= 0 \qquad \text{Fixed point gauge} \end{aligned}$$

The last gauge has been introduced by Fock. It is easy to check that the potential

$$A_\mu(x) = \int_0^1 \lambda \mathrm{d}\lambda\; x^\nu F_{\mu\nu}(\lambda x)$$

satisfies the gauge condition $x_\nu A^\mu = 0$ and that $\partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}$.

5.8 Solving 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[33]. Consider

$$\frac{\partial F^{\mu\nu}}{\partial x^\nu} = -\frac{4\pi}{c} j^\mu \,,$$
$$\frac{\partial}{\partial x^\nu}\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = \frac{\partial^2}{\partial x^\nu \partial x_\mu} A^\nu - \frac{\partial^2}{\partial x^\nu \partial x_\nu} A^\mu = -\frac{4\pi}{c} j^\mu \,.$$

Imposing the Lorenz condition

$$\frac{\partial A^\nu}{\partial x^\nu} = 0 \,,$$

we obtain from the previous equation

$$\frac{\partial^2}{\partial x^\nu \partial x_\nu} A^\mu = \frac{4\pi}{c} j^\mu \,.$$

The last equation can be split into two

$$\Delta \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} = -\frac{4\pi}{c}\vec{j} \,,$$
$$\Delta \varphi - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} = -4\pi\rho \,.$$

These wave equations represent a structure, which is already familiar to us, namely

$$\Delta\psi - \frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2} = -4\pi f\left(\vec{x}, t\right) \,. \tag{5.35}$$

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

$$\left(\Delta_x - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) G\left(\vec{x}, t; \vec{x}', t'\right) = -4\pi\delta\left(\vec{x} - \vec{x}'\right)\delta\left(t - t'\right) \,. \tag{5.36}$$

[33]The motion of the charges has to be strictly defined, i.e. even though the charges produce an electromagnetic field, their motion will not be influenced by the presence of external electromagnetic fields.

Note that $G(\vec{x},t;\vec{x}',t')$ 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 (5.35) reads

$$\psi(\vec{x},t) = \int G(\vec{x},t;\vec{x}',t')\, f(\vec{x}',t')\, \mathrm{d}^3x'\mathrm{d}t\,.$$

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 into the integral - two delta functions will emerge by virtue of (5.36), which upon integration will turn $f(\vec{x}',t')$ into $f(\vec{x},t)$. In what follows we will need the Fourier transforms of all the elements of equation (5.36)

$$\delta(\vec{x}-\vec{x}')\,\delta(t-t') = \frac{1}{(2\pi)^4}\int_{-\infty}^{\infty}\mathrm{d}^3k\int_{-\infty}^{\infty}\mathrm{d}\omega\; e^{i\vec{k}\cdot(\vec{x}-\vec{x}')}e^{-i\omega(t-t')}\,,$$
$$G(\vec{x},t;\vec{x}',t') = \int_{-\infty}^{\infty}\mathrm{d}^3k\int_{-\infty}^{\infty}\mathrm{d}\omega\; g\left(\vec{k},\omega\right)e^{i\vec{k}\cdot(\vec{x}-\vec{x}')-i\omega(t-t')}\,.$$

Plugging these into the equation, we obtain

$$g(\vec{k},\omega)\left(-k^2+\frac{\omega^2}{c^2}\right) = -4\pi\frac{1}{(2\pi)^4} = -\frac{1}{4\pi^3}\,,$$

which amounts to

$$g(\vec{k},\omega) = \frac{1}{4\pi^3}\frac{1}{\vec{k}^2-\frac{\omega^2}{c^2}}\,.$$

From this one can find an integral expression for $G(\vec{x},t;\vec{x}',t')$

$$G(\vec{x},t;\vec{x}',t') = \frac{1}{4\pi^3}\int_{-\infty}^{\infty}\mathrm{d}^3k\int_{-\infty}^{\infty}\mathrm{d}\omega\frac{e^{i\vec{k}\cdot(\vec{x}-\vec{x}')-i\omega(t-t')}}{\vec{k}^2-\frac{\omega^2}{c^2}}\,.$$

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 $\omega = \pm c\left|\vec{k}\right|$. 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 $G = 0$ if $t < t'$

Advanced Green function states $G = 0$ if $t > t'$

Notice that the difference of the two $G_{adv} - G_{ret}$, called the Pauli Green's function G_{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 $R = \left|\vec{R}\right|$. Then we have

$$e^{-i\omega(t-t')} \sim e^{\Im\omega\tau} ,$$

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 half plane when performing the analytic continuation. According to this prescription, the Green's function is specified as follows

$$G(\vec{x}, t; \vec{x}', t') = \frac{1}{4\pi^3} \int \mathrm{d}^3k \int \mathrm{d}\omega \frac{e^{i\vec{k}R - i\omega\tau}}{k^2 - \frac{1}{c^2}(\omega + i\varepsilon)^2} .$$

We can conveniently rewrite the previous statement, by making use of partial fractions

$$\begin{aligned} G(\vec{x}, t; \vec{x}', t') &= \\ &= \frac{1}{4\pi^3} \int_{-\infty}^{\infty} \mathrm{d}^3k \int_{-\infty}^{\infty} \mathrm{d}\omega e^{i\vec{k}R} \frac{c}{2k} \left[\frac{1}{ck - i\varepsilon - \omega} - \frac{1}{-ck - i\varepsilon - \omega} \right] e^{-i\omega\tau} . \end{aligned} \tag{5.37}$$

In the limit $\varepsilon \to 0$, using Cauchy's theorem[34], we find

$$
\begin{aligned}
G\left(\vec{x},t;\vec{x}',t'\right) &= \frac{1}{4\pi^3}\int_{-\infty}^{\infty} \mathrm{d}^3k e^{i\vec{k}\cdot\vec{R}} 2\pi i \frac{c}{2k}\left[e^{-ick\tau} - e^{ick\tau}\right] && (5.38)\\
&= \frac{c}{2\pi^2}\int_{-\infty}^{\infty} \mathrm{d}^3k \frac{e^{i\vec{k}\cdot\vec{R}}}{k}\sin(ck\tau)\\
&= \frac{c}{2\pi^2}\int_0^{\infty} \mathrm{d}k\, k\, \sin(ck\tau)\int_0^{\pi}\sin\theta \mathrm{d}\theta \int_0^{2\pi} d\varphi\, e^{ikR\cos\theta}\\
&= \frac{c}{\pi}\int_0^{\infty} \mathrm{d}k\, k\, \sin(ck\tau)\int_{-1}^{1} \mathrm{d}x\, e^{ikRx}\\
&= \frac{2c}{\pi R}\int_0^{\infty} \mathrm{d}k \sin(kR)\sin(ck\tau) && (5.39)\\
&= \frac{1}{\pi R}\int_{-\infty}^{\infty} \mathrm{d}\left(ck\right)\sin\left(\frac{\left(ck\right)R}{c}\right)\sin\left(\left(ck\right)\tau\right) && (5.40)\\
&= -\frac{1}{4\pi R}\int_{-\infty}^{\infty} \mathrm{d}x\left(e^{ix\frac{R}{c}} - e^{-ix\frac{R}{c}}\right)\left(e^{ix\tau} - e^{-ix\tau}\right) && (5.41)\\
&= \frac{1}{2\pi R}\int_{-\infty}^{\infty} \mathrm{d}x\left(e^{ix\left(\tau-\frac{R}{c}\right)} - e^{ix\left(\tau+\frac{R}{c}\right)}\right)\\
&= \frac{1}{R}\delta\left(\tau - \frac{R}{c}\right) - \frac{1}{R}\delta\left(\tau + \frac{R}{c}\right) && (5.42)\\
&= \frac{1}{R}\delta\left(\tau - \frac{R}{c}\right) && (5.43)
\end{aligned}
$$

Note that in the meantime we have used: partial fractions (5.37), the Cauchy theorem in (5.37-5.38), switched to spherical coordinates and integrated over the angles(5.39), substituted $ck = x$ (5.40), expanded the trigonometric functions in terms of their complex exponentials (5.41), and identified Fourier transforms of delta functions (5.42). 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

$$
G_{ret}\left(\vec{x},t;\vec{x}',t'\right) = \frac{\delta\left(t' + \frac{|\vec{x}-\vec{x}'|}{c} - t\right)}{|\vec{x}-\vec{x}'|}\,.
$$

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

[34]Cauchy integral formula reads

$$
f(a) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}\mathrm{d}\,z\,,
$$

where a function $f(z)$ is holomorphic inside the region surrounded by a contour C and integration is performed in counter-clockwize direction.

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

$$\begin{aligned} \varphi(\vec{x}, t) &= \int \frac{\delta\left(t' + \frac{|\vec{x}-\vec{x}'|}{c} - t\right)}{|\vec{x} - \vec{x}'|} \rho(\vec{x}', t') \, \mathrm{d}^3x' \mathrm{d}t' + \varphi_0 \\ &= \int \frac{\rho\left(\vec{x}', t - \frac{|\vec{x}-\vec{x}'|}{c}\right)}{|\vec{x} - \vec{x}'|} \mathrm{d}^3x' + \varphi_0 \,, \end{aligned} \tag{5.44}$$

$$\begin{aligned} \vec{A}(\vec{x}, t) &= \frac{1}{c} \int \frac{\delta\left(t' + \frac{|\vec{x}-\vec{x}'|}{c} - t\right)}{|\vec{x} - \vec{x}'|} \vec{j}(\vec{x}', t') \, \mathrm{d}^3x' \mathrm{d}t' + \vec{A}_0 \\ &= \frac{1}{c} \int \frac{\vec{j}\left(\vec{x}', t - \frac{|\vec{x}-\vec{x}'|}{c}\right)}{|\vec{x} - \vec{x}'|} \mathrm{d}^3x' + \vec{A}_0 \,, \end{aligned} \tag{5.45}$$

where φ_0 and $\vec{A}_0$ are the solutions of the homogeneous d'Alambert equations (those corresponding to the free electromagnetic field).

Note that for φ in the case of time-independent ρ and $\vec{j}$ we have

$$\varphi = \int \frac{\rho(\vec{x}')}{|\vec{x} - \vec{x}'|} \mathrm{d}^3x' .$$

This is just the electrostatic formula for the scalar potential. Moreover, if the current $\vec{j}$ is time-independent, we obtain

$$\vec{A}(\vec{x}) = \frac{1}{c} \int \frac{\vec{j}(\vec{x}')}{|\vec{x} - \vec{x}'|} \mathrm{d}^3x' .$$

This potential defines the following magnetic field

$$\vec{H} = \mathrm{rot}_x \vec{A} = \frac{1}{c} \int \left[\frac{\mathrm{rot}_x \vec{j}(\vec{x}')}{|\vec{x} - \vec{x}'|} + \vec{\nabla}_x \frac{1}{|\vec{x} - \vec{x}'|} \times \vec{j}(\vec{x}') \right] \mathrm{d}^3x' . \tag{5.46}$$

Note the use above of the following identity

$$\mathrm{rot}(\varphi \vec{a}) = \varphi \, \mathrm{rot} \, \vec{a} + \vec{\nabla}\varphi \times \vec{a} .$$

The first term in (5.46) vanishes, because curl is taken with respect to coordinates x, while the current $\vec{j}$ depends on x'. This leaves

$$\vec{H} = -\frac{1}{c} \int \frac{\vec{R} \times \vec{j}(\vec{x}')}{R^3} \mathrm{d}^3x' = \frac{1}{c} \int \frac{\left[\vec{j}(\vec{x}'), \vec{x} - \vec{x}'\right]}{|\vec{x} - \vec{x}'|^3} \mathrm{d}^3x' .$$

This is the famous *law of Biot-Savart*, which relates magnetic fields to their source currents.

Let us now show that G_{ret} is Lorentz invariant. We write

$$G_{ret}\left(\vec{x},t;\vec{x}',t'\right)=\Theta\left(t-t'\right)\frac{\delta\left(t'+\frac{|\vec{x}-\vec{x}'|}{c}-t\right)}{|\vec{x}-\vec{x}'|}\,.$$

Here the extra term $\Theta\left(t-t'\right)$ ensures that $G_{ret}\left(\vec{x},t;\vec{x}',t'\right)=0$ for $t<t'$, because

$$\Theta\left(t-t'\right)=\begin{cases}0,\,t<t'\\1,\,t\geq t'\end{cases}$$

When we use

$$\delta\left(f\left(x\right)\right)=\sum_i\frac{\delta\left(x-x_i\right)}{\left|f'\left(x_i\right)\right|}\,.$$

In the last formula the derivative is evaluated at the set of points x_i, such that $f\left(x_i\right)=0$. Realizing that for a wave propagating at the speed of light $\mathrm{d}s^2=0$ and using some algebraic trickery[35], we get

$$\begin{aligned}G_{ret}\left(\vec{x},t;\vec{x}',t'\right)&=2c\Theta\left(t-t'\right)\frac{\delta\left(|\vec{x}-\vec{x}'|-c\left(t-t'\right)\right)}{2\left|\vec{x}-\vec{x}'\right|}\\&=2c\Theta\left(t-t'\right)\frac{\delta\left(|\vec{x}-\vec{x}'|-c\left(t-t'\right)\right)}{|\vec{x}-\vec{x}'|+c\left(t-t'\right)}\\&=2c\Theta\left(t-t'\right)\delta\left(\left|\vec{x}-\vec{x}'\right|^2-c^2\left(t-t'\right)^2\right)\,,\end{aligned}$$

where the argument of the delta function is the 4-interval 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 orthochronical (ones that maintain causality) Lorentz transformations.

5.9 Causality principle

A quick word on intervals. A spacetime interval we have already defined as

$$\mathrm{d}s^2=c^2\mathrm{d}t^2-\mathrm{d}x_i^2 \tag{5.47}$$

We refer to them differently depending on the sign of $\mathrm{d}s^2$:

[35] Introduce $u=|\vec{x}-\vec{x}'|-c\left(t-t'\right)$. Then

$$\delta\left(\left|\vec{x}-\vec{x}'\right|^2-c^2\left(t-t'\right)^2\right)=\delta\Big(u(u+2c(t-t'))\Big)=\delta\Big(u^2+2uc(t-t'))\Big)\,.$$

Thus, we introduce $f(u)=u^2+2uc(t-t')$ 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,

$$\delta\left(\left|\vec{x}-\vec{x}'\right|^2-c^2\left(t-t'\right)^2\right)=\frac{\delta\left(|\vec{x}-\vec{x}'|-c\left(t-t'\right)\right)}{(2u+2c(t-t'))|_{u=0}}=\frac{\delta\left(|\vec{x}-\vec{x}'|-c\left(t-t'\right)\right)}{2c(t-t')}\,.$$

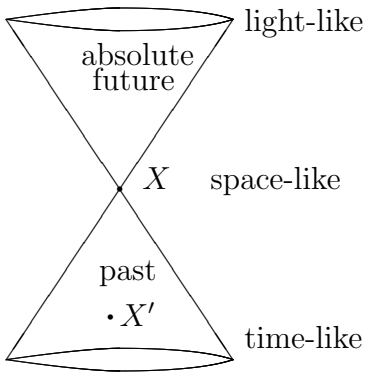


Figure 15: At every point in time every observer has his past light cone, which is a set of all events that could have influenced his presence, and a future light cone, the set of events which the observer can influence. The boundaries of the light cones also define the split between different kinds 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.

time-like intervals if $\mathrm{d}s^2 > 0$

space-like intervals if $\mathrm{d}s^2 < 0$

light-like intervals (also called null intervals) if $\mathrm{d}s^2 = 0$

Consider Figure 10 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,

$$\frac{\delta A(X)}{\delta j(X')} \sim G(X - X') = 0 \tag{5.48}$$

for $X < X'$ or points X' that are space-like to X. Hence, the causality principle for the Green function is

$$G(X' - X) = 0\,, \tag{5.49}$$

in terms of the conditions described above. The retarded Green's function is the only relativistic Green's function which has this property.

6. Radiation

The last part of these lectures will treat two classical radiation problems: Liénard-Wiechert potentials and the dipole radiation. Before studying the radiation problems,

we consider the field produced by an electric charge which moves uniformly with velocity $\vec{v}$.

