ФИЗИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

THE PRINCIPLES OF QUANTUM FIELD THEORY

АРУТЮНОВ
ГЛЕБ ЭДУАРДОВИЧ

ФИЗФАК МГУ

БЛАГОДАРИМ ЗА ПОДГОТОВКУ КОНСПЕКТА
СТУДЕНТА ФИЗИЧЕСКОГО ФАКУЛЬТЕТА МГУ
СОЛОВЫХ АЛЕКСАНДРА АЛЕКСЕЕВИЧА

Contents

Lecture 1. Classical Fields and Symmetries

The history and main challenge of QFT

Quantum field theory (QFT), perhaps, is one of the greatest achievements of theoretical physicists of 20s century. It is arose from an attempt to combine two already existent and experimentally verified theories: theory of special relativity and quantum mechanics. QFT is the unique tool which is used to describe the physics of elementary particles. QFT is characterized by numerous precision experiments or tests such as collider experiments like, for instance, those which people carry out in the Large Hadron Collider at Cern or at Fermi National Accelerator Laboratory in USA.

Methods of QFT has an universal character and this theory is not only applied in sphere of particle physics, but also widely used in, for example, condensed matter theory. Thus, QFT applied for dynamical systems with large or infinite degrees of freedom.

In spite of well developed state of QFT, starting from 1928 discovery of the electron by Dirac, it has many undiscovered peaces of knowledge.

The fact that in QFT scientists deal with dynamical systems with large or infinite degrees of freedom leads to a list of well-known problems. Some of the physical observables or quantities which you would like to measure or predict will be given by divergent integrals. It is a so-called problem of infinities. Perhaps, most sharply the problem of infinities shows itself when people try to merge QFT with gravitational theory.

All of today's experience shows that, perhaps, the degrees of freedom that can be obtained in QFT are not a suitable starting point for building a quantum theory. This means that, probably, there is some hidden theory on the next level where degrees of freedom should be reconsidered or recovered and understood by another point of view. One of such next-level theory is a string theory. There is a theory of one-dimensional extended objects and it offers very intriguing possibility how to think about quantum fields in a very different way and, hopely, to construct the quantum approach to avoid the problem of infinities. The problem of infinities is also known under the name renormalization.

Concerning particle physics, the main goal is to be able to ultimately compute and predict physical observables such as, for instance, cross-ratios in scattering experiments, decays of unstable particles and the spectra of bound states. All of these are very difficult theoretical questions because there is no direct access to local losses during particle interaction. It is unknown how to access this loss, but as one of the possible ways is

to build a theoretical model of interaction loss between elementary particles and then test this predictions, for instance, in colliding experiments, when cross-ratios or decay rates of unstable particles can be experimentally measured.

Discussion of symmetries

Symmetries play an extremely important role not in a QFT, but in physics in general and in our lives. It essentially means that the loss of nature is not arbitrary from one space-time point to another. If you have a certain knowledge of measure at one space-time point and you move away and you made measurements in the other points of space-time, you better find something which in agreement with your previous experiments, but not something arbitrary subjected to new loss. So, symmetries provide a certain order in our world, allowing us to talk about universal laws. That is why the first theme will be about symmetries and their relations to conservation laws. This will be the so-called Noether's first theorem. he transition will be made to the traditional or canonical quantization of 3 basic fields: scalar, spinor and electromagnetic, and then to the scattering matrix (or S-matrix). So, the gotten knowledge about quantized fields will be applied to evaluation of the scattering matrix of the quantum electrodynamics (QED) theory, which is used to describe electrons and photons.

The lecturer suggests people to read some books for better understanding of the course. There are:

1) Michael E. Pestin and Daniel V. Schroeder „An introduction to quantum field theory“ (canonical book, which is used in many universities)

2) A. Zee „Quantum Field Theory in a Nutshell“

3) L.D. Landau and E.M. Lifshits „Theoretical physics. Quantum electrodynamics“

4) N.N. Bogolyubov and D.V. Shirkov „Quantum fields“

5) Silvan S. Schweber „An Introduction to Relativistic Quantum Field Theory“

6) L.A. Takhtajyan „Quantum mechanics for mathematicians“

7) Brian C. Hall „Quantum Theory for Mathematicians“

8) Faddeev and Yakubovsky „Lectures on quantum mechanics for mathematics students“

Classical fields and symmetries

The discussion of classical fields and symmetries includes lagrangian description of classical systems with an infinite number of degrees of freedom and Noether's first theorem that allows to construct dynamical invariants or quantities that remain invariants during dynamical evolution of the system.

Let's remind how in classical mechanics the transition from discrete to continuum takes place. To describe continuous systems like, for instance, solid, it is necessary to make a transition from finite number of degrees of freedom to infinite. This means that it is necessary to specify at least coordinates of all points and the number of this points is infinite. By the fact continuous can be reached by taking appropriate limit of a system with a finite number of discrete coordinates. For instance, let's consider an elastic rod of fixed length which undergoes small longitudinal vibrations (fig. 1.1).

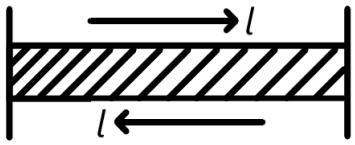

Fig. 1.1. An elastic rod of fixed length l

Let's describe this system in the framework of classical mechanics. The rod can be approximated as a system of equal mass and particles with equal distances between neighbor points.

If the rod is represented as a system of n particles connected with each other by a springs (fig. 1.2), then the distance between each particle will be $\Delta a = l/n$, where we denoted a distance between particles in equilibrium position as Δa.

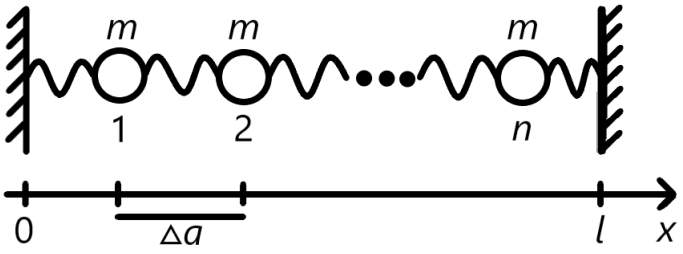

Fig. 1.2. The rod is represented as a system of n particles connected with each other by a springs

From fig. 2 it can be seen that system has $n+1$ springs and that is why the length of rod can be written as:

$$l = (n+1)\Delta a \ . \tag{1.1}$$

Each spring has such characteristic as Hooke's constant k.

The displacement of i^{th} particle from it's equilibrium position will be described by the coordinate ϕ_i. Displacement means that the particle is taken and moved a little bit away from it's equilibrium position.

The kinetic energy of all particles can be easily found from:

$$T = \sum_{i=1}^{n} \frac{m}{2} \dot{\phi}^2 \ . \tag{1.2}$$

It is also known that the presented system has a potential energy stored in springs, and such an energy may be computed. For this reason, the coordinate system should be introduced.

The left end of i^{th} particle will have coordinate:

$$x_i = i\Delta a + \phi_i \ . \tag{1.3}$$

At the same time the right end of i^{th} particle will have coordinate:

$$x_{i+1} = (i+1)\Delta a + \phi_{i+1} \ . \tag{1.4}$$

The potential energy stored in the i^{th} spring will be:

$$U_i = \frac{k}{2} \left(x_{i+1} - x_i\right)^2 \ . \tag{1.5}$$

With the use of (1.3) and (1.4), (1.5) can be simplified to:

$$U_i = \frac{k}{2} (\Delta a + \phi_{i+1} - \phi_i)^2 \ . \tag{1.6}$$

The total potential energy of the system can be gotten by summarizing of impacts from each spring:

$$U = \sum_{i=0}^{n} \frac{k}{2} \left(\Delta a + \phi_{i+1} - \phi_i\right)^2 \ . \tag{1.7}$$

Formula (1.7) can be simplified by opening the brackets:

$$U == \frac{k}{2} \sum_{i=0}^{n} \left(\phi_{i+1} - \phi_i\right)^2 + \frac{k}{2} \Delta a^2 \left(n+1\right) + k\Delta a \sum_{i=0}^{n} \left(\phi_{i+1} - \phi_i\right) \ , \tag{1.8}$$

where the last expression represents cross member.

Let's associate

$$\phi_0 = \phi_{n+1} = 0 \ , \tag{1.9}$$

because the left end of 0^{th} spring and the right end of $(n+1)^{th}$ spring are not movable.

Because of the statement (1.9) and the fact that all displacements will compensate each other, the last term of the (1.8) will be equal to zero.

It is known that in classical mechanics the constant shift of the potential plays no role, because the force is given by the minus gradient of the potential energy:

$$F_i = -\frac{\partial U_i}{\partial \phi_i} \ . \tag{1.10}$$

Such a way formula (1.9) can be simplified to:

$$U = \frac{1}{2} k \sum_{i=0}^{n} \left(\phi_{i+1} - \phi_i\right)^2 \ . \tag{1.11}$$

If we place (1.11) into (1.10), we will get that

$$F_i = k \left(\phi_{i+1} + \phi_{i-1} - 2\phi_i\right) \ . \tag{1.12}$$

The force F_i presented in (1.12) consists of difference between forces that acts on i^{th} particle from the right and from the left sides. It can be seen from:

$$F_i = \underbrace{k \left(\phi_{i+1} - \phi_i\right)}_{\text{from the right}} - \underbrace{k \left(\phi_i - \phi_{i-1}\right)}_{\text{from the left}} \ . \tag{1.13}$$

This means that kinetic and potential energies of the system are known and, therefore, the lagrangian of the system may be formed. From course of theoretical mechanics it is well known that:

$$L = T - U \ , \tag{1.14}$$

Then it is necessary to place (1.2) and (1.11) into (1.14):

$$L = \sum_{i=1}^{n} \frac{m}{2} \dot{\phi}_i^{\,2} - \frac{k}{2} \sum_{i=0}^{n} \left(\phi_{i+1} - \phi_i\right)^2 \ . \tag{1.15}$$

As can be seen from (1.15), we have a system with finite number of degrees of freedom and these degrees of freedom can be described by ϕ_i.

Then it is necessary to take a continuum limit by sending the number of particles to infinity. At the same time, the number of springs also will tend to infinity. The distance between particles in the opposite side will tend to zero:

$$n \to \infty \, , \ \Delta a \to 0 \ . \tag{1.16}$$

From (1.1) it can be clearly seen that n and Da changes with the same speed and that is why l remains constant. Despite this fact that there is an another problem which is linked with the fact that each particle in system has a mass. With increasing of particles number it will tend to infinity and therefore it is necessary to keep it finite. This target can be reached if mass of each particle m will decrease with the same speed as Δa. Let's define a mass density μ:

$$\mu = \frac{m}{\Delta a} \ . \tag{1.17}$$

From (1.17) it can be seen that the mass density is constant.

Besides the mass of each particle it is also necessary to keep constant forces between particles. For this goal the Hook's constant with the Δa change will be scaled as follows:

$$k\Delta a = const = \varkappa \ , \tag{1.18}$$

where k is now a function of Δa and

$$k \sim \frac{\varkappa}{\Delta a} \ . \tag{1.19}$$

Then it is necessary to rewrite lagrangian in terms of Δa:

$$L = \frac{1}{2}\sum_{i=1}^{n} \Delta a \left(\frac{m}{\Delta a}\right) \dot{\phi_i}^2 - \frac{1}{2}\sum_{i=0}^{n} \Delta a \left(k\Delta a\right) \left(\frac{\phi_{i+1} - \phi_i}{\Delta a}\right)^2 \ . \tag{1.20}$$

Using (1.17) and (1.18), formula (1.20) can be written as:

$$L = \frac{1}{2}\sum_{i=1}^{n} \Delta a \mu \dot{\phi_i}^2 - \frac{1}{2}\sum_{i=0}^{n} \Delta a \varkappa \left(\frac{\phi_{i+1} - \phi_i}{\Delta a}\right)^2 \ . \tag{1.21}$$

When the limit will be taken, Δa will tend to zero. Due to this fact the discrete index i of displacement variable can be replaced by a continuum variable x. In other worlds, the variable ϕ_i will be replaced by $\phi\left(i\Delta a\right)$. Due to the replacement of i, $i\Delta a$ may be changed by x and as a result:

$$\phi_i \to \phi\left(x\right) \ . \tag{1.22}$$

It should be noticed, that with $n \to \infty$ $\phi(x)$ becomes a continuous field.

After the transition to the continuum

$$\frac{\phi_{i+1} - \phi_i}{\Delta a} \to \frac{\phi\left(x + \Delta a\right) - \phi(x)}{\Delta a} \underset{n\to\infty}{\longrightarrow} \partial_x \phi\left(x\right) \ . \tag{1.23}$$

Using expressions (1.22) and (1.23), (1.21) can be written as:

$$L = \frac{1}{2}\int_0^l \mathrm{d}x \left[\mu \dot{\phi}^2\left(x\right) - \varkappa \left(\partial_x \phi\left(x\right)\right)^2\right] \ , \tag{1.24}$$

where sums from $i = 0$ or 1 to n were replaced by integrals from $x = 0$ to l.

The limiting procedure with equations of motion can also be performed. For this task it is necessary to get them from lagrangian with finite n and then go to the limit. The equations of motion for discrete system can be written as:

$$\frac{m}{\Delta a}\ddot{\phi}_i - k\Delta a\frac{\phi_{i+1} + \phi_{i-1} - 2\phi_i}{(\Delta a)^2} = 0 \ . \tag{1.25}$$

The second expression in (1.25) can be replaced by second derivative when Δa tends to zero:

$$\lim_{\Delta a\to 0}\frac{\phi_{i+1} + \phi_{i-1} - 2\phi_i}{(\Delta a)^2} = \frac{\partial^2\phi}{\partial x^2} = \partial_{xx}\phi \ . \tag{1.26}$$

As a result the final view of equation of motion (1.25) will be:

$$\mu\ddot{\phi}(x) - \varkappa\partial_{xx}\phi(x) = 0 \ . \tag{1.27}$$

Equation (1.27) also can be gotten from (1.24), but for this target it is necessary to represent (1.27) in the next view:

$$L = \int_0^l \mathrm{d}x\,\mathcal{L} \ , \tag{1.28}$$

where $\mathcal{L}$ is a lagrangian density and equal

$$\mathcal{L} = \frac{1}{2}\left(\mu\dot{\phi}^2 - \varkappa\left(\partial_x\phi\right)^2\right) \tag{1.29}$$

for the presented situation.

In theoretical mechanics there is one more important term such as action, which can be defined as an integral of the lagrangian with respect to time. By this moment it is not needed to pass the general picture, which says that the case is being conducted with an action for our continuum dynamical system.

The action will be indicated as S and it will be a functional of displacement field $\phi(x,t)$:

$$S[\phi] = \int_{t_1}^{t_2} L\mathrm{d}t \ . \tag{1.30}$$

Lagrangian at the same time can be written as presented at (1.29) and, therefore, (1.30) will have a view:

$$S[\phi] = \int_{t_1}^{t_2}\mathrm{d}t\int_0^l \mathrm{d}x \times \mathcal{L}\left(\phi(x,t)\,,\,\dot{\phi}(x,t)\,,\,\partial_x\phi(x,t)\right) \ . \tag{1.31}$$

To obtain equations of motion directly from lagrangian (1.28) it is needful to understand how the action changes under an infinitesimal change of the field. For this purpose it is necessary to replace $\phi(x,t)$ with $\phi(x,t)+\delta\phi(x,t)$. This procedure is called variation because the field is varied to observe how the action changes.

The derivatives of ϕ can be also written as:

$$\begin{cases} \frac{\partial}{\partial t}\phi\left(x,t\right) \to \frac{\partial}{\partial t}\phi\left(x,t\right)+\frac{\partial}{\partial t}\delta\phi\left(x,t\right) \\ \frac{\partial}{\partial x}\phi\left(x,t\right) \to \frac{\partial}{\partial x}\phi\left(x,t\right)+\frac{\partial}{\partial x}\delta\phi\left(x,t\right) \end{cases} \tag{1.32}$$

Then it is needed to see how the action reacts on the infinitesimal change of the field. So the result was calculated as follows:

$$\delta S\left[\phi\right] = S\left[\phi+\delta\phi\right]-S\left[\phi\right] \ . \tag{1.33}$$

Formula (1.33) gives the following:

$$\delta S\left[\phi\right] = \int_{t_1}^{t_2}\mathrm{d}t\int_0^l \mathrm{d}x\left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi+\frac{\partial\mathcal{L}}{\partial\dot{\phi}}\partial_t\delta\phi+\frac{\partial\mathcal{L}}{\partial\left(\partial_x\phi\right)}\partial_x\delta\phi\right] \ . \tag{1.34}$$

Integral in (1.34) can be evaluated by parts:

$$\begin{gathered} \delta S\left[\phi\right] = \int_{t_1}^{t_2}\mathrm{d}t\int_0^l\mathrm{d}x\left[\frac{\partial\mathcal{L}}{\partial\phi}-\partial_t\frac{\partial\mathcal{L}}{\partial\dot{\phi}}-\partial_x\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\right]\delta\phi+ \\ +\int_0^l\mathrm{d}x\frac{\partial\mathcal{L}}{\partial(\partial_t\phi)}\delta\phi\Big|_{t=t_1}^{t=t_2}+\int_{t_1}^{t_2}\mathrm{d}t\frac{\partial\mathcal{L}}{\partial(\partial_x\phi)}\delta\phi\Big|_{x=0}^{x=l} \end{gathered} \tag{1.35}$$

To find the trajectories (get equations of motion) of the dynamical systems it is necessary to minimize the action which means that it is needed put to zero the variational derivative:

$$\frac{\delta S\left[\phi\right]}{\delta\phi} = 0 \ . \tag{1.36}$$

It should be noticed that it was needed to minimize the action under the condition that values of the field in the initial and finite times are unaffected. In other words:

$$\delta\phi\left(x,t_1\right) = \delta\phi\left(x,t_2\right) = 0 \ , \tag{1.37}$$

that means that initial and final profile of the field must be kept in tact.

Expression (1.37) can be clearly understood if return to classical mechanics. In classical case there is a particle which moves from one point to another and it makes some actual

trajectory. The particle moves in the some potential and interacts with this one which lead to formation of the definite trajectory. In the variational principle, the fixed initial and final positions of the particle cannot be changed, and all other trajectories that are in the vicinity of the actual trajectory are tried. For all of this trajectories the principle of least action telling us that the value of the action must be largen than the value of the action on the actual trajectory (fig. 1.3). But comparison of values of the action on the virtual and the actual trajectories is made under the assumption that the initial and the final points taken by a particle keep fixed.

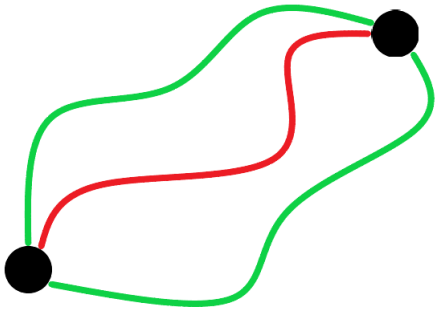

Fig. 1.3. Possible trajectories of the particle

The same thing happens for a field. Suppose that there is an initial moment of time that equals to t_1 and the ϕ function can be drawn as represented on (fig. **??**). Then when time grows, the field starts to move and at the time moment equals to t_2 has another profile. Such a way trajectory of the field forms a surface.

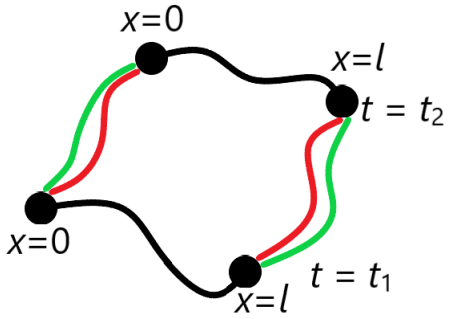

Fig. 1.4. The field evolution from $t = t_1$ to $t = t_2$

Due to formula (1.37) the integral

$$\int_0^l \mathrm{d}x \frac{\partial \mathcal{L}}{\partial \left(\partial_t \phi\right) \delta\phi}\Bigg|_{t=t_1}^{t=t_2} \tag{1.38}$$

will be equal to zero.

Then it is necessary to find the third integral in formula (1.35). This term can be put to zero or not put to zero due to physical grounds. For instance, if we say that we are dealing with the rod which is clamped between two walls which means that the profile of the field at the left and right points is fixed (fig. 1.5), the $\delta\phi$ will equal to zero and therefore third integral in formula (1.35) will be equal to zero.

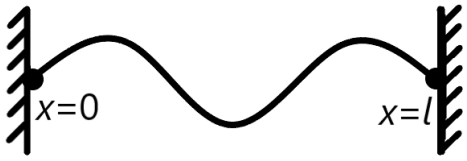

Fig. 1.5. The profile of the field of the rod which is clamped between two walls

Such a way in formula (1.35) only one expression stayed. For the actual trajectory the vanishing of δS for any variation $\delta\phi$ by the basic lemma of a variational calculus represents the equation of motion:

$$\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_t \phi\right)}\right)+\frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_x \phi\right)}\right)-\frac{\partial \mathcal{L}}{\partial \phi}=0 \ . \tag{1.39}$$

If we apply equations of motion (1.39) to the concrete problem and place the lagrangian density (1.29) into (1.39), we will get equations of motion similar to (1.27):

$$\ddot{\phi}-c^2 \partial_{xx} \phi=0 \ , \tag{1.40}$$

where c is a next ratio between μ and $\varkappa$:

$$c=\sqrt{\frac{\varkappa}{\mu}} \ . \tag{1.41}$$

The value c can interpreted as a speed which characterizes propagation of vibrations through the rod.

Equations of motions gotten for the presented dynamical system are very easy because of their linearity. As it well known, for the linear equations the superposition principle works. So it is possible to solve equations by creating of the next linear combination:

$$\phi\left(x,t\right)=e^{ipx} a_p\left(t\right)+e^{-ipx} b_p\left(t\right) \ . \tag{1.42}$$

Then it is necessary to apply boundary conditions for the gotten solution. For the left wall of the rod, the boundary condition will give:

$$\phi\left(0,t\right)=0 \Rightarrow \ a_p\left(t\right)+b_p\left(t\right)=0 \rightarrow b_p\left(t\right)=-a_p\left(t\right) \ . \tag{1.43}$$

Expression (1.43) leads to the simplification of the (1.42):

$$\phi(x,t) = a_p(t)\left(e^{ipx} - e{-}ipx\right) \to a_p(t)\sin(px) \ , \tag{1.44}$$

where the multiplier $2i$ was entered in $a_p(t)$.

At the same time for the right border of the rod there will be the following expression:

$$\phi(l,t) = a_p(t)\sin(pl) = 0 \ . \tag{1.45}$$

In expression (1.45) it is necessary to equate to zero one of two multipliers. It can be clearly seen that $a_p(t) = 0$ corresponds to $\phi(x,t) = 0$. That is why it is necessary to equate to zero $\sin(pl)$ and get the condition for p-value.

$$\sin(pl) = 0 \Rightarrow \ pl = \pi n \to p = p_n = \frac{\pi n}{l} \ , \ n \in \mathbb{Z} \ . \tag{1.46}$$

Then it is possible to take the gotten anzac and plug it into equations of motion. In this way there will be an equation which will describe a time evolution of the a_p coefficient:

$$\ddot{a}_p + c^2 p^2 a_p(t) = 0 \ . \tag{1.47}$$

The solution of the (1.47) is well known and equal to:

$$a_p(t) = e^{i\omega_p t} a_p \ , \tag{1.48}$$

where ω_p is equals to:

$$\omega_p = \pm cp \ . \tag{1.49}$$

For any p_n the solution of equation of motion can be found, but for general solution $\phi(x,t)$ it is required to find a sum of all possible n values:

$$\phi(x,t) = \sum_{n=0}^{\infty} \sin(p_n x)\left(A_n \cos(\omega_n t) + B_n \sin(\omega_n t)\right) \ , \tag{1.50}$$

where sine and cosine replace exponents $+i\omega_n t$ and $-i\omega_n t$, because ω_p can be positive or negative due to formula (1.49). That is why it is enough to sum only from $n = 0$ to ∞. It should be noticed that there is a zero-solution when $n = 0$.

Formula (1.50) has two undetermined coefficients A_n and B_n which can be determined if specify the initial value of the field and it's first derivative at the time moment $t = 0$. It should be noticed that the initial conditions must be represented as a sum of cosines and sines of $\omega_n t$ and then we can equate coefficients between this cosines and sines with coefficients A_n and B_n from (1.50).

Summarizing the previous information, a description of the continuum system in classical mechanics was made and the expression for the displacement field was found.

Lecture 2. Tensor Fields, Euler-Lagrange Equations, Lorentz Transformations

Classical fields in many dimensional case

The concept of the displacement field can be further generalized to the case of multidimensional systems. For instance, in two dimensions it is possible to consider two-dimensional lattice of springs. In such a way it is necessary to present a plane of particles with mass m and distance between neighbors Δa (fig. 2.1).

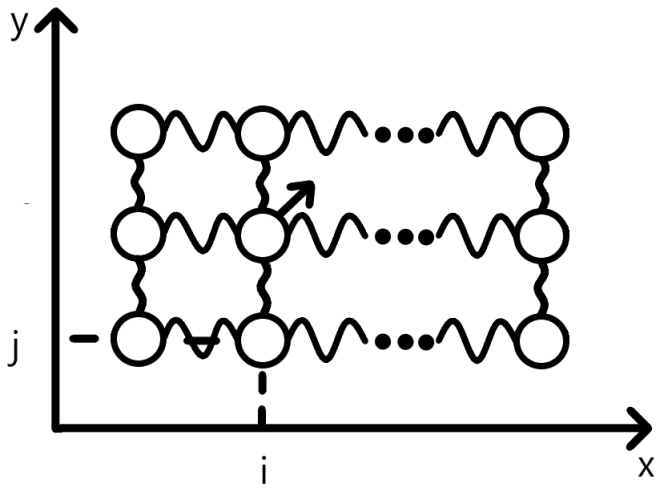

Fig. 2.1. A two-dimensional lattice of springs

Then it is necessary to allow longitudinal fluctuations of particles connected with the springs existence. In this case a natural quantity to discuss is the displacement field ϕ depending on coordinates of a point at the equilibrium position. This coordinates can be written as i and j for x and y respectively. In such a way the displacement field ϕ will be a vector with the following notation:

$$\vec{\phi}_{(i,j)} \ . \tag{2.1}$$

The continuum approximation of the presented system is called a membrane. In this case a longitudinal fluctuations of the membrane will be searched.

Further a transmission to three-dimensional case can be completed. For this purpose it is required to represent the system as it shown on fig. 2.2.

The displacement field of such system will be a three-dimensional vector $\vec{\phi}_{(i,j,k)}$ coordinates of which can be specified by indexes i, j and k for x, y and z respectively.

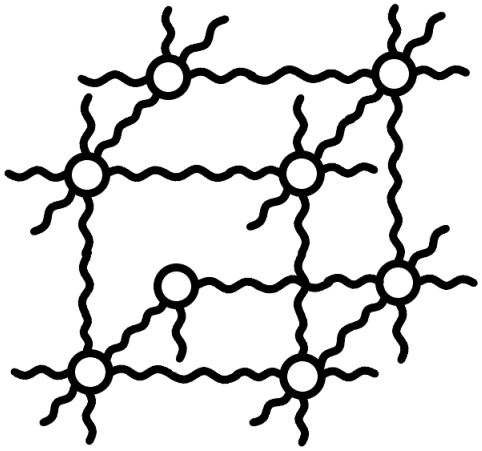

Fig. 2.2. A three-dimensional lattice of springs

Then it is necessary to complete absolutely the same derivation procedure which have been already done for the case of the one-dimensional system. At the end it is needed to end up with the following differential equation which describes the evolution of the three-dimensional vector $\vec{\phi}$. Any component of the displacement field vector will undergo the following dynamics:

$$\ddot{\vec{\phi}} - c_1\partial_{xx}\vec{\phi} - c_2\partial_{yy}\vec{\phi} - c_3\partial_{zz}\vec{\phi} - c_4\partial_{xy}\vec{\phi} - c_5\partial_{yz}\vec{\phi} - c_6\partial_{xz}\vec{\phi} = 0 \ , \tag{2.2}$$

where $\ddot{\phi}$ is the second time derivative and coefficients $c_1, \ldots, c_6$ describe a properties of the observed solid body. It also can be seen that many quantities which appear in the description are fields depending on space-time variables and this fields behaves themselves as tensors or more generally tensor fields.

Therefore it is very natural to introduce and discuss the concept of a tensor field from the beginning.

Tensors and tensor fields

As it well known tensors in any coordinate system represent a set of numbers which transforms with changing from one coordinate system to another in a definite way.

For tensor fields we are asking about transformation of the set of numbers under general coordinate transformations. Consider a transformation of coordinates of this type:

$$x^\mu \to x'^\mu\left(x^\nu\right) \ . \tag{2.3}$$

Such transformations are assumed to be invertible and this transformations belong to

the class which is directly defined as general coordinate transformations. Let's assume that the work is done in d dimensions. That is why index μ can take values from 1 to d.

Then a tensor field

$$\phi^{\mu_1\dots\mu_p}_{\nu_1\dots\nu_q}(x) \tag{2.4}$$

is a tensor field of rank (p, q).

Under general coordinate transformations it transforms in the following way. In the new coordinate system it is described as

$$\phi'^{\mu_1\dots\mu_p}_{\nu_1\dots\nu_q}(x') = \frac{\partial x'^{\mu_1}}{\partial x^{\lambda_1}}\dots\frac{\partial x'^{\mu_p}}{\partial x^{\lambda_p}}\frac{\partial x^{\rho_1}}{\partial x'^{nu_1}}\dots\frac{\partial x^{\rho_q}}{\partial x'^{\nu_q}}\phi^{\lambda_1\dots\lambda_p}_{\rho_1\dots\rho_q}(x) \ , \tag{2.5}$$

where over indexes $\lambda_1 \dots \lambda_p$ and indexes $\rho_1 \dots \rho_q$ are used for summation due to Einstein's convention.

Partial derivatives

$$\frac{\partial x'^{\mu}}{\partial x^{\lambda}} = J_{\mu\lambda} \tag{2.6}$$

can be represented as elements of the matrix which is called Jacobian matrix. This matrix is non degenerate and therefore it is the element of the group $GL(d, \mathbb{R})$, where d means dimension and $\mathbb{R}$ means that the Jacobian matrix is real. The group $GL(d, \mathbb{R})$ is a group of all invertible $d \times d$ matrices. d is a positive integer number which specifies the range of the coordinates

$$\mu,\ \lambda,\ \nu,\ \rho = 1,\ \dots,\ d\ . \tag{2.7}$$

Further restrictions on possible transformations of coordinates can be imposed by physical requirements. For instance, if you are talking about Galilean invariants then you restrict general coordinate transformations to that of the rotation group. The transition from x to x' in this situation will be only rotation.