6.1 Fields of a uniformly moving charge

The Lorentz transformation of $\vec{x}$ is

$$\vec{x}' = \vec{x} - a\vec{v}t + \frac{a-1}{v^2}\vec{v}(\vec{v}\vec{x})\,.$$

In what follows we need to know x'^2. We compute

$$\begin{aligned} x'^2 &= x^2 + a^2v^2t^2 + \frac{(a-1)^2}{v^2}(\vec{v}\vec{x})^2 - 2a(\vec{v}\vec{x})t + 2\frac{a-1}{v^2}(\vec{v}\vec{x})^2 - 2a(a-1)(\vec{v}\vec{x})t \\ &= x^2 + a^2v^2t^2 + \frac{(\vec{v}\vec{x})^2}{v^2}\Big(\underbrace{(a-1)^2 + 2(a-1) + 1}\,-1\Big) - 2a^2(\vec{v}\vec{x})t \\ &= x^2 + a^2v^2t^2 + \frac{a^2-1}{v^2}(\vec{v}\vec{x})^2 - 2a^2(\vec{v}\vec{x})t = x^2 + a^2v^2t^2 + \frac{a^2}{c^2}(\vec{v}\vec{x})^2 - 2a^2(\vec{v}\vec{x})t\,, \end{aligned}$$

since $\frac{a^2-1}{v^2} = \frac{a^2}{c^2}$. Then, we proceed as follows

$$\begin{aligned} x'^2 &= \underbrace{x^2 - 2(\vec{v}\vec{x})t + v^2t^2} + (a^2-1)(v^2t^2 - 2(\vec{v}\vec{x})t + x^2 - x^2) + \frac{a^2}{c^2}(\vec{v}\vec{x})^2 \\ &= (\vec{x} - \vec{v}t)^2 + a^2\frac{v^2}{c^2}(\vec{x} - \vec{v}t)^2 - \frac{a^2}{c^2}(v^2x^2 - (\vec{v}\vec{x})^2) \\ &= (\vec{x} - \vec{v}t)^2 \underbrace{\Big(1 + a^2\frac{v^2}{c^2}\Big)}_{a^2} - \frac{a^2}{c^2}(v^2x^2 - (\vec{v}\vec{x})^2) \end{aligned}$$

Thus,

$$x'^2 = a^2(\vec{x} - \vec{v}t)^2 - \frac{a^2}{c^2}(v^2x^2 - (\vec{v}\vec{x})^2) = a^2(\vec{x} - \vec{v}t)^2 - \frac{a^2}{c^2}[\vec{v},\vec{x}]^2\,.$$

We further note that in the vector product $[\vec{v},\vec{x}]$ one can replace $\vec{x}$ for $\vec{x} - \vec{v}t$ without changing the result. The final answer we need reads as

$$x'^2 = a^2(\vec{x} - \vec{v}t)^2 - \frac{a^2}{c^2}(v^2x^2 - (\vec{v}\vec{x})^2) = a^2\Big((\vec{x} - \vec{v}t)^2 - \frac{1}{c^2}[\vec{v},\vec{x} - \vec{v}t]^2\Big)\,.$$

The electric and magnetic fields in the stationary frame are

$$\begin{aligned} \vec{E} &= a\vec{E}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{E}') - \frac{a}{c}[\vec{v},\vec{H}'] = a\vec{E}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{E}')\,, \\ \vec{H} &= a\vec{H}' - \frac{a-1}{v^2}\vec{v}(\vec{v}\cdot\vec{H}') + \frac{a}{c}[\vec{v},\vec{E}'] = \frac{a}{c}[\vec{v},\vec{E}']\,, \end{aligned} \tag{6.1}$$

because in the moving frame $\vec{H}' = 0$. The electric field is $\vec{E}' = e\frac{\vec{x}'}{x'^3}$. Thus, we compute

$$\vec{E} = \frac{e}{x'^3}\left(a\vec{x} - a^2\vec{v}t + \frac{a(a-1)}{v^2}\vec{v}(\vec{v}\vec{x}) - \frac{a-1}{v^2}\vec{v}\Big((\vec{v}\vec{x}) - av^2t + (a-1)(\vec{v}\vec{x})\Big)\right)\,,$$

which results into a very simple formula

$$\vec{E}(x,t) = \frac{e(\vec{x}-\vec{v}t)}{a^2\Big((\vec{x}-\vec{v}t)^2 - \frac{1}{c^2}[\vec{v},\vec{x}-\vec{v}t]^2\Big)^{3/2}}\,.$$

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

$$\vec{E}(x,t) = \frac{e\vec{R}}{R^3}\frac{\Big(1-\frac{v^2}{c^2}\Big)}{\Big(1-\frac{v^2}{c^2}\sin^2\theta\Big)^{3/2}}\,.$$

As to the magnetic field, one gets

$$\vec{H} = \frac{a}{c}\Big[\vec{v},\frac{e}{x'^3}(\vec{x}-a\vec{v}t+\frac{a-1}{v^2}\vec{v}(\vec{v}\vec{x}))\Big] = \frac{1}{c}\Big[\vec{v},\frac{ae}{x'^3}\vec{x}\Big]\,.$$

Obviously, the last expression can be written as

$$\vec{H}(x,t) = \frac{1}{c}[\vec{v},\vec{E}]\,.$$

The corresponding energy flux is

$$\vec{S}(x,t) = \frac{c}{4\pi}[\vec{E},\vec{H}] = \frac{1}{4\pi}[\vec{E}[\vec{E},\vec{v}]] = \frac{1}{4\pi}(\vec{v}E^2 - \vec{E}(\vec{E}\cdot\vec{v}))\,.$$

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.

6.2 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

$$\varphi(\vec{x},t) = \int \frac{\delta\Big(t'+\frac{|\vec{x}-\vec{x}'|}{c}-t\Big)}{|\vec{x}-\vec{x}'|}e\delta(\vec{x}'-\vec{r}(t'))\,\mathrm{d}^3x'\mathrm{d}t' \tag{6.2}$$

Let us take $\vec{x}' = \vec{r}(t')$, because only then the integrand is non-zero. Then eq.(6.2) can be integrated over $\vec{x}'$ and we get

$$\varphi(\vec{x},t) = e\int \frac{\delta\left(t' + \frac{|\vec{x}-\vec{r}(t')|}{c} - t\right)}{|\vec{x}-\vec{r}(t')|}\mathrm{d}t' \,. \tag{6.3}$$

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

$$\varphi(\vec{x},t) = \frac{e}{R}\frac{1}{1-\frac{1}{c}\frac{\vec{R}\cdot\vec{v}}{R}} = \frac{e}{R - \frac{\vec{R}\cdot\vec{v}}{c}}\,. \tag{6.4}$$

We can use the same line of reasoning to show

$$\vec{A}(\vec{x},t) = \frac{e}{c}\frac{\vec{v}}{(R - \frac{\vec{R}\cdot\vec{v}}{c})}\,. \tag{6.5}$$

The formulae (6.4) and (6.5) 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,

$$\frac{\partial R(t')}{\partial t} = \frac{\partial R(t')}{\partial t'}\frac{\partial t'}{\partial t} = -\frac{\vec{R}\cdot\vec{v}}{R}\frac{\partial t'}{\partial t} = c\left(1 - \frac{\partial t'}{\partial t}\right)\,. \tag{6.6}$$

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 (6.6) 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

$$(\vec{v}\cdot\nabla)\vec{R} \;=\; (\vec{v}\cdot\nabla)\vec{x} - (\vec{v}\cdot\nabla)\vec{r}(t').$$

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

$$(\vec{v}\cdot\nabla)\vec{R} \;=\; \vec{v} - (\vec{v}\cdot\nabla t')\vec{v}.$$

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

$$\vec{E} = \frac{e}{(R-\frac{\vec{R}\cdot\vec{v}}{c})^3}\Big((\vec{R}-\frac{\vec{v}R}{c})(1-\frac{v^2}{c^2})+\frac{1}{c^2}(\vec{R}(\vec{R}\cdot\dot{\vec{v}})-R^2\dot{\vec{v}}) - \frac{R}{c^3}(\vec{v}(\vec{R}\cdot\dot{\vec{v}})-\dot{\vec{v}}(\vec{R}\cdot\vec{v}))\Big).$$

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

$$\vec{E} = \frac{e}{(R-\frac{\vec{R}\cdot\vec{v}}{c})^3}\Big((\vec{R}-\frac{\vec{v}R}{c})(1-\frac{v^2}{c^2})+\frac{1}{c^2}(\vec{R}\times((\vec{R}-\frac{R\vec{v}}{c})\times\dot{\vec{v}}))\Big).$$

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

$$\vec{H} = \frac{\vec{R}}{R}\times\frac{e}{(R-\frac{\vec{R}\cdot\vec{v}}{c})^3}\Big((-\frac{\vec{v}R}{c})(1-\frac{v^2}{c^2})+\frac{1}{c^2}(-R^2\dot{\vec{v}}) - \frac{R}{c^3}(\vec{v}(\vec{R}\cdot\dot{\vec{v}})-\dot{\vec{v}}(\vec{R}\cdot\vec{v}))\Big).$$

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

$$\vec{H} = \frac{1}{R}\left[\vec{R},\vec{E}\right],$$
$$\vec{E} = e\frac{\left(1-\frac{v^2}{c^2}\right)\left(\vec{R}-\frac{\vec{v}}{c}R\right)}{\left(R-\frac{\vec{R}\cdot\vec{v}}{c}\right)^3} + \frac{e\left[\vec{R},\left[\vec{R}-\frac{\vec{v}}{c}R,\dot{\vec{v}}\right]\right]}{c^2\left(R-\frac{\vec{R}\cdot\vec{v}}{c}\right)^3}.$$

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.

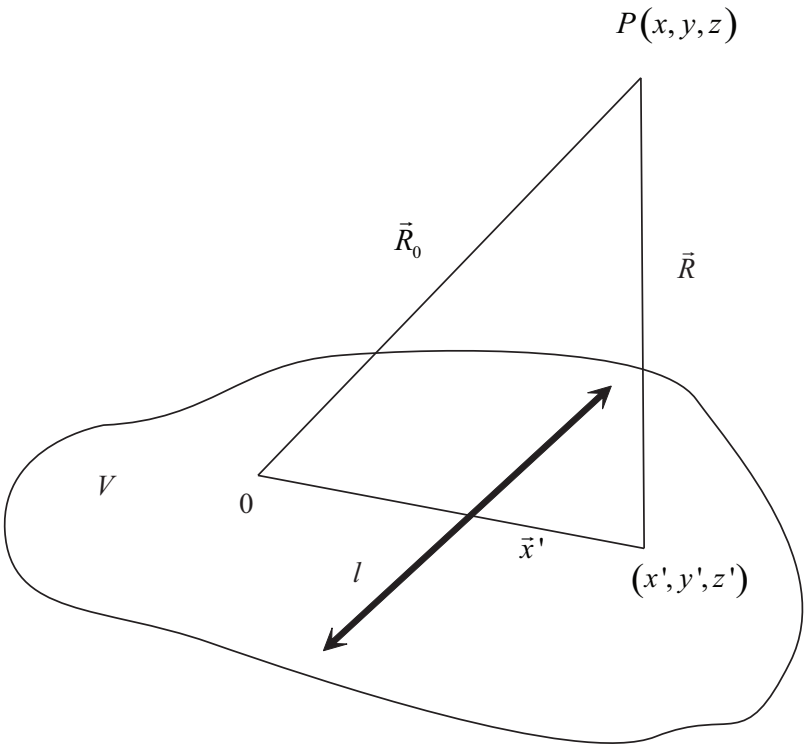


Figure 16: A diagrammatic representation of a dipole

6.3 Dipole Radiation

Field of a neutral system is expressed with the help of the so-called electric moment given in its discretized form as

$$\vec{d} = \sum_{i=1}^{N} e_i \vec{R}_i \,, \tag{6.7}$$

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 16 using Pythagorean theorem and assuming that $\vec{l} \ll R_0$, l being the characteristic size, we

get[36]

$$R = \sqrt{\left(\vec{R}_0 - \vec{R}'\right)^2} = \sqrt{\vec{R}_0^2 - 2\vec{R}_0 \cdot \vec{R}' + \vec{R}'^2} \approx$$
$$\approx \sqrt{\vec{R}_0^2 \left(1 - 2\frac{\vec{R}_0 \cdot \vec{R}'}{\vec{R}_0^2}\right)} \approx R_0 \left(1 - \frac{\vec{R}_0 \cdot \vec{R}'}{\vec{R}_0^2}\right) = R_0 - \frac{\vec{R}_0 \cdot \vec{R}'}{R_0}\,.$$

By using (5.44), 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 (6.7)

$$\vec{d}\left(t - \frac{R_0}{c}\right) = \int \mathrm{d}^3x' \vec{R}' \rho\left(x', t - \frac{R_0}{c}\right). \tag{6.8}$$

Therefore[37],

$$\varphi = -\frac{\vec{R}}{R} \cdot \frac{\partial}{\partial R} \frac{\vec{d}\left(t - \frac{R}{c}\right)}{R}\,.$$

Further, we find

$$\mathrm{div} \frac{\vec{d}\left(t - \frac{R}{c}\right)}{R} = \vec{d} \cdot \vec{\nabla} \frac{1}{R} + \frac{1}{R} \mathrm{div}\, \vec{d} = -\frac{\vec{d} \cdot \vec{R}}{R^3} + \frac{1}{R} \mathrm{div}\, \vec{d},$$
$$\mathrm{div}\, \vec{d} = \frac{\partial d_i}{\partial x^i} = \frac{\partial d_i}{\partial R} \frac{\partial R}{\partial x^i} = \left(\frac{\vec{R}}{R} \cdot \frac{\partial \vec{d}}{\partial R}\right),$$

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,

$$\varphi = \frac{\vec{d} \cdot \vec{R}}{R^3} - \frac{\vec{R}}{R^2} \frac{\partial \vec{d}}{\partial R}\,.$$

[36] Here $\vec{R}' \equiv (x', y', z')$.

[37] To simplify our further treatment, the have changed the notation $R_0 \to R$.

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 (5.45), 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 (6.8), 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[38]

$$\vec{A}=\frac{1}{cR}\cdot\frac{\partial}{\partial t}\vec{d}\left(t-\frac{R}{c}\right).$$

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

$$\vec{P}\,(t,R)=\frac{\vec{d}\left(t-\frac{R}{c}\right)}{R}. \tag{6.9}$$

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

$$\varphi(\vec{R},t)=-\mathrm{div}\vec{P}\,(t,R)\;,\qquad \vec{A}(\vec{R},t)=\frac{1}{c}\frac{\partial\vec{P}\,(t,R)}{\partial t};$$

[38]Here we again changed the notation $R_0\to R$.

The potentials are spherically symmetric, *i.e.* they depend on the distance R only. For the electromagnetic fields we have

$$\begin{aligned}\vec{H} &= \mathrm{rot}\vec{A}\,(t) = \frac{1}{c}\frac{\partial}{\partial t}\mathrm{rot}\vec{P}\,(t,R)\;;\\ \vec{E} &= -\frac{1}{c}\frac{\partial\vec{A}\,(t)}{\partial t} - \vec{\nabla}\varphi = -\frac{1}{c^2}\frac{\partial^2\vec{P}\,(t,R)}{\partial t^2} - \vec{\nabla}\mathrm{div}\vec{P}\,(t,R)\\ &= -\frac{1}{c^2}\frac{\partial^2\vec{P}\,(t,R)}{\partial t^2} + \vec{\nabla}^2\vec{P}\,(t,R) + \mathrm{rot}\;\mathrm{rot}\;\vec{P}\,(t,R)\;.\end{aligned}$$

On the last line the sum of the first two terms is equal to zero by virtue of the wave equation. This results in

$$\vec{E} = \mathrm{rot}\;\mathrm{rot}\;\vec{P}\,(t,R)\;. \tag{6.10}$$

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

$$\begin{aligned}\vec{P}\,(t,R) &= \vec{d}_0\frac{f\left(t-\frac{R}{c}\right)}{R}\,,\\ \mathrm{rot}\;\vec{P} &= \frac{f}{R}\mathrm{rot}\;\vec{d}_0 + \left[\vec{\nabla}\frac{f}{R},\vec{d}_0\right] = \frac{\partial}{\partial R}\left(\frac{f\left(t-\frac{R}{c}\right)}{R}\right)\left[\frac{\vec{R}}{R},\vec{d}_0\right] =\\ &= \frac{1}{R}\frac{\partial}{\partial R}\left(\frac{f\left(t-\frac{R}{c}\right)}{R}\right)\left[\vec{R},\vec{d}_0\right]\end{aligned}$$

as $\mathrm{rot}\,\vec{d}_0 = 0$. In the spherical coordinate system we compute the corresponding components

$$\begin{aligned}\left|\left[\vec{R},\vec{d}_0\right]\right| &= Rd_0\sin\theta\,,\\ \left[\vec{R},\vec{d}_0\right]_R &= \left[\vec{R},\vec{d}_0\right]_\theta = 0\,,\\ \left[\vec{R},\vec{d}_0\right]_\phi &= -Rd_0\sin\theta\,.\end{aligned}$$

and get[39]

$$\begin{aligned}\left(\mathrm{rot}\;\vec{P}\right)_R &= \left(\mathrm{rot}\;\vec{P}\right)_\theta = 0\,,\\ \left(\mathrm{rot}\;\vec{P}\right)_\phi &= -d_0\sin\theta\frac{\partial}{\partial R}\left(\frac{f\left(t-\frac{R}{c}\right)}{R}\right) = -\sin\theta\frac{\partial}{\partial R}P\,(t,R)\;.\end{aligned}$$

[39]Note that P here is the numerical value of the Herz vector $\vec{P}$.

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 (6.9), where we find

$$\begin{aligned} H_R &= H_\theta = 0 \\ H_\phi &= -\sin\theta \frac{1}{c} \frac{\partial^2 P\left(t,R\right)}{\partial t\, \partial R}\,. \end{aligned}$$

The components of curl of any vector field $\vec{a}$ in spherical coordinates are given by

$$\begin{aligned} \left(\text{rot }\vec{a}\right)_R &= \frac{1}{R\sin\theta}\left(\frac{\partial}{\partial\theta}\left(\sin\theta a_\phi\right) - \frac{\partial a_\theta}{\partial R}\right)\,; \\ \left(\text{rot }\vec{a}\right)_\theta &= \frac{1}{R\sin\theta}\left(\frac{\partial a_R}{\partial\phi} - \frac{\partial}{\partial R}\left(R\sin\theta a_\phi\right)\right)\,; \\ \left(\text{rot }\vec{a}\right)_\phi &= \frac{1}{R}\left(\frac{\partial}{\partial R}\left(R a_\theta\right) - \frac{\partial a_R}{\partial\theta}\right)\,. \end{aligned}$$

Using these formulae together with equation (6.10), 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

$$\vec{d}\left(t - \frac{R}{c}\right) = \vec{d}_0 e^{i\omega\left(t-\frac{R}{c}\right)}\,. \tag{6.11}$$

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 (6.11). Similarly in this limit the electric

field components become

$$E_R = \frac{2\cos\theta}{R^2}P = \frac{2\cos\theta}{R^3}d\left(t\right)\,;$$
$$E_\theta = \frac{\sin\theta}{R^2}P = \frac{\sin\theta}{R^3}d\left(t\right)\,.$$

At any given moment t, this is a field of a static dipole. For the magnetic field we find

$$\vec{H} = \frac{1}{cR^3}\left[\frac{\partial\vec{d}\left(t\right)}{\partial t},\vec{R}\right] = \frac{J}{cR^3}\left[\vec{\ell},\vec{R}\right]\,.$$

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*[40].

Far away from the oscillator

Let us now consider the region far away from the oscillator, i.e. the region where

$$R \gg \frac{\lambda}{2\pi} \qquad \text{or} \qquad \frac{1}{R} \ll \frac{\omega}{c} = \frac{2\pi}{\lambda}\,.$$

Distances greater than the wavelength are called *wave-zone*. In this particular limit our field components become

$$H_\phi = -\frac{\omega^2}{c^2}\sin\theta P = -\frac{\omega^2}{c^2}\sin\theta\frac{d\left(t-\frac{R}{c}\right)}{R}\,;$$
$$E_R = 0\,;$$
$$E_\theta = -\frac{\omega^2}{c^2}\sin\theta\frac{d\left(t-\frac{R}{c}\right)}{R} = H_\phi\,.$$

Thus summarizing we get

$$E_R = E_\phi = H_R = H_\theta = 0\,,$$

and

$$E_\theta = H_\phi = -\frac{\omega^2\sin\theta}{c^2R}d_0\cos\omega\left(t-\frac{R}{c}\right)\,,$$

or

$$E_\theta = H_\phi = \frac{\sin\theta}{c^2R}\frac{\partial^2 d\left(t-\frac{R}{c}\right)}{\partial t^2}\,.$$

This last result is valid for any arbitrary $d\left(t\right)$, not necessarily $d_0 f\left(t\right)$, because we can always perform a harmonic Fourier decomposition of any function. Thus in the

[40]Note that $E \sim \frac{1}{R^3}$ and $H \sim \frac{1}{R^2}$.

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[41]. 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

$$S = \frac{c}{4\pi}\left|\left[\vec{E},\vec{H}\right]\right| = \frac{c}{4\pi}EH = \frac{1}{4\pi}\frac{\sin^2\theta}{c^3R^2}\left(\frac{\partial^2 d\left(t-\frac{R}{c}\right)}{\partial t^2}\right)^2\,,$$

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

$$\Sigma = \int_0^{2\pi}\int_0^{\pi} SR^2\sin\theta \mathrm{d}\phi\mathrm{d}\theta = \\ = \int_0^{2\pi}\int_0^{\pi}\frac{1}{4\pi}\frac{\sin^2\theta}{c^3R^2}\left(\frac{\partial^2 d\left(t-\frac{R}{c}\right)}{\partial t^2}\right)^2 R^2\sin\theta\mathrm{d}\phi\mathrm{d}\theta = \frac{2}{3c^3}\left[\frac{\partial^2 d\left(t-\frac{R}{c}\right)}{\partial t^2}\right]^2 = \frac{2}{3c^3}\ddot{d}^2\,.$$

For $d\left(t-\frac{R}{c}\right) = d_0\cos\omega\left(t-\frac{R}{c}\right)$ the flux for one period is

$$\int_0^T \Sigma\,\mathrm{d}t = \frac{2}{3c^3}d_0^2\omega^4\int_0^T\cos^2\omega\left(t-\frac{R}{c}\right)\mathrm{d}t = \\ = \frac{d_0^2\omega^4 T}{3c^3} = \frac{2\pi d_0^2\omega^3}{3c^3} = \frac{2\pi d_0^2}{3}\left(\frac{2\pi}{\lambda}\right)^3\,.$$

The averaged radiation in a unit time is then

$$\langle\,\Sigma\,\rangle = \frac{1}{T}\int_0^T \Sigma\mathrm{d}t = \frac{cd_0^2}{3}\left(\frac{2\pi}{\lambda}\right)^4\,. \tag{6.12}$$

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

[41]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.

telegraphing one should use waves of relatively short wavelengths[42] (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[43] 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

$$\begin{aligned}\langle\, J^2 \,\rangle &= \frac{1}{T}\int_0^T J^2 \mathrm{d}t = \frac{1}{T\ell^2}\int_0^T \left(\frac{\partial \vec{p}\left(t-\frac{R}{c}\right)}{\partial t}\right)^2 \mathrm{d}t = \\ &= \frac{1}{T\ell^2}\int_0^T d_0^2\omega^2 \sin^2\omega\left(t-\frac{R}{c}\right)\mathrm{d}t = \frac{d_0^2\omega^2}{T\ell^2}\frac{\pi}{\omega} = \frac{\pi d_0^2\omega^2}{\ell^2\frac{2\pi}{\omega}\omega} = \frac{d_0^2\omega^2}{2\ell^2}\,.\end{aligned}$$

Using the result (6.12), it is now easy to find R_λ

$$R_\lambda = \frac{c d_0^2}{3}\left(\frac{2\pi}{\lambda}\right)^4 \frac{2\ell^2}{d_0^2\omega^2} = \frac{2c}{3\ell^2}\left(\frac{2\pi}{\lambda}\right)^4 \frac{1}{\left(\frac{2\pi}{\lambda}c\right)^2} = \frac{2}{3c}\left(\frac{2\pi\ell}{\lambda}\right)^2\,.$$

6.4 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

$$U = \frac{1}{2}\int \mathrm{d}V \rho(x)\varphi(x)\,.$$

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)

$$mc^2 \sim \frac{e^2}{r_e}\,,$$

[42] Generally these range from tens of meters to tens of kilometers.

[43] In this case charge polarized chemical bonds between the atoms in the particles in the atmosphere act as little oscillators.

where m and e are the mass and the charge of electron. Thus, we can define *the classical radius of electron*

$$r_e = \frac{e^2}{mc^2} \sim 2.818 \cdot 10^{-15}\,\mathrm{m}\,.$$

In SI units it reads as $r_e = \frac{1}{4\pi\epsilon_0}\frac{e^2}{mc^2}$. 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*, which is the fundamental limitation on measuring the position of a particle taking both quantum mechanics and special relativity into account. Its theoretical value is given by

$$\frac{\hbar}{mc} \sim 137\, r_e \sim 10^{-13}\,\mathrm{m}\,,$$

where $\alpha = \frac{1}{137} = \frac{e^2}{\hbar c}$ is the fine structure constant for electromagnetism. The most recent experimental measurement of campton wavelenght (CODATA 2002) is one order of magnitude larger and is approximately equal to $2.426 \cdot 10^{-12}\,\mathrm{m}$.

6.5 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

$$\varphi^{(3)} = -\frac{1}{6c^3}\frac{\partial^3 t}{\partial t^3}\int \mathrm{d}^3x'\; R^2\rho\,, \qquad \vec{A}^{(2)} = -\frac{1}{c^2}\frac{\partial}{\partial t}\int \mathrm{d}^3x'\vec{j}\,.$$

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

$$\begin{aligned}
\vec{A}'^{(2)} &= -\frac{1}{c^2}\frac{\partial}{\partial t}\int \mathrm{d}^3x'\vec{j} - \frac{1}{6c^2}\frac{\partial^2}{\partial t^2}\vec{\nabla}\int \mathrm{d}^3x'\ R^2\rho \\
&= -\frac{1}{c^2}\frac{\partial}{\partial t}\int \mathrm{d}^3x'\vec{j} - \frac{1}{3c^2}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\ \vec{R}\rho\ = \\
&= -\frac{1}{c^2}\sum e\dot{\vec{v}} - \frac{1}{3c^2}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\ (\vec{R}_0 - \vec{r})\rho\ = -\frac{2}{3c^2}\sum e\dot{\vec{v}}\,. \qquad (6.13)
\end{aligned}$$

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.*

7. Advanced magnetic phenomena

Magnetic properties of all substances admit a clear and logical systematization. At high temperatures all of the substances are either *diamagnetics* or *paramagnetics*.

If some stuff is put between the poles of a magnet, the magnetic lines change in comparison to the situation when the staff is absent. Under applying magnetic field, all the substances get magnetized. This means that every piece of volume behave itself as a magnetic, while the magnetic moment of the whole body is a vector sum of magnetic moments of all volume elements. A measure of magnetization is given by $\vec{\mathcal{M}}$ which is the magnetic moment density (the magnetic dipole moment per unit volume). The product $\vec{\mathcal{M}}V$, where V is the volume, gives a total magnetic moment of a body $\vec{M} = \vec{\mathcal{M}}V$.

A non-zero $\vec{\mathcal{M}}$ appears only when external magnetic field is applied. When magnetic field is not very strong, $\vec{\mathcal{M}}$ changes linearly with the magnetic field $\vec{H}$:

$$\vec{\mathcal{M}} = \chi \vec{H}\,.$$

Here χ is called magnetic susceptibility (it is a dimensionless quantity). Then

- Paramagnetics are the substances for which $\chi > 0$
- Diamagnetics are the substances for which $\chi < 0$
- Substances with $\chi = 0$ are absent in Nature

Magnetic properties of substances are often described not by χ but rather by the *magnetic permeability*:

$$\kappa = 1 + 4\pi\chi\,.$$

For paramagnetics $\kappa > 1$ and for diamagnetics $\kappa < 1$. Introduce the magnetic induction $\vec{B}$:

$$\vec{B} = \vec{H} + 4\pi\vec{M}\,.$$

Then, $\vec{B} = \kappa\vec{H}$ and $\kappa = 1+4\pi\chi$. Although vector $\vec{B}$ is called by a vector of magnetic induction and $\vec{H}$ by a vector of magnetic field, the actual sense of $\vec{B}$ is that it is $\vec{B}$ (but not $\vec{H}$!) is the average magnetic field in media.

For $\chi = -1/4\pi$ we have $\kappa = 0$. This is the situation of an ideal diamagnetic, in which the average magnetic field $\vec{B} = 0$. Ideal magnetics do exists – they are superconductors. Absence of a magnetic field inside a superconductor is known as the Meissner-Ochsenfeld effect (1933).

In 1895 Pierre Curie discovered that magnetic susceptibility is inversely proportional to the temperature. The behavior of $\chi = \chi(T)$ is well described by the following Curie-Weiss law

$$\chi(T) = \frac{C}{T - T_c}\,,$$

where C is a constant and T_c is known as the paramagnetic Curie temperature.

7.1 Exchange interactions

Identical particles behave very differently in classical and quantum mechanics. Classical particles move each over its own trajectory. If positions of all the particles were fixed at the initial moment of time, solving equations of motion one can always identify the positions of particles at later times. In quantum mechanics the situation is different, because the notion of trajectory is absent. If we fix a particle at a given moment of time, we have no possibility to identify it among other particles at later moments of time. In other words, in quantum mechanics identical particles are absolutely indistinguishable. This principle implies that permutation of two identical particles does not change a quantum state of a system.

Consider a wave-function of two particles $\Psi(1,2)$. Under permutation $\Psi(1,2) \to \Psi(2,1)$ a state of a system should not change. This means that

$$\Psi(2,1) = e^{i\alpha}\Psi(1,2)\,,$$

where $e^{i\alpha}$ is a phase factor. Applying permutation again, we get $e^{2i\alpha} = 1$, i.e. $e^{i\alpha} = \pm 1$. Thus, there are two types of particles:

1. $\Psi(1,2) = \Psi(2,1)$ which corresponds to the Bose-Einstein statistics
2. $\Psi(1,2) = -\Psi(2,1)$ which corresponds to the Fermi-Dirac statistics

Furthermore, an internal property which defines to which class/statistics a particle belongs is the spin. Particles with zero or integer spin obey the Bose-Einstein statistics, particles with half-integer spin obeys the Fermi-Dirac statistics.

Spin of electron is $1/2$, and, therefore, electrons are fermions. As such, they obey the Pauli exclusion principle – in each quantum state one can find only one electron.

Consider a system consisting of two electrons which interact only electrostatically. Neglecting magnetic interaction between the electrons means neglecting the existence of spins. Let $\psi(\vec{r_1}, \vec{r_2})$ be the orbital wave function. Here $\vec{r}_1$ and $\vec{r}_2$ are coordinates of electrons. One cannot completely forget about spins because the total wave function

$$\Psi(1,2) = S(\sigma_1, \sigma_2)\psi(\vec{r_1}, \vec{r_2})$$

must be anti-symmetric. Here $S(\sigma_1, \sigma_2)$ is the spin wave function which describes a spin state of electrons. For two electrons there are four states which lead to either anti-symmetric wave function with the total spin $S = 0$:

$$S = 0\,, \qquad \uparrow\downarrow - \downarrow\uparrow$$

or symmetric wave function with $S = 1$:

$$\begin{aligned} s_z &= -1 \quad \downarrow\downarrow \\ s_z &= 0 \quad \uparrow\downarrow + \downarrow\uparrow \\ s_z &= 1 \quad \uparrow\uparrow \end{aligned}$$

Here s_z is the projection of spin on z-axis. For two electrons

$$\vec{S} = \vec{s}_1 + \vec{s}_2$$

and taking square (quantum mechanically!) we obtain

$$S(S+1) = s_1(s_1+1) + s_2(s_2+1) + 2\vec{s_1}\cdot\vec{s_2}$$

so that

$$\vec{s_1}\cdot\vec{s_2} = \frac{1}{2}(S(S+1) - s_1(s_1+1) - s_2(s_2+1))$$

From this formula we therefore find that

$$\vec{s_1}\cdot\vec{s_2} = \begin{cases} -\frac{3}{4} & \text{for} \quad S=0 \\ \frac{1}{4} & \text{for} \quad S=1 \end{cases}$$

Returning back to the wave function we conclude that

$$\begin{aligned} \text{for} \quad S=0 \quad \psi(\vec{r}_1,\vec{r}_2) = \psi_s \quad -- \quad \text{symmetric function} \\ \text{for} \quad S=1 \quad \psi(\vec{r}_1,\vec{r}_2) = \psi_a \quad -- \quad \text{anti-symmetric function} \end{aligned}$$

Symmetric and anti-symmetric functions describe different orbital motion of electrons and therefore they correspond to different values of energies. Which energy is realized depends on a problem at hand. For instance, for a molecule of H_2 the minimal energy corresponds to the symmetric wave function and, as a result, the electron spin S is equal to zero.

$$\begin{aligned} E_s &\Longleftrightarrow S=0 \\ E_a &\Longleftrightarrow S=1 \end{aligned}$$

Spin Hamiltonian

$$H_s = \frac{1}{4}(E_s + 3E_a) + (E_a - E_s)\vec{s_1}\cdot\vec{s_2}$$

Here the first term $\frac{1}{4}(E_s+3E_a) \equiv \bar{E}$ does not depend on spin and represents the energy averaged over all spin states (three states for $S=1$ and one state for $S=0$). The second term depends on spins of electrons. Introducing $A = E_a - E_s$, we can write

$$H_s = \bar{E} - A\vec{s_1}\cdot\vec{s_2}$$

This allows to relate energetic preference of states with $S=0$ and $S=1$ with the sign of A. For $A<0$ the "anti-parallel" configuration of spins is preferred, while for $A>0$ – "parallel". The parameter A is called an *exchange integral*. The Hamiltonian H_s describes the so-called exchange interaction.