Analogously, if you implement Einstein's relativity principle, which allows only for Lorentz transformation, than in this case you will restrict the general coordinate transformation to that of Lorentz transformations.

A simple example of a tensor is given by a scalar. Scalar in this case is an object which transform under general coordinate transformations according to the formula:

$$\phi'(x') = \phi(x) \ . \tag{2.8}$$

The scalar does not have any indexes and therefore the Jacobian matrices do not appear. It is a tensor of rank $(0, 0)$.

It's important to realize that the point in different coordinates can be specified by the coordinate x and x', but physically it is the one point.

Summarizing, the scalar field is a function, which under the change of coordinates behaves itself in the way written by formula (2.8).

It should be noticed that the scalar field is not something which behaves like (2.8). It is not true. If you take any function like, for instance, $\sin x$ and then make a next coordinate transformation:

$$x \to x'^2 \tag{2.9}$$

you will get

$$\sin x \to \sin x'^2 \ . \tag{2.10}$$

But, of course, now it is necessary to realize that the written in (2.10) expression will not the old function depending on x', because if you write

$$\phi(x) = \sin x \tag{2.11}$$

and complete a transformation written in (2.9), the $\phi(x')$ will be described by formula:

$$\phi(x') = \sin x' \ . \tag{2.12}$$

That is why it is needed to make transformed in the (2.10) expression equal to $\phi'(x')$.

Next to the scalar field is a vector field, which carries one index up. An example of a vector field would be, for instance, a velocity vector which can be also made as a vector field. For instance, if you discuss the flow of a liquid in hydrodynamics, then you need tensor field of velocities of volume elements of the liquid.

There are also some examples of a co-vector, which means a tensor with a lower index. This tensor would be a gradient of a function. More generally you might have, for instance, symmetric on the manifold which would be a tensor with two lower indices and so on. It will also be seen in the development of the course, a few other examples of tensor fields.

Action principle and Euler-Lagrange equations

Then it is necessary to return back to the action principle in general. Action principle that was already discussed plays an important role in describing of the dynamics.

In the action principle the work takes place with with the functional S, which is a functional of field ϕ. This functional has the following form:

$$S[\phi] = \int \mathrm{d}x \mathcal{L}(x, \phi_i, \partial_\mu \phi_i) \ , \tag{2.13}$$

where the $\mathrm{d}x$ can be defined as

$$\mathrm{d}x := \mathrm{d}x^0 \mathrm{d}x^1 \dots \mathrm{d}x^{d-1} \ , \tag{2.14}$$

where d is a number of space-time dimensions, where x^0 signifies the time direction.

The x^0 element can be represented by the next formula:

$$x^0 = ct \ , \tag{2.15}$$

where c is the speed of light.

Other directions such as x^1, ..., x^{d-1} are spatial directions.

Finally, action is an integral over d dimensional space-time coordinates of the Lagrangian density, which is considered in general to be a function of x, fields ϕ_i, and the first derivatives.

It should be noticed that Lagrangian density depends on only first derivatives because non-degenerate or what is called unitary theories is considered, where equations of motion, which would be derived from this section, contain only second derivatives in time of the field ϕ and second derivatives in the spatial directions. The theories with derivatives of the third and higher orders of the field ϕ are much more complicated in the classical and especially the quantum case.

An example of the theory limited by the second order is a very natural generalization of the Newtonian mechanics, because in Newtonian mechanics deriving equations have a second order in time. It means that to specify the trajectory of a particle it is needed to specify the initial coordinate and initial velocity and it is enough to solve the Koshi problem for such dynamical system which linked with the fact that a unique solution is needed. This principle was inherited from standard Newtonian mechanics, and our goal is to apply it to fields as well.

Further, it will also be required that there is no explicit dependence on x. All dependence on x will come only through the field ϕ and its derivatives:

$$S\left[\phi\right] = \int \mathrm{d}x \mathcal{L}\left(\phi_i, \partial_\mu \phi_i\right) \ , \tag{2.16}$$

where the index i has double meaning. It will combine labels of different fields when the discussion will be about fields of different nature and also it will be uniform label for all possible tensor indices of one field. In formula (2.16) it is possible to consider the Lagrangian not only for a scalar field, but, for instance, Lagrangian for a vector field, as it can be, for instance, in electrodynamics where the electromagnetic field is a vector. Also

by index i, for example, it is possible to mean the index of a vector of the one vector field, but it is also possible to mean a collection of scalars by index i.

From the action S described by formula (2.16), equations of motion are derived by using the principle of the least action. In fact, an example how this is done with the help of variational principle where variational derivative was used has already been seen. As it was described in the previous lecture the variational derivative is a variation of the action with respect to variation of the field ϕ:

$$\varepsilon_i = \frac{\delta S}{\delta \phi_i} = \frac{\partial \mathcal{L}}{\partial \phi_i} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi_i\right)}\right) = 0 \ . \tag{2.17}$$

where the gotten expression that is known as Euler-Lagrange equations. Solving of the Euler-Lagrange equations it is possible to find the actual dynamical trajectories of the system.

Then it is necessary to differentiate the Lagrangian derivative with respect to x^μ in an explicit way by taking into account that Lagrangian density is a function of ϕ_i and its derivative. It is possible to write for this Lagrangian derivative:

$$\frac{\partial \mathcal{L}}{\partial \phi_i} - \frac{\partial^2 \mathcal{L}}{\partial \phi_{i,\mu} \partial \phi_j}\phi_{j,\mu} - \frac{\partial^2 \mathcal{L}}{\partial \phi_{i,\mu} \partial \phi_{j,\nu}}\phi_{j,\mu\nu} = 0 \ , \tag{2.18}$$

where

$$\phi_{i,\mu} = \frac{\partial \phi_i}{\partial x^\mu} \ . \tag{2.19}$$

Then it is important to realize that in fact Lagrangian is not uniquely defined, which leads to the one and the same equations of motion. It is practicable to have different Lagrangians, which lead to the one and the same equations of motion. This acts because the Lagrangian density may always be changed by adding to it total derivative:

$$\mathcal{L} \to \mathcal{L} + \partial_\mu \Lambda^\mu \ , \tag{2.20}$$

where Λ^μ is considered as a function

$$\Lambda^\mu = \Lambda^\mu\left(\phi_i\right) \ . \tag{2.21}$$

which depends on ϕ_i but not of their derivatives, because if you also assume that the Λ^μ depends on the derivative of ϕ_i, then it would lead to the appearance of the second derivative of ϕ in the Lagrangian. This is something which was prohibited because it was required that the Lagrangian depends on the first derivative of ϕ.

Then it is necessary to find out why the equations of motion will not be changed under the addition. This acts because of the next expression:

$$\partial_\mu \Lambda^\mu = \frac{\partial \Lambda^\mu}{\partial \phi} \partial_\mu \phi \ , \tag{2.22}$$

where it was considered that we deal with one scalar field ϕ.

Then it is possible to trace how under this addition the Lagrangian derivative changes. It is known that Lagrangian derivative is given by expression (2.17). If we want to observe how the Lagrangian derivative changes, we need to substitute the additional term instead of Lagrangian in (2.17) and see what happens. This means that the Lagrangian derivative will get an extra contribution depending on Λ:

$$\varepsilon \to \varepsilon + \delta\varepsilon\left(\Lambda\right) \ , \tag{2.23}$$

where the extra contribution $\delta\varepsilon\left(L\right)$ will be given by:

$$\delta\varepsilon\left(L\right) = \frac{\partial}{\partial\phi}\left(\frac{\partial\Lambda^\lambda}{\partial\phi}\partial_\lambda\phi\right) - \frac{\partial^2}{\partial\phi\partial\phi_{,\mu}}\left(\frac{\partial\Lambda^\lambda}{\partial\phi}\phi_{,\lambda}\right)\phi_{,\mu} - \frac{\partial^2}{\partial\phi_{,\mu}\partial\phi_{,\nu}}\left(\frac{\partial\Lambda^\lambda}{\partial\phi}\phi_{,\lambda}\right)\phi_{,\mu\nu} = 0 \ . \tag{2.24}$$

It can be seen from (2.24) that the additional contribution $\delta\varepsilon\left(L\right)$ is zero by itself. First of all, if you look at the last term in (2.24), you will see that the Λ^λ depends on ϕ we have its first derivative. If we differentiate the expression in the brackets

$$\left(\frac{\partial\Lambda^\lambda}{\partial\phi}\phi_{,\lambda}\right) \tag{2.25}$$

with respect to derivatives, since we have two derivatives before brackets, the result of the differentiation will be zero, because of the fact that expression in the brackets involve the first derivative of ϕ only once. Remaining two terms will cancel each other and it can be checked by opening the brackets and evaluating of the derivatives. Finally, the contribution to the Euler-Lagrangian equations from the total derivative presented in (2.24) is actually zero. That is why an additional term does not contribute to equations of motion.

Lorentz and Poincare groups

After the general remarks presented in previous section it is now possible to come to the discussion of Lorentz and Poincare groups.

As it was already mentioned in the first lecture, symmetries play an extremely important role in the description of nature. For instance, the translational symmetry implies that if

we know physical loss at one space-time point, then we know them at any other which means that in fact, the physical losses are not random from point to point. Similar remarks concern, for instance, rotational symmetry, which means that losses in different directions can be related to each other. Among all possible symmetries, there is one which has a universal character and this is Poincare symmetry, which combines Lorentz symmetry or Lorentz transformations with shifts of space-time coordinates.

From mathematical point of view, these transformations have a structure of a Lie group. First of all, before describing of the Poincare symmetry, it is needed to introduce the notion of the Lie group which will be defined as G. The group G is a set of elements of any nature, which satisfies the following set of axioms:

1) For any two elements G_1 and G_2, which belong to group G, one can define their product, $G1 \cdot G2$, which also belongs to the group G. The product of G_1 and G_2 is associative, which means that it does not matter how to put the brackets:

$$(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3) \ . \tag{2.26}$$

2) Group G must include a unit element e such that:

$$g \cdot e = e \cdot g = g \ , \tag{2.27}$$

where a unit element can be also called as identity.

3) For any element g from the group G, there exists an inverse element g^{-1} such that:

$$g \cdot g^{-1} = g^{-1} \cdot g = e \ . \tag{2.28}$$

An important class of groups constitute Lie groups. A smooth manifold G of dimension n is called Lie group, if G is supplied with a structure of a group and compatible with a structure of a smooth manifold, which means that the group operations are smooth.

A physical important example of Lie group, which can be well known from courses on classical mechanics is a rotation group. The rotation group is the group of rotational matrices in three dimensions. The explicit discussion of this group will be held a little bit later, because it is a subgroup of the Lorentz group.

Another example, which is primarily important, is a Lorentz group. Now consider a d dimensional Minkowski space $\mathbb{R}^{1,d-1}$. The Lorentz transformations by definition are linear coordinate transformations of Minkowski space of the form:

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu} \ , \tag{2.29}$$

where over index ν there is a summation in the Einstein's term.

These transformations of a Minkowski space are such, that they preserve the interval between two events in a Minkowski space. The infinitesimal relativistic interval ds is defined through the quadratic form:

$$ds^2 = \eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu = c^2 \mathrm{d}t^2 - \left(\mathrm{d}x^1\right)^2 - \left(\mathrm{d}x^2\right)^2 - \left(\mathrm{d}x^3\right)^2 \ , \tag{2.30}$$

where $\eta_{\mu\nu}$ is Minkowski metric, which is a diagonal matrix and indexes μ and ν are running in general from 0 to 3, if the four dimensional Minkowski space is discussing:

$$\eta_{\mu\nu} = \mathrm{diag}\left(+1,\, -1,\, -1,\, -1\right) \ , \ \mu\,, \nu \, = 0,\, 1,\, 2,\, 3 \ . \tag{2.31}$$

Now it can be seen that x^μ is a vector with components x^0, x^1, x^2, x^3:

$$x^\mu = \left(x^0,\, x^1,\, x^2,\, x^3\right) \ . \tag{2.32}$$

If you treat x^μ as a vector, the Lorentz transformations can be written in the matrix form

$$x' = \Lambda x \ , \tag{2.33}$$

cover vector x and it transforms passing from one Lorentz frame into another, simply with matrix Λ like:

$$\Lambda = |\Lambda^\mu_\nu| \ , \ \mu,\, \nu \, = \, 0,\, 1,\, 2,\, 3 \ . \tag{2.34}$$

The requirement of Λ to preserve the quadratic form is an explicit form written in the following way:

$$\eta_{\mu\nu} \Lambda^\mu \alpha \Lambda^\nu_\beta = \eta_{\alpha\beta} \ . \tag{2.35}$$

In other words, Lorentz transformations are transformations which leave our Minkow's metric invariant. Formula (2.35) defines the class of matrices Λ, which can be associated with Lorentz transformations.

The transformation from formula (2.35), which preserves Minkowsky metric can also be written in the matrix form:

$$\Lambda^t \eta \Lambda = \eta \ , \tag{2.36}$$

where η and Λ is a four by four matrix.

Definition presented in (2.36) is a generalization of the standard condition

$$\Lambda^t \Lambda = \mathbb{1} \ , \tag{2.37}$$

which is known for orthogonal matrices representing the rotation group. Because of the fact that Minkowsky metric is used in expression (2.36), the condition in (2.36) can be called as pseudoorthogonality condition arising from the fact that now the work with Minkowsky space takes place rather than the work with Euclidean d dimensional space.

It can be showed that matrices, which satisfies condition (2.36), form a group. It is clear that as a product in this group, it is possible to consider the usual product of matrices.

It should be also noticed that condition (2.36) works only for Minkowsky space. It is a condition which is directly related with the definition of something that is meant by Lorentz transformations. Lorentz transformations are transformations which preserve Minkowsky metric η. It is also practicable to define a group which preserves this metric in general relativity, or in special relativity, such a group is called a group of isometrics. Group of isometrics is a group which preserves a given metric. That is why it is possible to say in this respect that the Lorentz group is a group of isometrics of Minkowsky metric. If we want a different constant metric instead of η, it is possible to use, if we have a different manifold, not Minkowsky space. Then it is necessary to consider what a different space with a different metric and also define a group of isometrics in a similar way, which was completed before.

Let's take two matrices, Λ_1 and Λ_2, which satisfy the condition (2.36) and then consider the product of these two matrices. It is needed also to show that the condition (2.36) works for Λ_1 and Λ_2:

$$\begin{cases} \Lambda_1^t \eta \Lambda_1 = \eta \\ \Lambda_2^t \eta \Lambda_2 = \eta \end{cases} \tag{2.38}$$

meaning that the expression

$$(\Lambda_1 \Lambda_2)^t \eta \Lambda_1 \Lambda_2 = \eta \tag{2.39}$$

will be satisfied.

Expression (2.39) would be also a Lorentz transformation. It is easy to proof that (2.39) is right:

$$\Lambda_2^t \underbrace{\Lambda_1^t \eta \Lambda_1}_{=\eta} \Lambda_2 = \Lambda_2^t \eta \Lambda_2 = \eta \ . \tag{2.40}$$

Finally, from (2.40) it is clearly seen that the product of two Lorentz transformations is a Lorentz transformation. Identity matrix would be a trivial Lorentz transformation, which is a identity matrix. Then it is needed to show that any Λ has an inverse, which is also a Lorentz transformation. To complete this it is possible to take and compute the

determinant of both sides:

$$\det\left(\Lambda^t\eta\Lambda\right) = \left(\det\Lambda\right)^2\det\eta\ , \tag{2.41}$$

where the rule that under transposition the determinant is unchanged was used.

On the other hand, by the definition of Lorentz transformation, expression (2.41) is the same as the determinant of η. The determinant of η is not equal to zero by the definition and equal to minus one. That is why

$$\left(\det\Lambda\right)^2 = 1\ . \tag{2.42}$$

From formula (2.42), it was concluded that all matrices of Lorentz transformations always have

$$\det\Lambda = \pm 1\ . \tag{2.43}$$

But this means that matrix of Lorentz transformation is non-degenerate and therefore it is invertible. The inverse is easy to find from the defining relation of the Lorentz transformations:

$$\Lambda^t\eta\Lambda = \eta \Rightarrow\ \Lambda^t\eta = \eta\Lambda^{-1} \Rightarrow\ \eta\Lambda^t\eta = \eta^2\Lambda^{-1} = \Lambda^{-1}\ . \tag{2.44}$$

Finally, we got that

$$\Lambda^{-1} = \eta\Lambda^t\eta\ . \tag{2.45}$$

Therefore to show that Λ^{-1} is also Lorentz transformation, it is necessary to check the following:

$$\left(\Lambda^{-1}\right)^t\eta\Lambda^{-1} = \eta\ . \tag{2.46}$$

It may be done by substituting the expression for Λ^{-1}:

$$\left(\Lambda^{-1}\right)^t\eta\Lambda^{-1} = \left(\eta\Lambda^t\eta\right)^t\underbrace{\eta\eta}_{\mathbb{1}}\Lambda^t\eta = \eta\underbrace{\Lambda\eta\Lambda^t}_{\eta}\eta = \eta\underbrace{\eta\eta}_{\mathbb{1}} = \eta\ , \tag{2.47}$$

where the statement that

$$\Lambda\eta\Lambda^t = \eta \tag{2.48}$$

was used. Expression (2.48) follows from the defining relation

$$\Lambda^t\eta\Lambda = \eta \tag{2.49}$$

Expression (2.48) means that Λ^t is also a matrix of Lorentz transformation. By this moment it is clear that if Λ is a matrix of Lorentz transformations, Λ^t is also a matrix of Lorentz transformation. And this can be gotten again from the same expression for Λ^{-1}:

$$\Lambda^{-1} = \eta\Lambda^t\eta \mid \Lambda \cdot \tag{2.50}$$

$$\mathbb{1} = \Lambda\eta\Lambda^t\eta \mid \cdot\,\eta \tag{2.51}$$

$$\eta = \Lambda\eta\Lambda^t\eta^2 = \Lambda\eta\Lambda^t \ . \tag{2.52}$$

In other words, we got that if Λ is a Lorentz transformation, is a Lorentz transformation, then the matrices Λ^{-1}, Λ^t and $\Lambda^{t\,-1}$ are matrices of Lorentz transformations. According to definition of the group, Lorentz transformations for described group can be associated with the usual matrix multiplication.

As can be seen, defining the relation of the Lorentz group implies the following important statement: if the condition of preservation of the Lorentz metric is taken and the low index of Λ is specified to be zero, then the following will be obtained:

$$\eta_{\mu\nu}\Lambda^\mu_0\Lambda^\nu_0 = \left(\Lambda^0_0\right)^2 - \left(\Lambda^i_0\right)^2 = \eta_{00} = 1 \ . \tag{2.53}$$

From expression (2.53) can be seen that:

$$\left(\Lambda^0_0\right)^2 = 1 + \left(\Lambda^i_0\right)^2 \geq 1 \ . \tag{2.54}$$

Expression in (2.54) is always bigger or equal to one, because a non-negative contribution $\left(\Lambda^i_0\right)^2$ is added to one. Therefore there are two possibilities for the component Λ^0_0:

$$\Lambda^0_0 \geq 1 \ \textit{or} \ \Lambda^0_0 \leq -1 \ . \tag{2.55}$$

The correspondent Lorentz group is a six dimensional non-compact Lie group whose mathematical name is $O\,(1,\,3)$, where O stands for orthogonal. This group consists of four topologically separated spaces.

Now it is possible discuss the following statements. First of all, Lorentz transformations may reverse the direction of time or not reverse the direction of time, e.g. keep the direction of time. In other words, they can transform a future point in time like vector into a past point in one. In other words, it's related to the effect if condition (2.55) completes.

The second thing is that Lorentz transformation reverse or not reverse the four dimensional reference frame. This related to the effect if

$$\det\Lambda = 1 \ \textit{or} \ \ \det\Lambda = -1 \ . \tag{2.56}$$

We have two operations, which are also Lorentz transformations, and there are discrete operations. One of them is parity and another is time reversal. Time reversal operation will reverse direction of time and the parity will change the orientation of the frame and transform the matrices with a determinant equal to one with two matrices with determinant equal to minus one and vice versa.

Λ's for which $\Lambda^0_0 \geq 1$ preserve the direction of time. For this reason they are called orthochronous. The product of two orthochronous transformations is also an orthochronous transformation. It can be seen easily if the following is noticed

$$\left|\Lambda^0_0\right| > \left\|\Lambda^i_0\right\| \ . \tag{2.57}$$

This implies that modulus of Λ^0_0 is bigger than the norm of the vector Λ^i_0.

Now Λ^i_0 is a 3-dimensional vector and then $\|\Lambda^i_0\|$ is the square of its norms and it's naturally follows from the relation (2.54).

Analogously, if in the relation (2.54) $\Lambda \to \Lambda^t$ is changed, which is also Lorentz transformation, the following will be obtained:

$$\left|\Lambda^0_0\right| > \left\|\Lambda^0_i\right\| \ . \tag{2.58}$$

Therefore, if the product of two Lorentz transformations Λ and Λ' is taken, it will be:

$$(\Lambda\Lambda')^{\,0}_{\,0} = \Lambda^0_0\Lambda'^0_{\ 0} + \Lambda^0_i\Lambda'^i_{\ 0} \ . \tag{2.59}$$

Now if the Cauchy-Bonyakowski-Schwarz inequality is used, which essentially tells that for any two vectors x and y the following is found:

$$|(x,y)| \leq \|x\|\,\|y\| \ . \tag{2.60}$$

It will be obtained that:

$$\left|\Lambda^0_i\Lambda'^i_{\ 0}\right| \leq \left\|\Lambda^0_i\right\|\left\|\Lambda'^i_{\ 0}\right\| \ . \tag{2.61}$$

Therefore, if it is assumed that Λ^0_0 and $\Lambda'^0_{\ 0}$ are both positive, it will be concluded that:

$$(\Lambda\Lambda')^{\,0}_{\,0} > 0 \ , \tag{2.62}$$

which means that Λ^0_0 must be bigger or equal to one, because any Lorentz transformation with positive Λ^0_0 must have the property that it's zero-zero component is bigger or equal to one.

In other words, if we have two orthochronous transformations, then the product of these two orthochronous transformations is also an orthochronous transformation, which means that orthochronous transformations by itself form a group.

It can be also shown that the inverse of the orthochronous transformation is an orthochronous transformation. And this group or subgroup of the Lorentz group, which consists of orthochronous transformations is denoted as $O^+(1,3)$, where plus is related to the fact that we preserve the direction of time.

Now concerning the determinant, the Lorentz transformations, which preserve orientation are called proper. These are transformations with determinant equals to $+1$. The transformations with determinant equals to -1 are called improper.

The proper Lorentz transformations form a subgroup of the Lorentz group, which is denoted as $SO(1,3)$, where the letter S as usual means special.

Now it is possible to combine these two properties of being orthochronous and being proper and get a subgroup, which is called a restricted Lorentz group. A restricted Lorentz group denoted as $SO^+(1,3)$ and consists of proper orthochronous Lorentz transformations. Actually identity matrix is a proper orthochronous Lorentz transformation.

Now it is possible to get the picture of how topologically the Lorentz group looks like:

1) A component which consists of all Lorentz transformations, which have a $\Lambda^0_0 \geq 1$, and which have $\det\Lambda = 1$. This is the restricted Lorentz group $SO^+(1,3)$.

2) A component which contain $\Lambda^0_0 \leq -1$ and $\det\Lambda = -1$. This component presents improper transformations, inverting the direction of time. The passage from restricted Lorentz group to this component is done by application of what is called time reversal. Time reversal is a particular Lorentz transformation, which looks like:

$$T = \mathrm{diag}\,(-1,\,1,\,1,\,1)\ . \tag{2.63}$$

3) A component which contain $\Lambda^0_0 \geq 1$ and $\det\Lambda = 1$. This component is obtained from the proper orthochronous subgroup by means of application, what is called parity or space inversion denoted by P:

$$P = \mathrm{diag}\,(1,\,-1,\,-1,\,-1)\ . \tag{2.64}$$

As can be seen from (2.64), parity changes orientation of the three spatial coordinate axes by multiplying them by minus one.

4) A component which contain $\Lambda_0^0 \leq -1$ and $\det \Lambda = 1$. That is why all transformations of the component are not orthochronous. The component is obtained from restricted Lawrence subgroup by means of combined application of a space inversion with the time inversion. Combined application keeps the determinant equals to one, but since the time reversal is employed, then what will be obtained are matrices of non-orthochronous Lawrence transformations.

Discrete operations, which relates this topologically distinguished or topologically different components of the Lawrence group, are the following discrete operations:

$$\{e,\, P,\, T,\, PT\} \ . \tag{2.65}$$

Presented discrete transformations also form what is called a discrete group or discrete subgroup of the Lawrence group $O\,(3,1)$, which allows the movement among these different topologically different components.

Described list of components suggests that only component with $\Lambda_0^0 \geq 1$, and $\det \Lambda = 1$ here is an actual subgroup of the Lawrence group. Other components, if we take them by themselves, do not form a subgroup. There is only one subgroup and this subgroup is a subgroup of proper orthochronous Lawrence transformations because only this component contains an identity group. The connected subgroup of the Lorentz group of dimension six is $SO^+\,(1,3)$.

The Lorentz group is six dimensional Lie group because of the fact that Lawrence transformations combine three rotations around three coordinate axes: x, y and z. In addition, there are three boosts, which involve one special direction and the time direction. If you have a four dimensional Minkowsky space, the picture of possible transformations will look like presented on fig. 2.3.

Then it is necessary to remind how any rotation can be parameterized. It is possible to do it by specifying the rotation axis $\vec{n}$ and the rotation can be done by specifying a unit vector. If $\vec{n}$ is a unit vector, it is also needed to specify a rotation angle around this unit vector, and the rotation angle will be θ.

If the following three by three matrices are defined

$$a_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\ ,\ a_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}\ ,\ a_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\ , \tag{2.66}$$

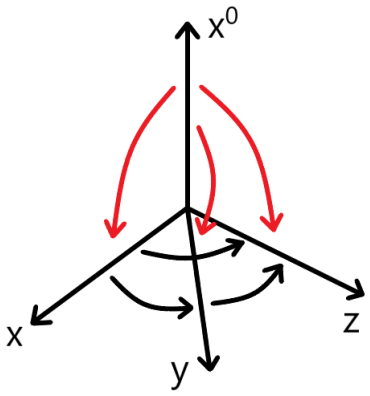

Fig. 2.3. The picture of possible transformations of a four dimensional Minkowsky space

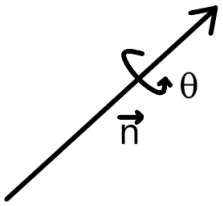

Fig. 2.4. The rotation axis $\vec{n}$ and the rotation angle θ

then such a transformation is described by a matrix

$$R\left(\vec{n},\theta\right)=e^{\theta\vec{n}\vec{a}}\ , \tag{2.67}$$

where $\vec{a}$ comprises three matrices a_1, a_2, a_3:

$$\vec{a}=\left(a_1,a_2,a_3\right)\ . \tag{2.68}$$

$R\left(\vec{n},\theta\right)$ is a three by three matrix of rotations specified by the direction $\vec{n}$ and the angle θ. It is possible to compute this matrix quite explicitly. In fact, the scalar product $\vec{n}\vec{a}$ can be written as:

$$\left(\vec{n}\vec{a}\right)_{ij}=-\varepsilon_{ijk}n_k\ . \tag{2.69}$$

Then it is practicable to exponentiate the product simplified by formula (2.69) and this gives us an explicit three by three matrix with the following matrix elements. After computation of this matrix elements the following expression will be obtained

$$R_{ij}\left(\vec{n},\theta\right)=\cos\theta\,\delta_{ij}+\left(1-\cos\theta\right)n_in_j-\sin\theta\,\varepsilon_{ijk}n_k\ . \tag{2.70}$$

It can be easily checked that the matrix R is orthogonal. It satisfies the properties that

$$R^t\cdot R=R\cdot R^t=1\ . \tag{2.71}$$

It means that this matrix is an element of the group of rotation $SO(3)$.

It also should be said that

$$-\pi < \theta \leq \pi \ , \tag{2.72}$$

which covers the whole interval 2π.

But in fact, it is possible to restrict the rotation angle to run from 0 to π, because the negative values of θ from $-\pi$ to 0 are equivalent to changing the direction of vector $\vec{n}$ to the opposite. The vector $\vec{n}$ can be allowed to be an identity or a unit vector, which actually runs over θ or run over the two dimensional sphere.

Very often people use this model for a topological description of the rotation group (fig. 2.5). They say that the angle can be identified with the lens of the vector $\vec{n}$. This vector is not a unit one, but the wall of a radius π is felt. Any point inside this wall of radius π will represent a rotation in the direction, specified by the line connecting a zero point with a point inside the ball. The rotation angle will be given by the lens of the interval connecting this point with the origin. A bit non-trivial in this description of the rotations is that diametrally opposite points on a sphere should be identified because of the fact that a rotation on angle π and $-\pi$ will give the same rotation.

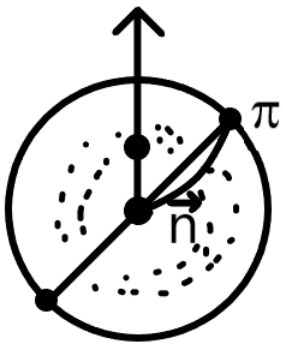

Fig. 2.5. The model for a topological description of the rotation group

Now it is possible to embed the rotation group into the Lorenz group and embedding is done in the following way:

$$\Lambda(n,\theta) = \begin{pmatrix} 1 & 0 \\ 0 & R(\vec{n},\theta) \end{pmatrix} \ , \tag{2.73}$$

where the time direction is untouched and represented by 1.

It also can be checked about the matrix R that it satisfies the property that

$$R(\vec{n},\theta+2\pi) = R(\vec{n},\theta) \ . \tag{2.74}$$

Boosts

Now let's come to boosts. Boosts will be represented by four by four matrices:

$$b_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; b_1 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; b_1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} . \tag{2.75}$$

Now there will be rotation, but around what is called hyperbolic angle. The corresponding Lorentz transformations are now specified again by the axis around which the rotation is happened and by the rotation angle, which is called ϑ:

$$\Lambda\left(\vec{n}, \vartheta\right) = e^{\vartheta \vec{n} \vec{b}} , \tag{2.76}$$

where $\vec{b}$ is equal to:

$$\vec{b} = \left(b_1, \, b_2, \, b_3\right) \; . \tag{2.77}$$

Lecture 3. Lorentz and Poincare Groups, Noether's Theorem

Repeating of the rotation group

At the last discussion it was considered that any rotation can represented by the rotation $R(\vec{n},\theta)$ given by the formula:

$$R(\vec{n},\theta) = e^{\theta\vec{n}\cdot\vec{a}} \ . \tag{3.1}$$

If the corresponding exponent is computed, then a rotation matrix will be obtained

$$R_{ij}(\vec{n},\theta) = \cos\theta\,\delta_{ij} + (1-\cos\theta)\,n_i n_j - \sin\theta\varepsilon_{ijk}n_k \ , \tag{3.2}$$

where i and j runs from 1 to 3.