7.2 One-dimensional Heisenberg model of ferromagnetism

Here we will study in detail so-called one-dimensional spin-$\frac{1}{2}$ Heisenberg model of ferromagnetism. We will solve it exactly by a special technique known as coordinate Bethe ansatz.

Consider a discrete circle which is a collection of ordered points labelled by the index n with the identification $n \equiv n + L$ reflecting periodic boundary conditions. Here L is a positive integer which plays the role of the length (volume) of the space. The numbers $n = 1, \dots, L$ form a fundamental domain. To each integer n along the chain we associate a two-dimensional vector space $V = \mathbb{C}^2$. In each vector space we pick up the basis

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We will call the first element "spin up" and the second one "spin down". We introduce the spin algebra which is generated by the spin variables S_n^α, where $\alpha = 1, 2, 3$, with commutation relations

$$[S_m^\alpha, S_n^\beta] = i\hbar\epsilon^{\alpha\beta\gamma} S_n^\gamma \delta_{mn} \,.$$

The spin operators have the following realization in terms of the standard Pauli matrices: $S_n^\alpha = \frac{\hbar}{2}\sigma^\alpha$ and the form the Lie algebra $\mathfrak{su}(2)$. Spin variables are subject to the periodic boundary condition $S_n^\alpha \equiv S_{n+L}^\alpha$.

The Hilbert space of the model has a dimension 2^L and it is

$$\mathscr{H} = \prod_{n=1}^{L} \otimes V_n = V_1 \otimes \cdots \otimes V_L$$

This space carries a representation of the global spin algebra whose generators are

$$S^\alpha = \sum_{n=1}^{L} \mathbb{I} \otimes \cdots \otimes \underbrace{S_n^\alpha}_{n-\text{th place}} \otimes \cdots \otimes \mathbb{I} \,.$$

The Hamiltonian of the model is

$$H = -J \sum_{n=1}^{L} S_n^\alpha S_{n+1}^\alpha \,,$$

where J is the coupling constant. More general Hamiltonian of the form

$$H = -\sum_{n=1}^{L} J^\alpha S_n^\alpha S_{n+1}^\alpha \,,$$

where all three constants J^α are different defines the so-called XYZ model. In what follows we consider only XXX model. The basic problem we would like to solve is to find the spectrum of the Hamiltonian H.

The first interesting observation is that the Hamiltonian H commutes with the spin operators. Indeed,

$$[H, S^\alpha] = -J \sum_{n,m=1}^{L} [S_n^\beta S_{n+1}^\beta, S_m^\alpha] = -J \sum_{n,m=1}^{L} [S_n^\beta, S_m^\alpha] S_{n+1}^\beta + S_n^\beta [S_{n+1}^\beta, S_m^\alpha]$$
$$= -i\hbar \sum_{n,m=1}^{L} \left(\delta_{nm} \epsilon^{\alpha\beta\gamma} S_n^\beta S_{n+1}^\gamma - \delta_{n+1,m} \epsilon^{\alpha\beta\gamma} S_n^\beta S_{n+1}^\gamma \right) = 0 \,.$$

In other words, the Hamiltonian is central w.r.t all $\mathfrak{su}(2)$ generators. Thus, the spectrum of the model will be degenerate – all states in each $\mathfrak{su}(2)$ multiplet have the same energy.

In what follows we choose $\hbar = 1$ and introduce the raising and lowering operators $S_n^\pm = S_n^1 \pm i S_n^2$. They are realized as

$$S^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} , \qquad S^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} .$$

The action of these spin operators on the basis vectors are

$$\begin{aligned} &S^+|\uparrow\rangle = 0 \,, && S^+|\downarrow\rangle = |\uparrow\rangle \,, && S^3|\uparrow\rangle = \tfrac{1}{2}|\uparrow\rangle \,, \\ &S^-|\downarrow\rangle = 0 \,, && S^-|\uparrow\rangle = |\downarrow\rangle \,, && S^3|\downarrow\rangle = -\tfrac{1}{2}|\downarrow\rangle \,. \end{aligned}$$

This indices the action of the spin operators in the Hilbert space

$$\begin{aligned} &S_k^+|\uparrow_k\rangle = 0 \,, && S_k^+|\downarrow_k\rangle = |\uparrow_k\rangle \,, && S_k^3|\uparrow_k\rangle = \tfrac{1}{2}|\uparrow_k\rangle \,, \\ &S_k^-|\downarrow_k\rangle = 0 \,, && S_k^-|\uparrow_k\rangle = |\downarrow_k\rangle \,, && S_k^3|\downarrow_k\rangle = -\tfrac{1}{2}|\downarrow_k\rangle \,. \end{aligned}$$

The Hamiltonian can be then written as

$$H = -J \sum_{n=1}^{L} \tfrac{1}{2}(S_n^+ S_{n+1}^- + S_n^- S_{n+1}^+) + S_n^3 S_{n+1}^3 \,,$$

For $L = 2$ we have

$$H = -J\Big(S^+ \otimes S^- + S^- \otimes S^+ + 2S^3 \otimes S^3 \Big) = -J \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} .$$

This matrix has three eigenvalues which are equal to $-\frac{1}{2}J$ and one which is $\frac{3}{2}J$. Three states

$$v_{s=1}^{\mathrm{hw}} = \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\text{h.w.}} , \qquad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} , \qquad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

corresponding to equal eigenvalues form a representation of $\mathfrak{su}(2)$ with spin $s = 1$ and the state

$$v^{\rm hw}_{s=0} = \underbrace{\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}}_{\rm h.w.}$$

which corresponds to $\frac{3}{2}J$ is a singlet of $\mathfrak{su}(2)$. Indeed, the generators of the global $\mathfrak{su}(2)$ are realized as

$$S^+ = \begin{pmatrix} 0&1&1&0 \\ 0&0&0&1 \\ 0&0&0&1 \\ 0&0&0&0 \end{pmatrix}, \qquad S^- = \begin{pmatrix} 0&0&0&0 \\ 1&0&0&0 \\ 1&0&0&0 \\ 0&1&1&0 \end{pmatrix}, \qquad S^3 = \begin{pmatrix} 1&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&-1 \end{pmatrix}.$$

The vectors $v^{\rm hw}_{s=1}$ and $v^{\rm hw}_{s=0}$ are the highest-weight vectors of the $s = 1$ and $s = 0$ representations respectively, because they are annihilated by S^+ and are eigenstates of S^3. In fact, $v^{\rm hw}_{s=0}$ is also annihilated by S^- which shows that this state has zero spin. Thus, we completely understood the structure of the Hilbert space for $L = 2$.

In general, the Hamiltonian can be realized as $2^L \times 2^L$ symmetric matrix which means that it has a complete orthogonal system of eigenvectors. The Hilbert space split into sum of irreducible representations of $\mathfrak{su}(2)$. Thus, for L being finite the problem of finding the eigenvalues of H reduces to the problem of diagonalizing a symmetric $2^L \times 2^L$ matrix. This can be easily achieved by computer provided L is sufficiently small. However, for the physically interesting regime $L \to \infty$ corresponding to the *thermodynamic limit* new analytic methods are required.

In what follows it is useful to introduce the following operator:

$$P = \frac{1}{2}\Big(\mathbb{I} \otimes \mathbb{I} + \sum_\alpha \sigma^\alpha \otimes \sigma^\alpha\Big) = 2\Big(\frac{1}{4}\mathbb{I} \otimes \mathbb{I} + \sum_\alpha S^\alpha \otimes S^\alpha\Big)$$

which acts on $\mathbb{C}^2 \otimes \mathbb{C}^2$ as the permutation: $P(a \otimes b) = b \otimes a$.

It is appropriate to call S^3 the operator of the total spin. On a state $|\psi\rangle$ with M spins down we have

$$S^3|\psi\rangle = \Big(\frac{1}{2}(L - M) - \frac{1}{2}M\Big)|\psi\rangle = \Big(\frac{1}{2}L - M\Big)|\psi\rangle\,.$$

Since $[H, S^3] = 0$ the Hamiltonian can be diagonalized within each subspace of the full Hilbert space with a given total spin (which is uniquely characterized by the number of spins down).

Let $M < L$ be a number of overturned spins. If $M = 0$ we have a unique state

$$|F\rangle = |\uparrow \cdots \uparrow\rangle.$$

This state is an eigenstate of the Hamiltonian with the eigenvalue $E_0 = -\frac{JL}{4}$:

$$H|F\rangle = -J\sum_{n=1}^{L} S_n^3 S_{n+1}^3 |\uparrow\cdots\uparrow\rangle = -\frac{JL}{4}|\uparrow\cdots\uparrow\rangle\,.$$

Let M be arbitrary. Since the M-th space has the dimension $\frac{L!}{(L-M)!M!}$ one should find the same number of eigenvectors of H in this subspace. So let us write the eigenvectors of H in the form

$$|\psi\rangle = \sum_{1\le n_1<\cdots<n_M\le L} a(n_1,\ldots,n_M)|n_1,\ldots,n_M\rangle$$

with some unknown coefficients $a(n_1,\ldots,n_M)$. Here

$$|n_1,\ldots,n_M\rangle = S_{n_1}^- S_{n_2}^- \ldots S_{n_M}^- |F\rangle$$

and non-coincident integers describe the positions of the overturned spins. Obviously, the coefficients $a(n_1,\ldots,n_M)$ must satisfy the following requirement of periodicity:

$$a(n_2,\ldots,n_M,n_1+N) = a(n_1,\ldots,n_M)\,.$$

The coordinate Bethe ansatz postulates the form of these coefficients (Hans Bethe, 1931)

$$a(n_1,\ldots,n_M) = \sum_{\pi\in S_M} A_\pi \exp\Big(i\sum_{j=1}^{M} p_{\pi(j)} n_j\Big)\,.$$

Here for each of the M overturned spins we introduced the variable p_j which is called *pseudo-momentum* and S_M denotes the permutation group over the labels $\{1,\ldots,M\}$. To determine the coefficients A_π as well as the set of pseudo-momenta $\{p_j\}$ we have to use the eigenvalue equation for H and the periodicity condition for $a(n_1,\ldots,n_M)$. It is instructive to work in detail the cases $M=1$ and $M=2$ first.

For $M=1$ case we have

$$|\psi\rangle = \sum_{n=1}^{L} a(n)|n\rangle\,, \qquad a(n) = Ae^{ipn}\,.$$

Thus, in this case

$$|\psi\rangle = A\sum_{n=1}^{L} e^{ipn}|n\rangle$$

is nothing else but the Fourier transform. The periodicity condition leads to determination of the pseudo-momenta

$$a(n+L) = a(n) \qquad \Longrightarrow \qquad e^{ipL} = 1\,,$$

i.e. the $\frac{L!}{(L-1)!1!} = L$ allowed values of the pseudo-momenta are

$$p = \frac{2\pi k}{L} \qquad \text{with} \qquad k = 0, \cdots, L-1\,.$$

Further, we have the eigenvalue equation

$$H|\psi\rangle = -\frac{JA}{2}\sum_{m,n=1}^{L} e^{ipm}\Big[S_n^+ S_{n+1}^- + S_n^- S_{n+1}^+ + 2S_n^3 S_{n+1}^3\Big]\,|m\rangle = E(p)|\psi\rangle\,.$$

To work out the l.h.s. we have to use the formulae

$$S_n^+ S_{n+1}^-|m\rangle = \delta_{nm}|m+1\rangle\,, \qquad S_n^- S_{n+1}^+|m\rangle = \delta_{n+1,m}|m-1\rangle$$

as well as

$$\begin{aligned} 2S_n^3 S_{n+1}^3|m\rangle &= \frac{1}{2}|m\rangle\,, \qquad \text{for} \quad m \neq n, n+1\,,\\ 2S_n^3 S_{n+1}^3|m\rangle &= -\frac{1}{2}|m\rangle\,, \quad \text{for} \quad m = n, \quad \text{or} \quad m = n+1\,. \end{aligned}$$

Taking this into account we obtain

$$\begin{aligned} H|\psi\rangle \;=\; -\frac{JA}{2}\Big[\sum_{n=1}^{L}\Big(e^{ipn}|n+1\rangle + e^{ip(n+1)}|n\rangle\Big) + \frac{1}{2}\sum_{m=1}^{L}\Big(\sum_{\substack{n=1\\ n\neq m,m-1}}^{L}\Big)e^{ipm}|m\rangle \\ -\;\frac{1}{2}\sum_{n=1}^{L} e^{ipn}|n\rangle - \frac{1}{2}\sum_{n=1}^{L} e^{ip(n+1)}|n+1\rangle\Big]\,. \end{aligned}$$

Using periodicity conditions we finally get

$$H|\psi\rangle = -\frac{JA}{2}\sum_{n=1}^{L}\Big(e^{ip(n-1)} + e^{ip(n+1)} + \frac{L-4}{2}e^{ipn}\Big)|n\rangle = -\frac{J}{2}\Big(e^{-ip} + e^{ip} + \frac{L-4}{2}\Big)|\psi\rangle\,.$$

From here we read off the eigenvalue

$$E - E_0 = J(1-\cos p) = 2J\sin^2\frac{p}{2}\,,$$

where $E_0 = -\frac{JL}{4}$. Excitation of the spin chain around the pseudo-vacuum $|F\rangle$ carrying the pseudo-momentum p is called a *magnon*[44]. Thus, magnon can be viewed

[44]The concept of a magnon was introduced in 1930 by Felix Bloch in order to explain the reduction of the spontaneous magnetization in a ferromagnet. At absolute zero temperature, a ferromagnet reaches the state of lowest energy, in which all of the atomic spins (and hence magnetic moments) point in the same direction. As the temperature increases, more and more spins deviate randomly from the common direction, thus increasing the internal energy and reducing the net magnetization. If one views the perfectly magnetized state at zero temperature as the vacuum state of the ferromagnet, the low-temperature state with a few spins out of alignment can be viewed as a gas of quasi-particles, in this case magnons. Each magnon reduces the total spin along the direction of magnetization by one unit of and the magnetization itself by, where g is the gyromagnetic ratio. The quantitative theory of quantized spin waves, or magnons, was developed further by Ted Holstein and Henry Primakoff (1940) and Freeman Dyson (1956). By using the formalism of second quantization they showed that the magnons behave as weakly interacting quasi-particles obeying the Bose-Einstein statistics (the bosons).

as the pseudo-particle with the momentum $p = \frac{2\pi k}{L}$, $k = 0, \dots, L-1$ and the energy

$$E = 2J \sin^2 \frac{p}{2} \,.$$

The last expression is the dispersion relation for one-magnon states.

Let us comment on the sign of the coupling constant. If $J < 0$ then $E_k < 0$ and $|F\rangle$ is not the ground state, i.e. a state with the lowest energy. In other words, in this case, $|F\rangle$ is not a vacuum, but rather a pseudo-vacuum, or "false" vacuum. The true ground state in non-trivial and needs some work to be identified. The case $J < 0$ is called the anti-ferromagnetic one. Oppositely, if $J > 0$ then $|F\rangle$ is a state with the lowest energy and, therefore, is the true vacuum. Later on we will see that the anti-ferromagnetic ground state corresponds $M = \frac{1}{2}L$ and, therefore, it is spinless. The ferromagnetic ground state corresponds to $M = 0$ and, therefore, carries maximal spin $S^3 = \frac{1}{2}L$.[45]

Let us now turn to the more complicated case $M = 2$. Here we have

$$|\psi\rangle = \sum_{1 \le n_1 < n_2 \le L} a(n_1, n_2) |n_1, n_2\rangle \,,$$

where

$$a(n_1, n_2) = A e^{i(p_1 n_1 + p_2 n_2)} + B e^{i(p_2 n_1 + p_1 n_2)} \,.$$

The eigenvalue equation for H imposes conditions on $a(n_1, n_2)$ analogous to the $M = 1$ case. Special care is needed, however, when two overturned spins are sitting next to each other. Thus, we are led to consider

$$\begin{aligned} H|\psi\rangle &= -\frac{J}{2} \sum_{1 \le n_1 < n_2 \le L} a(n_1, n_2) \sum_{m=1}^{L} \Big[S_m^+ S_{m+1}^- + S_m^- S_{m+1}^+ + 2 S_m^3 S_{m+1}^3 \Big] |n_1, n_2\rangle \\ &= \Bigg\{ -\frac{J}{2} \Big[\sum_{\substack{1 \le n_1 < n_2 \le L \\ n_2 > n_1 + 1}} a(n_1, n_2) \Big(|n_1 + 1, n_2\rangle + |n_1, n_2 + 1\rangle + |n_1 - 1, n_2\rangle + |n_1, n_2 - 1\rangle \Big) \\ &\qquad + \frac{L-4}{2} \sum_{\substack{1 \le n_1 < n_2 \le L \\ n_2 > n_1 + 1}} a(n_1, n_2) |n_1, n_2\rangle - \frac{1}{2} 4 \sum_{\substack{1 \le n_1 < n_2 \le L \\ n_2 > n_1 + 1}} a(n_1, n_2) |n_1, n_2\rangle \Big] \Bigg\} + \\ &+ \Bigg\{ -\frac{J}{2} \sum_{1 \le n_1 \le L} a(n_1, n_1 + 1) \Big[|n_1, n_1 + 2\rangle + |n_1 - 1, n_1 + 1\rangle + \big(\frac{L-2}{2} - 1 \big) |n_1, n_1 + 1\rangle \Big] \Bigg\} \,. \end{aligned}$$

[45] Many crystals possess the ordered magnetic structure. This means that in absence of external magnetic field the averaged quantum-mechanical magnetic moment in each elementary crystal cell is different from zero. In the ferromagnetic crystals (Fe, Ni, Co) the averaged values of magnetic moments of all the atoms have the same orientation unless the temperature does not exceed a certain critical value called the Curie temperature. Due to this, ferromagnets have a spontaneous magnetic moment, i.e. a macroscopic magnetic moment different from zero in the vanishing external field. In more complicated anti-ferromagnetic crystals (carbons, sulfates, oxides) the averaged values of magnetic moments of individual atoms compensate each other within every elementary crystal cell.

Here in the first bracket we consider the terms with $n_2 > n_1+1$, while the last bracket represents the result of action of H on terms with $n_2 = n_1 + 1$. Using periodicity conditions we are allowed to make shifts of the summation variables n_1, n_2 in the first bracket to bring all the states to the uniform expression $|n_1, n_2\rangle$. We therefore get

$$H|\psi\rangle = -\frac{J}{2}\Bigg\{\sum_{n_2>n_1} a(n_1-1,n_2)|n_1,n_2\rangle + \sum_{n_2>n_1+2} a(n_1,n_2-1)|n_1,n_2\rangle$$
$$+\sum_{n_2>n_1+2} a(n_1+1,n_2)|n_1,n_2\rangle + \sum_{n_2>n_1} a(n_1,n_2+1)|n_1,n_2\rangle + \frac{L-8}{2}\sum_{n_2>n_1+1} a(n_1,n_2)|n_1,n_2\rangle\Bigg\}$$
$$-\frac{J}{2}\Bigg\{\sum_{1\le n_1\le L} a(n_1,n_1+1)\Big[|n_1,n_1+2\rangle + |n_1-1,n_1+1\rangle + \frac{L-4}{2}|n_1,n_1+1\rangle\Big]\Bigg\}\,.$$

Now we complete the sums in the first bracket to run the range $n_2 > n_1$. This is achieved by adding and subtracting the missing terms. As the result we will get

$$H|\psi\rangle =$$
$$-\frac{J}{2}\Bigg\{\sum_{n_2>n_1}\Big(a(n_1-1,n_2)+a(n_1,n_2-1)+a(n_1+1,n_2)+a(n_1,n_2+1)+\frac{L-8}{2}a(n_1,n_2)\Big)|n_1,n_2\rangle$$
$$-\sum_{1\le n_1\le L}\Big(a(n_1,n_1)|n_1,n_1+1\rangle + a(n_1+1,n_1+1)|n_1,n_1+1\rangle +$$
$$+\underbrace{a(n_1,n_1+1)|n_1,n_1+2\rangle}+\underbrace{a(n_1,n_1+2)|n_1,n_1+2\rangle}+\frac{L-8}{2}a(n_1,n_1+1)|n_1,n_1+1\rangle\Big)\Bigg\}$$
$$-\frac{J}{2}\Bigg\{\sum_{1\le n_1\le L} a(n_1,n_1+1)\Big[\underbrace{|n_1,n_1+2\rangle + |n_1-1,n_1+1\rangle}+\frac{L-4}{2}|n_1,n_1+1\rangle\Big]\Bigg\}\,.$$

The underbraced terms cancel out and we finally get

$$H|\psi\rangle =$$
$$-\frac{J}{2}\Bigg\{\sum_{n_2>n_1}\Big(a(n_1-1,n_2)+a(n_1,n_2-1)+a(n_1+1,n_2)+a(n_1,n_2+1)+\frac{L-8}{2}a(n_1,n_2)\Big)|n_1,n_2\rangle\Bigg\}$$
$$+\frac{J}{2}\Bigg\{\sum_{1\le n_1\le L}\Big(a(n_1,n_1)+a(n_1+1,n_1+1)-2a(n_1,n_1+1)\Big)|n_1,n_1+1\rangle\Bigg\}\,.$$

If we impose the requirement that

$$a(n_1,n_1)+a(n_1+1,n_1+1)-2a(n_1,n_1+1)=0 \tag{7.1}$$

then the second bracket in the eigenvalue equation vanishes and the eigenvalue problem reduces to the following equation

$$2(E-E_0)a(n_1,n_2) = J\big[4a(n_1,n_2)-\sum_{\sigma=\pm1} a(n_1+\sigma,n_2)+a(n_1,n_2+\sigma)\big]\,. \tag{7.2}$$

Substituting in eq.(7.1) the Bethe ansatz for $a(n_1, n_2)$ we get

$$Ae^{(p_1+p_2)n} + Be^{i(p_1+p_2)n} + Ae^{(p_1+p_2)(n+1)} + Be^{i(p_1+p_2)(n+1)}$$
$$-2\Big(Ae^{i(p_1n+p_2(n+1))} + Be^{i(p_2n+p_1(n+1))}\Big) = 0\,.$$

This allows one to determine the ratio

$$\frac{B}{A} = -\frac{e^{i(p_1+p_2)} + 1 - 2e^{ip_2}}{e^{i(p_1+p_2)} + 1 - 2e^{ip_1}} \,.$$

Problem. Show that for real values of momenta the ratio $\frac{B}{A}$ is the pure phase:

$$\frac{B}{A} = e^{i\theta(p_2,p_1)} \equiv S(p_2,p_1) \,.$$

This phase is called the S-matrix. We further note that it obeys the following relation

$$S(p_1,p_2)S(p_2,p_1) = 1 \,.$$

Thus, the two-magnon Bethe ansatz takes the form

$$a(n_1,n_2) = e^{i(p_1n_1+p_2n_2)} + S(p_2,p_1)e^{i(p_2n_1+p_1n_2)} \,,$$

where we factored out the unessential normalization coefficient A.

Let us now substitute the Bethe ansatz in eq.(7.2). We get

$$\begin{aligned} 2(E-E_0)\Big(Ae^{i(p_1n_1+p_2n_2)} + Be^{i(p_2n_1+p_1n_2)}\Big) &= J\Big[4\Big(Ae^{i(p_1n_1+p_2n_2)} + Be^{i(p_2n_1+p_1n_2)}\Big) - \\ &- \Big(Ae^{i(p_1n_1+p_2n_2)}e^{ip_1} + Be^{i(p_2n_1+p_1n_2)}e^{ip_2}\Big) - \Big(Ae^{i(p_1n_1+p_2n_2)}e^{-ip_1} + Be^{i(p_2n_1+p_1n_2)}e^{-ip_2}\Big) \\ &- \Big(Ae^{i(p_1n_1+p_2n_2)}e^{ip_2} + Be^{i(p_2n_1+p_1n_2)}e^{ip_1}\Big) - \Big(Ae^{i(p_1n_1+p_2n_2)}e^{-ip_2} + Be^{i(p_2n_1+p_1n_2)}e^{-ip_1}\Big)\Big] \,. \end{aligned}$$

We see that the dependence on A and B cancel out completely and we get the following equation for the energy

$$E - E_0 = J\Big(2 - \cos p_1 - \cos p_2\Big) = 2J\sum_{k=1}^{2} \sin^2\frac{p_k}{2} \,.$$

Quite remarkably, the energy appears to be additive, i.e. the energy of a two-magnon state appears to be equal to the sum of energies of one-magnon states! This shows that magnons essentially behave themselves as free particles in the box.

Finally, we have to impose the periodicity condition $a(n_2, n_1 + L) = a(n_1, n_2)$. This results into

$$e^{i(p_1n_2+p_2n_1)}e^{ip_2L} + \frac{B}{A}e^{ip_1L}e^{i(p_2n_2+p_1n_1)} = e^{i(p_1n_1+p_2n_2)} + \frac{B}{A}e^{i(p_2n_1+p_1n_2)}$$

which implies

$$e^{ip_1L} = \frac{A}{B} = S(p_1,p_2) \,, \qquad\qquad e^{ip_2L} = \frac{B}{A} = S(p_2,p_1) \,.$$

The last equations are called "Bethe equations". *They are nothing else but the quantization conditions for momenta p_k.*

Let us note the following useful representation for the S-matrix.

We have

$$\begin{aligned} S(p_2,p_1) &= -\frac{e^{ip_2}\big(e^{ip_1}-1\big)+1-e^{ip_2}}{e^{ip_1}\big(e^{ip_2}-1\big)+1-e^{ip_1}} = -\frac{e^{ip_2}e^{\frac{i}{2}p_1}\big(e^{\frac{i}{2}p_1}-e^{-\frac{i}{2}p_1}\big)+e^{\frac{i}{2}p_2}\big(e^{-\frac{i}{2}p_2}-e^{\frac{i}{2}p_2}\big)}{e^{ip_1}e^{\frac{i}{2}p_2}\big(e^{\frac{i}{2}p_2}-e^{-\frac{i}{2}p_2}\big)+e^{\frac{i}{2}p_1}\big(e^{-\frac{i}{2}p_1}-e^{\frac{i}{2}p_1}\big)} \\ &= -\frac{e^{\frac{i}{2}p_2}\sin\frac{p_1}{2}-e^{-\frac{i}{2}p_1}\sin\frac{p_2}{2}}{e^{\frac{i}{2}p_1}\sin\frac{p_2}{2}-e^{-\frac{i}{2}p_2}\sin\frac{p_1}{2}} = \frac{\big(\cos\frac{p_2}{2}+i\sin\frac{p_2}{2}\big)\sin\frac{p_1}{2}-\big(\cos\frac{p_1}{2}-i\sin\frac{p_1}{2}\big)\sin\frac{p_2}{2}}{\big(\cos\frac{p_1}{2}+i\sin\frac{p_1}{2}\big)\sin\frac{p_2}{2}-\big(\cos\frac{p_2}{2}-i\sin\frac{p_2}{2}\big)\sin\frac{p_1}{2}} \\ &= -\frac{\cos\frac{p_2}{2}\sin\frac{p_1}{2}-\cos\frac{p_1}{2}\sin\frac{p_2}{2}+2i\sin\frac{p_1}{2}\sin\frac{p_2}{2}}{\cos\frac{p_1}{2}\sin\frac{p_2}{2}-\cos\frac{p_2}{2}\sin\frac{p_1}{2}+2i\sin\frac{p_1}{2}\sin\frac{p_2}{2}} = \frac{\frac{1}{2}\cot\frac{p_2}{2}-\frac{1}{2}\cot\frac{p_1}{2}+i}{\frac{1}{2}\cot\frac{p_2}{2}-\frac{1}{2}\cot\frac{p_1}{2}-i}\,. \end{aligned}$$

Thus, we obtained

$$S(p_1,p_2) = \frac{\frac{1}{2}\cot\frac{p_1}{2}-\frac{1}{2}\cot\frac{p_2}{2}+i}{\frac{1}{2}\cot\frac{p_1}{2}-\frac{1}{2}\cot\frac{p_2}{2}-i}\,.$$

It is therefore convenient to introduce the variable $\lambda = \frac{1}{2}\cot\frac{p}{2}$ which is called *rapidity* and get

$$S(\lambda_1,\lambda_2) = \frac{\lambda_1-\lambda_2+i}{\lambda_1-\lambda_2-i}\,.$$

Hence, on the rapidity plane the S-matrix depends only on the difference of rapidities of scattering particles.

Taking the logarithm of the Bethe equations we obtain

$$Lp_1 = 2\pi m_1 + \theta(p_1,p_2)\,, \qquad Lp_2 = 2\pi m_2 + \theta(p_2,p_1)\,,$$

where the integers $m_i \in \{0,1,\dots,L-1\}$ are called *Bethe quantum numbers*. The Bethe quantum numbers are useful to distinguish eigenstates with different physical properties. Furthermore, these equations imply that the total momentum is

$$P = p_1 + p_2 = \frac{2\pi}{L}(m_1+m_2)\,.$$

Writing the equations in the form

$$p_1 = \underbrace{\frac{2\pi m_1}{L}} + \frac{1}{L}\theta(p_1,p_2)\,, \qquad p_2 = \underbrace{\frac{2\pi m_2}{L}} + \frac{1}{L}\theta(p_2,p_1)\,,$$

we see that the magnon interaction is reflected in the phase shift θ and in the deviation of the momenta p_1, p_2 from the values of the underbraced one-magnon wave numbers. What is very interesting, as we will see, that *magnons either scatter off each other or form the bound states.*

The first problem is to find all possible Bethe quantum numbers (m_1, m_2) for which Bethe equations have solutions. The allowed pairs (m_1, m_2) are restricted to

$$0 \le m_1 \le m_2 \le L - 1\,.$$

This is because switching m_1 and m_2 simply interchanges p_1 and p_2 and produces the same solution. There are $\frac{1}{2}L(L+1)$ pairs which meet this restriction but only $\frac{1}{2}L(L-1)$ of them yield a solution of the Bethe equations. Some of these solutions have real p_1 and p_2, the others yield the complex conjugate momenta $p_2 = p_1^*$.

The simplest solutions are the pairs for which one of the Bethe numbers is zero, e.g. $m_1 = 0$, $m = m_2 = 0, 1, \ldots, L-1$. For such a pair we have

$$Lp_1 = \theta(p_1, p_2)\,, \qquad Lp_2 = 2\pi m + \theta(p_2, p_1)\,,$$

which is solved by $p_1 = 0$ and $p_2 = \frac{2\pi m}{L}$. Indeed, for $p_1 = 0$ the phase shift vanishes: $\theta(0, p_2) = 0$. These solutions have the dispersion relation

$$E - E_0 = 2J \sin^2 \frac{p}{2}\,, \qquad p = p_2$$

which is the same as the dispersion for the one-magnon states. These solutions are nothing else but $\mathfrak{su}(2)$-descendants of the solutions with $M = 1$.

One can show that for $M = 2$ all solutions are divided into three distinct classes

$$\underbrace{\text{Descendents}}_{L}\,, \qquad \underbrace{\text{Scattering States}}_{\frac{L(L-5)}{2}+3}\,, \qquad \underbrace{\text{Bound States}}_{L-3}$$

so that

$$L + \frac{L(L-5)}{2} + 3 + L - 3 = \frac{1}{2}L(L-1)$$

gives a complete solution space of the two-magnon problem.

The most non-trivial fact about the Bethe ansatz is that many-body (multi-magnon) problem reduces to the two-body one. It means, in particular, that the multi-magnon S-matrix appears to be expressed as the product of the two-body ones. Also the energy is additive quantity. Such a particular situation is spoken about as "Factorized Scattering". In a sense, factorized scattering for the quantum many-body system is the same as integrability because it appears to be a consequence of existence of additional conservation laws. For the M-magnon problem the Bethe equations read

$$e^{ip_k L} = \prod_{\substack{j=1 \\ j \neq k}}^{M} \mathrm{S}(p_k, p_j)\,.$$

The most simple description of the bound states is obtained in the limit when $L \to \infty$. If p_k has a non-trivial positive imaginary part then $e^{ip_k L}$ tends to ∞ and this means that the bound states correspond in this limit to poles of the r.h.s. of the Bethe equations. In particular, for the case $M = 2$ the bound states correspond to poles in the two-body S-matrix. In particular, we find such a pole when

$$\tfrac{1}{2}\cot\tfrac{p_1}{2} - \tfrac{1}{2}\cot\tfrac{p_2}{2} = i\,.$$

This state has the total momentum $p = p_1 + p_2$ which must be real. These conditions can be solved by taking

$$p_1 = \frac{p}{2} + iv\,, \qquad p_2 = \frac{p}{2} - iv\,.$$

The substitution gives

$$\cos\tfrac{1}{2}\Big(\tfrac{p}{2}+iv\Big)\sin\tfrac{1}{2}\Big(\tfrac{p}{2}-iv\Big) - \cos\tfrac{1}{2}\Big(\tfrac{p}{2}-iv\Big)\sin\tfrac{1}{2}\Big(\tfrac{p}{2}+iv\Big) = 2i\sin\tfrac{1}{2}\Big(\tfrac{p}{2}+iv\Big)\sin\tfrac{1}{2}\Big(\tfrac{p}{2}-iv\Big)\,,$$

which is

$$\cos\frac{p}{2} = e^v\,.$$

The energy of such a state is

$$E = 2J\Big(\sin^2\frac{p_1}{2} + \sin^2\frac{p_2}{2}\Big) = 2J\Big(\sin^2\Big(\frac{p}{4}+i\frac{v}{2}\Big) + \sin^2\Big(\frac{p}{4}-i\frac{v}{2}\Big)\Big)\,.$$

We therefore get

$$E = 2J\Big(1 - \cos\frac{p}{2}\cosh v\Big) = 2J\Big(1 - \cos\frac{p}{2}\frac{\cos^2\frac{p}{2}+1}{2\cos\frac{p}{2}}\Big) = J\sin^2\frac{p}{2}\,.$$

Thus, the position of the pole uniquely fixes the dispersion relation of the bound state.