It can be easily checked that the matrix R is orthogonal, which means that

$$R^t R = RR^t = \mathbb{1} \ . \tag{3.3}$$

Also, the matrix R is periodic:

$$R(\vec{n},\theta+2\pi) = R(\vec{n},\theta) \ . \tag{3.4}$$

As can be seen, R represents three dimensional rotations and embed into a six dimensional Lawrence group in the following way:

$$\Lambda(\vec{n},\theta) = \begin{pmatrix} 1 & 0 \\ 0 & R(\vec{n},\theta) \end{pmatrix} \ , \tag{3.5}$$

where Λ is a matrix of Lorentz transformation.

Lorentz boosts

When we speak about Lorentz boosts, it is necessary to consider a similar parametrization with the rotation group, but instead of three by three matrices it is needed to consider four by four matrices b_i which have a form described in previous lection.

Exponentiating these matrices, but now multiplied with the angle ϑ, will give the matrix of Lawrence boosts:

$$\Lambda(\vec{n},\vartheta) = e^{\vartheta\vec{n}\vec{b}} \ . \tag{3.6}$$

If the corresponding exponentiation is performed, the following matrix will be found

$$\Lambda\left(\vec{n},\vartheta\right)=\begin{pmatrix}\cosh\vartheta & -n_i\sinh\vartheta\\ -n_i\sinh\vartheta & \delta_{ij}\left(\cos\vartheta-1\right)n_in_j\end{pmatrix}\ , \tag{3.7}$$

where Λ now it is parameterized by the hyperbolic angle and the name hyperbolic is related with the hyperbolic functions: cosh and sinh.

It is also should be noticed that the hyperbolic angle ϑ is not running from 0 to 2π, but it runs from $-\infty$ to $+\infty$. Therefore, ϑ is usually called rapidity because it is related to a velocity with which we boost the Lorentz system with respect to the reference system.

It can be seen that $\cosh\vartheta$ is always bigger or equal to one

$$\cosh\vartheta\geq 1\ . \tag{3.8}$$

By the property, it is a variable which takes values from 1 to $+\infty$ and, for this reason, it is possible to parameterize it as

$$\cosh\vartheta=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\ , \tag{3.9}$$

where v is a three dimensional velocity vector and c is the speed of light. In other words, the rapidity is related to velocity by means of formula (3.9).

Now, the transformations, as can be seen from (3.9), are now correspond to the orthochronous transformations. Clearly this means that the component Λ_{00} of the Lawrence transformation is bigger than 0, which means that the work takes place with orthochronous Lorentz transformations.

Equation (3.9) can be solved for the modulus of velocity and if we do that, the following answer will be found

$$|\vec{v}|=\pm c\,\tanh\vartheta\ . \tag{3.10}$$

Let's take from (3.10) the expression with the plus sign. If the plus sign is selected, it is necessary to restrict the variable ϑ to run from 0 to $+\infty$ because in this case tanh will be positive and we will get on the right side the non negative expression.

In this case, if this choice is made, it will be seen that the variable $\sinh\vartheta$ will be given by

$$\sinh\vartheta=\frac{|\vec{v}|}{c\sqrt{1-\frac{\vec{v}^2}{c^2}}}\ . \tag{3.11}$$

Now it is possible to use the parametrization of cosh and sinh in the formula for Lorentz transformation

$$\Lambda\left(\vec{v}\right)=\begin{pmatrix} \frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}} & -\frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}}\frac{\vec{v}^+}{c} \\ -\frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}}\frac{\vec{v}}{c} & \mathbb{1}+\left(\frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}}-1\right)\frac{\vec{v}\otimes\vec{v}^+}{\vec{v}^2} \end{pmatrix}\ , \tag{3.12}$$

where Λ is four by four matrix and depends on $\vec{v}$, which represented in (3.12) as a column of size 3. It also should be noticed that $\vec{v}^+$ is a row of size 3.

Variable v in formula (3.12) is the velocity with which a moving frame is boosted with respect to the original inertial frame.

The expression $\vec{v}\otimes\vec{v}^+$ means that $\vec{v}$ is multiplied by transposed $\vec{v}$ in terms of the tensor multiplication

$$\vec{v}\otimes\vec{v}^+=\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}\otimes\begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix}\ , \tag{3.13}$$

where for an element with indexes ij we will get

$$\left(\vec{v}\otimes\vec{v}^+\right)_{ij}=v_iv_j\ . \tag{3.14}$$

Now let's formulate the general form of Lorentz transformations. A four dimensional vector $\mathbb{X}$ can be transformed to a vector $\mathbb{X}'$ by a rotation matrix and also by a matrix of Lawrence boost:

$$\mathbb{X}'=\underbrace{\Lambda\left(\vec{n},\theta\right)\Lambda\left(\vec{v}\right)}_{\text{generic Lorentz transformation}}\mathbb{X}\ . \tag{3.15}$$

From formula (3.15) there are three parameters related to rotations and three parameters in the velocity vector related to Lawrence boost. All together this parameters form a six parametric group of transformations.

If expression (3.15) is written more explicitly, it will be obtained that

$$\begin{pmatrix} ct' \\ x^{1\prime} \\ x^{2\prime} \\ x^{3\prime} \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & R\left(\vec{n},\theta\right) \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}} & -\frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}}\frac{\vec{v}^+}{c} \\ -\frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}}\frac{\vec{v}}{c} & \mathbb{1}+\left(\frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}}-1\right)\frac{\vec{v}\otimes\vec{v}^+}{\vec{v}^2} \end{pmatrix}\begin{pmatrix} ct \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}\ . \tag{3.16}$$

If expression (3.16) is written for components by multiplication of the vector on the right side by two matrices, it will be possible to derive the transformation loss for

individual components of the four dimensional vector. For instance, for t' it will be found that t' is related to the original coordinates by the following formula

$$t' = \frac{t - \frac{(\vec{x}\cdot\vec{v})}{c^2}}{\sqrt{1-\frac{\vec{v}^2}{c^2}}} \ . \tag{3.17}$$

The three dimensional vector $\vec{x}'$ is related to the original vector $\vec{x}$ and time t by the formula

$$\vec{x}' = R\left(\vec{n},\theta\right)\left[\vec{x} - \frac{\vec{v}t}{\sqrt{1-\frac{\vec{v}^2}{c^2}}} + \left(\frac{1}{\sqrt{1-\frac{\vec{v}^2}{c^2}}} - 1\right)\frac{\vec{v}\cdot(\vec{v}\cdot\vec{x})}{\vec{v}^2}\right] \ . \tag{3.18}$$

We may now recognize that expressions in (3.17) and (3.18) are standard formulas of special relativity. (3.17) represents a transformation of time coordinate t, and (3.18) represents a transformation of the spatial coordinates x under Lawrence transformations governed by the rotation parameters, θ and by the velocity vector. These formulas are standard formulas which may be recognized from the course of special relativity.

As an exercise formulas can be used, for instance, to derive the law of addition of velocities in special relativity. For that it is necessary to consider 2 successive Lawrence transformations, which are boosts for velocities v^1 and v^2, and use this formulas in order to see how the velocity add in special relativity.

Now, it is possible to extend the Lorentz group to the Poincare group by including shifts of space time coordinates of the following form

$$x^\mu \to \xi'^\mu = x^\mu + a^\mu \ , \tag{3.19}$$

where a^μ is a constant shift which is a constant four dimensional vector.

Lie algebra

Lie algebra is important notion, which is related to the notion of a Lie group. And a Lie algebra is an infinitesimal version of a Lie group.

Mathematically a Lie algebra is a linear space, which is supplied an operation called commutator, which is bilinear, skew symmetric and satisfies a Jacobi identity.

If two elements of a vector space v and w are applied, the commutator of this two vectors will have the next property:

$$[v,w] = -\left[w,v\right] \ . \tag{3.20}$$

The Jacobi identity means that if three elements v, w and u are taken, next cyclic permutation will be obtained:

$$[[v,w],u]+[[u,v],w]+[[w,u],v]=0\ , \tag{3.21}$$

where the operation written in (3.21) must be bilinear. Such a linear space is called a Lie algebra.

In fact, every Lie group has a Lie algebra, and in particular a Poincare group has a Lie algebra. And this Lie algebra is spent by generators, which we will denote by $M_{\mu\nu}$, where μ and ν running from 0 to 3. The $M_{\mu\nu}$ will be Lorenz generators. In addition to Lorenz generators, there will be also generators responsible for shifts, which will be denoted as P_μ and call as momentum generators.

It will be seen a little bit later that space-time shifts related to momentum of our dynamical system and it's very natural to call the P_μ momentum generators.

The Lie algebra relations between generators $M_{\mu\nu}$ and generators P_μ, which make a Lie algebra of the Poincare group, are given by the following commutation relations:

$$\begin{cases}[M_{\mu\nu},M_{\rho\sigma}]=i\left(-\eta_{\nu\rho}M_{\mu\sigma}+\eta_{\mu\rho}M_{\nu\sigma}+\eta_{\nu\sigma}M_{\mu\rho}-\eta_{\mu\sigma}M_{\nu\rho}\right)\\ [M_{\mu\nu},P_\sigma]=\left(\eta_{\mu\sigma}P_\nu-\eta_{\nu\sigma}P_\mu\right)\\ [P_\mu,P_\nu]=0\end{cases} \tag{3.22}$$

Exponentiating of generators of the Lie algebra will give the corresponding group elements and, in fact, this was already done , for instance, for a case of the rotation group. It was seen that the group element is constructed by means of exponentiating of matrices a, but the matrices a are elements of the Lie algebra of the rotation group $SO\,(3)$ and the corresponding commutator of a_i and a_j of matrices is easy to compute:

$$[a_i,a_j]=\varepsilon_{ijk}a_k\ . \tag{3.23}$$

In quantum theory, the Poincare generators become Hermitian operators. Hermitian operators means that $M_{\mu\nu}$ and P_μ are represented as operators, which satisfy the condition (3.22).

A Hermitian operators also satisfy the following condition:

$$M^+_{\mu\nu}=M_{\mu\nu}\ ,\ P^+_\mu=P_\mu\ . \tag{3.24}$$

The presented generators will be explicitly realized on a certain space, called the Fock space. We will come to the discussion of the Fock space a little bit later.

A scheme which we want to achieve in quantum field theory connected with searching of so called unit area and realization of unitary representations of the Poincare group.

As it can be known unitary representations play an important role in quantum mechanics and also as a consequence in the quantum field theory because unitarity is naturally related to the conservation of probabilities.

In quantum field theory it is also necessary to achieve unitary representations of our symmetry groups and one of the important symmetry groups is the Poincare group, which is a symmetry group of the space-time.

Therefore, the main goal will be to obtain the unitary representation of the Poincare group. Finally, we can also introduce important relation between the proper orthochronous Lorentz group, whose name is $SO^{+}(1,3)$, and the special linear group of complex 2 by 2 matrices, whose mathematical name is $SL(2,\mathbb{C})$.

There are three Pauli matrices σ_i, which are well known from quantum mechanics

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \ \sigma^1 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \ \sigma^1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{3.25}$$

Pauli matrices presented in (3.25) can be also combined with the unit matrix σ^0, which is the unit 2 by 2 matrix

$$\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ . \tag{3.26}$$

All Pauli matrices can be combined in one definition of a matrix σ^μ in the following way

$$\sigma^\mu = \left(\sigma^0, \sigma^i\right) \ , \tag{3.27}$$

where i can be equal to 1, 2 and 3.

Simultaneously, a notion of matrices can introduced

$$\overline{\sigma}^\mu = \left(\sigma^0, -\sigma^i\right) \ . \tag{3.28}$$

σ^μ and $\overline{\sigma}^\mu$ are collections of 2 by 2 matrices. Then, having a four dimensional vector x^μ, a 2 by 2 Hermitian matrix can be defined by means of the following construction:

$$X = \sum_{\mu=0}^{3} x^\mu \sigma^\mu = \sum_{\mu=0}^{3} \eta_{\mu\nu} x^\mu \overline{\sigma}^\nu \ . \tag{3.29}$$

If the matrices σ in (3.29) is substituted , it will be seen that explicitly, we construct the 2 by 2 matrices of the following type:

$$X = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} \ . \tag{3.30}$$

It's easy to see that since x^μ are real, the matrix X is Hermitian, which means that under commission conjugation it stays invariant:

$$X^+ = X \ . \tag{3.31}$$

If there is such a commission matrix, it is possible to uniquely determine the original components x^μ by taking a trace of x with a metric σ^μ:

$$x^\mu = \frac{1}{2}\mathrm{Tr}\left(X \cdot \sigma^\mu\right) \ . \tag{3.32}$$

It can be easily verified that computing a trace from (3.32) will lead to turning back to the original coordinates x^μ.

Finally, it also can be seen that if the determinant of the commission matrix X is computed, it will be

$$\det X = \left(x^0 + x^3\right)\left(x^0 - x^3\right) - \left(x^1 - ix^2\right)\left(x^1 + ix^2\right) \ . \tag{3.33}$$

If (3.33) is simplyfied, it will be obtained that

$$\det X = \left(x^0\right)^2 - \left(x^1\right)^2 - \left(x^2\right)^2 - \left(x^3\right)^2 \ . \tag{3.34}$$

Expression (3.34) is nothing else as a four interval, which is Lawrence invariant.

Now it is possible to define on the commission matrix X the action of the group SL. Let's take a matrix g from $SL\left(2,\mathbb{C}\right)$ and allow it to be multiplied on the commission matrix X in the following way:

$$X \to X' = gXg^+ \ , \tag{3.35}$$

It is clear that such transformation will not destroy hermiticity and will preserves it, because due to the rules of hermitian conjugation the following expression will be obtained

$$X'^+ = \left(gXg^+\right)^+ = \left(g^+\right)^+ X^+ g^+ = gX^+g^+ \ . \tag{3.36}$$

It also can be seen that if we look at the determinant of X', this determinant will be equal to

$$\det X' = \det gXg^+ = \det g \underbrace{\det g^+}_{=\det g} \det X \ , \tag{3.37}$$

where the determinant of g^+ is determinant of g^t but determinant does not changes with transportation of the matrix. Since the talk is about special linear group, determinant of g is equal to 1 by the definition of something that is called special linear group.

Let's remind that the special linear group is a group of 2 by 2 complex mattresses with a unit determinant. That is why

$$\det X' = \det X \ . \tag{3.38}$$

Let's conclude that since these determinants are equal to one, we get determinant of X and determinant of X is a four interval. Therefore, completed transformations by matrices from $SL\left(2,\mathbb{C}\right)$ group preserve the four interval. Once they preserve the four interval it is possible to say that they just perform on the vector x^μ Lorentz transformations.

It can be seen that the conservation of the vector length is guaranteed by the transformation. But this transformation may be written more explicitly

$$x'^\mu = \frac{1}{2}\mathrm{Tr}\left(X'\sigma^\mu\right) = \frac{1}{2}\mathrm{Tr}\left(gXg^+\sigma^\mu\right) \ . \tag{3.39}$$

Then it is possible to simplify (3.39) using the definition of x from (3.29):

$$x'^\mu = \frac{1}{2}\mathrm{Tr}\left(\sigma^\mu g\overline{\sigma}^\rho g^+\right)\eta_{\rho\nu}x^\nu \ . \tag{3.40}$$

Formula (3.40) relates the transformed coordinates to the original coordinates and it is known how this transformation should look like in terms of matrix of Lorentz transformation

$$x'^\mu = \Lambda^\mu_\nu x^\nu \tag{3.41}$$

and this means that the matrix Λ^μ is found in an explicit way and it's given by the formula

$$\Lambda\left(g\right)^\mu_\nu = \frac{1}{2}\mathrm{Tr}\left(\sigma^\mu g\overline{\sigma}^\rho g^+\right)\eta_{\rho\nu} \ , \tag{3.42}$$

where Λ is parameterized by an element g of the $SL\left(2,\mathbb{C}\right)$.

In fact, formula (3.42) gives a map from $SL\left(2,\mathbb{C}\right)$ to the group $SO^+\left(1,3\right)$.

It should be noticed why $SO^+\left(1,3\right)$ is written. This is because, if the element Λ^0_0 of the gotten matrix is written, it will be obtained that

$$\Lambda\left(g\right)^0_0 = \frac{1}{2}\eta_{00}\mathrm{Tr}\left(\sigma^0 g\overline{\sigma}^0 g^+\right) \ , \tag{3.43}$$

where $\eta_{00}=1$, σ^0 and $\overline{\sigma}^0$ are identity matrices.

That is why $\Lambda\left(g\right)^0_0$ is bigger than zero and therefore, it was realized by means of the construction

$$\Lambda\left(g\right)^0_0 = \frac{1}{2}\mathrm{Tr}\left(gg^+\right) > 0 \tag{3.44}$$

the orthochronous transformation.

It also should be pointed that all nondiagonal elements of the matrix Λ will be 0, because the metric η is diagonal.

Let's prove that the transformation presented in (3.42) is proper. That means that it is necessary to prove that determinant of the matrix is equal to one. It is possible to compute it easily by using the definition of Λ^{μ}_{ν} and it will be found that the determinant of Λ will be reduced to the determinant of g.

As a result it will be obtained that

$$\det \Lambda = (\det g)^2 \cdot \left(\det g^{+}\right)^2 = 1 \ . \tag{3.45}$$

What is very interesting is that the map from $SL\,(2,\mathbb{C})$ to $SO^{+}\,(1,3)$ is a map 2 to 1. It means that, in fact, there are two matrices of $SL\,(2,\mathbb{C})$ which go to one matrix of Lorentz transformation. For instance, two elements g and $-g$, which both belong to $SL\,(2,\mathbb{C})$, go to one element of the Lorentz group. If g is changed to $-g$, the matrix Λ, which is construct by formula (3.42) will not be changed.

And moreover, it is known that the group $SL\,(2,\mathbb{C})$ is simply connected while $SO^{+}\,(1,3)$ is not simply connected. The term simply connected means that if $SO^{+}\,(1,3)$ is considered as a topological manifold and a closed loop such as presented at (fig. 3.1) is considered, it will be seen that this loop cannot be contracted to a point all the time, returning inside is $SO^{+}\,(1,3)$.

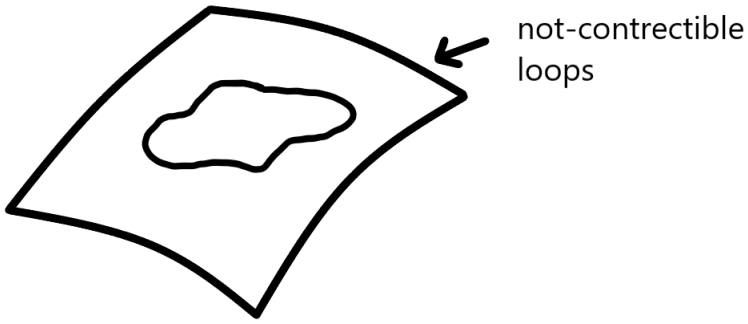

Fig. 3.1. A topological manifold with a closed loop

In other words, in $SO^{+}\,(1,3)$ there are non contractible loops. In a way, $SO^{+}\,(1,3)$ looks like a circle.

If a circle is considered, which is the simplest group $U\,(1)$ by multiplication, then, if any loop is considered, it will be clear that it is not possible to contract it.

In fact, we discovered here that $SL\,(2,\mathbb{C})$ represents a simply connected double cover of $SO^{+}\,(1,3)$ (fig. 3.2).

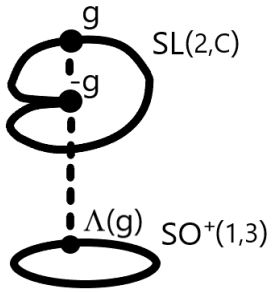

Fig. 3.2. A picture how $SL(2,\mathbb{C})$ double cover of $SO^+(1,3)$

It also can be introduced that any simply connected manifold which covers non simply connected manifold is called the universal cover. So, $SL(2,\mathbb{C})$ is the universal cover of $SO^+(1,3)$.

As an exercise, it can be checked that indeed the matrix Λ^μ_ν, which is constructed in by formula (3.42), is a matrix of Lorentz transformation, it preserves the Minkowsky metric. It means that you need to take Minkowsky metric $\eta_{\mu\nu}$ and need to apply to it Lorentz transformation $\Lambda^\mu_\rho\Lambda^\nu_\lambda$ and then you can compute this expression by using formula (3.42).

In fact, this has already been shown in a simple way by noting that transformations of the type, which was considered, preserve the determinant and determinant is a four-interval.

It also should be noticed that in general, if non-trivial background takes place, which is not Minkowsky space, then this background is characterized by certain metric, which can not be put on it. This metric might have or might not have what is called isometries. Isometries typically would provide a symmetries of a quantum field theory on a curved background, when the background is not flat like in Minkowsky space. It can be seen that quantum particles not only in Minkowsky space, but, for instance, quantum particles in the background of black hole. This means that space-time is very curved. This means that different background requires the different metric. The possible group of isometries is a group which preserves metric. In fact, for each manifold it is possible to define the notion of a tangent space. Even if there is a curved manifold, there is a tangent space. Then in this tangent space, there will be a group of transformations of tangent vectors which are attached to this space and this space will be linear. Therefore, there will be a group of linear transformations. If this tangent space is supplied or inherits Minkowsky metric, it will be a Lorenz group of the tangent space. Such a way Lorenz group becomes associated with a tangent space, but not to the whole manifold moving from point to point. There

will be a different sort of local Lorenz groups. In this case, in general, if there is no trivial metric, it is possible to talk about different morphisms which preserve this metric and these different morphisms are called isometries and they form the group of isometries.

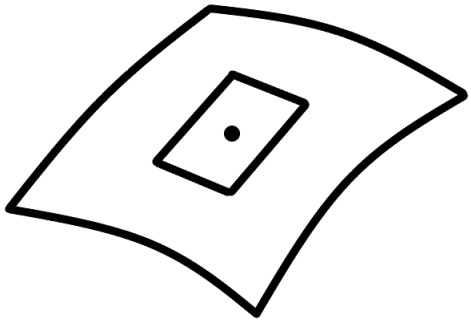

Fig. 3.3. The tangent space in very curved space-time

Noether's theorem

Noether's theorem has been established by me Noether in 1980. This theorem is important, because it relates symmetries to conservation laws. If a dynamical system are given, then a special role is played by the so called dynamical invariances or quantities, which are invariant with respect to dynamics of the system. Very often dynamical invariances are called as conservation laws. This means that if there is some dynamical systems, for instance, a system of particles, which are characterized by coordinates and momenta, when the dynamical system develops in time, the coordinates and momenta undergo changes, but it is possible to create some complicated or simple quantities, which stay invariant while time develops.

So, x_i and p_i of individual particles of the system change with time, but dynamical invariance remains unchanged. For this reason they are called conservation laws.

If the goal is to describe a dynamical system, it is not enough to know solutions of equations of motion. It is also important to be able to express the basic physical characteristics of a system like, for instance, conserved energy or conserved momentum. It is necessary to to know the energy and momentum or angular momentum of the system as a whole and for this reason it is needed to be able to find expressions for this quantities like energy momentum, angular momentum and so on in terms of individual dynamical coordinates and momenta, in other words phase space variables of the system. This goal is precisely achieved by means of Noether's first theorem.

Noether's theorem tells the following: let the action of a dynamical system be invariant, up to a boundary term with respect to a finite parametric, that is dependent on s constant parameters, continuous transformations of fields and space time coordinates. Then

1) there exists s linearly independent currents J_n^μ, where n runs from 1 to s with divergences equal to certain linear combinations of Lagrangian derivatives.

2) on-shell, that is on solutions of equations of motion these currents are divergenceless which means that $\partial_\mu J_n^\mu = 0$, and they give rise to dynamical invariance or conservation loss, which are conserved functionals of fields and their derivatives.

The symmetries of the action, which leaves the action invariant up to a boundary term are called variational symmetries. The symmetries we are talking about in the Noether's theorem can be formulated in the following form: transformation of the coordinates is involved as

$$x^\mu \to x'^\mu = x'^\mu(\phi, x) \tag{3.46}$$

where x'^μ depending on fields or several fields, which might have a different nature and can be scalars, vectors, tensors, and so on and depending on original coordinates.

Together with this, we also need to introduce transformations of fields

$$\phi_i(x) \to \phi_i'(x') = \phi_i'(\phi, x) \ . \tag{3.47}$$

It is assumed that the work is carried out in d-dimensional space and that is why index μ is allowded to run from 0 to $d-1$. Derivatives of fields under this transformations will change accordingly

$$\partial_\mu \phi_i \to \partial'_\mu \phi_i' \ . \tag{3.48}$$

In general, such transformations can be non-trivial, because not only space-time coordinates are being transformed, but also fields. Moreover, space-time coordinates may be required to be transformed by involving fields.

Let's now prove the Noether's theorem. First of all, the theorem requires that the symmetries of the action may be seen up to a boundary term. It means that if the action functional is taken, which is an integral over d-dimensional space-time of the Lagrangian density. Then in new coordinates, the action will look like

$$S' = \int_\Omega \mathrm{d}^d x' \mathcal{L}\left(\phi_i'(x'), \partial'_\mu \phi_i'(x')\right) \ , \tag{3.49}$$

where the Lagrangian density depends on transformed fields $\phi'_i(x')$, and derivatives $\partial'_\mu \phi'_i(x')$. Integration will also be done over the image of the original region Ω under the coordinate transformation.

Originally there is an integration region Ω, which can be arbitrary. Under the transformation $x \to x'$ it is mapped to some other region, which is called Ω' (fig. 3.4).

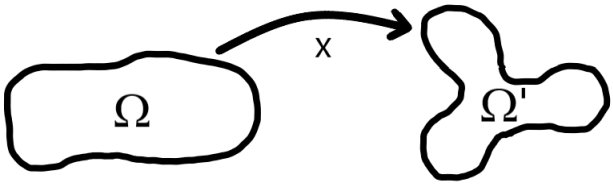

Fig. 3.4. The transformation $x \to x'$

The condition of the theorem requires that the new action linked with the old action by the formula

$$S' = \int_\Omega \mathrm{d}^d x \left(\mathcal{L}\left(\phi_i(x), \partial_\mu \phi_i(x)\right) + \underbrace{\frac{\mathrm{d}\Lambda^\mu}{\mathrm{d}x^\mu}}_{\text{boundary term}} \right) . \tag{3.50}$$

In the last lecture, it has been discussed that adding a derivative $\frac{\mathrm{d}\Lambda^\mu}{\mathrm{d}x^\mu}$, where Λ^μ is a function of fields ϕ does not influence on equations of motion. Variational symmetries allow the change of the action up to a boundary term.

The condition (3.50) is required in the Noether's theorem and the gotten invariance must be evaluated on the arbitrary region Ω.

It also should be noticed, that for the Noether's theorem it is not important that the Λ^μ depends only on field ϕ. In general, Λ^μ might depend on derivatives of ϕ, but the problem is that, in this case, this addition will lead to the problem that the physical requirement to have a unique solution of the standard Kashi problem will not preserve, because the new Lagrangian would involve derivatives higher than the first derivative. But the Noether's theorem is telling us that any boundary term is allowed.

Now it is necessary to use arbitrariness of the volume Ω and compare two integrals for S'. For that, it is possible to make a change of coordinates from $x' \to x$ and pass to integration over the original variable x over the region Ω.

Under the change of coordinates it is necessary to use a Jacobian, which is a determinant of the Jacoby matrix. Such a way the following will be obtained

$$\mathrm{d}^d x' = \left| \frac{\partial x'^\alpha}{\partial x^\beta} \right| \times \mathrm{d}^d x \ . \tag{3.51}$$

Then there will be two integrals over the same volume Ω and since Ω is arbitrary it is possible to change it arbitrarily. Left hand side should be equal to the right hand side and then this is only possible if the integrals are equal to each other.

So, the condition of the equality of the integrals can be written in the following way:

$$\mathcal{L}\left(\phi'\left(x'\right),\partial'_{\mu}\phi'\left(x'\right)\right)\left|\frac{\partial x'^{\alpha}}{\partial x^{\beta}}\right|=\mathcal{L}\left(\phi_{i}\left(x\right),\partial_{\mu}\phi_{i}\left(x\right)\right)+\frac{\mathrm{d}\Lambda^{\mu}}{\mathrm{d}x^{\mu}}\ . \tag{3.52}$$

The infinitesimal version of this equation is considered. This means that infinitesimal transformations that infinitesimal transformations are necessary, in which x^{μ} goes to x'^{μ} and a small variation of x^{μ} is infinitesimally added to x^{μ}.

$$x^{\mu}\rightarrow x'^{\mu}=x^{\mu}+\delta x^{\mu}\ . \tag{3.53}$$

For fields infinitesimal version $phi_{i}\left(x\right)$ goes to $\phi'_{i}\left(x'\right)$, which is

$$\phi_{i}\left(x\right)\rightarrow\phi'_{i}\left(x'\right)=\phi_{i}\left(x\right)+\delta\phi_{i}\left(x\right)\ , \tag{3.54}$$

where again the edition to phi_{i} is an infinitesimal version of the transformation.

It is also needed to parameterize transformations with the help of individual parameters. As it was described in the Noether's theorem, we are talking about s independent constant parameters, which parameterize transformations. Transformations are linear and continuous functions of parameters and, therefore, it is possible to write

$$\delta x^{\mu}=\sum_{1\leq n\leq s}X_{n}^{\mu}\cdot\underbrace{\epsilon^{n}}_{\text{parameters}}\ , \tag{3.55}$$

where ϵ^{n} are parameters parameters, which are constant in a sense that they do not depend on space-time points or fields and they just numbers, X_{n}^{μ} are quantities which describe response of x^{μ} on the change by infinitesimal parameters ϵ^{n}.

Also for fields it is possible to write $\delta\phi_{i}\left(x\right)$ is given by a response of fields:

$$\delta\phi_{i}\left(x\right)=\sum_{1\leq n\leq s}\Phi_{i,n}\epsilon^{n}\ , \tag{3.56}$$

where $\Phi_{i,n}$ are responses of fields on infinitesimal transformation with parameters ϵ_{n}.

Then let's look at how Lagrangian density transforms under symmetry transformations. For this reason it is necessary to consider $\phi'_{i}\left(x'\right)$. So infinitesimally it is written in the following way

$$\phi'_{i}\left(x'\right)=\phi'_{i}\left(x+\delta x\right)\ . \tag{3.57}$$

Now, it is possible to expand variable in (3.57) in a Taylor series in the following way

$$\phi_i'(x') = \phi_i'(x) + \partial_\mu \phi_i'(x)\, \delta x^\mu + \dots \ , \tag{3.58}$$

where instead of $\dots$ it is practicable to continue a Taylor series, but there is no interest in that, because infinitesimal transformations are considered.

Then it is possible to simplify expression (3.58) and get

$$\phi_i'(x') = \phi_i'(x) + \partial_\mu \phi_i'(x)\, X_n^\mu \epsilon^n + \dots \ . \tag{3.59}$$

Now it can be seen that ϕ_i' will be different from ϕ_i also by order ϵ, but one order ϵ is already in expression (3.59) and this means that there will not be a mistake at the leading order in ϵ if there is a replacement in the $\partial_\mu \phi_i'$ with the original field ϕ_i. In this case, there will not be a at a given order ϵ because the difference between ϕ' and ϕ is of order ϵ and it will contribute to the order ϵ^2.

$$\phi_i'(x') = \phi_i'(x) + \partial_\mu \phi_i(x)\, X_n^\mu \epsilon \ . \tag{3.60}$$

At this point, it is important to realize that the variation, which is applied to the field ϕ, δ does not commute with derivative $\frac{\partial}{\partial x^\mu}$. That is why, because variation δ an effect of both changing coordinates and fields. This variation is due to both the change of the form of the field, and also due to the change of the argument of the field. The variation $\delta\phi$ incodes both variation of coordinates and the variation of the form of the field. In order to distinguish variation of the form from the general variation, which also involves variation of the coordinates, the variation of the form of the field is introduced, which will not be denoted by $\bar{\delta}$. By definition this is

$$\bar{\delta}\phi_i = \phi_i'(x) - \phi_i(x) \ . \tag{3.61}$$

In other words, it can be seen how the variation of the form of the field changes, keeping this field at the same point. So space-time point is kept untouched. Such a way the difference between primed and unprimed fields is considered. It is possible to find what this variation is, because this is nothing else as the following thing:

$$\bar{\delta}\phi_i = (\Psi_{i,n} - \partial_\mu \phi_i X_n^\mu) \cdot \epsilon^n \ . \tag{3.62}$$

Expression for (3.62) can be gotten from the next condition:

$$\phi_i(x) + \Phi_{i,n}\epsilon^n = \phi'(x) + \partial_\mu \phi_i X_n^\mu \cdot \epsilon^n \ , \tag{3.63}$$

where then it is necessaru to subtract $\phi_i(x)$ from $\phi_i'(x)$:

$$\bar{\delta}\phi_i = \phi_i'(x) - \phi_i(x) = (\Phi_{i,n} - \partial_\mu \phi_i X_n^\mu)\,\epsilon^n \ . \tag{3.64}$$

From (3.64) it can be seen that variation of the form coincides with the full variation only if $\delta x^\mu = 0$.

Now it is possible to compute the variation of the Lagrangian density and for that the Lagrangian is considered as a function of the variable x.

The new Lagrangian density is equal to

$$\mathcal{L}'(x') = \mathcal{L}\left(\phi'(x'), \partial'_\mu \phi'(x')\right) \ . \tag{3.65}$$

As can be seen from (3.65) the new Lagrangian is an old Lagrangian density evaluated on the new coordinates and new fields.

So now it is obtained that

$$\mathcal{L}'(x') = \mathcal{L}'(x) + \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x^\mu}\delta x^\mu \ . \tag{3.66}$$

It is possible to add the Lagrangian density and then subtract it

$$\mathcal{L}'(x') = \mathcal{L}(x) + \mathcal{L}'(x) - \mathcal{L}(x) + \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x^\mu}\delta x^\mu \tag{3.67}$$

which gives us variational form of the Lagrangian plus total derivative:

$$\mathcal{L}'(x') = \mathcal{L} + \bar{\delta}\mathcal{L} + \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x^\mu}\delta x^\mu \ . \tag{3.68}$$

The variation of the Lagrangian is computed in the usual way as

$$\bar{\delta}\mathcal{L}(x) = \frac{\partial \mathcal{L}}{\partial \phi_i}\bar{\phi}_i + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\partial_\mu \bar{\delta}\phi_i \ , \tag{3.69}$$

where it is possoble to use the effect that the variation of the form commutes with the space-time derivative and, therefore, this gives us the following expression:

$$\bar{\delta}\mathcal{L}(x) = \underbrace{\left(\frac{\partial \mathcal{L}}{\partial \phi_i} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\right)\right)}_{\varepsilon_i}\bar{\delta}\phi_i + \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\bar{\delta}\phi_i\right) \ . \tag{3.70}$$

The last thing to take into account is the infinitesimal expression for integration measure. On the next lecture it will be shown that the integration measure is given by the following expression

$$\left|\frac{\partial x'^\alpha}{\partial x^\beta}\right| = \left|\frac{\partial}{\partial x^\beta}\left(x^\alpha + \delta x^\alpha\right)\right| = \left|\delta^\alpha_\beta + \frac{\partial \delta x^\alpha}{\partial x^\beta}\right| \ . \tag{3.71}$$

It can be seen that if there is no goal to restrict yourself to the leading order in ϵ, then expression (3.71) in the leading order gives us

$$\left|\frac{\partial x'^{\alpha}}{\partial x^{\beta}}\right| = 1 + \frac{\mathrm{d}}{\mathrm{d}x^{\mu}}\delta x^{\mu} + \dots \ . \tag{3.72}$$

Now let's return back to the original equation (3.52), which is required in the Noether's theorem. The infinitesimal version of this expression will be considered. That is why Lagrangian will be replaced by Lagrangian plus its first variation. As a result on the left side of the (3.52) it wiil be obtained that

$$\left(\mathcal{L} + \varepsilon_i\bar{\delta}\phi_i + \frac{\mathrm{d}}{\mathrm{d}x^{\mu}}\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_{\mu}\phi_i\right)}\bar{\delta}\phi_i + \dots\right) + \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x^{\mu}}\delta x^{\mu}\right)\left(1 + \frac{\mathrm{d}}{\mathrm{d}x^{\mu}}\delta x^{\mu} + \dots\right) \ . \tag{3.73}$$

On the right hand side it will be obtained that

$$\mathcal{L} + \frac{\mathrm{d}\delta\Lambda^{\mu}}{\mathrm{d}x^{\mu}} \ . \tag{3.74}$$

Then it is necessary to open the brackets on the left side and collect the terms at leading order.

It should be noticed that $\delta\Lambda^{\mu}$ is understood as $\Lambda^{\mu}_{n}\cdot\epsilon^{n}$ and the leading order is epsilon.

After collecting the terms of order ϵ, the following expression wiil be obtained

$$\frac{\mathrm{d}}{\mathrm{d}x^{\mu}}\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_{\mu}\phi_i\right)}\bar{\delta}\phi_i + \mathcal{L}\delta x^{\mu} - \delta\Lambda^{\mu}\right) = \frac{\mathrm{d}}{\mathrm{d}x^{\mu}}\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_{\mu}\phi_i\right)} - \frac{\partial\mathcal{L}}{\partial\phi_i}\right)\bar{\delta}\phi_i \tag{3.75}$$

From expression (3.75) it is possible to read off that there is a current exist, which will be called J^{μ} and can be written as

$$J^{\mu} = \frac{\partial\mathcal{L}}{\partial\left(\partial_{\mu}\phi_i\right)}\bar{\delta}\phi_i + \mathcal{L}\delta x^{\mu} - \delta\Lambda^{\mu} \ , \tag{3.76}$$

The current written in (3.76) has a property that:

$$\boxed{\frac{\mathrm{d}J^{\mu}}{\mathrm{d}x^{\mu}} = -\varepsilon_i\bar{\delta}\phi_i} \ , \tag{3.77}$$

i.e. the divergence of the current J^{μ} is given by a linear combination of Lagrangian derivatives ε_i and this linear combination involves the variation of the form of the field ϕ_i.

Index n will be introduced in the following way: if we plug in variations, explicitly, a variation of the form of the field, variation of the coordinates and variation of Λ^{μ}, it will not be possible to see that for since parameters, ε^{n} is independent. For each of these

parameters, tere will not be a conserved current. This conserved current J_n^μ, where index n runs from 1 to s, will be given by the following expression

$$\boxed{J_n^\mu = -\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi_i\right)}\left(\Phi_{i,n} - \partial\_\nu \phi_i X_n^\nu\right) - \mathcal{L}\cdot X_n^\mu + \Lambda_n^\mu}\,. \tag{3.78}$$

Expression from (3.78) can be called as Noether's current. On equations of motion or, in other words, on shell all Lagrangian derivatives are equal to 0 and this is why currents are conserved. In other words, their divergences are equal to 0

$$\frac{\mathrm{d}J_n^\mu}{\mathrm{d}x^\mu} = 0 \text{ for } \ \forall n = 1, \ldots, s\ . \tag{3.79}$$

Expression (3.79) is basically completes the proof of the Noether's theorem. The only point that was not explained is the construction of the dynamical invariance, which relies on the divergenceless currents, but this will be explained in the next lecture.

Lecture 4. Applications of Noether's Theorem. Conservation Laws and Symmetries

During the previous lecture it has been found an expression for the conserved current, which follows from the Noether's theorem. This current has a following structure

$$J_n^\mu = -\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi_i\right)}\left(\Phi_{i,n} - \partial_\nu \phi_i X_n^\nu\right) - \mathcal{L} X_n^\mu + \Lambda_n^\mu \ , \tag{4.1}$$

where in the gotten expression there are two indexes μ and n, where μ runs the space-time coordinates from 0 to $d-1$. At the same time index n represents a range of independent parameters and runs from 1 to s. Variable s can be characterized as a set of independent variation parameters.

With respect to index μ there is a conservation law of the next view

$$\partial_\mu J_n^\mu = 0 \ , \tag{4.2}$$

where this condition is right on shell, which means on solutions of equations of motion (Euler-Lagrange equations).

The local conservation laws (4.2) can be used to define the integral invariance. It is necessary to assume that all fields vanish at spatial infinity and then defines the following integral

$$J_n = \int_{\mathcal{B}} \mathrm{d}\vec{x} J_n^0 \ , \tag{4.3}$$

which is denoted as an integral of the zero component of the current, where zero-component means time component. It also should be noticed that it is necessary to integrate over $d-1$-dimensional hyper surface, which is orthogonal to the time direction.

Formula (4.3) can be visualized as presented at (fig. 4.1).

Then when such a quantity J_n is constructed, the statement is that this quantity is a conserved charge, which sometimes can be called as a Noether's charge. It is true, because if J_n is differentiated over time, there will be obtained that

$$\frac{\mathrm{d}J_n}{\mathrm{d}x^0} = \frac{\partial}{\partial x^0}\int_{\mathcal{B}} \mathrm{d}\vec{x} J_n^0 = \int_{\mathcal{B}} \mathrm{d}\vec{x}\frac{\partial J_n^0}{\partial x^0} \ . \tag{4.4}$$

where the differentiation over x^0 is moved inside the integral, because we are integrating over spatial directions.

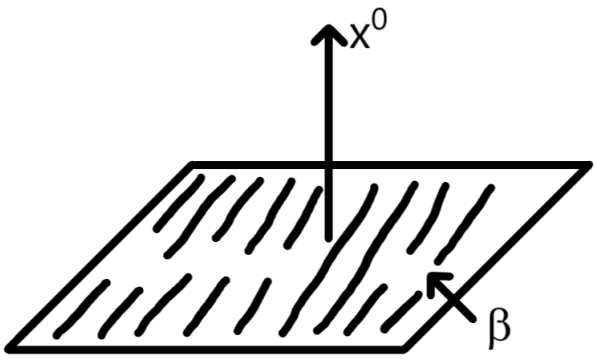

Fig. 4.1. Integration over $d-1$-dimensional hyper surface orthogonal to the time direction

Now, the current which depends on time and on the spatial components becomes the partial derivative, but then from the conservation law it follows that in terms of indexes

$$\frac{\partial J^0}{\partial x^0} + \frac{\partial J^i}{\partial x^i} = 0 \ , \tag{4.5}$$

where i runs over spatial indexes from 1 to $d-1$.

The derivatives with indexes i can be replaced by the divergence. Such a way it is possible to write a new expression for $\frac{\mathrm{d}J_n}{\mathrm{d}x^0}$:

$$\frac{\mathrm{d}J_n}{\mathrm{d}x^0} = -\int_{\mathcal{B}} \frac{\partial J^i}{\partial x^i} \mathrm{d}\vec{x} \ , \tag{4.6}$$

where the total derivative in the integral was obtained and such a way it is possible to reduce integral to the integral over the boundary of the space region (fig. 4.2).

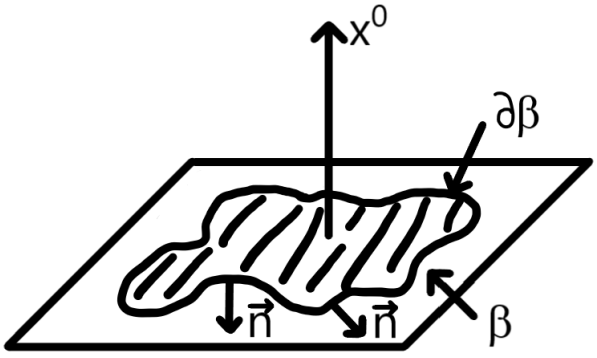

Fig. 4.2. Integration over the boundary of the space region

A boundary of the space region will be denoted as $\partial\mathcal{B}$ and then it is possible to write (4.6) as

$$\frac{\mathrm{d}J_n}{\mathrm{d}x^0} = -\int_{\partial\mathcal{B}} \mathrm{d}\left(\vec{J}_n \cdot \vec{n}\right) \ , \tag{4.7}$$

where $\vec{n}$ is a normal vector to the boundary of the region.

The expression in the brackets of the (4.7) is a flux of the vector J_n through the boundary.

Then it is necessary to tend boundaries to infinity. If the boundary ia taken, for instance, as a sphere, surrounding the space as a $d-1$-dimensional ball, the boundary will be a $d-2$ dimensional sphere surrounding the ball. If our fields are assumed to have a vanishing behavior at infinity, then in the limit where this boundary tends to infinity the integral will tends to zero, because fields are assumed to die at special infinity.

In this case, if the boundary conditions described above is assumed, then it will be seen that there is a next conservation law:

$$\frac{\mathrm{d}J_n}{\mathrm{d}t} = 0 \ . \tag{4.8}$$

Then it is possible to apply the general theorem of Noether to the discussion of concrete examples of symmetry transformations and derive for this examples the corresponding conserved currents.

First of all, let's start with the case of an internal symmetry.

Internal symmetry

This is a case where the space-time variable x^μ is not touched and do not do anything with x^μ. That is why:

$$\delta x^\mu = 0 \ . \tag{4.9}$$

From (4.9) it is clearly seen that the only variation that is possible to be done is a variation of the field. That is why it's called internal symmetry, because it does not involve change of the space-time.

This variation, as it is known, is written down as

$$\delta\phi'\left(x'\right) = \sum_{1\leq n\leq s} \Phi_{i,n}\delta\epsilon^n \ , \tag{4.10}$$

where the only non-trivial response is Φ and therefore, if we look at the general expression (4.1), it will be seen that only the term $\Phi_{i,n}$ contributes, because X_n^ν response on variation of space-time is absent.

Therefore, for this case the current looks rather simple:

$$J_n^\mu = -\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi_i\right)}\Phi_{i,n} \ . \tag{4.11}$$

The second very important example is an example of a conserved quantity, which is called energy-momentum tensor.

Energy-momentum tensor

Sometimes people also can call it as stress-energy tensor. This corresponds to infinitesimal space-time transformations, which can be firstly considered from the case of a single scalar field.

The transformations which will be performed are translations of space time coordinates:

$$x'^{\mu} = x^{\mu} + \delta x^{\mu} \ , \tag{4.12}$$

where δx^{μ} is a shift, which can be written in the following form:

$$\delta x^{\mu} = \delta^{\mu}_{\nu} \varepsilon^{\nu} \ , \tag{4.13}$$

where $delta^{\mu}_{\nu}$ is a Kronecker symbol, ε^{ν} are constant shifts.

From formula (4.13) it can be clearly seen that the response

$$X^{\mu}_{\nu} = \delta_n u^{\mu} \ . \tag{4.14}$$

For the scalar field it is known that transformation property is

$$\phi' (x') = \phi (x) + \delta\phi \ . \tag{4.15}$$

Therefore, in this case, the $\delta\phi (x)$ is actually vanishes, and therefore response

$$\delta\phi = 0 \to \Phi_{i,n} = 0 \ . \tag{4.16}$$

That is why the formula for conserved current now looks as follows:

$$T^{\mu}_{\nu} = \frac{\partial \mathcal{L}}{\partial (\partial_{\mu}\phi)} \partial_{\nu}\phi - \delta^{\mu}_{\nu} \mathcal{L} \ , \tag{4.17}$$

where for the stress energy tensor people usually use a special notation for the current T^{μ}_{ν}. It also should be noticed that index n was replaced by space-time index ν.

So, there is a tensor, which is directly derived by applying the general expression of the Noether's theorem. Expression (4.17) is called canonical stress energy tensor for a scalar field.

Then the next question should be asked. How this tensor will be looking, if the vector field is placed instead of the scalar field. Let's look at the field denoted by ϕ^λ, where lambda is a vector index, and consider the same transformations

$$x'^\mu = x^\mu + \epsilon^\mu = x^\mu + \delta^\mu_\nu \epsilon^\nu \ . \tag{4.18}$$

The question is: what is the transformation for the field now? What is the response of the vector field on coordinate shifts?

It turns out that this response vanishes in the same way as it was for the case of a scalar field. As it can be seen from the following formula if the variation of vector field is considered, then this is by definition should be response of a vector field by:

$$\delta\phi^\lambda = \Phi^\lambda_\mu \cdot \epsilon^\mu = 0 \ . \tag{4.19}$$

On the other hand, the $\delta\phi^\lambda$ is the transformation which comes from the general infinitesimal version of the transformation of a vector under the coordinate shift. This formula is known, because it was discussed in the first lecture.

$$\phi'^\lambda(x') = \frac{\partial x'^\lambda}{\partial x^\rho}\phi^\rho(x) \ . \tag{4.20}$$

Expression (4.20) shows how the general formula looks like. Then the infinitesimal transformation will be written in the following way:

$$x'^\lambda = x^\lambda + \varepsilon^\lambda \ . \tag{4.21}$$

It can be seen that in this case, that the infinitesimal change of the derivative

$$\frac{\partial \delta x'^\lambda}{\partial x^\rho} = \frac{\partial \varepsilon^\lambda}{\partial x^\rho} = 0 \ , \tag{4.22}$$

where ε^λ is constant and that is why the derivative will vanish.

As a result, the variation $\delta\phi^\lambda$ vanishes. Therefore for the vector field the response Φ^λ_μ vanishes, as it was for the case of a scalar field. That is why, it is seen that the stress energy tensor for a vector field will have a similar form to what was for the scalar field:

$$\boxed{T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi^\lambda)}\partial_\nu \phi^\lambda - \delta^\mu_\nu \mathcal{L}} \ . \tag{4.23}$$

There is only one difference that it is necessary to sum over vector components.

Now a bit more sophisticated transformations are considered, which are related to angular momentum and to the conservation loss related to the conservation of angular momentum.

Angular momentum

For this case it is necessary to consider infinitesimal rotations. For example, for infinitesimal rotations of Minkowski space there are transformations of the following type

$$x'^{\mu} = x^{\mu} + x_{\nu} \cdot \varepsilon^{\mu\nu} \ , \tag{4.24}$$

where parameters of this transformations $\varepsilon^{\mu\nu}$ satisfy the following property

$$\varepsilon^{\mu\nu} = -\varepsilon^{\nu\mu} \ . \tag{4.25}$$

In other words, these parameters constitute a second rank tensor, which is antisymmetric with respect to permutation of indices μ and ν.

If our work takes place in four-dimensional space, such a tensor will have six independent components. That is because μ runs in this case from 0 to 3 and, therefore, antisymmetric tensor will have the number of components equal to 4 multiplied by 3 and divided by 2, which is 6.

This six components will present six Lawrence transformations, which include three usual rotations of space-time coordinates, plus three boosts, where there will be a rotation between one of the spatial coordinates and the time.

Now it is necessary to find the response of coordinates on such a transformation. It is needful to take into account that only parameters ε^{μ} with $\mu < \nu$ are independent because parameters with $\nu > \mu$ can be found from the condition of antisymmetry (4.25). This means that it is possible to represent the transformation of coordinates in the following way:

$$\delta x^{\lambda} = x_{\nu}\varepsilon^{\lambda\nu} = x_{\nu}\delta^{\lambda}_{\mu} \cdot \varepsilon^{\mu\nu} \ . \tag{4.26}$$

As can be seen from (4.26), δ^{λ}_{μ} was introduced, but if there is a sum over μ, then it will be replaced by λ and we will return back to the formula:

$$\delta x^{\lambda} = x_{\nu}\varepsilon^{\lambda\nu} \ . \tag{4.27}$$

Now, there will be a sum over all μ and ν:

$$\delta x^{\lambda} = \sum_{\mu<\nu} x_{\nu}\delta^{\lambda}_{\mu}\varepsilon^{\mu\nu} + \sum_{\mu>\nu} x_{\nu}\delta^{\lambda}_{\mu}\varepsilon^{\mu\nu} \ . \tag{4.28}$$

Then in the second sum it will be necessary to change the index μ for ν, because μ and ν are dummy summation indexes. Then it is needed to use expression (4.25) and the

resulting formula can be written in the following way:

$$\delta x^{\lambda} = \sum_{\mu<\nu} \left(x_{\nu}\delta^{\lambda}_{\mu} - x_{\mu}\delta^{\lambda}_{\nu} \right) \varepsilon^{\mu\nu} \ , \tag{4.29}$$

where some now goes over independent components.

Then it is possible to write the response of the coordinates on such a variation

$$X^{\lambda}_{\mu\nu} = x_{\nu}\delta^{\lambda}_{\mu} - x_{\mu}\delta^{\lambda}_{\nu} \ . \tag{4.30}$$

where $\mu < \nu$.

Let's consider again a case of a scalar field for which there is

$$\phi'(x') = \phi(x) \ . \tag{4.31}$$

Again, from expression (4.31) we will get that $\delta\phi = 0$ and, again, response of variation of fields is actually zero ($\phi_n = 0$).

Now the general form of the Noether's current will be obtained and the non-trivial response will be substituted and the following formula for the current will be obtained. Now the current actually will have two indexes μ and ν instead of one index n, which was before. The current for this case is usually denoted by the letter M.

$$M^{\lambda}_{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_{\lambda}\phi)} \left(\partial_{\mu}\phi x_{\nu} - \partial_{\nu}\phi x_{\mu} \right) + \mathcal{L}\left(x_{\mu}\delta^{\lambda}_{\nu} - x_{\nu}\delta^{\lambda}_{\mu} \right) \ . \tag{4.32}$$

The conservation law will read now in this way:

$$\boxed{\partial_{\lambda} M^{\lambda}_{\mu\nu} = 0} \tag{4.33}$$

for any $\mu < \nu$.

In arbitrary dimension the tensor $M^{\lambda}_{\mu\nu}$ will have

$$\frac{d(d-1)}{2} \tag{4.34}$$

independent components.

In fact, it can be seen that it contains in addition to the usual angular momentum, which can be associated with the spatial rotations, also conservation laws which are related to Lorentz boosts, because as it was discussed in four dimensions the components μ and ν provide in total six independent conserved quantities. Three of these quantities correspond to the standard angular momentum related to the rotations between spatial directions and three other correspond to Lorentz boosts. It will be necessary to associate

the corresponding conserved quantity with which one of the six possible transformations according to the Noether's theorem.

The formula for the angular momentum tensor can be rewritten in the following way: there will be a small rearrangement of the terms. As a result, $M^{\lambda}_{\mu\nu}$ will be written in the following way:

$$M^{\lambda}_{\mu\nu} = x_\nu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\lambda \phi\right)} \partial_\mu \phi - \mathcal{L}\delta^{\lambda}_{\mu} \right) - x_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\lambda \phi\right)} \partial_\nu \phi - \mathcal{L}\delta^{\lambda}_{\nu} \right) \ . \tag{4.35}$$

Then, if a comparison of what is written in the brackets is performed, it will be obvious that what we see is the stress energy tensor for the scalar field. Therefore, expression (4.35) can be written as

$$M^{\lambda}_{\mu\nu} = x_\nu T^{\lambda}_{\mu} - x_\mu T^{\lambda}_{\nu} \ , \tag{4.36}$$

where T^{λ}_{μ} and T^{λ}_{ν} are stress tensors for the scalar field.

Further it will be needed to place the gotten expression into the conservation law and it will be possible to see what it implies:

$$\partial_\lambda M^{\lambda}_{\mu\nu} = \partial_\lambda \left(x_\nu T^{\lambda}_{\mu} - x_\mu T^{\lambda}_{\nu} \right) = \eta_{\lambda\nu} T^{\lambda}_{\mu} + x_\nu \partial_\lambda T^{\lambda}_{\mu} - \eta_{\lambda\mu} T^{\lambda}_{\nu} - x_\mu \partial_\lambda T^{\lambda}_{\nu} \ , \tag{4.37}$$

where due to the conservation of a stress tensor two terms $\partial_\lambda T^{\lambda}_{\mu}$ and $\partial_\lambda T^{\lambda}_{\nu}$ will vanish. Using the rule of raising of lower indexes with the help of the Minkowski metric the following will be obtained in formula (4.37):

$$\partial_\lambda M^{\lambda}_{\mu\nu} = T_{\nu\mu} - T_{\mu\nu} \ . \tag{4.38}$$

For the moment, expression (4.38) was derived, which actually shows us that for the case of a Noether's theorem, there is the divergence of angular momentum tensor vanishing $(\partial_\lambda M^{\lambda}_{\mu\nu})$. Then the stress energy tensor for a case of a scalar field must be automatically symmetric:

$$T_{\mu\nu} = T_{\nu\mu} \ . \tag{4.39}$$

The question is: what happens if, for instance, the energy momentum tensor $M^{\lambda}_{\mu\nu}$ for a case of a vector field is considered? There is an trivial response of a vector on space time rotations and this response can be found from the following:

$$\delta\phi^{\lambda} = \sum_{\mu<\nu} \phi^{\lambda}_{\mu\nu} \varepsilon^{\mu\nu} \ , \tag{4.40}$$

where μ and ν represent independent parameters of Lawrence transformations (six parameters in four dimensions).

On the other hand, $\delta\phi^\lambda$ should be given by the transformation law of a vector on the coordinate transformations which are in this case

$$\delta\phi^\lambda = \frac{\partial \delta x^\lambda}{\partial x^\rho}\phi^\rho(x) = \underline{\frac{\partial}{\partial x^\rho}\left(\sum_{\mu\nu}\left(x_\nu \delta^\lambda_\mu - x_\mu \delta^\lambda_\nu\right)\right)} \ . \tag{4.41}$$

where the underlined expression can be interpreted as a coordinate response.

The field response will have the next structure:

$$\Phi^\lambda_{\mu\nu} = \left(\eta_{\rho\nu}\delta^\lambda_\mu - \eta_{\rho\mu}\delta^\lambda_\nu\right)\phi^\rho = \phi_\nu \delta^\lambda_\mu - \phi_\mu \delta^\lambda_\nu \ . \tag{4.42}$$

Now, if the response from (4.42) is substituted together with a response of coordinates into the general formula for the current, it will be possible to find for the tensor of angular momentum of a vector field the following formula:

$$M^\lambda_{\mu\nu} = -\frac{\partial \mathcal{L}}{\partial\left(\partial_\lambda \phi^\rho\right)}\left[\phi_\nu \delta^\rho_\mu - \phi_\mu \delta^\rho_\nu + \left(x_\mu \partial_\nu \phi - x_\nu \partial_\mu \phi\right)\right] - \mathcal{L}\left(x_\nu \delta^\lambda_\mu - x_\mu \delta^\lambda_\nu\right) \ . \tag{4.43}$$

If the gotten expression is compared with stress tensor of a vector field, it will be possible to see that expression (4.43) may be written in the following way:

$$M^\lambda_{\mu\nu} = x_\nu T^\lambda_\mu - x_\mu T^\lambda_\nu + \left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\lambda \phi^\nu\right)}\phi_\mu - \frac{\partial \mathcal{L}}{\partial\left(\partial_\lambda \phi^\mu\right)}\phi_\nu\right) \ . \tag{4.44}$$

If we look at expression (4.44), it will be obvious that it does not reduce to the previous case of a scalar field, because for a scalar field we had only first two terms, but now there is an addition term.

It should be noticed that the first two term, which was in a case of the scalar field, is called orbital momentum. The additional piece characterizes polarization properties of the field and relates to the notion of spin. Such a way the part in the brackets of expression (4.44) is called a spin part.

From the discussion of the conservation law it is known that

$$\partial_\lambda M^\lambda_{\mu\nu} = T_{\nu\mu} - T_{\mu\nu} + \partial_\lambda\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\lambda \phi^\nu\right)}\phi_\mu - \frac{\partial \mathcal{L}}{\partial\left(\partial_\lambda \phi^\mu\right)}\phi_\nu\right) \ . \tag{4.45}$$

From expression (4.45) it is not possible to conclude that the spin part and the orbital part are separately conserved. It can be seen that this parts will be separately conserved only in the case, while the stress tensor is symmetric. In other words, while

$$T_{\mu\nu} = T_{\nu\mu} \tag{4.46}$$

the spin part will be separately conserved.

In general, it is unknown if the stress tensor is symmetric and it is possible to find the non-symmetric tensor. In this case, there is no possibility to define two separate quantities, because there is no separate conservation of the orbital part and the spin part. In total, when the orbital part is added up to spin part, it is always conserved because by the Noether's theorem, it comes from a unique quantity namely the tensor of angular momentum, which according to the Noether's theorem should have vanishing divergence. In particular, the case of Dirac field will be studied, where it will be found out that the stress energy is not symmetric and therefore there is no separate conservation of spin and orbital parts.

Lagrangians of „Wess-Zumino“

The last example of conservation laws and the application of the Noether's theorem will concern Lagrangians of so called „Wess-Zumino“ type. Wess and Zumino are scientists who are responsible for the contribution to the discovery of super-symmetry.

In some cases, the Lagrangians have the property that they are invariant, mainly Lagrangians which have super symmetry, only up to boundary terms. Since the appearance of boundary terms is allowed by Noether's theorem, there are examples where there are conservation laws associated to the existence or appearance of the boundary terms.

The simplest Lagrangian density, which exhibit the property to be invariant up to boundary terms, is the Lagrangian which has the next form:

$$\mathcal{L} = C^{\mu}_{ij}\phi^i\partial_\mu\phi^j \ , \tag{4.47}$$

where the Lagrangian wrote for a number of scalar fields ϕ^i, index i is running from 1 to the number m and μ is a space-time index.