8. Non-linear phenomena in media

Remarkably, there exist certain differential equations for functions depending on two variables (x, t) which can be treated as integrable Hamiltonian systems with infinite number of degrees of freedom. This is an (incomplete) list of such models

- The Korteweg-de-Vries equation

$$\frac{\partial u}{\partial t} = 6uu_x - u_{xxx}\,.$$

- The non-linear Schrodinger equation

$$i\frac{\partial\psi}{\partial t} = -\psi_{xx} + 2\kappa|\psi|^2\psi\,,$$

  where $\psi = \psi(x,t)$ is a complex-valued function.

- The Sine-Gordon equation

$$\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} + \frac{m^2}{\beta}\sin\beta\phi = 0$$

- The classical Heisenberg magnet

$$\frac{\partial\vec{S}}{\partial t} = \vec{S}\times\frac{\partial^2\vec{S}}{\partial x^2}\,,$$

  where $\vec{S}(x,t)$ lies on the unit sphere in $\mathbb{R}^3$.

The complete specification of each model requires also boundary and initial conditions. Among the important cases are

1. *Rapidly decreasing case.* We impose the condition that

$$\psi(x,t) \to 0 \quad \text{when} \quad |x| \to \infty$$

   sufficiently fast, i.e., for instance, it belongs to the Schwarz space $\mathscr{L}(\mathbb{R}^1)$, which means that ψ is differentiable function which vanishes faster than any power of $|x|^{-1}$ when $|x| \to \infty$.

2. *Periodic boundary conditions.* Here we require that ψ is differentiable and satisfies the periodicity requirement

$$\psi(x+2\pi,t) = \psi(x,t)\,.$$

The soliton was first discovered by accident by the naval architect, John Scott Russell, in August 1834 on the Glasgow to Edinburg channel.[46] The modern theory originates from the work of Kruskal and Zabusky in 1965. They were the first ones to call Russel's solitary wave a solition.

[46]Russel described his discovery as follows: "I believe I shall best introduce this phenomenon by describing the circumstances of my own first acquaintance with it. I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped-not so the mass of the water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet along and a foot or foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon which I have called the Wave of Translation, a name which it now very generally bears.

8.1 Solitons

Here we discuss the simplest cnoidal wave type (periodic) and also one-soliton solutions of the KdV and SG equations For the discussion of the cnoidal wave and the one-soliton solution of the non-linear Schrodinger equation see the corresponding problem in the problem set.

Korteweg-de-Vries cnoidal wave and soliton

By rescaling of t, x and u one can bring the KdV equation to the canonical form

$$u_t + 6uu_x + u_{xxx} = 0\,.$$

We will look for a solution of this equation in the form of a single-phase periodic wave of a permanent shape

$$u(x,t) = u(x - vt)\,,$$

where $v = \text{const}$ is the phase velocity. Plugging this ansatz into the equation we obtain

$$-vu_x + 6uu_x + u_{xxx} = \frac{d}{dx}\Big(- vu + 3u^2 + u_{xx}\Big) = 0\,.$$

We thus get

$$-vu + 3u^2 + u_{xx} + e = 0\,,$$

where e is an integration constant. Multiplying this equation with an integrating factor u_x we get

$$-vuu_x + 3u^2u_x + u_xu_{xx} + eu_x = \frac{d}{dx}\Big(- \frac{v}{2}u^2 + u^3 + \frac{1}{2}u_x^2 + eu\Big) = 0\,,$$

We thus obtain

$$u_x^2 = k - 2eu + vu^2 - 2u^3 = -2(u - b_1)(u - b_2)(u - b_3)\,,$$

where k is another integration constant. In the last equation we traded the integration constants e, k for three parameters $b_3 \geq b_2 \geq b_1$ which satisfy the relation

$$v = 2(b_1 + b_2 + b_3)\,.$$

Equation

$$u_x^2 = -2(u - b_1)(u - b_2)(u - b_3)\,,$$

describes motion of a "particle" with the coordinate u and the time x in the potential $V = 2(u - b_1)(u - b_2)(u - b_3)$. Since $u_x^2 \geq 0$ for $b_2 \leq u \leq b_3$ the particle oscillates between the end points b_2 and b_3 with the period

$$\ell = 2\int_{b_2}^{b_3} \frac{\mathrm{d}u}{\sqrt{-2(u - b_1)(u - b_2)(u - b_3)}} = \frac{2\sqrt{2}}{(b_3 - b_2)^{1/2}}\mathrm{K}(m)\,,$$

where m is an elliptic modulus $0 \le m = \frac{b_3-b_2}{b_3-b_1} \le 1$.

The equation

$$u_x^2 = -2(u-b_1)(u-b_2)(u-b_3)\,,$$

can be integrated in terms of Jacobi elliptic cosine function $\mathrm{cn}(x,m)$ to give

$$u(x,t) = b_2 + (b_3-b_2)\,\mathrm{cn}^2\Big(\sqrt{(b_3-b_1)/2}(x-vt-x_0), m\Big)\,,$$

where x_0 is an initial phase. This solution is often called as *cnoidal wave*. When $m \to 1$, i.e. $b_2 \to b_1$ the cnoidal wave turns into a solitary wave

$$u(x,t) = b_1 + \frac{A}{\cosh^2\Big(\sqrt{\frac{A}{2}}(x-vt-x_0)\Big)}\,.$$

Here the velocity $v = 2(b_1+b_2+b_3) = 2(2b_1+b_3) = 2(3b_1+b_3-b_1)$ is connected to the amplitude $A = b_3 - b_1$ by the relation

$$v = 6b_1 + 2A\,.$$

Here $u(x,t) = b_1$ is called a background flow because $u(x,t) \to b_1$ as $x \to \pm\infty$. One can further note that the background flow can be eliminated by a passage to a moving frame and using the invariance of the KdV equation w.r.t. the Galilean transformation $u \to u + d$, $x \to x - 6dt$, where d is constant.

To sum up the cnoidal waves form a three-parameter family of the KdV solutions while solitons are parametrized by two independent parameters (with an account of the background flow).

Sine-Gordon cnoidal wave and soliton

Consider the Sine-Gordon equation

$$\phi_{tt} - \phi_{xx} + \frac{m^2}{\beta}\sin\beta\phi = 0\,,$$

where we assume that the functions $\phi(x,t)$ and $\phi(x,t) + 2\pi/\beta$ are assumed to be equivalent. Make an ansatz

$$\phi(x,t) = \phi(x-vt)$$

which leads to

$$(v^2-1)\phi_{xx} + \frac{m^2}{\beta}\sin\beta\phi = 0\,.$$

This can be integrated once

$$C = \frac{v^2-1}{2}\phi_x^2 - \frac{m^2}{\beta^2}\cos\beta\phi = \frac{v^2-1}{2}\phi_x^2 + \frac{2m^2}{\beta^2}\sin^2\frac{\beta\phi}{2} - \frac{m^2}{\beta^2}\,.$$

where C is an integration constant. This is nothing else as the conservation law of energy for the mathematical pendulum in the gravitational field of the Earth! We further bring equation to the form

$$\phi_x^2 = \frac{2}{v^2-1}\Big(C + \frac{m^2}{\beta^2} - \frac{2m^2}{\beta^2}\sin^2\frac{\beta\phi}{2}\Big)\,. \tag{8.1}$$

As in the case of the pendulum we make a substitution $y = \sin\frac{\beta\phi}{2}$ which gives

$$(y')^2 = \frac{m^2}{(v^2-1)}(1-y^2)\left(\frac{C+\frac{m^2}{\beta^2}}{\frac{2m^2}{\beta^2}} - y^2\right)\,.$$

This leads to solutions in terms of elliptic functions which are analogous to the cnoidal waves of the KdV equation. However, as we know the pendulum has three phases of motion: oscillatory (elliptic solution), rotatory (elliptic solution) and motion with an infinite period. The later solution is precisely the one that would correspond to the Sine-Gordon soliton we are interested in. Assuming $v^2 < 1$ we see[47] that such a solution would arise from (8.1) if we take $C = -\frac{m^2}{\beta^2}$. In this case equation (8.1) reduces to

$$\phi_x = \frac{2m}{\beta\sqrt{1-v^2}}\sin\frac{\beta\phi}{2}\,.$$

This can be integrated to[48]

$$\phi(x,t) = -\epsilon_0\frac{4}{\beta}\arctan\exp\Big(\frac{m(x-vt-x_0)}{\sqrt{1-v^2}}\Big)\,.$$

Here $\epsilon_0 = \pm 1$. This solution can be interpreted in terms of relativistic particle moving with the velocity v. The field $\phi(x,t)$ has an important characteristic – topological charge

$$Q = \frac{\beta}{2\pi}\int \mathrm{d}x\frac{\partial\phi}{\partial x} = \frac{\beta}{2\pi}(\phi(\infty)-\phi(-\infty))\,.$$

On our solutions we have

$$Q = \frac{\beta}{2\pi}\Big(-\epsilon_0\frac{4}{\beta}\Big)(\frac{\pi}{2}-0) = -\epsilon_0\,,$$

because $\arctan(\pm\infty) = \pm\frac{\pi}{2}$ and $\arctan 0 = 0$. In addition to the continuous parameters v and x_0, the soliton of the SG model has another important discrete characteristic – topological charge $Q = -\epsilon_0$. Solutions with $Q = 1$ are called solitons (kinks), while solutions with $Q = -1$ are called ani-solitons (anti-kinks).

[47] Restoring the speed of light c this condition for the velocity becomes $v^2 < c^2$, i.e., the center of mass of the soliton cannot propagate faster than light.

[48] From the equation above we see that if $\phi(x,t)$ is a solution then $-\phi(x,t)$ is also a solution.

Here we provide another useful representation for the SG soliton, namely

$$\phi(x,t) = \epsilon_0 \frac{2i}{\beta} \log \frac{1 + ie^{\frac{m(x-vt-x_0)}{\sqrt{1-v^2}}}}{1 - ie^{\frac{m(x-vt-x_0)}{\sqrt{1-v^2}}}} .$$

Indeed, looking at the solution we found we see that we can cast it in the form $\arctan \alpha = z \equiv -\frac{\beta}{4\epsilon_0}\phi(x,t)$ or $\alpha = \tan z = -i\frac{e^{2iz}-1}{e^{2iz}+1}$, where $\alpha = e^{\frac{m(x-vt-x_0)}{\sqrt{1-v^2}}}$. From here $z = \frac{1}{2i}\log\frac{1+i\alpha}{1-i\alpha}$ and the announced formula follows.

*R*emark. The stability of solitons stems from the delicate balance of "nonlinearity" and "dispersion" in the model equations. Nonlinearity drives a solitary wave to concentrate further; dispersion is the effect to spread such a localized wave. If one of these two competing effects is lost, solitons become unstable and, eventually, cease to exist. In this respect, solitons are completely different from "linear waves" like sinusoidal waves. In fact, sinusoidal waves are rather unstable in some model equations of soliton phenomena.

Sine-Gordon model has even more sophisticated solutions. Consider the following

$$\phi(x,t) = \frac{4}{\beta} \arctan \frac{\omega_2}{\omega_1} \frac{\sin\left(\frac{m\omega_1(t-vx)}{\sqrt{1-v^2}} + \phi_0\right)}{\cosh\left(\frac{m\omega_2(x-vt-x_0)}{\sqrt{1-v^2}}\right)} .$$

This is solution of the SG model which is called a *double-soliton or breather*. Except motion with velocity v corresponding to a relativistic particle the breather oscillates both in space and in time with frequencies $\frac{mv\omega_1}{\sqrt{1-v^2}}$ and $\frac{m\omega_1}{\sqrt{1-v^2}}$ respectively. The parameter ϕ_0 plays a role of the initial phase. In particular, if $v = 0$ the breather is a time-periodic solution of the SG equation. It has zero topological charge and can be interpreted as the bound state of the soliton and anti-soliton.

9. Appendices

9.1 Appendix 1: Trigonometric formulae

Some important trigonometric formulae

$$\begin{aligned}
\sin(x \pm y) &= \sin x \cos y \pm \sin y \cos x \\
\cos(x \pm y) &= \cos x \cos y \mp \sin x \sin y \\
\sin x \pm \sin y &= 2 \sin \frac{x \pm y}{2} \cos \frac{x \mp y}{2} \\
\cos x + \cos y &= 2 \cos \frac{x + y}{2} \cos \frac{x - y}{2} \\
\cos x - \cos y &= -2 \sin \frac{x + y}{2} \sin \frac{x - y}{2}
\end{aligned}$$

9.2 Appendix 2: 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

$$\left(\frac{\mathrm{d}z^1}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}z^n}{\mathrm{d}t} \right) = (\eta^1, \ldots, \eta^n)\,.$$

In the other coordinate system we will have

$$\left(\frac{\mathrm{d}x^1}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}x^n}{\mathrm{d}t} \right) = (\xi^1, \ldots, \xi^n)\,.$$

Obviously,

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = \sum_{j=1}^{n} \frac{\partial x^i}{\partial z^j} \frac{\mathrm{d}z^j}{\mathrm{d}t}\,.$$

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 $f(x^1, \ldots, x^n)$:

$$\nabla f = \left(\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n} \right) = (\xi_1, \ldots, \xi_n) \,.$$

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 $A^i_j = \frac{\partial x^i}{z^j}$. 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

$$\begin{aligned} \text{Velocity vector} \qquad & \xi \;\; = A\eta \,, \\ \text{Gradient} \qquad & \eta^t = \xi^t A \,. \end{aligned}$$

After taking transposition of the second line, we get

$$\begin{aligned} \text{Velocity vector} \qquad & \xi = A\eta \,, \\ \text{Gradient} \qquad & \eta = A^t \xi \,. \end{aligned}$$

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_i^j 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. Finally, assuming that there is an object $\phi_{i_1\ldots i_p}^{j_1\ldots j_q}$ with p upper indices transforming in the vector fashion and q lower indices transforming in the co-vector one, we arrive at the general definition of a (p,q)-type tensor presented in section 1.4.

9.3 Appendix 3: 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

$$\begin{aligned}
\langle \delta F, \eta \rangle &= \lim_{\epsilon \to 0} \frac{d}{d\epsilon} \frac{1}{2} \int \mathrm{d}t \, g_{ij}(x(t) + \epsilon\eta)(\dot{x}^i + \epsilon\dot{\eta}^i)(\dot{x}^j + \epsilon\dot{\eta}^j) = \\
&= \lim_{\epsilon \to 0} \frac{d}{d\epsilon} \frac{1}{2} \int \mathrm{d}t \Big[g_{ij}(x) + \epsilon \frac{\partial g_{ij}}{\partial x^k} \eta^k + \ldots \Big] \Big[\dot{x}^i \dot{x}^j + 2\epsilon \dot{x}^i \dot{\eta}^j + \ldots \Big] \\
&= \int \mathrm{d}t \, \Big[- \frac{d}{dt}(g_{ik}\dot{x}^k) + \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \dot{x}^i \dot{x}^j \Big] \eta^k \,.
\end{aligned}$$

Thus, for the corresponding variational derivative we find

$$\frac{\delta F}{\delta x^k(t)} = - \frac{d}{dt}(g_{ik}\dot{x}^k) + \frac{1}{2} \frac{\partial g_{ij}}{\partial x^k} \dot{x}^i \dot{x}^j \,.$$

Vanishing of this functional derivative gives an extremality condition for the corresponding functional.

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

$$\frac{\delta u(x)}{\delta u(y)} = \delta(x - y) \,.$$

9.4 Appendix 4: 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_q^α 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_q^α and x_p^α are defined. It is required that any of these two coordinate systems is expressible via the other by a differentiable map:

$$\begin{aligned} x_p^\alpha &= x_p^\alpha(x_q^1, \cdots x_q^n)\,, \qquad \alpha = 1, \cdots n \\ x_q^\alpha &= x_q^\alpha(x_p^1, \cdots x_p^n)\,, \qquad \alpha = 1, \cdots n \end{aligned} \tag{9.1}$$

  Then the Jacobian $\det\left(\frac{\partial x_p^\alpha}{\partial x_q^\beta}\right)$ is different from zero. The functions (9.1) are called *transition functions* from coordinates x_q^α to x_p^α 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

$$C'^{k}_{ij} = A^r_i A^s_j C^m_{rs} (A^{-1})^k_m \,. \tag{9.2}$$

Skew-symmetry and the Jacobi identity for the commutator imply that the tensor C^k_{ij} defines the Lie algebra if and only if

$$C^k_{ij} = -C^k_{ij} \,, \qquad C^m_{p[i} C^p_{jk]} = 0 \,.$$

Classify all Lie algebras means in fact to find all solutions of these equations modulo the equivalence relation (9.2).