Alsp it should be noticed that coefficients C^{μ}_{ij} are anti symmetric and this means that

$$C^{\mu}_{ij} = -C^{\mu}_{ji} \ . \tag{4.48}$$

This is needed otherwise the Lagrangian will be a total derivative. So this anti-symmetry property guarantees that there is no possibility to take the derivative out.

The infinitesimal transformation of fields ϕ_i, which leaves the Lagrangian invariant up to the boundary term is a simple shift. This means that a field phi_i is taken and it is shifted by constant parameter ε^i.

It can be seen that the variation of the Lagrangian density in this case is

$$\delta\mathcal{L} = C^{\mu}_{ij}\varepsilon^i\partial_\mu\phi^j \ . \tag{4.49}$$

Because of the constant C and ε it is possible to rewrite (4.49) as

$$\delta\mathcal{L} = \partial_\mu \left(C^\mu_{ij} \varepsilon^i \phi^j \right) \ . \tag{4.50}$$

As can be seen from (4.50) the gotten variation looks like the total derivative. It is possible to rewrite it as

$$\delta\mathcal{L} = \partial_\mu \Lambda^\mu \ , \tag{4.51}$$

where Λ^μ is

$$\Lambda^\mu = C^\mu_{ij} \varepsilon^i \phi^j \ . \tag{4.52}$$

It is also doable to write Λ^μ as the sum of components Λ^μ_i:

$$\Lambda^\mu = \Lambda^\mu_i \varepsilon^i \ . \tag{4.53}$$

That is why

$$\Lambda^\mu_i = C^\lambda_{ij} \phi^j \ . \tag{4.54}$$

It is atteinable now to use the formula for the Noether's current and write

$$J^\mu_i = -\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu \phi^i\right)} + C^\mu_{ij} \phi^j \ , \tag{4.55}$$

If the Lagrangian from expression (4.47) is taken and placed into (4.53), the following will be obtained

$$J^\mu_i = -C^\mu_{ji} \phi^i + C^\mu_{ij} \phi^i = 2C^\mu_{ij} \phi^i \ . \tag{4.56}$$

It is possible to immediately test the conservation law. For this case, it is necessary to take the current from (4.56) and compute its divergence. As a result there will be received that

$$\partial_\mu J^\mu_i = 2\underline{\partial_\mu \left(C^\mu_{ij} \phi^j \right)} \ , \tag{4.57}$$

where the underlined equation is nothing else as an equation of motion for a field ϕ^i.

Solutions of equations of motion as should be according to the Noether's theorem will be the divergence of the corresponding current and it will vanish. Finally there will be obtained

$$\partial_\mu J^\mu_i = 0 \ . \tag{4.58}$$

The first remark, which we should to make is so-called improvement procedure.

Improvement procedure

The point of this discussion is that, in principle, the conserved current J_n^μ is not uniquely defined. As it has already been seen, the Noether's theorem gives us a very concrete expression for the current and this current is called canonical. But it turns out that it is possible to change the current by adding to it so called improvement term:

$$J_n^\mu \to \underbrace{J_n^\mu}_{\text{canonical}} + \underbrace{\partial_\nu \chi_n^\mu}_{\text{topological}} \ , \tag{4.59}$$

where χ is a tensor which is assumed to be skew symmetric and then

$$\chi_n^{\mu\nu} = -\chi_n^{\nu\mu} \ . \tag{4.60}$$

Then it can be seen that if such a term is added up, the conservation law will not be changed or influenced, because if there is a performance with a derivative ∂_μ and compute the divergence of this extra contribution, the following expression will be obtained:

$$\partial_\mu \partial_\nu \chi_n^{\mu\nu} = 0 \ , \tag{4.61}$$

where there is a zero because derivatives ∂_μ and ∂_ν commute, while the tensor χ is skew symmetric.

In other words, it is clear that adding the topological term does not break the conservation law. Moreover, if χ is a function of fields, which vanish at infinity, it also does not break the conservation of the Noether's charge, because

$$\int_\beta \mathrm{d}\vec{x} \partial_\nu \chi^{0\nu} \ . \tag{4.62}$$

It is possible to write the derivative in (4.62) as

$$\partial_\nu \chi^{0\nu} = \partial_0 \chi^{00} + \partial_k \chi^{0k} \ , \tag{4.63}$$

where the variable $\partial_0 \chi^{00}$ simply vanishes and therefore it is possible to write (4.62) in the following way:

$$-\int_\beta \mathrm{d}\vec{x} \partial_k \chi^{0k} \ . \tag{4.64}$$

From (4.64) it is clearly seen that, as before, there will be a flux of the corresponding vector χ^{0k} over the boundary of the integration area β. Since the boundary tends to infinity the integral (4.64) tends to zero.

Then the next question may be asked: is it possible to make the stress tensor, which might be not symmetrical, to become symmetric by adding a proper improvement term? In other words, it is neccessary to take a $T^{\mu\nu}$ and add to it the improvement term:

$$T^{\mu\nu} \to T^{\mu\nu} + \partial_\rho \chi^{\rho\mu\nu} \ , \tag{4.65}$$

where according to the previous discussion

$$\chi^{\rho\mu\nu} = -\chi^{\mu\rho\nu} \ . \tag{4.66}$$

If the tensor could be symmetric then, as it was discussed,the conservation of the orbital part and the spin part will be gotten.

The answer to that question is that it turns out that there exists an improved symmetric stress tensor $\mathcal{J}^{\mu\nu}$, which is symmetric only if and only if the anti-symmetric part of the canonical stress tensor $T^{\mu\nu}$ is the total derivative.

Let's assume that the anti-symmetric part is a total derivative. This means that

$$T^{\mu\nu} - T^{\nu\mu} = -\partial_\rho \Omega^{\rho\mu\nu} \ , \tag{4.67}$$

where Ω satisfies the anti-symmetry property with respect to indices μ and ν:

$$\Omega\rho\mu\nu = -\Omega\rho\nu\mu \ . \tag{4.68}$$

Let's prove that it is always possible to improve such a tensor by adding to it the improvement term. To construct the improvement term it is needed to do the following: the improvement term $\chi^{\rho\mu\nu}$ is taken in the following form:

$$\chi^{\rho\mu\nu} = \frac{1}{2}\left(\Omega\rho\mu\nu + \Omega\mu\nu\rho - \Omega\nu\rho\mu\right) \ . \tag{4.69}$$

Then it can be seen that, in fact, the tensor χ has the desired symmetry:

$$\chi^{\rho\mu\nu} = -\chi^{\mu\rho\nu} \ , \tag{4.70}$$

where it is possible to prove this statement if indexes are changed in the (4.69):

$$\chi^{\rho\mu\nu} = \frac{1}{2}\left(\Omega\mu\rho\nu + \Omega\rho\nu\mu - \Omega\nu\mu\rho\right) \ . \tag{4.71}$$

Then, because the Ω has a property of skew symmetry with respect to the last two indices, it is possible to rewrite expression (4.71) as

$$\chi^{\rho\mu\nu} = \frac{1}{2}\left(-\Omega\mu\nu\rho - \Omega\rho\mu\nu + \Omega\nu\rho\mu\right) \ . \tag{4.72}$$

If terms from the (4.72) are compared with terms from the (4.69), it will be clear that the (4.70) is true.

The second statement is that if χ defined by the (4.69) is taken and $\chi^{\rho\nu\mu}$ is subtracted, the following will be obtained:

$$\chi^{\rho\mu\nu} - \chi^{\rho\nu\mu} = \Omega^{\rho\mu\nu} \ . \tag{4.73}$$

Therefore, in the anti symmetric part (4.67) it is possible to substitute instead of the tensor $\Omega^{\mu\nu}$ the next difference:

$$T^{\mu\nu} - T^{\nu\mu} = -\partial_\rho \left(\chi^{\rho\mu\nu} - \chi^{\rho\nu\mu}\right) \ . \tag{4.74}$$

Then it is possible to open the brackets and get that:

$$T^{\mu\nu} - T^{\nu\mu} = -\partial_\rho \chi^{\rho\mu\nu} + \partial_\rho \chi^{\rho\nu\mu} \ . \tag{4.75}$$

Let's combine T with derivatives in the following way:

$$T^{\mu\nu} + \partial_\rho \chi^{\rho\mu\nu} = T^{\nu\mu} + \partial_\rho \chi^{\rho\nu\mu} \ , \tag{4.76}$$

where it is attainable to conclude that the terms on the left and the right sides are symmetric tensors $\mathcal{J}^{\mu\nu}$ and $\mathcal{J}^{\nu\mu}$ respectively. It is automatically means that

$$\mathcal{J}^{\mu\nu} = \mathcal{J}^{\nu\mu} \ . \tag{4.77}$$

Indeed, it has been proved that if anti symmetric part of the stress tensor is a total derivative, then the stress energy tensor can always be improved by adding a proper improvement term.

In the other way, it is necessary to assume the opposite, that there exists an improved symmetric tensor, which is the tensor $\mathcal{J}^{\mu\nu}$ of the next view:

$$\mathcal{J}^{\mu\nu} = T^{\mu\nu} + \partial_\rho \chi^{\rho\mu\nu} \ . \tag{4.78}$$

But then from the (4.78), it automatically follows that the anti-symmetric part of the $T^{\mu\nu}$ is a total divergence, because

$$\mathcal{J}^{\nu\mu} = T^{\nu\mu} + \partial_\rho \chi^{\rho\nu\mu} \tag{4.79}$$

and it is possible to subtract (4.79) from expression (4.78):

$$T^{\mu\nu} - T^{\nu\mu} = \partial_\rho \left(\chi^{\rho\mu\nu} - \chi^{\rho\nu\mu}\right) \ , \tag{4.80}$$

which in (4.80) it is shown that anti symmetric part is a total divergence. Therefore, the opposite statement is also approved.

The tensor $\chi^{\rho\mu\nu}$ which is skew symmetric in indexes ρ and μ and whose derivative can be added to improve the stress tensor has a name of „Bellinfante“ tensor.

The final remark concerning this improvement procedure is that even improvement is not unique. Let's suppose thatoer start was from a canonical tensor and then we improved it by adding a proper „Bellinfante“ tensor and made the tensor symmetric. But even if the tensor is symmetric, it is still possible to add to it the following improvement term:

$$T_{\mu\nu} \to T_{\mu\nu} + \partial^\rho \partial^\sigma W_{\mu\rho\nu\sigma} \ , \tag{4.81}$$

where W is a tensor of the fourth rank, where it is required from this tensor W the same symmetries as the symmetries of the Riemann tensor. It's known that the Riemann tensor has the following symmetries:

$$\begin{cases} W_{\mu\rho\nu\sigma} = -W_{\rho\mu\nu\sigma} \\ W_{\mu\rho\nu\sigma} = -W_{\mu\rho\sigma\nu} \\ W_{\mu\rho\nu\sigma} = W_{\nu\sigma\mu\rho} \end{cases} \tag{4.82}$$

So if there is such a tensor as W and it is added up with two derivatives on W, then there will still be a symmetric tensor.

Finally, it is possible to find the symmetric tensor directly from a gravitational theory. There is a procedure which produce straightforwardly a symmetric stress energy tensor. But this one has to consider the action for field, which we are interested in coupled to an arbitrary gravitational background.

It is possible to take our "favorite"field, for instance, scalar field, vector field and so on and couple it couple our field to arbitrary gravitational background, which essentially means that our field is coupled to the metric. Then the action has standard form. For instance, in four dimensions it would be derivative of the Lagrangian density:

$$S = \int \mathrm{d}^4 x \mathcal{L} \ . \tag{4.83}$$

The Lagrangian density in addition to the field, that is considered, also involves the space time metric $g_{\mu\nu}$ with the following determinant:

$$g = \det g_{\mu\nu} \tag{4.84}$$

It's well known that under different morphisms

$$x'^{\mu} \to x^{\mu} + \xi^{\mu} \ , \tag{4.85}$$

where ξ^{μ} is a vector field. ξ^{μ} by definition is a vector which depends on the coordinates of space-time

$$\xi^{\mu} := \xi^{\mu}(x) \ . \tag{4.86}$$

The metric is known to transform in the following way:

$$\delta g^{\mu\nu} = \nabla^{\mu}\xi^{\nu} + \nabla^{\nu}\xi^{\mu} \ . \tag{4.87}$$

The transformation of the metric induces a transformation of the action. If a variation of the action is performed, the action response in the following way:

$$\delta S = \int \mathrm{d}^4 x \frac{\delta \mathcal{L}}{\delta g_{\mu\nu}} \left(\nabla^{\mu}\xi^{\nu} + \nabla^{\nu}\xi^{\mu}\right) \ . \tag{4.88}$$

Then the quantity can be introduced, which is stress energy tensor in the presence of the gravitational field. By definition, this is

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta \mathcal{L}}{\delta g^{\mu\nu}} \ . \tag{4.89}$$

In the gravity theory it's proved that the following identity takes place:

$$\partial_{\mu} \underbrace{\left(\sqrt{-g} T^{\mu\nu} \xi_{\nu}\right)}_{\text{divergence}} = \sqrt{-g} \nabla_{\mu} \left(T^{\mu\nu}\xi_{\nu}\right) \ . \tag{4.90}$$

It is now possible to open the brackets in the (4.90) and write that

$$\partial_{\mu}\left(\sqrt{-g}T^{\mu\nu}\xi_{\nu}\right) = \sqrt{-g}\nabla_{\mu}T^{\mu\nu}\xi_{\nu} + \sqrt{-g}T^{\mu\nu}\nabla_{\mu}\xi_{\nu} \ . \tag{4.91}$$

Formula (4.91) means that in fact, it is now possible to come back to formula (4.88) and rewrite the variation δS as

$$\delta S = \int \mathrm{d}^4 x \sqrt{-g} T^{\mu\nu} \nabla_{\mu} \xi_{\nu} \ , \tag{4.92}$$

where the integral presented in (4.92) can be evaluated by parts:

$$\delta S = \int \mathrm{d}^4 x \partial_{\mu}\left(\sqrt{-g}T^{\mu\nu}\xi_{\nu}\right) - \int \mathrm{d}^4 x \sqrt{-g} \nabla_{\mu} T^{\mu\nu} \xi_{\nu} \ . \tag{4.93}$$

If it is assumed that fields ξ_{ν} vanish at infinity, then the total derivative term does not contribute and gets out. Since our interest is the action to be invariant with respect to

default morphisms, then it means that the variation δS must vanish for any ξ_ν. The first term of the (4.93) will vanish and therefore the one possible variant is

$$\nabla_\mu T^{\mu\nu} = 0 \ . \tag{4.94}$$

There is the expression in the (4.94), which looks like the conservation law, but it's not true, because the derivative in (4.94) is not the usual derivative but it's a covariant derivative.

If we work on the flat background and make the metric in expression (4.94) Minkowski, then the variant derivative turns into the usual derivative and $T^{\mu\nu}$ will be conserved in the standard sense.

$$\nabla_\mu = \partial_\mu \to \partial_\mu T^{\mu\nu} = 0 \ . \tag{4.95}$$

On the other hand, from expression (4.89) it can be seen that since metric is symmetric with respect to indexes μ and ν, the object which will be obtained by means of the Lagrangian variation with respect to the symmetric metric will be automatically symmetric.

$$T^{\mu\nu} = T^{\nu\mu} \ . \tag{4.96}$$

There is a way to obtain a symmetric tensor by coupling the field to the general gravitational background with an arbitrary metric and then upon this object is derived, it can be seen that it will have a meaning of the stress energy tensor with the usual conservation law, if we return back in the final equation (4.94 to the flat Minkowski metric.

Casimir operators

Tensors are objects, which transform themselves under Lorentz transformations. They are conserved quantities with respect to the time evolution of the dynamical system. If there is some kind of Lagrangian or Hamiltonian driving the dynamics, then these quantities are indeed conserved. The numbers, which are computed as components of tensors, depend on the Lorentz frame in which we compute them. If a rotation of the Lorentz frame is performed or if our level is boosted, for instance, with respect to the original frame, then numbers represented by the components of, for instance, $T_{\mu\nu}$ or $M^\lambda_{\mu\nu}$ will be changed. They will go to other numbers, which also will conserve with the time. Because of the dynamics of our system, they always been conserved, but they will be changed because these objects are tensors.

From this point of view, you may ask yourself what are the quantities which actually characterize our dynamical system in such a way that they are independent on the choice of the Lorentz frame. In representation theory these quantities are known as Casimir's or Casimir operators.

The Casimir's objects commute with all generators of a Poincare group:

$$[C, P_\mu] = 0 \ , \tag{4.97}$$

where C is a Casimir's object.

Expression (4.97) means that if C is taken to be a Casimir, then it must commute with all generators of translations P_μ, which are components constructed as conserved charges related to stress tensor. So

$$P_\mu = \int T^0_\mu \mathrm{d}\vec{x} \ , \tag{4.98}$$

where it is necessary to integrate over spatial directions. Analogously, Casimir must commute with all generators of angular momentum

$$[C, M_{\mu\nu}] = 0 \ , \tag{4.99}$$

which are obtained by

$$M_{\mu\nu} = \int M^0_{\mu\nu} \mathrm{d}\vec{x} \tag{4.100}$$

with respect to the space.

There may be a question: what are the Casimir elements of Casimir operators for the case of the Poincare group? For the case of the Poincare group, there are two Casimirs. One Casimir is simply equal to

$$C_1 = P_\mu P^\mu = P_0^2 - \vec{P}^2 \ . \tag{4.101}$$

The gotten Casimir object takes the values m^2c^2 and therefore

$$P_0^2 - \vec{p}^2 = m^2c^2 \ . \tag{4.102}$$

If the terms in the (4.102) are placed to the one side, it will be possible to see that something that was obtained is called mass-shell condition for a relativistic particle, which is going to be studied in detail a little bit later.

$$P_0^2 - \vec{p}^2 - m^2c^2 = 0 \ . \tag{4.103}$$

If the commutation relations between C_1 and another P_σ or generator $M_{\mu\nu}$ is written, it will be found out that

$$\begin{cases} [P_\mu P^\mu, P_\sigma] = 0 \\ [P_\mu P^\mu, M_{\mu\nu}] = 0 \end{cases} \tag{4.104}$$

The second Casimir is constructed by using a certain special vector W_μ, which is constructed in the following way

$$W_\mu = \varepsilon_{\mu\nu\rho\lambda} P^\nu M^{\rho\lambda} \ , \tag{4.105}$$

where we got a fully anti-symmetric tensor. The second Casimir object is called „Pauli-Lubanski“ vector. The property of this vector is that it is orthogonal with respect to the Lorentz invariant scalar product:

$$W_\mu P^\mu = 0 \ . \tag{4.106}$$

The second property is that it may be used to build up the second Casimir of the Poincare group, which is given by a square of the „Pauli-Lubanski“ vector:

$$C_2 = W_\mu W^\mu \ . \tag{4.107}$$

It is possible to check by direct calculations that if C_2 is taken and it is commuted with P_σ, it will be found out that

$$[W_\mu W^\mu P_\sigma] = 0 \tag{4.108}$$

as well as it can be found with $M_{\rho\lambda}$:

$$[W_\mu W^\mu M_{\rho\lambda}] = 0 \ . \tag{4.109}$$

Values of the second Casimir in irreducible representations massive of the Poincare group will look like as follows

$$C_2 = m^2 c \cdot s\,(s+1) \ , \tag{4.110}$$

where s is a half integer numbers, starting from 0, $\frac{1}{2}$, 1, $\frac{3}{2}$ and so on, which is called the spin.

In other words, massive irreducible representations of the Poincare group are characterized by two numbers the mass and spin. In fact, this is in acted with characterization of elementary particles, because every elementary particle has a mass and spin or it is

massless, but massless is a separate case, it is a separate representations of the Poincare group, because as can be seen for massless representation C_1 and C_2 will be equal to zero.

In that case, there appears a new Casimir, which is called Helicity. Helicity coincides with a projection of a spin of a particle on the direction of motion:

$$\frac{\vec{p}\cdot\vec{s}}{|\vec{p}|}\ , \tag{4.111}$$

where (4.111) is a projection of spin on the direction of motion.

The talk about helicity will be in detail, when the discussion will be about the Dirac's theory, where it will be possible to realize how helicity appears, why it plays an important role and why helicity is applied only to the case of massless particles, but not the massive ones.

It is now necessary to introduce some definitions. First of all, let's start with representation. Representation means the representation of abstract generators of the Poincare algebra by concrete operators, acting in some concrete space of states.

Basically, it is needful to realize pure abstract algebraic relations of the Poincare group. Those commutators which have been written down in the previous lectures, by means of concrete operators realized in a concrete Hilbert space. It is necessaty to think about this objects as about a concrete operators, acting on states in some Hilbert space.

Then it is necessary to introduce a reducible representation. Representation is called the reducible if there will be no invariant subspaces, where invariant subspaces means that a vector is taken from the subspace and we act on it with generators. The result of this action remains belong to the same subspace and it is not possible to get out of the subspace.

If there are no subspaces except the trivial one and the whole space, then such representation of a Lie group is called irreducible.

And on the top, a notion of unitarity is added up. Representation is called unitary if the Lie algebra generators are realized by Hermitian operators. If the Lorentz generators by Hermitian operators are realized, the next conditions will be true:

$$\begin{cases} M^{+}_{\mu\nu} = M_{\mu\nu} \\ P^{+}_{\mu} = P_{\mu} \end{cases} \tag{4.112}$$

The unitary representations are important, because they can tell us that probabilities that is needed to be computed, will not depend on the particular Poincare frame in which

it is being computed, because unitarity preserves probability as it is known from courses on quantum mechanics.

In this way, the interest is in constructing unitary irreducible representations of Poincare group. It is known that since Poincare group is non compact, in particular, because it includes the Lorentz group which is non compact, then by the theorem from the group theory it is known that such unitary representations of non compact group must necessarily be infinity dimensional. That is why the spaces in which it is possible to realize such unitary irreducible representations of the Poincare group must be infinity dimensional, and that is one of the reasons why it is necessary to deal with fields, which in court, an infinite number of degrees of freedom, on which in quantum theory unitary irreducible representations of the Poincare group will be realized.

Lecture 5. Conserved Charges as Symmetry Generators. Klein-Gordon Field. Mass-Shell Condition

Conserved charges as the symmetry generators

Conserved charges that follow from Noether's theorem play a role of infinitesimal generators of symmetries. Symmetry generators means that they play a role of generators of symmetry, which is a correspond to according to the Noether's theorem.

To understand this statement it is necessary to recall the construction of the Hamiltonian formalism for the case of field theory. Normally the start should be from the Lagrangian description and introduced Lagrangian, Lagrangian density and the action. To pass to the Hamiltonian formalism, it is neceassary to introduce in addition to the canonical coordinates $\phi_i(x)$, which in this case is played by our fields, a canonical momenta

$$\pi_i(x) = \frac{\delta \mathcal{L}}{\delta \dot{\phi}_i} \ , \tag{5.1}$$

where the expression looks like an usual expression in classical mechanics. As can be known, a momentum is a derivative of the Lagrangian with respect to velocity. In this case, it is also possible to write (5.1) as

$$\pi_i(x) = \frac{\partial \mathcal{L}}{\partial\left(\partial_0 \phi_i(x)\right)} \ , \tag{5.2}$$

where ∂_0 is a derivative with respect to x^0, which is the same as

$$x^0 = ct \ . \tag{5.3}$$

The relation between the momentum and the derivative of the Lagrangian density allows to express the velocity of the field in terms of canonical momentum.

$$\dot{\phi}_i(x) \to \pi_i(x) \ . \tag{5.4}$$

Therefore, then it is necessary to construct the Hamiltonian using the standard prescription of classical mechanics. It is neeeded to construct it as

$$H = \int \mathrm{d}\vec{x} \left(\pi_i \dot{\phi}_i\right) - L \ , \tag{5.5}$$

where it is possible to write L as an integral of Lagrangian density and get the following expression:

$$H = \int \mathrm{d}\vec{x} \left[\pi_i \dot{\phi}_i - \mathcal{L}\right] = \int \mathrm{d}\vec{x} \mathcal{H} \ , \tag{5.6}$$

where $\mathcal{H}$ is an Hamiltonian density, which can be written as:

$$\mathcal{H} = \pi_i \dot{\phi}_i - \mathcal{L} \ . \tag{5.7}$$

Then absolutely the same procedure should be implemented, but for the case of fields. An important role in constructing the Hamiltonian formalism in the usual classical mechanics is played by an object, which is called Poisson's bracket, because this bracket allows to formulate Hamilton's equations of motion in a compact and a simple way. In general, it is known that a Poisson's bracket is an operation on a space of functions or a phase space. If there is a phase space, which is parameterized by the coordinates p and q, and then on this phase space, there will be a space over the phase space, which will be a space of functions, where it is possiblet to define an operation, which is called Poisson's bracket, which is a map cross functions into functions (fig. 5.1).

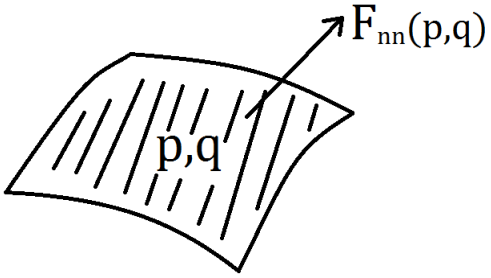

Fig. 5.1. A phase space, parameterized by the coordinates p and q, with space of functions over the phase space

Any two functions on a phase space it put in correspondence as an a function on the phase space and the Poisson's bracket satisfies the following conditions:

1) the Poisson's bracket is skew symmetric:

$$\{f, g\} = -\{g, f\} \ . \tag{5.8}$$

2) the Poisson's bracket is bilinear:

$$\{\lambda f + \mu g, h\} = \lambda\{f, h\} + \mu\{g, h\} \ . \tag{5.9}$$

3) the Poisson's bracket must satisfy Jacobi identity. If a Poisson's bracket of f and g is taken and then bracket with another function h and then add up a cyclic permutations of these functions f, g and h, then the result will be equal to:

$$\{\{f, g\}, h\} + \text{cyclic permutations} = 0 \ . \tag{5.10}$$

Since our work is done with fields, for any two local in time functionals, for instance, $F[\pi, \phi]$ and $G[\pi, \phi]$ depending on momenta and coordinates the Poisson's bracket will be defined. Let's define the Poisson's bracket as a following functional:

$$\{F, G\} = \sum_i \int \mathrm{d}\vec{x} \left[\frac{\delta F}{\delta \pi_i(x)} \frac{\delta G}{\delta \phi_i(x)} - \frac{\delta F}{\delta \phi_i(x)} \frac{\delta G}{\delta \pi_i(x)} \right] . \tag{5.11}$$

In the (5.11) a definition of the canonical Poisson's bracket over the space of functionals was written, which now parameterized to depend on canonical coordinates and momenta.

Before functions of p and q were studied, where p and q are coordinates on the classical phase space of classical mechanics. To any point a value, which is the value of the function at this point, is put in correspondence.

Then it is workable to go on with this definition and compute the Poisson's brackets between sort of coordinates on the infinite dimensional phase space, which are fields and their momenta. A simple exercise to see that if two fields were taken at different space points and at the same time t, then the result of evaluation of the Poisson's bracket between the fields will be equal to zero:

$$\{\phi_i(t, \vec{x}), \phi_j(t, \vec{y})\} = 0 . \tag{5.12}$$

That is simply, because for the bracket to be non-zero it is necessary to vary one functional with respect to momenta and the other with respect to the coordinate, but in the (5.12) there are only coordinates and therefore the Poisson's bracket would vanish.

The same expression will be right for momenta. If there is a Poisson's bracket of two momenta at different space points, then the result will be zero:

$$\{\pi_i(t, \vec{x}), \pi_j(t, \vec{y})\} = 0 . \tag{5.13}$$

Finally, if the bracket between momenta and coordinates is evaluated at different space points and at the same time t, then the result will be equal to:

$$\{\pi_i(t, \vec{x}), \phi_j(t, \vec{y})\} = \delta_{ij} \delta(\vec{x} - \vec{y}) . \tag{5.14}$$

It is possible to group expressions (5.12), (5.13) and (5.14) under the name of generalization of the canonical Poisson's bracket in classical mechanics to the case of fields.

It is also important that in the Hamiltonian setting the Poisson's brackets are equal-time. It can be seen, because the right hand side of expressions (5.12), (5.13) and (5.14) does not depend on time.

It also should be noticed that if different times are selected, it will not be possible to know what will be the outcome and therefore there is a way to solve equations of motion for the fields and compute the Poisson's bracket on the phase space.

If there is such equal time Poisson's structure as in formulas above, it is possible to define the time evolution of any functional. For instance, there is a way to say that F under the Hamiltonian flow set up by our preference choice of the Hamiltonian. Some function is selected on a phase space, which will be regarded as the Hamiltonian and then evolution of any function on a phase space can be described by means of the following formula:

$$\boxed{\dot{F} = \{H, F\}}\,. \tag{5.15}$$

In fact, formula (5.15) is the standard equation of classical mechanics, which goes without any modification to the field theory.

If the time evolution of any functional is defined, it is necessary to compute the equal-time Poisson's bracket of H with F.

For the canonical Poisson's bracket of q and p it is possible to write similar to (5.12)-(5.14) expressions:

$$\begin{cases} \{q_i, q_j\} = 0 \\ \{p_i, p_j\} = 0 \\ \{p_i, q_j\} = \delta_{ij} \end{cases} \tag{5.16}$$

If, for instance, coordinate q_i is taken, the $\dot{q}$ with the help of the Hamiltonian H will be equal to:

$$\dot{q}_i = \{H, q_i\} = \frac{\partial H}{\partial p_i}\ . \tag{5.17}$$

At the same time, if the time derivative of the momentum is computed, then according to equation (5.15) in the usual mechanics the following can be found

$$\dot{p}_i = \{H, p_i\} = -\frac{\partial H}{\partial q}\ . \tag{5.18}$$

Such a way, we got in (5.17) and (5.18) nothing else as Hamilton's equations of motion. Therefore (5.15) represents Hamilton's equations in field theory.

Finally, it is practicable to generalize the procedure for evaluation of the (5.15). It is necessary to evaluate the Poisson's bracket in (5.15) and for this purpose it is needed to evaluate (5.11) and pick up a functional, which will be regarded as a Hamiltonian, and then for any functional F Hamilton's equations will be obtained. Such a way recalled a standard Hamiltonian formalism.

If we remember the general expression for the Noether's charge, which was recieved from the Noether's theorem and which is given by the following formula

$$J_n = \int \mathrm{d}\vec{y} J_n^0 (y) \ , \tag{5.19}$$

where it is necessary to integrate over spatial directions.

Formula (5.19) can be transformed into:

$$J_n = \int \mathrm{d}\vec{y} \left[-\pi_j \left(\vec{\Phi}_{j,n} (\phi) - \partial_\mu \phi_i X_n^\mu (\phi) \right) - \mathcal{L} X_n^0 (\phi) + \Lambda_n^0 (\phi) \right] \ . \tag{5.20}$$

Then let's try to compute a canonical Poisson's bracket in the space of fields for evaluating the bracket between the field the conserved Noether's charge. For this case, it is needed to introduce one more property of the Poisson's bracket:

$$\{ f \cdot g, h \} = \{ f, h \} \cdot g + \{ g, h \} \cdot f \ . \tag{5.21}$$

Before evaluation of the Poisson's bracket we also need to remember that:

$$\pi_j = \frac{\partial \mathcal{L}}{\partial (\partial_0 \phi_j)} \ . \tag{5.22}$$

Due to formulas (5.21) and (5.23) we can write that:

$$\{ \phi_i (x), J_n \} = \int \mathrm{d}\vec{y} \left(\{ \pi_j (y), \phi_i (x) \} (\Phi_{j,n} (\phi) - \partial_\mu \phi_i X_n^\mu (\phi)) + \pi_j (y) \{ \phi_i (x), \partial_\mu \phi_j \} X_n^\mu (\phi) - \{ \phi_i (x), \mathcal{L} (\tag{5.23}$$

where the nontrivial impact to the Poisson's bracket was recieved only from first two terms in the bracket of (5.20)/

Then it is needful to compute the Poisson's bracket between $\phi_i (x)$ and the Lagrangian density

$$\{ \phi_i (x), \mathcal{L} \} = \{ \phi_i (x), \mathcal{L} (\phi, \partial_\mu \phi) \} \ , \tag{5.24}$$

where Lagrangian density is a function of ϕ and $\partial_\mu \phi$. Therefore, the rule presented in (5.21) can be used and Lagrangian density can be differentiated with respect to ϕ:

$$\{ \phi_i (x), \mathcal{L} \} = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \{ \phi_i (x), \partial_\mu \phi_j \} \ . \tag{5.25}$$

It also should be noticed that since $\mathcal{L}$ is a Lagrangian density and it is known that this is a function of ϕ_j and the derivative of ϕ_j, where index j means that, in general, there are many fields, it can be assumed that the bracket acts on any function on the phase

space in the same way as differentiation. For instance, we would compute the Poisson's bracket for the usual case of mechanics, if there are q_i and a function of q_j and p_j, as

$$\{q_i, f(q_j, p_j)\} = \frac{\partial f}{\partial q_j}\{q_i, q_j\} + \frac{\partial f}{\partial \rho_j}\{q_i, p_j\} \ , \tag{5.26}$$

where it was supposed that the bracket acts as a derivative and there is the rule that if there is a derivative of a composite function $f(g(x))$, it could be simplified into

$$\frac{\mathrm{d}}{\mathrm{d}x} f(g(x)) = \frac{\partial f}{\partial g}\frac{\partial g}{\partial x} \ . \tag{5.27}$$

Due to rule (5.27), (5.26) can also be written as:

$$\{q_i, f(q_j, p_j)\} = \frac{\partial f}{\partial q_j}\frac{\partial q_j}{\partial q_i} + \frac{\partial f}{\partial \rho_j}\frac{\partial \rho_j}{\partial q_i} \ . \tag{5.28}$$

If we return to (5.25), it will be seen that the non-trivial contribution might come only if μ equal to zero and it is possible to get that:

$$\{\phi_i(x), \mathcal{L}\} = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi_j(x))}\{\phi_i(x), \partial_0 \phi_j\} \ . \tag{5.29}$$

Let's look at the Poisson's bracket in (5.29). It is clear that this break will be non-trivial only if i equals to j.

Then it is possible to find that

$$\{\phi_i(x), \mathcal{L}\} = \pi_j\{\phi_i(x), \partial_0 \phi_j(y)\} \ . \tag{5.30}$$

Now it is possible to substitute the gotten result back into the original expression (5.23) and see, that there is a cancelation between two terms. As the result it will be obtained that:

$$\{\phi_i(x), J_n\} = \int \mathrm{d}\vec{y}\{\pi_j(y), \phi_i(x)\}\left(\Phi_{j,n}(\phi) - \partial_\mu \phi_i X_n^\mu(\phi)\right) \ , \tag{5.31}$$

where the second term in (5.23) also non-trivial only if $\mu = 0$.

Then (5.14) can be used to simplify formula (5.31):

$$\{\phi_i(x), J_n\} = \int \mathrm{d}\vec{y}\delta_{ij}\delta(\vec{x} - \vec{y})\left(\Phi_{j,n}(\phi) - \partial_\mu \phi_i X_n^\mu(\phi)\right) \ . \tag{5.32}$$

Important to mention that delta-functions $\delta(\vec{x} - \vec{y})$ and δ_{ij} also can be used in order to get the final expression:

$$\{\phi_i(x), J_n\} = \Phi_{i,n} - \partial_\mu \phi_i X_n^\mu \ . \tag{5.33}$$

What is written down in the (5.33) is nothing else as variation of the form of the field:

$$\{\phi_i(x), J_n\} = \overline{\delta}_n \phi_i \ . \tag{5.34}$$

Sometimes the variation of the form (5.34) in classical differential geometry is called the Lie derivative.

If generators J_n are summed with parameters ε^n of symmetry transformations, it will be obtained that:

$$\{\phi_i(x), J_n \varepsilon^n\} = (\Phi_{i,n} - \partial_\mu \phi_i X_n^\mu)\, \varepsilon^n = \overline{\delta}\phi_i \ , \tag{5.35}$$

where we got on the right side just a variation of the form of the field ϕ_i under infinitesimal transformations. And this is exactly what was meant by a statement that infinitesimal variations of the form or infinitesimal symmetry transformations are generated by the conserved Noether's charges with respect to the canonical Poisson's structure.

$$\overline{\delta}\phi_i = \{\phi_i(x), J\} \ . \tag{5.36}$$

It is possible to give a concrete important example, which is a demonstration of the property that conserved charges shows themselves up as symmetry generators.

Let's consider, for instance, time translation. It is known that this symmetry corresponds to the conservation of energy. Time variable gets translated by adding to x^0 infinitesimal parameter ε:

$$x^0 \to x'^0 = x^0 + \varepsilon \ . \tag{5.37}$$

Let's also consider for simplicity just single scalar field. It is known that there is a conserved current corresponding to the symmetry and this current is component of a stress energy tensor T_0^μ. As it well known this component is conserved and the local conservation of energy for this component is

$$\partial_\mu T_0^\mu = 0 \ . \tag{5.38}$$

The corresponding Noether's charge for this component has the next form:

$$J = \int \mathrm{d}\vec{x} T_0^0(x) \ , \tag{5.39}$$

where there was an integration over the spatial directions.

It also should be noticed that if J is differentiated with respect to x^0, the following will be recieved

$$\frac{\mathrm{d}}{\mathrm{d}x^0} J = 0 \ . \tag{5.40}$$

Let's remember the expression for the stress-energy tensor for the case of a single scalar field. It is found that

$$T_{\nu}^{\mu} = \frac{\partial \mathcal{L}}{\partial \left(\partial_{\mu}\phi\right)}\partial_{\nu}\phi - \delta_{\nu}^{\mu}\mathcal{L} \ . \tag{5.41}$$

Our takk is about T_0^0 component and due to this fact it is possible to place 0 and 0 instead of indexes μ and ν and get that

$$T_{0}^{0} = \frac{\partial \mathcal{L}}{\partial \left(\partial_{0}\phi\right)}\partial_{0}\phi - \mathcal{L} \ . \tag{5.42}$$

Then the following may be done. It is known that the expression

$$\frac{\partial \mathcal{L}}{\partial \left(\partial_{0}\phi\right)} \tag{5.43}$$

is momentum of the field ϕ and therefore in expression (5.42) the momentum multiplied by velocity minus Lagrangian density is used, but according to the Hamiltonian formalism the gotten expression is nothing else as a Hamiltonian density of the field:

$$\mathcal{H} = \frac{\partial \mathcal{L}}{\partial \left(\partial_{0}\phi\right)}\partial_{0}\phi - \mathcal{L} = \pi\partial_{0}\phi - \mathcal{L} \ . \tag{5.44}$$

Therefore, the Hamiltonian density conserve Noether's charge.

In this case, J can be expressed as

$$J = \int \mathrm{d}\vec{x}\left(\pi\partial_{0}\phi - \mathcal{L}\right) = \int \mathrm{d}x\mathcal{H} \ . \tag{5.45}$$

If the goal is to have the proper normalization constants, it is necessary to suggest that

$$J = \frac{1}{c}H \ . \tag{5.46}$$

So, next expression is recieved

$$H = c\int \mathrm{d}x\mathcal{H} \ . \tag{5.47}$$

So, Hamiltonian appears to coincide with a Noether's charge corresponding to time translations. This means that in fact,

$$\left\{\phi\left(x\right), J\right\} = -\partial_{0}\phi\left(x\right) \ , \tag{5.48}$$

where the fact that for a scalar field the response Φ on translations of coordinates and time is equal to zero and X_n^{μ} is proportional to δ_0^{μ} was used.

It also should be noticed that other components responsible for the conservation of the total momentum are related to the

$$P_i = \int \mathrm{d}\vec{x} T_i^0 \ , \tag{5.49}$$

where i is running from 1 to 3. This would be the next example which will be studied, but before this it is necessary to use formula (5.46) to get that

$$\{\phi(x), H\} = -\frac{\partial}{\partial t}\phi(x) \ . \tag{5.50}$$

If the place of ϕ and H is changed, according to the rule of replacement of the Poisson's bracket the next formula will be obtained:

$$\{H, \phi(x)\} = \dot{\phi}(x) \ . \tag{5.51}$$

The exact manifestation of Hamilton's equations of motion was received in the (5.51), which was declared at the beginning of lecture. It can be seen that the infinitesimal generator of time translations according to this formalism is nothing else as the Hamiltonian for the scalar field. It is possible to make a conclusion that the Hamiltonian generates time translation. On the other hand, translation of time is the symmetry, such that the corresponding conserved Noether's charge coincides with the Hamiltonian according to the Noether's theorem and the Hamiltonian generates infinitesimal translations in time.

It is possible to do more general setting without split into spatial components. Corresponding to momentum conservation and energy conservation it is possible to construct the generator P_μ, which is a generator of space-time shifts, where index μ running from 1 to 3. According to this discussion it will be found out that the P_μ will be equal to

$$P_\mu = \int \mathrm{d}\vec{x} \left(\pi^\rho \partial_\mu \phi_\rho - \mathcal{L}\delta_\mu^0\right) \ . \tag{5.52}$$

It is also attainable to ask the same question about space-time rotations. It is known that they are generated by the components of the angular momentum $M_{\mu\nu}$:

$$M_{\mu\nu} = \int \mathrm{d}\vec{x} M_{\mu\nu}^0 \tag{5.53}$$

A very similar computation should be completed and it can be completed as an additional exercise. As a result the following expression should be found:

$$M_{\mu\nu} = \int \mathrm{d}\vec{x} \left(\phi_\mu \pi_\nu - \phi_\nu \pi_\mu + x_\nu \left(\pi^\rho \partial_\mu \phi_\rho - \mathcal{L}\delta_\mu^0\right) - x_\mu \left(\pi^\rho \partial_\nu \phi_\rho - \mathcal{L}\delta_\nu^0\right)\right) \ . \tag{5.54}$$

Such a way we got expressions for two generators of Poincare group (5.52) and (5.54) as functionals on a phase space generated by fields ϕ_μ and momentums π_μ. Even more to that, it is possible to compute Poisson's brackets between the gotten functionals by using fundamental Poisson's brackets. And it will be obtained that with respect to Poisson's brackets generators will reproduce the standard relations between the generators of the Poincare algebra:

$$\begin{cases} \{M_{\mu\nu}, P_\sigma\} = \eta_{\nu\sigma} P_\mu - \eta_{\mu\sigma} P_\nu \\ \{M_{\mu\nu}, M_{\rho\sigma}\} = \eta_{\nu\rho} M_{\mu\sigma} - \eta_{\mu\rho} M_{\nu\sigma} - \eta_{\nu\sigma} M_{\mu\rho} + \eta_{\mu\sigma} M_{\nu\rho} \end{cases} \tag{5.55}$$

It should be noticed that in quantum theory all Poisson's brackets will be replaced by commutators. Such a way it will be possible to get relations between Poincare generators as operators, and then this will be a relations of the Lie algebra of the Poincare group.

Now let's come to the next important topic which is called Klein-Gordon field.

Klein-Gordon field

First of all, let's write the action for this field:

$$S[\phi] = \frac{1}{c} \int \mathrm{d}^4 x \left[\frac{1}{2} \partial_\mu \phi(x) \, \partial^\mu \phi(x) - \frac{1}{2} \left(\frac{mc}{\hbar} \right)^2 \phi^2(x) \right] , \tag{5.56}$$

where it was continued to denote the field as $\phi(x)$. It was recieved that the Klein-Gordon field is a massive relativistic scalar field. The first term in brackets of the (5.56) is a term, which contains kinetic energy and it has already been derived something similar in the first lecture. This term was derived from the discrete approach, when the model of masses connected by springs in the limiting case to infinity case and the length of springs tended to zero was considered. And on the top of that one may add something, which is proportional to ϕ^2 multiplied with the coefficient, which is designed in such a way that the physical dimension of the kinetic energy and the term which contains ϕ^2 is the same. Four-dimensional integration measure $\mathrm{d}^4 x$ is understood as

$$\mathrm{d}^4 x = \mathrm{d}x^0 \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3 , \tag{5.57}$$

where as it is known

$$\mathrm{d}x^0 = c\mathrm{d}t \tag{5.58}$$

and, therefore, the speed of light in front of the action stands to cancel.

The action is relativistic because derivatives ∂_μ and ∂^μ are paired in such a way that the corresponding term will be invariant under Lawrence transformations.

Field ϕ itself is a scalar and it is known how it transforms under coordinate transformations. From the action S it is attainable to straightforwardly derive equations of motion and they have the following form:

$$\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\phi\left(x\right) = 0\ . \tag{5.59}$$

Equation (5.59) is the Euler-Lagrange equation.

It is also possible to rewrite (5.59) more explicitly in the following way:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_i^2} + ()^2\right)\phi\left(x\right) = 0\ . \tag{5.60}$$

Equation (5.60) has a name of Klein-Gordon equation. Sometimes equation (5.59) also written in a way where the operator $\partial_\mu\partial^\mu$ is replaced by

$$\Box = \partial_\mu\partial^\mu\ . \tag{5.61}$$

It also can be noticed that in fact, the combination $\frac{mc}{\hbar}$ in an inverted way has a meaning of

$$\lambda = \frac{\hbar}{mc}\ , \tag{5.62}$$

where λ has a dimension of length and is called Compton wavelengths associated with the particle of mass m.

Usually people prefer to work in the natural units, where

$$[\hbar] = [c] = 1 \tag{5.63}$$

and therefore the term $\frac{mc}{\hbar}$ is simply becomes m. That is why the action for the field ϕ can be written as

$$S\left[\phi\right] = \int \mathrm{d}^4x\left(\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - m^2\phi^2\right)\ . \tag{5.64}$$

It can be seen that if standard units are selected, then it is possible to write the action in the following form:

$$S\left[\phi\right] = \int \mathrm{d}t\mathrm{d}x\left[\frac{1}{2c^2}\dot{\phi}^2 - \frac{1}{2}\left(\vec{\nabla}\phi\right)^2 - \frac{1}{2}\left(\frac{mc}{\hbar}\right)^2\phi^2\right]\ . \tag{5.65}$$

Now it is available to straightly develop the Hamiltonian formalism and this is something that is needed for quantization. So, it is possible to derive canonical momentum for the field ϕ:

$$\pi\left(x\right) = \frac{\delta L}{\partial\dot{\phi}\left(x\right)}\ . \tag{5.66}$$

By the way, what is written down in the (5.65) is an integral over time of Lagrangian:

$$S\left[\phi\right] = \int \mathrm{d}t L \ . \tag{5.67}$$

Also it can be seen that if the Lagrangian is differentiated or the variation of the Lagrangian is taken with respect to $\dot{\phi}$, it will be found that:

$$\pi\left(x\right) = \frac{\dot{\phi}\left(x\right)}{c^2} = \frac{1}{c}\partial_0\phi\left(x\right) \ . \tag{5.68}$$

If the canonical procedure of passing from Lagrangian to the Hamiltonian is used, the Hamiltonian for the Klein-Gordon field will be obtained, which has the following form:

$$H = \frac{1}{2}\int \mathrm{d}\vec{x}\left[c^2\pi^2 + \partial_i\phi\partial_i\phi + \left(\frac{mc}{\hbar}\right)^2\phi^2\right] \ . \tag{5.69}$$

It is also interesting to know what is the physical dimension of the field ϕ. So this ϕ will be denoted in the square brackets $[\phi]$ and this will be physical notion for the physical dimension. Such a notation is used for dimension of any physical quantity. It is just taken in square brackets and transferred to the physical dimension.

It is known that the physical dimension of the Hamiltonian is a dimension of energy:

$$[H] = \varepsilon \ . \tag{5.70}$$

It is possible to apply this approach to the (5.69) and get that

$$[\varepsilon] = l^3 \times \frac{1}{l^2} \times [\phi]^2 = l\,[\phi]^2 \ , \tag{5.71}$$

where the fast that $\mathrm{d}\vec{x}$ has a dimension of l^3 and the inverse Compton length $\left(\frac{mc}{\hbar}\right)^2$ has a dimension of $\frac{1}{l^2}$ was used. Such a way it will be recieved that

$$[\phi] = \sqrt{\frac{\varepsilon}{l}} \ . \tag{5.72}$$

Such a quantity in classical physics $\frac{\varepsilon}{l}$, which has a meaning of energy per unit length, is called tension. So field ϕ has a physical dimension of a square root of tension.

It is possible to use the same procedure to find the physical dimension of the momentum. By doing analogous computation it will be found that

$$[\pi] = \frac{1}{cl}\sqrt{\frac{\varepsilon}{l}} \ . \tag{5.73}$$

It is also attainable to rewrite (5.73) as

$$[\pi] = \sqrt{\frac{\varepsilon}{c^2l^3}} = \sqrt{\frac{mc^2}{c^2l^3}} = \sqrt{\frac{m}{l^3}} \ . \tag{5.74}$$

It was obtained that the momentum has a physical dimension of square root of mass divided by volume or square root of mass density. If the action of the field ϕ is considered, it will be seen that it is possible to compute the physical dimension of the action:

$$[S] = \frac{1}{c} \times l^4 \times \frac{1}{l^2} \times [\phi]^2 = \frac{l^2}{c}\frac{\varepsilon}{l} = \frac{l \times \varepsilon}{c} \tag{5.75}$$

The length divided by velocity was recieved in the (5.75), which gives us time and it can be finally computed that:

$$[S] = t \times \varepsilon \ . \tag{5.76}$$

Also it should be noticed that time multiplied by energy is nothing else as momentum multiplied by length, which is the same as as angular momentum:

$$[S] = p \times l = [\hbar] \ . \tag{5.77}$$

This is a very important fact that the physical dimension of action is always the same as a dimension of the Planck constant or a dimension of the angular momentum. It does not only concern a scalar field it concerns any field.

Another important fact is that if there are some two functionals of field momentum and the field itself, for instance, $F[\pi, \phi]$ and $G[\pi, \phi]$.

Then it will be possible to find the physical dimension of the equal time Poisson's bracket given by the formula that was discussed at the beginning of the lecture. The physical dimension of this bracket, which will be denoted as square bracket of $\{F, G\}$, will be offset from the physical dimension of the product:

$$[F] \cdot [G] \tag{5.78}$$

by $\hbar$. In other words, it is possible to write that:

$$[\{F, G\}] = \frac{[F] \cdot [G]}{\hbar} \ . \tag{5.79}$$

This can be easily understood from the definition of the Poisson's bracket. It is known that the Poisson's bracket is a differential operation, which explicitly given by the following formula:

$$\{F, G\} = \int \mathrm{d}\vec{x} \left[\frac{\delta F}{\delta \pi(x)}\frac{\delta G}{\delta \phi(x)} - \frac{\delta F}{\delta \phi(x)}\frac{\delta G}{\delta \pi(x)}\right] \ , \tag{5.80}$$

It has been already computed that dimensions of $\pi(x)$ and $\phi(x)$ are:

$$\begin{cases} [\pi(x)] = \frac{1}{cl}\sqrt{\frac{\varepsilon}{l}} \\ [\phi(x)] = \sqrt{\frac{\varepsilon}{l}} \end{cases} \tag{5.81}$$

Therefore, in order to compute the dimension of the right hand side of the (5.80) first of all it is necessary to find what is the dimension of the variational derivative. To complete this it is necessary to consider what is the variation δF:

$$\delta F = \int \mathrm{d}\vec{x} \frac{\delta F}{\delta \phi(x)} \delta\phi(x) \ . \tag{5.82}$$

Thus, the result is in (5.82) is what is called variation of functional F. From this formula it can be seen that the dimension of the variation is the same as a dimension of F and equal to:

$$[\delta F] = [F] = l^3 \times \left[\frac{\delta F}{\delta \phi}\right] \times [\phi] \ , \tag{5.83}$$

where the fact that the integral over spatial coordinates equal to l^3 and that dimension of $\delta\phi$ is the same as a dimension of ϕ was used.

It is also possible to simplify (5.83) and get that:

$$[\delta F] = \left[\frac{\delta F}{\delta \phi}\right] \times l^3 \times \sqrt{\frac{\varepsilon}{l}} \ . \tag{5.84}$$

It is practicable to express $\left[\frac{\delta F}{\delta \phi}\right]$ from the (5.84) and get that:

$$\left[\frac{\delta F}{\delta \phi}\right] = \frac{[F]}{l^{5/2} \times \varepsilon^{1/2}} \ . \tag{5.85}$$

Before evaluation of the dimension of (5.80) it is also necessary to compute variational derivative with respect to the momentum $\pi(x)$. To make this we need to find variation of G:

$$[\delta G] = [G] = l^3 \times \left[\frac{\delta G}{\delta \pi}\right] \times [\pi] = \left[\frac{\delta G}{\delta \pi}\right] l^3 \sqrt{\frac{\varepsilon}{c^2 l^3}} \ . \tag{5.86}$$

The final expression for the (5.86) will have a next view:

$$\left[\frac{\delta G}{\delta \pi}\right] = \frac{[G] \times c}{l^{3/2} \times \varepsilon^{1/2}} \ . \tag{5.87}$$

Now let's compute the dimension of the Poisson bracket. It can be seen that according to the definition (5.80), it can be obtained that

$$[\{F, G\}] = l^3 \times \frac{[F]}{l^{5/2} \times \varepsilon^{1/2}} \times \frac{[G] \times c}{l^{3/2} \times \varepsilon^{1/2}} \ . \tag{5.88}$$

If (5.88) is simplified, it will be found that:

$$[\{F, G\}] = \frac{[F] \times [G] \times c}{l \times \varepsilon} = \frac{[F] \times [G]}{t \times \varepsilon} = \frac{[F] \times [G]}{\hbar} \ . \tag{5.89}$$

It has been declared that the dimension of the Poisson's bracket of two functionals is not equal to the product of the dimensions of this functionals but it's offset from them by $1/\hbar$:

$$[\{F, G\}] = \frac{[F] \times [G]}{\hbar} \ . \tag{5.90}$$

Then let's return back to the Klein-Gordon equation and will write it down in the Hamiltonian form.

Let's write the evolution of the field and the momentum, which will be given by the Poisson's bracket of the field and the momentum with the Hamiltonian:

$$\begin{cases} \dot{\phi}(x) = \{H, \phi(x)\} \\ \dot{\pi}(x) = \{H, \pi(x)\} \end{cases} \tag{5.91}$$

Expressions in the (5.91) are Hamiltonian equations of motion. If this Poisson's brackets are evaluated , it will be obtained that:

$$\begin{cases} \dot{\phi} = c^2 \pi \\ \dot{\pi} = \partial_i^2 \phi - \left(\frac{mc}{\hbar}\right)^2 \phi \end{cases} \tag{5.92}$$

It is possible to exclude momentum of the field in the (5.92) by taking a second derivative of ϕ:

$$\ddot{\phi} = c^2 \dot{\pi} = c^2 \partial_i^2 \phi - c^2 \left(\frac{mc}{\hbar}\right)^2 \phi \tag{5.93}$$

If both sides of the (5.93) are sevided by c^2 and everything is transferred to the left hand side, thw dollowing will be obtained

$$\frac{1}{c^2} \ddot{\phi} - \partial_i^2 \phi + \left(\frac{mc}{\hbar}\right)^2 \phi = 0 \ , \tag{5.94}$$

which is exactly the Klein-Gordon equation.

Also it is possible to rewrite (5.94) in the following form:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_i^2} + \left(\frac{mc}{\hbar}\right)^2\right) \phi = 0 \ , \tag{5.95}$$

The next important point to discuss is so called mass-shell condition and in a way this scalar field on a mass-shell amounts just to solving the Klein-Gordon equation. So there is the Klein-Gordon equation and now our goal is to solve it. How to do it?

How to solve Klein-Gordon equation?

The most efficient way to look at the solutions and to understand how solutions look like is to go to the Fourier space. So it is just necessary to make a Fourier transform by using the Fourier integral:

$$\phi(x) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^4 k e^{ikx} \tilde{\phi}(k) \ , \tag{5.96}$$

where $\tilde{\phi}(k)$ is a Fourier image of the field ϕ.

The expression kx in the exponent of the (5.96) is a Lawrence invariant scalar product, which is by definition:

$$k \cdot x = k_\mu x^\mu = k^0 x^0 - \vec{k} \cdot \vec{x} \ . \tag{5.97}$$

Very often the value $\vec{k}$ from the product (5.97) has a name of wave vector. It is also possible to place x^0 into this equation and get that:

$$k \cdot x = ck^0 \cdot t - \vec{k} \cdot \vec{x} = \omega t - \vec{k} \cdot \vec{x} \ , \tag{5.98}$$

where ω is called frequency and from this equation:

$$k^0 = \frac{\omega}{c} \ . \tag{5.99}$$

It also should be noticed that if (5.98) is placed into the exponent, it will be obtained that

$$e^{ikx} = e^{i\left(\omega t - \vec{k}\cdot\vec{x}\right)} \tag{5.100}$$

and in quantum mechanics this combination is usually called plane wave.

So Fourier transform for the field $\phi(x)$ is in fact the decomposition of a plane waves. And it can be said that the integral in the (5.96) is an integral of a plane waves which is taken.

Then it is necessary to take this Fourier representation of field $\phi(x)$ and plug it in Klein-Gordon equation and see what will be recieved for the corresponding Fourier image.

One more thing that it should be noticed that very often people use in the books slightly different representation for the Fourier transform, namely, they use so-called energy type variables. It may be found very popular in quantum mechanics, where Einstein and de-Broil formulas are used, which relate energy, frequency and the momentum with a wave vector. It's like a transition between wave characteristics and the particle characteristics. This is done with the help of the Einstein formula where energy is equal to

$$E = \hbar\omega \ . \tag{5.101}$$

Formula (5.101) was discovered by Einstein, when he studied the photo effect.

The second formula, which is due to de-Broil is that momentum of a particle is related to the wave vector of the corresponding wave process, represented by the particle. According to the formula

$$\vec{p} = \hbar \vec{k} \ . \tag{5.102}$$

It is possible to express ω and $\vec{k}$ from formulas (5.101) and (5.102) and get that

$$\omega = \frac{E}{\hbar} \text{ and } \vec{k} = \frac{\vec{p}}{\hbar} \ . \tag{5.103}$$

When $\omega t - \vec{k} \cdot \vec{x}$ is replaced with the new variables, then the exponent will have the next form:

$$e^{i\left(\omega t - \vec{k} \cdot \vec{x}\right)} \to e^{i(E \cdot t - \vec{p} \cdot \vec{x})/\hbar} \ . \tag{5.104}$$

The same plane-wave was recieved in the (5.104), but written by using energy type variables. In this case also, when the expression for the field ϕ is written, the next expression will be obtained:

$$\phi(x) = \int \frac{\mathrm{d}E \mathrm{d}\vec{p}}{(2\pi)^{3/2} c\hbar^4} e^{i(Et - \vec{p}\vec{x})/\hbar} \tilde{\phi}(p) \ . \tag{5.105}$$

The final thing is that if one of representations is taken, for instance, the representation presented in the (5.105) and plug it into the Klein-Gordon equation, it will be discovered that Fourier image satisfies the following equation:

$$\left[\left(\frac{E}{c} \right)^2 - \vec{p}^2 - m^2 c^2 \right] \tilde{\phi}(p) = 0 \ , \tag{5.106}$$

where ϕ depends on p^μ:

$$p^\mu = \left(p^0, \vec{p}\right) = \left(\frac{E}{c}, \vec{p} \right) \ . \tag{5.107}$$

According to (5.107), it is possiblet to rewrite (5.106) in the following form:

$$\left[p_\mu p^\mu - m^2 c^2 \right] \tilde{\phi}(p) = 0 \ . \tag{5.108}$$

There may be a thought that equation (5.108) has only trivial solution $\tilde{\phi}(p) = 0$, because it is possible to think that the only thing to do is to just cancel the bracket before $\tilde{\phi}(p)$. This would be so if the usual functions are used, but, in fact, one of the important things, which happen in quantum field theory and also in classical field theory, is that, in general, there should not be a thought about the field as about usual function on

space-time. The right way is to think about the field as a distribution. So, if there is $\phi(x)$ or $\phi(p)$, they will be distributions on space-time. Another name for this distributions is generalized functions.

A very characteristic example of a distribution is, for instance, Dirac's delta function. This is an example of a function, which is should not be understood as a usual function, but rather it is a distribution. Such a way distributions are linear continuous functionals on a space of basic functions. In any case, if the field ϕ is understood in the distributional sense, then equation (5.108) has the following solution:

$$\tilde{\phi}(p) = \delta\left(p^2 - m^2c^2\right)\varphi(p) \ , \tag{5.109}$$

where $\varphi(p)$ is a good continuous function without having any zero, where $p^2 - m^2c^2 = 0$. The condition

$$p_\mu p^\mu = m^2c^2 \tag{5.110}$$

is called mass-shell condition, because it gives a relationship between the energy and the particle momentum. According to this statement (5.110) has next alternative form:

$$E^2 = \vec{p}^{\,2}c^2 + m^2c^4 \ . \tag{5.111}$$

The condition (5.111) is exactly the relation between energy and momentum of a single relativistic particle from the special relativity:

$$E = \pm\sqrt{\vec{p}^{\,2}c^2 + m^2c^4} \ . \tag{5.112}$$

Lecture 6. More on Klein-Gordon equation. Canonical Quantization

How to solve Klein-Gordon equation?