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

$$X^\dagger = -X\,, \qquad\qquad \mathrm{tr}X = 0\,.$$

- 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:

$$\mathrm{Ad}(h^{-1}) = (\mathrm{Ad}h)^{-1}\,, \qquad \mathrm{Ad}(h_1 h_2) = \mathrm{Ad}h_1 \mathrm{Ad}h_2\,.$$

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

$$[\mathrm{ad}x, \mathrm{ad}y]w = \mathrm{ad}x\ \mathrm{ad}yw - \mathrm{ad}y\ \mathrm{ad}xw = [x,[y,w]] - [y,[x,w]] = [x,[y,w]] + [y,[w,x]] = \\ - [w,[x,y]] = [[x,y],w] = [z,w] = \mathrm{ad}zw\,.$$

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:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \qquad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

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.,

$$[t_+, t_+] = 0, \quad [t_+, t_-] = 2t_z, \quad [t_+, t_z] = -t_+, \quad [t_+, u_+] = v_+, \quad [t_+, u_-] = 0,$$
$$[t_+, v_+] = 0, \quad [t_+, v_-] = -u_-, \quad [t_+, y] = 0.$$

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

$$\begin{aligned}
\mathrm{ad}_x t_+ &= a t_+ \\
\mathrm{ad}_x t_- &= a t_- \\
\mathrm{ad}_x t_z &= 0 t_z \\
\mathrm{ad}_x u_+ &= (-\frac{1}{2}a + b) u_+ \\
\mathrm{ad}_x u_- &= (\frac{1}{2}a - b) u_- \\
\mathrm{ad}_x v_+ &= (\frac{1}{2}a + b) v_+ \\
\mathrm{ad}_x v_- &= (-\frac{1}{2}a - b) v_- \\
\mathrm{ad}_x y &= 0 y \,.
\end{aligned}$$

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.*

10. Problem Set

The problems in the problem set are ordered in accord with the sections of the lecture material.

10.1 Problems to section 1

Problem 1. *Shortest distance between two points.* By using the action principle, determine the shortest distance between two points on a two-dimensional plane.

Problem 2. *Minimal surface of revolution.* Consider a surface obtained by rotation around y-axis of some curve passing through two given points (x_1, y_1) and (x_2, y_2). Find the equation for the curve for which the area of this surface is minimal.

Problem 3. *Poisson brackets.* Check that the Poisson bracket

$$\{F, G\}(x) = \sum_{j=1}^{n} \Big(\frac{\partial F}{\partial p_j} \frac{\partial G}{\partial q^j} - \frac{\partial F}{\partial q^j} \frac{\partial G}{\partial p_j} \Big) .$$

satisfies the following conditions

$$\begin{aligned} &\{F, G\} = -\{G, F\} , \\ &\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0 \end{aligned}$$

for arbitrary functions F, G, H.

Problem 4. Consider a point particle moving in the potential U of the form depicted in figure 1.

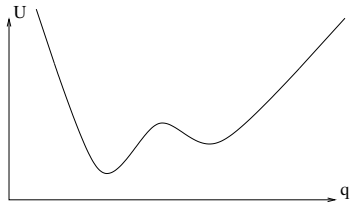


Fig. 1. Potential energy of a particle

Draw the phase curve of this particle. Hint: consult the case of the harmonic oscillator.

Problem 5. Consider a point particle moving in the potential U of the forms depicted in figure 2.

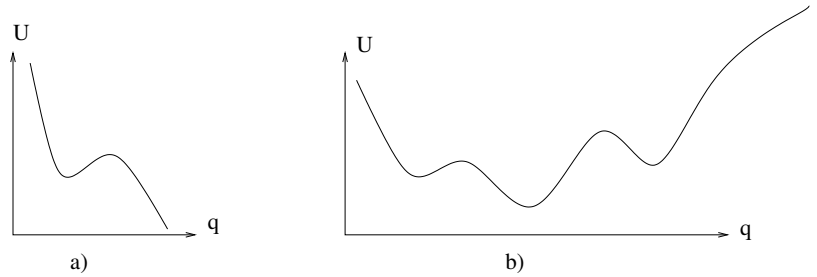


Fig. 2. Potential energies of a particle

Draw the corresponding phase curves.

Problem 6. Consider a point particle of unit mass m which moves in one dimension (the coordinate q and the momentum p) in the potential $U(q)$, where

- case 1:
$$U(q) = \frac{g^2}{q^2}\,, \qquad E > 0$$
and g^2 is a (coupling) constant.

- case 2:
$$U(q) = \frac{g^2}{\sinh^2 q}\,, \qquad E > 0$$

- case 3:
$$U(q) = -\frac{g^2}{\cosh^2 q}\,, \qquad -g^2 < E < 0$$

- case 4:
$$U(q) = -\frac{g^2}{\cosh^2 q}\,, \qquad E > 0$$

Solve equations of motion for each of these potentials. In which case the motion if finite?

Problem 7. *Harmonic oscillator.* The Lagrangian for the harmonic oscillator is

$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}m\omega^2 q^2\,.$$

Derive the corresponding equation of motion and write down its solution satisfying $q(t_1) = q_1$ and $q(t_2) = q_2$. Compute the value of the corresponding action on this solution.

Problem 8. Find the Hamiltonian and Hamilton's equations of motion for a system with two degrees of freedom with the following Lagrangian

$$L = \frac{m_1}{2}\dot{x}_1^2 + \frac{m_2}{2}\dot{x}_2^2 + B_{12}\dot{x}_1 x_2 + B_{21} x_1 \dot{x}_2 - U(x_1, x_2)\,.$$

Problem 9. The transformations below are not canonical. Find a simple way to modify them so that they become canonical

1. $P = p + q\,, \quad Q = 3q(e^{(p+q)^5} + 1) + p(2e^{(p+q)^5} - 1)\,.$
2. $P = pq\,, \quad Q = \ln(p^{2009} q^{2008})$

Problem 10. We know that the change of the Lagrangian by a total derivative of some function $f(q,t)$

$$L'(q, \dot{q}, t) = L(q, \dot{q}, t) + \frac{d}{dt} f(q,t)$$

does not change the Euler-Lagrange equations. Find the corresponding change in the momentum and the Hamiltonian. Show that it corresponds to performing a canonical transformation.

Problem 11. Use Noether's theorem to find conserved charges corresponding to the rotational symmetry of the Lagrangian of a particle moving in a n-dimensional flat space

$$L = \frac{1}{2} m v_i^2 - \kappa e^{-\alpha r}\,, \qquad r = \sqrt{x_i^2}.$$

How many independent charges are there?

Problem 12. Conserved charges corresponding to the rotational symmetry of the Lagrangian from the previous problem:

$$J_{ij} = p_i x_j - p_j x_i\,.$$

Compute the Poisson brackets of the charges with the momenta p_k, coordinates x_k and themselves, *i.e.*

$$\{J_{ij}, p_k\}\,, \qquad \{J_{ij}, x_k\}\,, \qquad \{J_{ij}, J_{kl}\}\,.$$

Problem 13. Consider a one-dimensional harmonic oscillator with the frequency ω and compute the area surrounded by the phase curve corresponding to the energy E. Show that the period of motion along this phase curve is given by $T = \frac{dS}{dE}$.

Problem 14. Let E_0 be the value of the potential at a minimum point ξ. Find the period $T_0 = \lim_{E \to E_0} T(E)$ of small oscillations in a neighborhood of the point ξ.

Problem 15. In the Kepler problem of planetary motion the three components J_i of the angular momentum have the following Poisson brackets

$$\{J_i, J_j\} = -\epsilon_{ijk} J_k$$

Show that there are three conserved quantities which commute between themselves with respect to the Poisson bracket (*i.e.* they are *in involution*)

$$H, \quad J_3, \quad J^2 = J_1^2 + J_2^2 + J_3^2$$

Here H is Kepler's Hamiltonian.

Problem 16. Introduce the polar coordinates for the Kepler problem

$$x_1 = r \sin\theta \cos\phi, \quad x_2 = r \sin\theta \sin\phi, \quad x_3 = r \cos\theta$$

Construct the canonical Hamiltonian H for the Kepler problem in the polar coordinates. Obtain the expressions for

$$J_3, \quad J^2 = J_1^2 + J_2^2 + J_3^2$$

in terms of the polar coordinates and the corresponding conjugate momenta.

Problem 17. The Euclidean space $\mathbf{R}^3$ can be endowed with the structure of a Poisson manifold, in local coordinates x^i, $i = 1, 2, 3$ the corresponding Poisson bracket is of the form

$$\{x^i, x^j\} = \epsilon^{ijk} x^k \,.$$

Show that

- the bracket is indeed Poisson;
- the bracket is degenerate and the function $(x^i)^2$ is in its center;
- reduce the Poisson bracket on the constant level manifold $(x^i)^2 = r^2 = \text{const}$ (a two-sphere); use for this spherical coordinates.
- write the corresponding symplectic form in spherical coordinate system.

Problem 18. Consider a string which perform small transverse oscillations in the (x, y)-plane around its equilibrium position coinciding with the x-axis. Denote by $y(x, t)$ the transversal displacement of the string from its equilibrium at the point x

at the moment of time t. Assuming that oscillations are small, derive the equation of motion for $y(x,t)$.

Problem 19. Consider the Lagrangian density of a real relativistic scalar field

$$\mathscr{L} = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2 \,.$$

Here $\eta^\mu = \eta_{\mu\nu} = \mathrm{diag}(+1,-1,-1,-1)$ is the Minkowski metric.

- Derive the corresponding Euler-Lagrange equations (the Klein-Gordon equation).
- Construct the corresponding Hamiltonian.
- By using Noether's theorem, derive the stress-energy tensor.

Problem 20. Consider the Lagrangian density of a complex relativistic scalar field

$$\mathscr{L} = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi^*\partial_\nu\phi - \frac{1}{2}m^2\phi^*\phi \,,$$

where $*$ means complex conjugation. Verify that the Lagrangian is invariant under the transformation $\phi \to e^{i\alpha}\phi$, where α is a real constant parameter. By using Noether's theorem, construct the corresponding conserved current and conserved charge.

Problem 21. Consider a theory which involves N scalar fields ϕ_a, all of them have the same mass, and has the following Lagrangian density

$$\mathscr{L} = \frac{1}{2}\sum_{a=1}^{N}\partial_\mu\phi_a\partial^\mu\phi_a - \frac{m^2}{2}\sum_{a=1}^{N}\phi_a^2 - g\Big(\sum_{a=1}^{N}\phi_a^2\Big)^2 .$$

Here the last term described non-linear interactions with the strength g (the coupling constant). Show that this Lagrangian density is invariant under the *non-abelian* symmetry group $G = O(N)$ (the group of orthogonal matrices).

Problem 22. *Sine-Gordon Lagrangian.* Consider the Sine-Gordon model with the Lagrangian density

$$\mathscr{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{m^2}{\beta^2}(1-\cos\beta\phi)$$

over two-dimensional Minkowski space-time. Using the canonical formalism construct the Hamiltonian (the generator of time translations) of the model. Using the Noether theorem construct the momentum P (the generator of space translations) and the generator K of Lorentz rotations.

Problem 23. *A Hamiltonian structure of the Korteweg-de-Vries equation* . Let $u(x)$ be a real-valued function on $\mathbf{R}$. Consider the following evolution equation

$$u_t = 6uu_x - u_{xxx} \, .$$

This is the so-called Korteweg-de-Vries equation (KdV). Consider the following Poisson bracket

$$\{u(x), u(y)\} = (u(x) + u(y))\delta'(x-y) + c\delta'''(x-y)$$

where c is a constant. Find the Hamiltonian which produces the KdV equation with respect this Poisson bracket.

Problem 24. Prove that if a second rank tensor $T_{\mu\nu}$ is anti-symmetric in one coordinate system, then it remains anti-symmetric in any other coordinate system. Verify the same property for the symmetric tensor.

Problem 25. Counting tensor components.

- How many independent components has a tensor $T^{\alpha\beta\cdots}$ of rank r, which has no symmetry properties, in an n-dimensional space?
- How many components has a tensor which is anti-symmetric in s indices?
- How many components has a tensor which is symmetric in s indices?

10.2 Problems to section 2

Problem 1. Prove the following formulae of vector analysis:

- For any vector $\vec{A}$:

$$\operatorname{div} \operatorname{rot} \vec{A} = 0.$$

- For any function f:

$$\operatorname{rot} \operatorname{grad} f = 0.$$

- For any two vectors $\vec{A}$ and $\vec{B}$:

$$\operatorname{grad}(\vec{A}\vec{B}) = (\vec{A}\nabla)\vec{B} + (\vec{B}\nabla)\vec{A} + \vec{B} \times \operatorname{rot} \vec{A} + \vec{A} \times \operatorname{rot} \vec{B} \, .$$

- For any function f and a vector $\vec{A}$:

$$\operatorname{rot} f\vec{A} = f \operatorname{rot} \vec{A} + \nabla f \times \vec{A} \, .$$

Problem 2. A conducter is a material inside of which electric charges can freely move under an electric field. In the electrostatic equilibrium a charged conducter has its charge distributed on the surface. By using the Gauss theorem together with

$$\int \vec{E} \cdot \mathrm{d}\vec{\ell} = 0$$

show that

- the electric field on the surface of a conducter is always normal to this surface;
- the value of the electric field on the surface is $E = 4\pi\sigma$, where σ is the surface charged densisty.

Problem 3. The simplest capacitor is made of two isolated conductors A and B placed close to each other. If we put on these conductors equal but sign opposite charges q and $-q$ then they will acquire certain potentials φ_A and φ_B. The ratio of the charge to the difference of the potentials is called a *capacitance*

$$C = \frac{q}{\varphi_A - \varphi_B}\,.$$

By using the Gauss theorem find the capacitance of

- two big plates of surface area S put on a distance d from each other;
- two concentric spheres with radii R_1 and R_2 ($R_2 > R_1$);
- two concentric cylinders of length L and radii R_1 and R_2 ($R_2 > R_1$).

Problem 4. Find the value $\alpha \equiv \alpha(d)$ for which a function

$$\varphi \sim \frac{1}{|x - x'|^{\alpha(d)}}$$

is harmonic in d dimensions (i.e. it is a solution of the Laplace equation $\nabla^2\varphi = 0$ outside $x = x'$).

Problem 4. Write the Laplace operator in spherical coordinates.

Problem 5. By using the Rodrigues formula, compute the norm of a Legendre polynomial

$$\int_{-1}^{1} \mathrm{d}x\, P_l(x)^2 = ?$$

Problem 6. Consider the following function $f(x)$ on $-1 \le x \le 1$:

$$f(x) = 1 \quad \text{for} \quad x > 0$$
$$f(x) = -1 \quad \text{for} \quad x < 0\,.$$

Expand this function into the series over Legendre polynomials.

Problem 7. Prove that for any single-valued function $\phi(x)$:

$$\delta(\varphi(x)) = \sum_i \frac{1}{|\varphi'(x_i)|}\delta(x - x_i)\,,$$

where $\delta(x)$ is the Dirac delta-function and x_i are the roots of the equation $\varphi(x) = 0$.

Problem 8. Consider two static electric charges of the values q and $-q$ separated by a small distance d. Find the scalar potential and the electric field on the distances much large than d.

Problem 9. An infinite plate of width a is charged homogeneously with the charge density ρ. Find the scalar potential and the corresponding electric field.

Problem 10. Electric charge is distributed in space with the following density $\rho = \rho_0 \cos(\alpha x)\cos(\beta y)\cos(\gamma z)$ making an infinite periodic lattice. Find the corresponding scalar potential.

Problem 11. Find the scalar potential potential and the electric field of a homogeneously charged ball. The radius of the ball is R and its charge is q.

Problem 12. Suppose it is known that all charges are contained in a volume of a size 1m^3. We want to know the potential on a distance 100m with accuracy up to 1/100 %, that is up to 10^{-4}. How many terms in the multipole expansion are suffice to keep?

Problem 13. Find an equation describing equipotential lines of a system of two point charges: charge $+q$ sitting at $z = a$ and charge $\pm q$, sitting at $z = -a$; draw these lines. *Hint:* Use the cylindrical coordinate system.

10.3 Problems to section 3

Problem 1. The vector potential produced by a magnetic dipole moment $\vec{M}$ is

$$\vec{A}(x) = \frac{\vec{M} \times \vec{R}}{|\vec{R}|^3}\,,$$

where $\vec{R} = \vec{x} - \vec{x}_0$, where $\vec{x}_0$ is a location of the magnetic moment and $\vec{x}$ is a point at which the vector potential is measured. Show that the corresponding magnetic field

is given by

$$\vec{H}(x) = \frac{3\vec{n}(\vec{n}\vec{M}) - \vec{M}}{|\vec{R}|^3}\,.$$

Problem 2. Consider an infinite long wire with the stationary current I. Find the value of the magnetic field in the surrounding space produced by this current distribution.