Let's continue with solving the Klein-Gordon equation. It has been already introduced the Fourier transform of the classical field $\phi(x)$. Such a way, the next transform is made:

$$\phi(x) \to \tilde{\phi}(p) \tag{6.1}$$

and it was found out that for the Fourier image of ϕ solution exists in the sense of generalized functions:

$$\tilde{\phi}(p) = \delta\left(p^2 - m^2c^2\right)\varphi(p) \ , \tag{6.2}$$

where $\tilde{\phi}(p)$ is a solution of the Klein-Gordon equation in the momentum space.

And the condition

$$p^2 - m^2c^2 = 0 \tag{6.3}$$

is called mass-shell condition for relativistic particle and it gives a standard relativistic relationship, because between the energy and the momentum of a particle.

If (6.3) is written with the help of a 3-dimensional vector of momentum and energy, then it will have the following form:

$$\left(\frac{E}{c}\right)^2 - \vec{p}^2 - m^2c^2 = 0 \ . \tag{6.4}$$

The solution of the (6.4) is given by

$$E = \pm c\sqrt{\vec{p}^2 + m^2c^2} \ . \tag{6.5}$$

There are two solutions with a positive and a negative energy. If the goal is to depict the function E as a function of the three dimensional momentum, a two sheeted hyperboloid as it shown on the (fig. 6.1) will be obtained.

Then let's take the solution (6.5) and return it back into the Fourier transform formula and write the field $\phi(x)$ as:

$$\phi(x) = \int \frac{\mathrm{d}E\mathrm{d}\vec{p}}{(2\pi)^{3/2}\, c\hbar^4}\delta p^2 - m^2c^2 e^{i(Et-\vec{p}\vec{x})/\hbar}\varphi(p) \ , \tag{6.6}$$

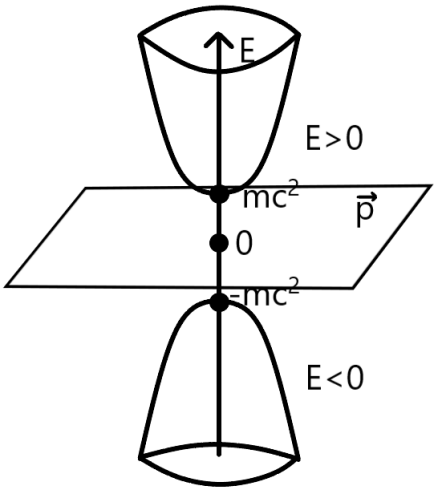

Fig. 6.1. The function E as a function of the three dimensional momentum

where it is also possible to write the (6.6) in the natural units and take

$$c\hbar^4 = 1 \ . \tag{6.7}$$

In the Fourier transform there is a delta-function and this delta-function is useful, because it is possible to use it in order to make an explicit integration. At least, it is possible to integrate over energy variable E and for that it is necessary to apply the formula of a delta-function of a composite variable:

$$\delta\left(f\left(x\right)\right) = \sum_i \frac{\delta\left(x - a_i\right)}{\left|f'\left(a_i\right)\right|} \ , \tag{6.8}$$

where it is apologized that:

$$\begin{cases} f\left(a_i\right) = 0 \\ f'\left(a_i\right) \neq 0 \end{cases} \tag{6.9}$$

Formula (6.8) can be called as a formula for change of variables in the delta-function.

Let's apply (6.8) for the next delta-function

$$\delta\left(\frac{E^2}{c^2} - \vec{p}^{\,2} - m^2c^2\right) \ , \tag{6.10}$$

where (6.10) can be gotten from the $\delta\left(p^2 - m^2c^2\right)$ using (6.5).

Then expression (6.8) will be used for the delta-function $\delta\left(p^2 - m^2c^2\right)$:

$$\delta\left(p^2 - m^2c^2\right) = \frac{c\delta\left(E - c\sqrt{\vec{p}^{\,2} + m^2c^2}\right)}{2\sqrt{\vec{p}^{\,2} + m^2c^2}} + \frac{c\delta\left(E + c\sqrt{\vec{p}^{\,2} + m^2c^2}\right)}{2\sqrt{\vec{p}^{\,2} + m^2c^2}} \ . \tag{6.11}$$

The (6.11) will be easy to understand, if the delta-function is represented as:

$$\delta\left(f\left(E\right)\right) \ , \tag{6.12}$$

where

$$f\left(E\right) = \frac{E^2}{c^2} - \vec{p}^{\,2} - m^2c^2 \ . \tag{6.13}$$

Such a way $f'\left(E\right)$ will be equal to:

$$f'\left(E\right) = \frac{2E}{c^2} \ . \tag{6.14}$$

The derivative $f'\left(E\right)$ should be evaluated at zeros of the function $f\left(E\right)$. In one case, there is

$$f'\left(E\right)\Big|_{E=c\sqrt{\vec{p}^{\,2}+m^2c^2}} = \frac{2\sqrt{\vec{p}^{\,2}+m^2c^2}}{c} \ . \tag{6.15}$$

The same should be done with a negative root and then the formula will be obtained:

$$f'\left(E\right)\Big|_{E=-c\sqrt{\vec{p}^{\,2}+m^2c^2}} = -\frac{2\sqrt{\vec{p}^{\,2}+m^2c^2}}{c} \ . \tag{6.16}$$

The answer, which will be obtained, if delta-function $\delta\left(p^2-m^2c^2\right)$ is integrated in (6.6) is:

$$\phi\left(x\right) = \frac{c}{\left(2\pi\right)^{3/2}\hbar^4}\left[\int\frac{\mathrm{d}\vec{p}}{2E}e^{i(Et-\vec{p}\vec{x})/\hbar}\right]\varphi\left(E,\vec{p}\right) + \int\frac{\mathrm{d}\vec{p}}{2E}e^{-i(Et+\vec{p}\vec{x})/\hbar}\varphi\left(-E,\vec{p}\right) \ , \tag{6.17}$$

where E here is not an independent variable anymore, but represents a positive branch of the dispersion relation (6.5):

$$E = c\sqrt{\vec{p}^{\,2}+m^2c^2} \ . \tag{6.18}$$

It is also possible to make a change $\vec{p}\to-\vec{p}$ in the second integral of the (6.17) and get the formula which people usually use:

$$\phi\left(x\right) = \frac{c}{\left(2\pi\right)^{3/2}\hbar^4}\left[\int\frac{\mathrm{d}\vec{p}}{2E}e^{i(Et-\vec{p}\vec{x})/\hbar}\right]\varphi\left(E,\vec{p}\right) + \int\frac{\mathrm{d}\vec{p}}{2E}e^{-i(Et+\vec{p}\vec{x})/\hbar}\varphi\left(-E,-\vec{p}\right) \ , \tag{6.19}$$

where the first term in the bracket is called positive frequency part of $\phi\left(x\right)$ and the second integral is called negative frequency of $\phi\left(x\right)$.

It also should be noticed that $\mathrm{d}\vec{p}$ does not change the sign, when $\vec{p}\to-\vec{p}$ is transformed, because Jacobian modulus is always positive.

The standard way to proceed further is to introduce the complex amplitudes. One of them is called $a^*(p)$ and this is taken to be

$$a^*(\vec{p}) = \frac{\varphi(E,\vec{p})}{\hbar^3\sqrt{2E}} \ . \tag{6.20}$$

Analogously, one defines an amplitude $a(p)$, which is by definition is:

$$a(\vec{p}) = \frac{\varphi(-E,-\vec{p})}{\hbar^3\sqrt{2E}} \ . \tag{6.21}$$

The gotten functions or amplitudes are only functions of momentum, because E is also a function of momentum and it comes from this function which was originally a function of 4-momentum p in the case of 4-dimensional Minkowski space. In other words, function of E and the function of 3-dimensional momentum $\vec{p}$. It is known that because the delta-function has been integrated, E in the (6.19) is not an independent variable, but it is a positive solution (6.18).

Then it is also possible to check that for a real scalar field amplitudes a and a^* right the following

$$\varphi^*(E,\vec{p}) = \varphi(-E,-\vec{p}) \ . \tag{6.22}$$

Expression (6.22) comes from considering the Fourier transform and using the fact that our talk is about the real scalar field. Therefore $\phi(x)$ is a real function. Then on Fourier amplitude there will be the relation (6.22) and this tells us that amplitudes a and a^* are simply complex conjugate of each other. That's why the star $(*)$ has a simple meaning of complex conjugation.

The Fourier transform, which is used to solve the Klein-Gordon equation, in terms of amplitudes a and a^* takes the following form:

$$\phi(x) = c\hbar^{1/2}\int \frac{\mathrm{d}\vec{p}}{(2\pi\hbar)^{3/2}}\frac{1}{\sqrt{2E}}\left[a^*(\vec{p})\,e^{i(Et-\vec{p}\vec{x})/\hbar} + a(\vec{p})\,e^{-i(Et+\vec{p}\vec{x})/\hbar}\right] \ . \tag{6.23}$$

Then obviously it is also possible to compute the expression for canonical momentum. It is known that canonical momentum is equal to

$$\pi(x) = \frac{\mathrm{d}\phi(x)}{\mathrm{d}t} \ . \tag{6.24}$$

Computing the time derivative from (6.24), the following will be obtained:

$$\pi(x) = \frac{i}{2c\hbar^{1/2}}\int \frac{\mathrm{d}\vec{p}}{(2\pi\hbar)^{3/2}}\sqrt{2E}\left[a^*(\vec{p})\,e^{i(Et-\vec{p}\vec{x})/\hbar} - a(\vec{p})\,e^{-i(Et+\vec{p}\vec{x})/\hbar}\right] \ . \tag{6.25}$$

The important point is the following: it will be seen that the Hamiltonian, which will be obtained, will be well defined and corresponds only to solutions with positive energy and this can be seen in the following way. The idea now is to take our original Hamiltonian H and rewrite it in terms of complex amplitudes a and a^*. What the thing that should be done is to take just the Hamiltonian H, which is equal to:

$$H = \frac{1}{2}\int \mathrm{d}\vec{x}\,\left[c^2\pi^2 + \partial_i\phi\partial_i\phi + \left(\frac{mc}{\hbar}\right)^2\phi^2\right]\ . \tag{6.26}$$

So what is necessary to make next is to take expressions for $\phi(x)$ (6.23) and for $\pi(x)$ and place it instead of $\phi(x)$ and for $\pi(x)$ in the (6.26).

When this substitution is done, one integration over $\vec{x}$ and two integration over momentum $\vec{p}$ will be recieved. It is possible to exchange the order of integration and it is practicable to integrate over $\vec{x}$ firstly. This will affect only on exponentials, because only they contain variable $\vec{x}$. So integration over $\vec{x}$ will give us a delta-function. Delta-function will depend on the difference of momentum related to one copy of ϕ and another will be related to another copy of ϕ. This means that it is possible to use the delta-function, which is obtained after integration over $\vec{x}$ to integrate over one momentum variable. At the end we will be left over with just one integration over momentum. In the process of evaluating H it is needed to use the expression for delta-function, which has the following form:

$$\delta(\vec{p}) = \int \frac{\mathrm{d}\vec{x}}{(2\pi\hbar)^3} e^{i\vec{p}\,\vec{x}/\hbar}\ . \tag{6.27}$$

Then, if the integration over x is done, the result for the Hamiltonian will look as follows:

$$H = \frac{1}{\hbar}\int \mathrm{d}\vec{p}\,E(\vec{p})\,a^*(\vec{p})\,a(\vec{p})\ \ , \tag{6.28}$$

where $E(\vec{p})$ has the form of (6.18).

The expression for the Hamiltonian (6.28) has been written down in terms of complex amplitudes and the Hamiltonian is real and expression $a^*(\vec{p})\,a(\vec{p})$ is positive and equal to:

$$a^*(\vec{p})\,a(\vec{p}) = |a(\vec{p})|^2\ . \tag{6.29}$$

It is also possible to rewrite the Hamiltonian in terms of frequency $\omega(p)$, which is related to the energy by Einstein formula:

$$E(\vec{p}) = \hbar\omega(\vec{p})\ . \tag{6.30}$$

Expression (6.30) will remove Plank's constant from expression (6.28). If (6.29) is rewritten with the help of (6.30), then the Hamiltonian takes the form:

$$H = \int \mathrm{d}\vec{p}\omega\left(\vec{p}\right) a^{*}\left(\vec{p}\right) a\left(\vec{p}\right) \ . \tag{6.31}$$

Further, it is possible to rewrite also an expression for $\phi\left(x\right)$ in terms of frequency:

$$\phi\left(x\right) = c\int \frac{\mathrm{d}\vec{k}}{\left(2\pi\right)^{3/2}} \frac{1}{\sqrt{2\omega\left(\vec{k}\right)}} \left(a\left(\vec{k}\right) e^{-i\left(\omega t - \vec{k}\,\vec{x}\right)} + a^{*}\left(\vec{k}\right) e^{i\left(\omega t - \vec{k}\,\vec{x}\right)}\right) \ . \tag{6.32}$$

The piece in the (6.32), which is related to $a^{*}\left(\vec{k}\right)$ is called positive frequency and a is called negative frequency. The positive and negative is a convention. Sometimes in the books, people use an opposite convention, they call what we call positive frequency negative and what we call negative, they call positive. The definition depends on which literature, you take. This is related to the fact that an opposite way to cause this amplitude is related to ideas that if it is possible to see the Schrodinger equation:

$$i\hbar\frac{\partial}{\partial t}e^{-i\omega t} = \underbrace{\hbar\omega}_{E} e^{-i\omega t} \ , \tag{6.33}$$

then the wave with a negative frequency $-i\omega t$ corresponds to the positive solution.

If the sign in exponent of (6.33) is changed, then the following expression will be obtained:

$$i\hbar\frac{\partial}{\partial t}e^{i\omega t} = -Ee^{i\omega t} \ . \tag{6.34}$$

That's why in the literature sometimes exponent of $-i\omega t$ associated oscillator to it is called positive frequency. But this convention will not be used since Schrodinger equation is not important in our case and Klein-Gordon equation is used. Therefore, positive frequency will be associated with positive exponent of $i\omega t$ correspondent in quantum theory to creation operators. The exponent of $-i\omega t$ will correspond negative frequency and annihilation operators.

Let's look at the physical meaning of amplitudes a^{*} and a. It can be seen that from the solution for field $\phi\left(x\right)$ it is useful to define the time-dependent amplitudes. So, $a^{*}\left(\vec{p},t\right)$ we can associate with $a^{*}\left(\vec{p}\right)$:

$$\begin{cases} a^{*}\left(\vec{p},t\right) = a^{*}\left(\vec{p}\right) e^{i\omega t} \\ a\left(\vec{p},t\right) = a\left(\vec{p}\right) e^{-i\omega t} \end{cases} \tag{6.35}$$

It can be seen that this time-dependent amplitudes can be obtained as a solution of the Hamilton's equations of motion for a and a^* and Hamilton's equations means that

$$\begin{cases} \frac{\mathrm{d}a}{\mathrm{d}t} = \{H, a\} \\ \frac{\mathrm{d}a^*}{\mathrm{d}t} = \{H, a^*\} \end{cases} \tag{6.36}$$

where H is a Hamiltonian, given by formula (6.31). The Poisson bracket is the usual oscillator bracket, which is equal to

$$\{a(\vec{p}), a^*(\vec{p}')\} = i\delta(\vec{p} - \vec{p}') \ , \tag{6.37}$$

where now there is an infinite collection of oscillators parameterized by the continuous parameter p. It is also possible to write following Poisson's brackets:

$$\{a(\vec{p}), a(\vec{p}')\} = 0 = \{a^*(\vec{p}), a^*(\vec{p}')\} \ . \tag{6.38}$$

Such a way a system in terms of complex amplitudes and its equations of motion according to the Hamiltonian formalism were described, where Hamiltonian is given by formula (6.31).

If the classical massive Klein-Gordon field is considered, it is possible to say that this is nothing is just an infinite collection of harmonic oscillators, which oscillate with the relativistic frequency $\omega(\vec{p})$.

Such a representation of a scalar Klein-Gordon field $\phi(x)$ in terms of complex amplitudes a and a^* is called holomorphic representation.

If the Poisson's bracket is used between the oscillators and use this expression is used for fields $\phi(x)$ and $\pi(x)$, which we have showed before, it is possible to compute the Poisson's bracket between $\phi(x)$ and $\phi(x')$, as well as between $\phi(x)$ and $\pi(x')$ and also between $\pi(x)$ and $\pi(x')$:

$$\begin{cases} \{\phi(x), \phi(x')\} = 0 \\ \{\phi(x), \pi(x')\} = i\delta(x - x') \\ \{\pi(x), \pi(x')\} = 0 \end{cases} \tag{6.39}$$

Such a way the standard formulas were restored for canonical Poisson's brackets between field and its momentum, that has been already discussed.

It is also available to write down expressions for other generators of the Poincare group. For instance, Hamiltonian is one of the generators of the Poincare group, but there are

other generators: spatial shifts, spatial rotations and Lorenz boosts. Let's now use this holomorphic representation to obtain expression for other generators of Poincare group in terms of amplitudes a and a^*.

In particular, let's write down the expression for generators of spatial shifts

$$P_i = \int \mathrm{d}\vec{x}\, \pi \partial_i \phi \tag{6.40}$$

in terms of oscillators, which takes the following form:

$$P_i = \frac{1}{\hbar} \int \mathrm{d}\vec{p} p_i a^* \left(\vec{p}\right) a \left(\vec{p}\right) \ . \tag{6.41}$$

Generators of spatial rotations M_{ij} is given by:

$$M_{ij} = \int \mathrm{d}\vec{x} \ \left(x_j \partial_i \phi - x_i \partial_j \phi\right) \ , \tag{6.42}$$

where this rotation generators were obtained from Noether's theorem. If solution for the field ϕ is substituted in terms of complex amplitudes, the following expression will be found:

$$M_{ij} = i \int \mathrm{d}\vec{p} a^* \left(\vec{p}\right) \left(p_i \partial_j - p_j \partial_i\right) a \left(\vec{p}\right) \ . \tag{6.43}$$

Then it is necessary to find expression for Lorenz boosts. It has the following form:

$$M_{0i} = \frac{1}{c} \int \mathrm{d}\vec{x}\, x_i \mathcal{H} - ctP_i \ . \tag{6.44}$$

So Lorenz boosts are the most complicated generators and they have explicit dependence on time that one of the features of this generators.

If (6.31) is rewritten in terms of oscillators, the following expression will be obtained:

$$M_{0i} = \frac{i}{2c} \int \mathrm{d}\vec{p} E \left(\vec{p}\right) \left(a^* \left(\vec{p}\right) \partial_i a \left(\vec{p}\right) - \partial_i a^* \left(\vec{p}\right) a \left(\vec{p}\right)\right) - ctP_i \ , \tag{6.45}$$

where P_i is the same thing as in the (6.41), which is given in terms of oscillators.

So Lorenz boost have an explicit time-dependence and this is a manifestation of the fact that in the Hamiltonian formulation the boost symmetries broken. This is not surprising, because in order to develop the Hamiltonian formalism it is necessary to fix the direction of time.

In the Hamiltonian formalism time plays always a distinguished role, while on the other hand, it is known that when Lorenz boost is done, the time direction is mixed with spatial directions.

If a boost is considered, then it is possible to get an explicit time dependence in the expression for this generator, but the conservation law provided by the Noether's theorem is of course the law, which holds. So, if a total derivative of boost generator over time is taken, it is known that in the Hamiltonian mechanics, if our generator has an explicit time dependence and it is needed to compute the total derivative accurately, first the partial derivative is computed, keeping all the other variables fixed and the boost is differentiated over its explicit time dependence and then a Poisson bracket of the Hamiltonian is added up with M_{0i}:

$$\frac{\mathrm{d}M_{0i}}{\mathrm{d}t} = \frac{\partial M_{0i}}{\partial t} + \{H, M_{0i}\} \ . \tag{6.46}$$

That is how time evolution equation should be written for any function, which has an explicit time dependence, and therefore, if this expression for the boost is differentiated with respect to the time variable, it will be obtained that:

$$\frac{\mathrm{d}M_{0i}}{\mathrm{d}t} = \{H, M_{0i}\} - cP_i \ . \tag{6.47}$$

In other words, if M_{0i} is conserved in time, then the total derivative $\frac{\mathrm{d}M_{0i}}{\mathrm{d}t}$ is equal to zero and from here it is available to get an expression that the Poisson bracket

$$\{H, M_{0i}\} = cP_i \ . \tag{6.48}$$

Expression (6.48) is nothing else as a part of the Poincare algebra. So H generates shifts in time and shifts in time commuting with boosts with respect to the Poisson bracket produce the generators of shifts in spatial directions. This is one way how relations of the Poincare algebra may be reproduced. But if the Poisson bracket between the oscillators is used and the Poisson relations of all these generators is computed between themselves, it will be discovered that they form nothing else as Poincare algebra with respect to this Poisson brackets for oscillator variables a and a^*.

The classical part is finished and the work with classical fields is also finished. The Klein-Gordon equation was solved and rewrote all physical quantities such as Hamiltonian and generators of the Poincare group were rewritten. Then the canonical quantization will be proceeded.

Canonical quantization

Canonical quantization consists in replacing the equal time Poisson bracket of classical fields with quantum Poisson brackets with what is called quantum Poisson bracket denoted

by the same expression as a classical Poisson bracket with a subscript $\hbar$:

$$\{,\} \to \{,\}_\hbar \ . \tag{6.49}$$

The realization of this quantum Poisson bracket is known and according to Dirac, this quantum Poisson bracket must be equal to:

$$\frac{i}{\hbar}[,] \ . \tag{6.50}$$

Expression (6.50) represents a Dirac way of quantizing fields and it's called canonical quantization.

On the one hand, one uses for the quantum Poisson bracket expression (6.49), while for the right hand side of this quantum Poisson bracket one uses the same expression as for the classical Poisson bracket. This means that when we quantize, we promote our classical field $\phi(x)$ becomes an operator. The same with momentum $\pi(x)$: it becomes an operator, acting in some space, which will be described a bit later.

The essential point is that this operators $\phi(x)$ and $\pi(x)$ have the following quantum Poisson brackets:

$$\{\phi(t,\vec{x}),\phi(t,\vec{y})\}_\hbar = \frac{i}{\hbar}[\phi(t,\vec{x}),\phi(t,\vec{y})] \ . \tag{6.51}$$

In expression (6.51) the equal time Poisson bracket is replaced with an equal time commutator. Since this canonical quantization is considered the quantum Poisson bracket is replaced by the value of the classical Poisson bracket which is zero, therefore:

$$\frac{i}{\hbar}[\phi(t,\vec{x}),\phi(t,\vec{y})] = 0 \ . \tag{6.52}$$

This means that fields commutes at different space points, but at the same moment of time. The same should be done for canonical momenta. Also there is zero on the right hand side:

$$\{\pi(t,\vec{x}),\pi(t,\vec{y})\}_\hbar = \frac{i}{\hbar}[\pi(t,\vec{x}),\pi(t,\vec{y})] = 0 \ , \tag{6.53}$$

while for the quantum Poisson bracket between momentum and field the following is obtained:

$$\{\pi(t,\vec{x}),\phi(t,\vec{y})\}_\hbar = \frac{i}{\hbar}[\pi(t,\vec{x}),\phi(t,\vec{y})] = \delta(\vec{x}-\vec{y}) \ . \tag{6.54}$$

Expression (6.54) can be rewritten as a equal time commutator of canonical momentum with the field is given by:

$$[\pi(t,\vec{x}),\phi(t,\vec{y})] = -i\hbar\delta(\vec{x}-\vec{y}) \ . \tag{6.55}$$

It's more convenient to write it as a following commutator:

$$[\phi(t,\vec{x}),\pi(t,\vec{y})] = i\hbar\delta(\vec{x}-\vec{y}) \ . \tag{6.56}$$

Relations between quantum fields are called canonical commutation relations. Similarly, upon quantization the classical amplitudes a^* and a are replaced by operators $a^\dagger$ and a:

$$\begin{cases} a^*(\vec{p}) \to a^\dagger(\vec{p}) \\ a(\vec{p}) \to a(\vec{p}) \end{cases} \tag{6.57}$$

They are understood as creation and annihilation operators. Meaning of creation and annihilation will be clarified a little bit later. So the creation and annihilation operators for harmonic oscillator labeled by this momentum variable $\vec{p}$ and the commutation relation between a and $a^\dagger$ are given by the formula:

$$\begin{cases} [a(\vec{p}),a(\vec{p}\,')] = 0 \\ \left[a^\dagger(\vec{p}),a^\dagger(\vec{p}\,')\right] = 0 \\ \left[a(\vec{p}),a^\dagger(\vec{p}\,')\right] = \hbar\delta(\vec{p}-\vec{p}\,') \end{cases} \tag{6.58}$$

In a way, it can be said that now an abstract algebra generated by generators $a(\vec{p})$ and $a^\dagger(\vec{p})$ is studied. This algebra satisfies the commutation relations (6.58).

The commutativity of fields at different space-time points, but at the same moment of time is important. In fact, from the quantum mechanical point of view commutativity of operators means that these operators can be measured simultaneously. The eigenvalues for this operators can be simultaneously measured. This fields then are independent observable. It is like the same field, but taken as the one and the same moment of time, but at different space points represent different observables, because the values of this field at the same time, but at different space time points commute.

Then, an important representation will be constructed for this moment abstract variables a and $a^\dagger$. Construction of representation means that the a and $a^\dagger$ will be identified as concrete operators acting in some space. It is needed that this space to be also the Hilbert space. The norm in this space will be introduced and the a and $a^\dagger$ will be realized as explicit operators acting on this space. The construction of this particularly important representation of quantum field theory is carried out in the following way: first of all, it is needed to introduce a state in the direct notation bra and ket to introduce a particular state, which can be called as vacuum state

$$|0\rangle \ , \tag{6.59}$$

which is a state without particles from physical point of view. Then it is assumed that this vacuum state is specified by the condition that, when any of the annihilation operators act on it, it gives zero:

$$a\left(\vec{p}\right)|0\rangle = 0 \ . \tag{6.60}$$

That's why these operators have a name of annihilation operators. So, annihilation operators annihilate the vacuum state. Acting on this state with creation operators, which are $a^{\dagger}\left(\vec{p}\right)$ will create a one particle state with momentum $\vec{p}$, which is denoted in this way as:

$$\sqrt{2\omega\left(\vec{p}\right)}a^{\dagger}\left(\vec{p}\right)|0\rangle = |\vec{p}\rangle \ \ , \tag{6.61}$$

where a normalization $\sqrt{2\omega\left(\vec{p}\right)}$ is indroduced.

In expression (6.61) it is recieved that $a^{\dagger}$ acting on the vacuum gives one particle states, which are labeled by the momentum $\vec{p}$.

So, our start was from the vacuum $|0\rangle$, then one particle state with momentum $|\vec{p}_1\rangle$ was constructed. Then it is possiblet to construct a state of two particles with momentums $\vec{p}_1$ and $\vec{p}_2$ - $|\vec{p}_1, \vec{p}_2\rangle$ and so on. More and more creation operators with different momentum are applied and then some number of particles are created, which in physical interpretation carries this momentum from $\vec{p}_1$ up to $\vec{p}_n$ - $|\vec{p}_1, \ldots, \vec{p}_n\rangle$. The space of all these states from vacuum up to n's particle state is called Fock space. Sometimes the representation of field operators in the Fock space is called representation of second quantization. This term will be explained a little bit later. Where does it come from? Why second? What does it mean second quantization? Why people talk about second quantization, when they talk about representation of field operators in the infinity dimensional Fock space?

So, the question is why this is a representation. In the Fock space operators a and $a^{\dagger}$ act in the following way: operator $a^{\dagger}$ simply adds a new particle with momentum $\vec{p}$. If the creation operator is taken and applied to a state, which has already n particles in it like, for instance, state $|\vec{p}_1, \ldots, \vec{p}_n\rangle$, then what this operator does, it's just adds a new particle:

$$a^{\dagger}\left(\vec{p}\right)|\vec{p}_1, \ldots, \vec{p}_n\rangle = \frac{1}{\sqrt{2\omega\left(\vec{p}\right)}}|\vec{p}, \vec{p}_1, \ldots, \vec{p}_n\rangle \ \ , \tag{6.62}$$

where to keep the normalization it is necessary to divide by the square root of $2\omega\left(\vec{p}\right)$.

The action operator or annihilation operator $a\left(\vec{p}\right)$ is a bit more complicated. When this operator acts on a state with n particles, it's supposed to annihilate:

$$a\left(\vec{p}\right)|\vec{p}_1, \ldots, \vec{p}_n\rangle = \hbar\sqrt{2\omega\left(\vec{p}\right)}\sum_{i=1}^{n}\delta\left(\vec{p} - \vec{p}_i\right)\left|\vec{p}_1, \ldots, \hat{\vec{p}}_i, \ldots, \vec{p}_n\right\rangle \ \ , \tag{6.63}$$

where $\hbar\sqrt{2\omega\left(\vec{p}\right)}$ is normalization and $\hat{\vec{p}}_i$ means that this particle is absent. And also it is needed to show distinguish action on the vacuum:

$$a\left(\vec{p}\right)|0\rangle = 0\ . \tag{6.64}$$

So, all the $a^\dagger$ operators commute between themselves, all a operators commute between themselves, that it is possible to verify that this actions are compatible with the non trivial property is that what happens when $a\left(\vec{p}\right)$ and $a^\dagger\left(\vec{p}\right)$ meets. Let's write:

$$a\left(\vec{p}\right)a^\dagger\left(\vec{p}\,'\right) \tag{6.65}$$

and then let's act with this operators on the n particle state and use formulas (6.62) and (6.63):

$$a\left(\vec{p}\right)a^\dagger\left(\vec{p}\,'\right)|\vec{p}_1,\ldots,\vec{p}_n\rangle = \frac{1}{\sqrt{2\omega\left(\vec{p}\,'\right)}}a\left(\vec{p}\right)|\vec{p}\,',\vec{p}_1,\ldots,\vec{p}_n\rangle\ . \tag{6.66}$$

Thenit is neessary to use $a\left(\vec{p}\right)$ operator:

$$\begin{gathered}a\left(\vec{p}\right)a^\dagger\left(\vec{p}\,'\right)|\vec{p}_1,\ldots,\vec{p}_n\rangle = \hbar\delta\left(\vec{p}-\vec{p}\,'\right)|\vec{p}_1,\ldots,\vec{p}_n\rangle + \\ +\hbar\sqrt{\tfrac{\omega(\vec{p})}{\omega(\vec{p}\,')}}\sum_{i=1}^n\delta\left(\vec{p}-\vec{p}_i\right)\left|\vec{p}\,',\vec{p}_1,\ldots,\hat{\vec{p}}_i,\ldots,\vec{p}_n\right\rangle\ .\end{gathered} \tag{6.67}$$

Now the goal is to act with an opposite order of operators:

$$a^\dagger\left(\vec{p}\,'\right)a\left(\vec{p}\right)|\vec{p}_1,\ldots,\vec{p}_n\rangle = \hbar\sqrt{2\omega\left(\vec{p}\right)}\sum_{i=1}^n\delta\left(\vec{p}-\vec{p}_i\right)a^\dagger\left(\vec{p}\,'\right)\left|\vec{p}_1,\ldots,\hat{\vec{p}}_i,\ldots,\vec{p}_n\right\rangle\ . \tag{6.68}$$

Finally, the operator $a^\dagger\left(\vec{p}\,'\right)$ will add one more particle and we will get:

$$a^\dagger\left(\vec{p}\,'\right)a\left(\vec{p}\right)|\vec{p}_1,\ldots,\vec{p}_n\rangle = \hbar\sqrt{\frac{\omega\left(\vec{p}\right)}{\omega\left(\vec{p}\,'\right)}}\sum_{i=1}^n\delta\left(\vec{p}-\vec{p}_i\right)\left|\vec{p}\,',\vec{p}_1,\ldots,\hat{\vec{p}}_i,\ldots,\vec{p}_n\right\rangle\ . \tag{6.69}$$

Then it is necessary to compare the results of two gotten expressions (6.67) and (6.69) for actions of operators a and $a^\dagger$, but in different order. If we subtract from the first actions, the second one, an expression for the commutator will be recieved. If a state with n particles is studied or, in other words, a more or less arbitrary state with any number of particles is considered, it will be clear that under the commutator the second term of (6.67) cancels the first term of (6.69). Such a way, we will get the following formula:

$$\left[a\left(\vec{p}\right),a^\dagger\left(\vec{p}\,'\right)\right]|\vec{p}_1,\ldots,\vec{p}_n\rangle = \hbar\delta\left(\vec{p}-\vec{p}\,'\right)|\vec{p}_1,\ldots,\vec{p}_n\rangle\ . \tag{6.70}$$

The gotten expression is valid for any state in the Fock space. The number of particles plays no role. Therefore, it can be said that the representation of algebra of operators a and $a^\dagger$ has been realized.