Problem 3. Determine the ratio of the magnetic and angular momenta for a system of two charged particles with the charges e_1 and e_2 and masses m_1 and m_2 assuming that their velocities are much less than the speed of light.

10.4 Problems to section 4

Problem 1. Verify that the unitary 2×2-matrices $g^\dagger g = \mathbb{1}$ form a group.

Problem 2. Prove that for any matrix A the following identity is valid

$$\det(\exp A) = \exp(\mathrm{tr} A)\,,$$

or, equivalently,

$$\exp(\mathrm{tr}\ln A) = \det A\,.$$

Remark. This is very important identity which enters into the proofs of many formulas from various branches of mathematics and theoretical physics. It must always stay with you. Learn it by heart by repeating the magic words "exponent trace of log is determinant".

Problem 3.
Let

$$A = \begin{pmatrix} 0 & -c_3 & c_2 \\ c_3 & 0 & -c_1 \\ -c_2 & c_1 & 0 \end{pmatrix}\,.$$

Show that the matrices

$$O(c_1, c_2, c_3) = (\mathbb{I} + A)(\mathbb{I} - A)^{-1}$$

belong to the group $SO(3)$. Show that the multiplication operation in $SO(3)$ written in coordinates (c_1, c_2, c_3) takes the form

$$O(c)O(c') = O(c'')$$

where

$$c'' = (c + c' + c \times c')/(1 - (c, c'))\,.$$

Here $c = (c_1, c_2, c_3)$ is viewed as three-dimensional vector.

Problem 4. Use the Lorentz transformation

$$dx = \frac{dx' + vdt'}{\sqrt{1 - \frac{v^2}{c^2}}}\,, \qquad dy = dy'\,, \qquad dz = dz'\,, \qquad dt = \frac{dt' + \frac{v}{c^2}dx'}{\sqrt{1 - \frac{v^2}{c^2}}}\,,$$

to derive the formulae for the Lorentz transformation of the velocity vector.

Problem 5. An inertial system M' moves with the velocity $\vec{v}$ with respect to a system M. In the system M' a bullet is shooted with the velocity $\vec{v}'$ under the angle θ'. Find the value of this angle in the system M. What happens if this bullet is a photon?

Problem 6. Let $\epsilon_{\mu\nu\rho\lambda}$ is a totally anti-symmetric tensor in Minkowski space. Compute

- $\epsilon_{\mu\nu\rho\lambda}\epsilon^{\mu\nu\rho\lambda} =?$
- $\epsilon_{\mu\nu\rho\lambda}\epsilon^{\mu\nu\rho\sigma} =?$
- $\epsilon_{\mu\nu\rho\lambda}\epsilon^{\mu\nu\tau\sigma} =?$
- $\epsilon_{\mu\nu\rho\lambda}\epsilon^{\mu\gamma\tau\sigma} =?$

Problem 7. Prove that for any tensor A^ν_μ:

$$\epsilon_{\alpha\beta\gamma\delta}A^\alpha_\mu A^\beta_\nu A^\gamma_\rho A^\delta_\lambda = \epsilon_{\mu\nu\rho\lambda}\mathrm{det}||A^\beta_\alpha||\,.$$

Problem 8. Find four linearly independent null vectors in Minkowski space. Is it possible to find four null vectors which are orthogonal?

Problem 9. Let system S' moves with respect to the system S with velocity v along the axis x. A clock resting in S' in a point (x'_0, y'_0, z'_0) in a moment t'_0 passes through a point (x_0, y_0, z_0) in the system S where the corresponding clock shows the time t_0. Write the Lorentz transformations witch relate two systems to each other.

Problem 10. (How velocities add in the relativistic case?) Prove the following formula

$$\sqrt{1 - \frac{\vec{v}^2}{c^2}} = \frac{\sqrt{1 - \frac{\vec{v}'^2}{c^2}}\sqrt{1 - \frac{\vec{V}^2}{c^2}}}{1 + \frac{(\vec{v}'\cdot\vec{V})}{c^2}}\,,$$

where $\vec{v}$ and $\vec{v'}$ are velocities of a particle in a systems S and S', respectively, and $\vec{V}$ is a velocity of S' with respect to S.

Problem 11. Prove that

$$|\vec{v}| = \frac{\sqrt{(\vec{v'} + \vec{V})^2 - \frac{(\vec{v'} \times \vec{V})^2}{c^2}}}{1 + \frac{(\vec{v'} \cdot \vec{V})}{c^2}},$$

where $\vec{v}$ and $\vec{v'}$ are velocities of a particle in a systems S and S', respectively, and $\vec{V}$ is a velocity of S' with respect to S.

Problem 12. Derive the addition formula for velocities assuming that the velocity $\vec{V}$ of S' with respect to S has an arbitrary direction. Represent the corresponding formula in vector notation.

Problem 13. Two beams of electrons collide with velocities $v = 0.9c$ measured with respect to the laboratory coordinate system. What is the relative velocity of electrons

- from the point of view of a laboratory observer?
- from the point of view of an observer moving with one of the beams?

Problem 14. (The Doppler effect). Find how the frequency ω and the wave vector $\vec{k}$ of a plane monochromatic wave transform form one inertial frame to another. The direction of the relative velocity $\vec{V}$ between two inertial frames is arbitrary.

Problem 15. Construct the Hamiltonian of a relativistic particle in a static gauge $t = \tau$, where τ is a world-line parameter.

Problem 16. Calculate the work which should be applied in order to accelerate an electron up to velocity 10^6 m/c. Treat an electron as a ball of the radius 2.8×10^{-15} meter, whose charge is homogeneously distributed over its surface.

Problem 17. Consider the motion of a charged (with charge e) relativistic massive particle (with mass m) in the Coulomb field with a potential $\phi = e'/r$. Solve the corresponding equations of motion.

Problem 18. Express momentum p of a relativistic particle via its kinetic energy.

Problem 19. Express the velocity v of a relativistic particle via its momentum p.

Problem 20. A relativistic particle with mass m_1 and velocity v_1 collides with a standing particle of mass m_2 so that both particles form now a single particle (a bound state). Find the mass M and the velocity V of this composite particle.

10.5 Problems to section 5

Problem 1. Consider a four-vector $A^\mu(x) \equiv A^\mu(t, \vec{x})$ in a four-dimensional space with the Minkowski metric. Construct the following Lagrangian density

$$\mathscr{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,, \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,.$$

- Derive the corresponding Euler-Lagrange equations (there are Maxwell's equations in a vacuum)

- Verify that the action remains invariant under constant shifts:

$$x^\mu \to x^\mu + \epsilon^\mu \,,$$

  and derive the stress-energy tensor of the electromagnetic field. This tensor obtained directly from Noether's method will not be symmetric $T^{\mu\nu} \neq T^{\nu\mu}$. Show that one can add a total derivative term $T^{\mu\nu} \to T^{\mu\nu} + \partial_\rho \chi^{\mu\nu\rho}$ which will make it symmetric.

- Verify that the action remains invariant under the following (Lorentz) transformations

$$\delta x^\mu \to \Lambda^{\mu\nu} x_\nu \,, \qquad \Lambda^{\mu\nu} = -\Lambda^{\nu\mu} \,, \qquad x^2 = \eta^{\mu\nu} x_\mu x_\nu \,.$$

  Derive the corresponding Noether current.

- Verify that the action remains invariant under rescalings (dilatation):

$$x^\mu \to \lambda x^\mu \,,$$

  where λ is a constant. Derive the Noether current corresponding to this symmetry.

- Consider special conformal transformations of the Minkowski space-time generated by the vector field

$$K_\mu = -\frac{i}{2}(x^2 \partial_\mu - 2 x_\mu x^\nu \partial_\nu), \qquad \delta x^\mu = \epsilon^\rho K_\rho \cdot x^\mu \,.$$

  Verify the invariance of the action under special conformal transformations and construct the corresponding Noether current.

Problem 2. Write the Lorentz invariants

$$F_{\mu\nu} F^{\mu\nu}$$
$$\epsilon^{\mu\nu\rho\lambda} F_{\mu\nu} F_{\rho\lambda}$$

via the electric and magnetic fields $\vec{E}$ and $\vec{B}$. How adding the second invariant to the action of the electromagnetic field will influence the corresponding equations of motion?

Problem 3. Write the stress-energy tensor $T^{\mu\nu}$ of the electromagnetic fields via electric and magnetic fields $\vec{E}$ and $\vec{B}$. Show that this tensor is conserved as a consequence of the Maxwell equations. Explain the physical meaning of various components of the stress-energy tensor.

Problem 4. Consider a charged particle which moves in a time-independent homogeneous electric field $\vec{E}$. Find the solution of the corresponding equations of motion.

Problem 5. Consider a charged particle which moves in a time-independent homogeneous magnetic field $\vec{H}$. Find the solution of the corresponding equations of motion.

Problem 6. Let the coordinate system S' moves with respect to the coordinate system S with velocity v along the axis x. Suppose in the system S there is a electromagnetic field described by the four-potential $A^\mu \equiv (A^0, A^1, A^2, A^3)$. Find the components of the four-potential A'^μ in the coordinate system S'.

Problem 7. Let the coordinate system S' moves with respect to the coordinate system S with velocity v along the axis x. Suppose in the system S there is a electromagnetic field $(\vec{E}, \vec{H})$. Find the electromagnetic field $(\vec{E}', \vec{H}')$ in the coordinate system S'.

Problem 8. Let the coordinate system S' moves with respect to the coordinate system S with velocity v along the axis x. In the framework of the previous problem, assume that $v << c$. Find the transformation law between $(\vec{E}, \vec{H})$ and $(\vec{E}', \vec{H}')$ up to terms of order $\frac{v}{c}$.

Problem 9. In the lecture notes the retarded Green function for the wave equation has been determined. Find the advanced Green function and the Pauli-Green function which is given by the difference of the advanced and retarded Green functions:

$$G_{\text{Pauli}} = G_{\text{adv}} - G_{ret}\,.$$

Show that the latter function obeys the homogeneous wave equation.

Problem 10. By using Cauchy's theorem calculate the following integrals

1.
$$\int_0^{2\pi} \frac{\mathrm{dt}}{(a + b\cos t)^2}\,, \qquad (a > b > 0);$$

2.
$$\int_0^{2\pi} \frac{(1+2\cos t)^n \cos nt}{1-a-2a\cos t}\mathrm{d}t\,, \qquad -1<a<\frac{1}{3};$$

3. (Laplace)
$$\int_0^{\infty} \frac{\cos x \mathrm{d}x}{x^2+a^2};$$

4. (Euler)
$$\int_0^{\infty} \frac{\sin x}{x}\mathrm{d}x\,.$$

Problem 11. Consider the stress-energy tensor of electromagnetic field:
$$T_{\mu\nu} = \frac{1}{4\pi}\Big(- F_{\mu\rho}F_{\nu}{}^{\rho} + \frac{1}{4}\eta_{\mu\nu}F_{\rho\lambda}F^{\rho\lambda}\Big)\,.$$
The following problems are in order

- Verify that $T_{\mu}{}^{\mu} = 0$
- Using Maxwell's equations with sources find the divergence
$$\frac{\partial T_{\mu\nu}}{\partial x_{\nu}} = ?$$

Problem 12. An electric dipole with the moment $\vec{p}$ moves with velocity $\vec{v}$ with respect to a rest coordinate system. Find electromagnetic field created by this dipole: $\varphi, \vec{A}$ and $\vec{E}$, $\vec{H}$.

Problem 13. Electromagnetic wave for which the electromagnetic field depends on one coordinate x (and on time) only, is called *flat* (plane). Assuming the gauge $\mathrm{div}\vec{A} = 0$, solve the Maxwell equations for flat (plane) electromagnetic waves.

Problem 14. Find the force acting on a wall which reflects (with the reflection coefficient R) a flat electromagnetic wave.

Problem 15. Suppose electromagnetic potential A_{μ} would a massive vector field with the action
$$S = -\frac{1}{4}\Big[\int \mathrm{d}^4x F_{\mu\nu}F^{\mu\nu} + m^2 A_{\mu}A^{\mu}\Big]\,.$$
Show that

- This action is not invariant under gauge transformations
- Derive an analogue of the Coulomb law implied by this massive vector field (in other words, find the electric field produced by a point-like static source).

Problem 16. Consider free (without charges) electomagnetic field in a finite volume of space, being a parallelepiped with sites equal to a,b and c.

- Write a Fourier decomposition of the vector potential
- Write the condition $\mathrm{div}\vec{A} = 0$ in the Fourier space
- Show that the Fourier coefficients of $\vec{A}$ satisfy the equations of motion of the harmonic oscillator
- Write the expressions for $\vec{E}$ and $\vec{H}$ via the Fourier coefficients of $\vec{A}$
- Find the Hamiltonian of free electomagnetic field in the Fourier space

10.6 Problems to section 6

Problem 1. Determine the Poynting vector for the fields produced by a charge e which moves with constant velocity $\vec{v}$. Show by explicit calculation that no energy is emitted by the charge during its motion.

Problem 2. Complete the derivation of the Lienart-Wiechert electric and magnetic fields

$$\vec{H} = \frac{1}{R}[\vec{R}, \vec{E}]\,,$$
$$\vec{E} = e\frac{\left(1 - \frac{v^2}{c^2}\right)\left(\vec{R} - \frac{\vec{v}}{c}R\right)}{\left(R - \frac{\vec{R}\cdot\vec{v}}{c}\right)^3} + e\frac{[\vec{R}, [\vec{R} - \frac{\vec{v}}{c}R, \dot{\vec{v}}]]}{c^2\left(R - \frac{\vec{R}\cdot\vec{v}}{c}\right)^3}$$

starting from the corresponding vector and scalar potential derived in the lecture notes.

Problem 3. Calculate the Poynting vector and the energy flux produced by an arbitrary moving charge by using the expressions for the Lienard-Wiechert fields.

Problem 4. An electric dipole with constant electric dipole moment magnitude is located at a point in the xy-plane and rotates with constant angular frequency.

- Determine the time-dependent electromagnetic fields at large distances from the dipole.
- Determine the radiated average power angular distribution and the total radiated power.

Problem 5. Let a positive charge $+e$ is concentrated at the origin of a coordinate system and negative charge $-e$ performs harmonic oscillations along the z-axis

around the positive charge. Find the radiation field of the corresponding system of charges. Compute the average radiated power.

Problem 6. Estimate the time of falling down of electron on the kernel in a hydrogen atom because of radiation of electromagnetic waves. Assume that electron moves on a circular orbit.

Problem 7. Determine the intensivity of dipole radiation of two charged particles (with charges e_1 and e_2 and masses m_1 and m_2) interacting by means of Coulomb potential.

Problem 8. Determine the average intensivity of dipole radiation of electron which moves over an elliptic orbit with respect to proton in a hydrogen atom.

Problem 9. Consider the following idealized situation with an infinitely long, thin, conducting wire along the z axis. For $t < 0$, it is current free, but at time $t = 0$ a constant current J is applied simultaneously over the entire length of the wire. Consequently, the wire carries the current

$$j(z) = \begin{cases} 0, \, t < 0 \\ J, \, t \geq 0 \end{cases}$$

It is assumed that the conductor is kept uncharged, *i.e.* $\rho = 0$. Determine $\vec{E}, \vec{H}$ and the Poynting vector $\vec{S}$ in the whole space.

10.7 Problems to section 7

Problem 1. Consider XXX Heisenberg model. For the chain of length $L = 3$ find the matrix form of the Hamiltonian as well as its eigenvalues. Construct the corresponding matrix representation of the global $\mathfrak{su}(2)$ generators. How many $\mathfrak{su}(2)$ multiplets the Hilbert space of the $L = 3$ model contains?

Problem 2. Carry out an explicit construction of the Bethe wave-function $a(n_1, n_2)$ for two-magnon states of the Heisenberg model. Derive the corresponding Bethe equations.

Problem 3. Show that L two-magnon states of the Heisenberg model with $p_1 = 0$ and $p_2 = \frac{2\pi m}{L}$ with $m = 0, 1, \ldots, L-1$ are $\mathfrak{su}(2)$-descendants of the one-magnon states.

Problem 4. Show that on the rapidity plane $\lambda = \frac{1}{2}\cot\frac{p}{2}$ the S-matrix of the Heisenberg model takes the form

$$S(\lambda_1, \lambda_2) = \frac{\lambda_1 - \lambda_2 + i}{\lambda_1 - \lambda_2 - i}\,.$$

Hence, it depends only on the difference of rapidities of scattering particles.

11. Recommended literature

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

4. B. Thiedé, *Electromagnetic Field Theory*, Upsilon Books, Uppsala, Sweden.

5. A. N. Vasil'ev, *Lectures on classical electrodynamics*, 2004. In Russian.