This way shows how to prove that representation of the oscillator algebra of creation and annihilation operators in the Fock space has been constructed.

An arbitrary state in the Fock space is a superposition of n particle states. An arbitrary state, which can be denoted by χ, will be given by:

$$|\chi\rangle = \sum_{n=0}^{\infty} \frac{1}{\sqrt{n!}} \int \prod_{i=1}^{n} \frac{\mathrm{d}\vec{p}_i}{\sqrt{2\omega\left(\vec{p}_i\right)}} \chi_n\left(\vec{p}_1, \ldots, \vec{p}_n\right) |\vec{p}_1, \ldots, \vec{p}_n\rangle \ . \tag{6.71}$$

Expression (6.71) represents a superposition, where states with a different number of particles are superposed and this is reflected by the term with sum from zero to infinity and also particles with different momentum are superposed.

The momentum could be different, but since momentum continues, summing over continuous variable is the same as to integrate over this variable and arbitrary coefficients in front of different momentum can be implemented by putting functions $\chi_n\left(\vec{p}_1, \ldots, \vec{p}_n\right)$.

Of course, it is possible to compute the norm of such a state. First of all, there is an assumption about the vacuum state that this is a state with a well defined norm equal to one. The norm of this state is scalar product of the state with itself and by definition is taken to be equal to one:

$$\langle 0|0\rangle = 1 \ . \tag{6.72}$$

It is possible to use the definition of the n particle states to compute the scalar product of states with different particles:

$$\langle \vec{q}_1, \ldots, \vec{q}_n|\vec{p}_1, \ldots, \vec{p}_m\rangle \ . \tag{6.73}$$

What is meant by definition is that it is understood that if a state is created by acting with a creation operator on the vacuum, then conjugation means that the conjugate state will be a state with bra vacuum acted by $a\left(\vec{p}\right)$:

$$a^{\dagger}\left(\vec{p}\right)|0\rangle \to \langle 0|\, a\left(\vec{p}\right) \ . \tag{6.74}$$

In other words, operators a and $a^{\dagger}$ are considered to be Hermitian conjugate to each other. Then, if then this assumption or this convention is used for how the conjugate states are understood, then it is possible to write in terms of annihilation operators the bra state $\langle \vec{q}_1, \ldots, \vec{q}_n|$ and with creation operators the ket state $|\vec{p}_1, \ldots, \vec{p}_n\rangle$. Then it is needed to move operators of annihilation through the creation operators to reach the right vacuum and produce zero:

$$(6.73) = \prod_{k=1}^{n} \sqrt{2\omega\left(\vec{q}_k\right)} \prod_{l=1}^{m} \sqrt{2\omega\left(\vec{p}_l\right)} \left\langle 0\middle|a\left(\vec{q}_1\right) \ldots a\left(\vec{q}_n\right) a^{\dagger}\left(\vec{p}_1\right) \ldots a^{\dagger}\left(\vec{p}_m\right)\middle|0\right\rangle \tag{6.75}$$

Then it is necessary to use commutation relations between a and $a^\dagger$ to move a through $a^\dagger$ to the right. Such a way a will reach the right vacuum and annihilate it and then analogously $a^\dagger$ move to the left and then annihilate the left vacuum. Every time a is commuted through $a^\dagger$ a delta-function will be recieved, due to the commutation relations between a and $a^\dagger$. If this computations are done, the following result will be obtained:

$$(6.73) = \delta_{mn} \sum_{\mathcal{P}} \prod_{i=1}^{n} \left(2\hbar\omega\left(\vec{p}_i\right)\right) \delta\left(\vec{p}_i - \vec{q}_{\mathcal{P}i}\right) \ . \tag{6.76}$$

where symbol $\mathcal{P}$ means a sum of all permutations. Permutation is understood as a map

$$\mathcal{P} = \begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix} \tag{6.77}$$

and permutation is a bijective map from $\{1, 2, \dots, n\}$ to itself. That is what is called a symmetric group. So, such bijective maps from 1 to n form a symmetric group, which mathematical name is S_n. And permutation is an element of the group S_n.

Expression (6.77) shows how the scalar product between particles with definite momentum looks like. If now the scalar product is used to compute the scalar product of an arbitrary state $\langle\psi|\chi\rangle$, which are built on the functions ψ and χ, the following result will be found:

$$\langle\psi|\chi\rangle = \sum_{n=0}^{\infty} \int \psi_n^*\left(\vec{p}_1, \dots, \vec{p}_n\right) \chi_n\left(\vec{p}_1, \dots, \vec{p}_n\right) \prod_{i=1}^{n} \mathrm{d}\vec{p}_i \ . \tag{6.78}$$

This is, in a way the standard scalar product, which will be used use in quantum mechanics to make the space of square integrable functions on a space R^n or simply on R to become a Hilbert space. The (6.78) represents a structure of the Hilbert space is introduced by means of the scalar product and in a sense this is the quantum mechanical generalization or simply generalization to the case of infinite number of particles of the standard quantum mechanical scalar product, which supplies the space of square integrable functions with the structure of the Hilbert space. So, Fock space is a Hilbert space. This also partially explains why the $\sqrt{2\omega\left(\vec{p}\right)}$ is introduced in the definition of states obtained by creation operator, it's explained by the fact that, if a scalar product of two states is considered, for instance, of two one particle states $\langle\vec{p}\,|\vec{q}\rangle$, then it gives us:

$$\langle\vec{p}\,|\vec{q}\rangle = 2\hbar\omega\left(\vec{p}\right)\delta\left(\vec{p} - \vec{q}\right) \ , \tag{6.79}$$

which is the same as

$$\langle\vec{p}\,|\vec{q}\rangle = 2cp^0\delta\left(\vec{p} - \vec{q}\right) \ , \tag{6.80}$$

if it is written in terms of zero component of the momentum, then the scalar product, because of the factor $2\hbar\omega\,(\vec{p})$, is relativistic invariant.

What does it mean? It means that if the Lorentz transformation with p and q or is done, in other words, go to another frame by applying Lorentz transformation on both p and q:

$$p'^{\mu} = \Lambda^{\mu}_{\nu} p^{\nu} \ , \ q'^{\mu} = \Lambda^{\mu}_{\nu} q^{\nu} \ , \tag{6.81}$$

then the product $\langle \vec{p} | \vec{q} \rangle$ will be relativistic invariant.

So, if $\langle \Lambda\vec{p} | \Lambda\vec{q} \rangle$ is done, where Λ is an arbitrary Lorentz transformation, then the scalar product remains invariant:

$$\langle \Lambda\vec{p} | \Lambda\vec{q} \rangle = \langle \vec{p} | \vec{q} \rangle \ . \tag{6.82}$$

Lecture 7. Second Quantization. Commutation and Green's Functions, Pauli-Jordan Function

Last lecture the procedure of canonical quantization was discussed, where the Poisson brackets for fields and the momenta of the fields have been essentially replaced with quantum Poisson brackets that are commutators and the construction of the Fock space has been also discussed, which is an infinitive dimensional Hilbert space constructed by a successive application of creation operators to the unique vacuum state.

In this lecture our goal is to develop these operator concepts.

Now it is possible to come to the question about the Hamiltonian. In the classical theory it is known that the Hamiltonian can be written in terms of creation and annihilation operators in the following way:

$$H = \int \mathrm{d}\vec{p}\,\omega\left(\vec{p}\right) a^{*}\left(\vec{p}\right) a\left(\vec{p}\right) \ . \tag{7.1}$$

The (7.1) was a classical expression for the Hamiltonian. If a similar computation for the Hamiltonian H in quantum field theory is done, of course, it is known that $a^{*}\left(\vec{p}\right)$ and $a\left(\vec{p}\right)$ will be replaced by $a^{\dagger}\left(\vec{p}\right)$ and $a\left(\vec{p}\right)$, but it should be done carefully, because these two operators: $a^{\dagger}\left(\vec{p}\right)$ and $a\left(\vec{p}\right)$, do not commute with each other and, therefore, the order of these operators in any operator expression, which was constructed with the help of the a and $a^{\dagger}$, matters. For the case of the Hamiltonian one accepts to write the following expression for the quantum Hamiltonian:

$$H = \int \mathrm{d}\vec{p}\,\omega\left(\vec{p}\right) a^{\dagger}\left(\vec{p}\right) a\left(\vec{p}\right) \tag{7.2}$$

meaning that operator $a^{\dagger}$ stands on the left from an operator a. This way of ordering of operators a and $a^{\dagger}$ is called normal ordering. In other words, normal ordering is a rule of ordering of operators a and $a^{\dagger}$ in such a way that operators $a^{\dagger}$ always go to the left or always stand on the left from the operator a.

In other words, if there is a certain expression from operators a and $a^{\dagger}$

$$aa^{\dagger}aa^{\dagger}a \tag{7.3}$$

and then the procedure of normal ordering is applied to this expression, then this expression should go to

$$a^{\dagger}a^{\dagger}aaa \ . \tag{7.4}$$

It should be recalled that the order of operators a between themselves and $a^\dagger$-s between themselves does not matter, because operators a commute between themselves and $a^\dagger$ also commute. Their relative order does not matter, but what matters is that all $a^\dagger$-s stand on the left from operators a. To symbolize this way of ordering people use the special notation, which is called normal ordering and they write double dots on the left from the expression, which will be put in normal order and to the right of it. So, expression (7.3) in the normal order form is given by the expression:

$$: aa^\dagger aa^\dagger a : \ . \tag{7.5}$$

The same idea of normal ordering is also applied to fields. If there is a set of fields $\{\phi_1(\vec{x}_1), \dots \phi_n(x_n)\}$ and our goal is to put their product in normal order, that means that in the resulting expression first of all all the fields will be written in terms of creation and annihilation operators and then all creation operators will be placed to the left from all annihilation operators.

The idea of adopting such ordering prescription goes to the fact that the energy of any state in Fock space becomes a well-defined quantity. In particular, vacuum carries zero energy, because if H is applied to the vacuum state then that would be

$$H|0\rangle = \int \mathrm{d}\vec{p}\,\omega(\vec{p})\, a^\dagger(\vec{p})\, a(\vec{p})\,|0\rangle \ . \tag{7.6}$$

Since vacuum is annihilated by operator $a(\vec{p})$, then the result of application of H to the vacuum state will give us zero. On the other hand, if the Hamiltonian is written in the opposite way or if anti-normal ordering is assumed, for instance, where $a^\dagger$ is placed to the right, then, in fact, our Hamiltonian will create first a one particle state out of the vacuum and then this one particle state will be annihilated by means of operator a:

$$a(\vec{p})\, a^\dagger(\vec{p})\,|0\rangle \ . \tag{7.7}$$

As can be seen, in order to compute the result of this action it is necessary to compute $a(\vec{p})$ through $a^\dagger(\vec{p})$, which will result into the $\delta(0)$, because it would be $\delta(\vec{p}-\vec{p})$. So, the $\delta(0)$ is actually infinite and in this case it can be seen that the action of the Hamiltonian on the vacuum state will give us a non-sensic equation, because the result of application will give us just zero. In order to prevent the problem with defining the Hamiltonian of any state in the Fock space and just to make it well defined, it is possible to apply the normal ordering prescription. It's very important to remember about this, that in quantum field theory for operators acting in the Fock space one usually uses the normal

order and prescription. In particular, this is done for the Hamiltonian in order to render the energy of Fock states well defined.

Now if our field ϕ is considered again, then in terms of creation and annihilation operators the field ϕ, which now is an operator has the following expansion. First of all there is the measure factor. If it is written with all $\hbar$ and c constants it will look like:

$$\phi(x) = c\int \frac{\mathrm{d}\vec{p}}{(2\pi\hbar)^{3/2}}\frac{1}{\sqrt{2\omega(\vec{p})}}\left(a^{\dagger}(\vec{p})\, e^{i(Et-\vec{p}\vec{x})/\hbar} + a(\vec{p})\, e^{-i(Et-\vec{p}\vec{x})/\hbar}\right) . \tag{7.8}$$

Let's now take the positive frequency part of the field $\phi(\vec{x})$, which contains $a^{\dagger}$. This piece will be denoted as ϕ^{+}:

$$\phi^{+}(\vec{x},0) = c\int \frac{\mathrm{d}\vec{p}}{(2\pi\hbar)^{3/2}} a^{+}(\vec{p})\, e^{-i\vec{p}\vec{x}/\hbar} . \tag{7.9}$$

Let's take the vacuum state and act on this vacuum state by operator ϕ^{+}. Such a way shows that:

$$\phi^{+}(\vec{x},0)\,|0\rangle = c\int \frac{\mathrm{d}\vec{p}}{(2\pi\hbar)^{3/2}}\frac{1}{\sqrt{2\omega(\vec{p})}} e^{-i\vec{p}\vec{x}/\hbar}\,|\vec{p}\rangle . \tag{7.10}$$

In particular, if another one particle state with a non-trivial momentum $\vec{p}$ is taken and a bra state $\langle\vec{p}|$ is taken, then it will be seen that

$$\big\langle\vec{p}\big|\phi^{+}(\vec{x},0)\big|0\big\rangle = \frac{c}{(2\pi\hbar)^{3/2}}\int \frac{\delta\vec{q}}{\sqrt{2\omega(\vec{q})}}\big\langle\vec{p}\big|a^{+}(\vec{q})\big|0\big\rangle\, e^{-i\vec{q}\vec{x}/\hbar} , \tag{7.11}$$

where the variable, over which the integration is, is denoted by $\vec{q}$ and $\vec{p}$ is a certain fixed momentum, which defines the one particle state $\langle\vec{p}|$. It is possible to compute the quantity described by (7.11) and it is necessary, first of all, to write down an expression for a state $\langle\vec{p}|$. It is known that it can be represented as:

$$\langle\vec{p}| = \langle 0|\, a(\vec{p})\,\sqrt{2\omega(\vec{p})} . \tag{7.12}$$

When (7.12)is plugged in the scalar product from (7.11), $a(\vec{p})$ will be commuted with $a^{\dagger}(\vec{q})$ and $a(\vec{p})$ will be moved to the right vacuum to annihilate it and as a result of this evaluation the delta function on $\vec{p}-\vec{q}$ will be obtained:

$$\big\langle 0\big|a(\vec{p})\, a^{+}(\vec{q})\big|0\big\rangle \sim \delta(\vec{p}-\vec{q})\,\langle 0|0\rangle \sim \delta(\vec{p}-\vec{q}) . \tag{7.13}$$

Then it is necessary to substitute expression (7.13) into the integral (7.11). Such a way, it will be seen that the result of the evaluation will become simply an integral of a delta function and when the delta function is integrated it will be found that:

$$\big\langle\vec{p}\big|\phi^{+}(\vec{x},0)\big|0\big\rangle = \frac{e^{-i\vec{p}\vec{x}/\hbar}}{(2\pi\hbar)^{3/2}} . \tag{7.14}$$

Expression (7.14) is recognized from quantum mechanics. It is known that this is a wave function of a free particle. Every particle moving with momentum $\vec{p}$. In other words, the procedure of applying the positive frequency part to the vacuum positive frequency part of our scalar field to the vacuum creates a particle with a standard quantum mechanical wave function, which corresponds to free particle moving with momentum $\vec{p}$.

Let's also make a comment about the procedure of constructing the Fock space, which also is known under the name second quantization.

Second quantization

The Klein-Gordon equation, which has been solved and already studied, was introduced by Schrodinger himself. His motivation for introducing the Klein-Gordon equation was rather simple. Basically it is known that there is the standard relativistic dispersion relation for relativistic particle:

$$\frac{E^2}{c^2} - \vec{p}^2 - m^2c^2 = 0 \ . \tag{7.15}$$

The relation (7.15) is called as Mass-Shell condition, but in fact this is a dispersion relation for relativistic particle, which relates energy of this particle with its momentum and Schrodinger's idea was simply to pass to quantum mechanics by replacing a three-momentum $\vec{p}$ with an operator of momentum, which equals to

$$\vec{p} = -i\hbar\frac{\partial}{\partial\vec{x}} \tag{7.16}$$

and replacing energy with

$$E \to i\hbar\frac{\partial}{\partial t} \ . \tag{7.17}$$

Energy is, in fact, an eigenstate of the hamiltonian and the hamiltonian generates time evolution. So, shifts in time, while momentum is responsible for the space shifts and it's realized as a derivative from (7.16). After this replacements on the left hand side of the dispersion relation the Klein-Gordon operator is discovered. If this little calculation is done, it will be seen that the operator will have the next form:

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_i^2} + \frac{m^2c^2}{\hbar^2} \ . \tag{7.18}$$

Instead of the usual mass-shell condition, which is formulated in terms of E and $\vec{p}$ now there is a second order differential operator and in a way the natural substitution

of the dispersion relation would be to take this differential operator and apply it to the wave function ϕ:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_i^2} + \frac{m^2c^2}{\hbar^2}\right)\phi = 0 \ . \tag{7.19}$$

In fact, then it is possible to say that the expression, which is in the (7.18), on the one hand mathematically is the kernel of the wave operator. On the other hand, from a physical point of view, written in the (7.19), is a Schrodinger equation for the wave function ϕ. And the fact that this equation contains second order derivative in time rather than first derivative in time is due to the fact that relativism is considered, which puts time and space variables. Since this variables are on equal footing and they are rotated into each other by means of Lorentz transformations, the second derivatives in the space directions are accompanied by second time derivatives in the time direction. In a way this is a relativistic version of the usual non-relativistic Schrodinger equation and this is an equation which defines the wave function ϕ. When the wave function is defined by means of equation (7.19) it is necessaryt to continue with the usual quantum mechanical analogy for the wave function ϕ. Then a number of problems arise immediately. Especially, in the case, when non-trivial interactions or self-interactions are considered for this theory. To clarify this let's introduce instead of a real field actually complex field, because complex field is more suitable to regard it as a wave function, because wave function in the standard quantum mechanics is a complex function. So, if a real scalar field is passed to complex scalar field, then the action for this field is written in the following way:

$$S\left[\phi\right] = \frac{1}{c}\int \mathrm{d}^4x \left[\partial_\mu\phi\partial^\mu\phi^* - \left(\frac{mc}{\hbar}\right)^2\phi^*\phi\right] \ . \tag{7.20}$$

Similar to what happens in the non-relativistic case, it is possible to introduce two important quantities. One of them is a probability density ρ and another is the probability current denoted as $\vec{j}$. These quantities are defined explicitly as

$$\rho = \frac{i\hbar}{2mc^2}\left(\phi^*\partial_t\phi - \phi\partial_t\phi^*\right) \tag{7.21}$$

and

$$\vec{j} = -\frac{i\hbar}{2m}\left(\phi^*\vec{\nabla}\,\phi - \phi\vec{\nabla}\,\phi^*\right) \ . \tag{7.22}$$

These quantities are introduced, because in the non-relativistic quantum mechanics it can be seen that these quantities satisfy the standard continuity equation

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\,\vec{j} = 0 \ . \tag{7.23}$$

The continuity equation is satisfied due to the Klein-Gordon equation, especially, it reduces to the Klein-Gordon equation for a field ϕ, when ρ and $\vec{j}$ are plugged into formula (7.23) will be seen that there will be zero, because the Klein-Gordon equation is satisfied.

Now it looks like the standard paradigm of usual quantum mechanics is followed. ρ and $\vec{j}$ were introduced, which satisfy the continuity equation as our goal is to have this quantity in the non-relativistic case. In fact, it is also possible to understand that what is written down in the (7.23) is nothing else as they're written down the covariant form of the current conservation

$$\partial_\mu J^\mu = 0 \ , \tag{7.24}$$

where J^μ is a 4 current with components

$$J^\mu = \left(cp, \vec{j}\right) \ . \tag{7.25}$$

The current J^μ is conserved in equations of motions and is simply a direct consequence of Noether's theorem, because the existence of this conserved quantity is due to the fact that the action for relativistic complex scalar field has the following symmetry

$$\phi \to e^{i\alpha}\phi \ , \tag{7.26}$$

where α is a constant parameter. From expression (7.26) it can be seen that if the field ϕ is changed by multiplying it with the phase and ϕ^* is simultaneously changed by multiplying it with a conjugate phase, then this phase disappears and the action is simply invariant with respect to these transformations. By Noether's theorem it is known that symmetry must correspond to a conserved current and this current exactly or the probability for current was written in expression (7.25). On equations of motion, which are a Klein-Gordon equations for ϕ and ϕ^* the current J^μ must be conserved and this is exactly what can be verified by direct means.

The problem is however that probability density is not positive. So, ρ is not positive in contrast to the expression for ρ which is in non-relativistic quantum mechanics. In non-relativistic quantum mechanics ρ is defined as

$$\rho = \phi^*\phi = |\phi|^2 \ . \tag{7.27}$$

This is an explicitly positive density. It can be given an interpretation of probability density: probability is is always non-negative.

For our case it can be seen that ρ is not positive from an explicit expression for ρ (7.21). That is because in the relativistic case the second order differential equation is

considered. Since we deal with the second order differential equation supplying it and solving it it is necessary to supply it with initial conditions and since it is of the second order, it is needed to supply as initial conditions not only the value of ϕ, but also the value of its time derivative at an initial moment of time. This values of ϕ and $\dot{\phi}$ can be chosen in an arbitrary way and if they can be chosen in an arbitrary way, so, ϕ and $\dot{\phi}$ might also be taken negative values and therefore ρ is not any more positive defined. In general actually ρ goes in the non-relativistic limit: when a speed of light tends to infinity, it goes to the density of non-relativistic quantum mechanics. This can be explicitly traced by using the expression for ϕ in the so-called WKB approximation and this expression for ϕ then starts from

$$\phi \sim e^{imc^2 t}\psi + \dots \ , \tag{7.28}$$

where $\dots$ are corrections to the expression and ψ is taken as a wave function from non-relativistic theory. Then it can be seen that, when expression (7.28) will be differentiated with respect to time, the first time derivative will be not acting on ψ. It will just

$$\phi_t \sim imc^2 e^{imc^2 t}\psi \ . \tag{7.29}$$

Then from expressions (7.28) and (7.29) it can be seen that at leading order in the limit $c \to \infty$, it will actually be obtained that:

$$\phi^* \partial_t \phi \sim \psi^* \psi \tag{7.30}$$

and time derivative will disappear. So, ρ is nice in all respects that it's a part of the conserved for current. It is a quantity, which has a correct non-relativistic limit, but it loses one important property, which allows to interpret it really as a probability density. Namely it's not positive anymore and, therefore, the whole quantum mechanical interpretation of the quantity ϕ as, in particular, a quantum mechanical wave function breaks. If the interaction is considered, thenit is possible to show that ρ cannot be served as a quantum mechanical probability density.

The correct interpretation of field ϕ has been found later by Pauli and Weiskopf. According to Pauli and Weiskopf, the field ϕ must be treated as an operator rather than the wave function. So, for us ϕ is not a wave function, but rather ϕ is an operator, which then should be used to define with its help the amplitudes and a consequence creation and annihilation operators, which are used to build the Fock space of the scalar field. We naturally come to the Fock space representation treating ϕ as an operator, because this operator should act in some space and the natural space, where it acts is a Fock space.

Such a treatment of ϕ as an operator acting in the Fock space is exactly what is called second quantization. This term is historical, because when second quantization is said, in a way it is suggested that under the first quantization act the classical scalar field is replaced by a wave function ϕ just like Schrodinger did, but in fact it's not a wave function and it is needed to quantize it again to promote it to an operator acting in some Hilbert space, which in this case is a Fock space.

Today, perhaps, it is not a good thing to to talk about first or second quantization, it is just known that it is necessary to treat ϕ from the very beginning as an operator acting in the Fock space and in this respect it is possible to refer to this as quantization of the scalar field, as a process by means of which we an operator is directly put in correspondence to a classical field and then for this operator a representation is constructed. It is represented in terms of creation and annihilation operators.

The next topic that is necessary to introduce is commutation and greens functions.

Commutation and Greens functions

In the theory of interacting fields even in the classical theory, it is known that solutions of inhomogeneous field equations with point-like sources play a special role. These solutions of such inhomogeneous equations with point-like sources have a special name, they are called Greens functions.

As it is known from the classical electrodynamics, greens functions can be of different nature. For instance, there could be retarded Greens functions, advanced Greens functions and, in particular, in quantum field theory the interest is in the so-called causal Greens functions, which are also known under the name of Feynman propagator.

Now let's to discuss how these functions are defined and explicitly constructed for the case of the Klein-Gordon field. Therefore, what is needed to consider here is the Klein-Gordon field.

Let's start from solutions of the Klein-Gordon equation without sources. This is related to such objects as commutators of quantum fields at different space-time points. Since our talk is about commutators, these functions sometimes are referred to as commutation functions.

It is necessary to answer a question: what is the commutator of quantum fields at different space-time points is? Not at the same time, where the commutation relations are canonical, because we have already discussed it, but at different space and time points.

The first very important commutation function, which will be discussed is the so-called Pauli-Jordan function. This function is defined in the following way

$$[\phi(\vec{x},t),\phi(\vec{x}',t')] = i\hbar c\mathcal{D}(x-x') \ , \tag{7.31}$$

where $\mathcal{D}$ is a Pauli-Jordan function and x and x' are points in the four-dimensional space-time. By definition a Pauli-Jordan function is a commutator of two quantum fields at different space-time points. As it will be seen $\mathcal{D}$ is a function and it is not an operator. Although the operator is being calculated, the result is a number c, which is a number or a function multiplied by the identification operator.

Before the derivation is sketched, an explicit form of what is going to found will be written. Since $\mathcal{D}(x-x')$ depends on one argument, it can be written as follows

$$\mathcal{D}(x) = -\frac{\epsilon(t)}{2\pi}\left[\delta\left(x^2\right) - \frac{1}{2}\left(\frac{mc}{\hbar}\right)^2\theta\left(x^2\right)\frac{J_1\left(\frac{mc}{\hbar}\sqrt{x^2}\right)}{\frac{mc}{\hbar}\sqrt{x^2}}\right] \ , \tag{7.32}$$

where $\theta(x^2)$ is Heaviside function, J_1 is a Bessel function. It should be also noticed that in our case instead of x^2 the following is used

$$(x-x')^2 = c^2(t-t')^2 - (\vec{x}-\vec{x}')^2 \ . \tag{7.33}$$

One more variable that it is needed to introduce is ϵ:

$$\epsilon(t) = \theta(t) - \theta(-t) = \begin{cases} +1\,, & t>0 \\ 0\,, & t=0 \\ -1\,, & t<0 \end{cases} \ . \tag{7.34}$$

Let's also remind what is $\theta(t)$:

$$\theta(t) = \begin{cases} +1\,, & t>0 \\ 0\,, & t\leq 0 \end{cases} \tag{7.35}$$

It can be seen from formula (7.32) that, because of the presence of the prefactor $\epsilon(t)$, $\mathcal{D}(x)$ turns out to be zero if $t=t'$. This means that the equal time commutator vanishes, which is good, because it shows that it's compatible with the fact that for the same time argument fields commute. So,

$$[\phi(\vec{x},t),\phi(\vec{x}',t)] = 0 \tag{7.36}$$

and this is the result of canonical quantization and this is how canonical Poisson bracket requires to have when to the quantum theory is passed under the procedure of canonical quantization.

The second feature is that it also can be seen that commutator of two local fields vanishes if their space time points are separated by the space-like interval. So, x^2 is space-like, if $x^2 < 0$. In our case our talk is about condition that if $(x - x')^2 < 0$, then Pauli-Jordan function vanishes. Therefore, fields, which are separated by such an interval are commute and this is in a way if our goal is implementation of causality principle in quantum field theory. This is an important fact, which is manifestation of causality principle in quantum field theory. Why it is so? That's because there is no signal known that can propagate faster than the speed of light and this means that if there are two events or fields which will be measured at points, which are separated by the space-like interval, then measurement of field in one point and the measurement of the field in the other point do not correlate, because these measurements are completely independent. Signal is not enough to propagate from the point x, where a measurement of the field ϕ is made to the point x', where it is implied that a measurement of the field at the point x' is made. In other words, field values separated by the space-like interval are independent observables and it is known that if there are two independent observables in order to be able to measure them according to the basic principles of quantum mechanics, they must commute. The last comment about the explicit expression, which is going to be found by means of computation is that expanding the Pauli-Jordan function around the light cone the following expression will be found:

$$\mathcal{D}(x) = -\frac{\epsilon(t)}{2\pi}\left[\delta\left(x^2\right) - \frac{1}{2}\left(\frac{mc}{\hbar}\right)^2 \theta\left(x^2\right) + \ldots\right] . \tag{7.37}$$

Essentially what we are expanding in the (7.37) is the piece of the (7.32) with J_1, where around light cone means that this function is expanded around $\sqrt{x^2} \approx 0$. So, when $\sqrt{x^2}$ tends to 0 the ratio J_1 to the $\sqrt{x^2}$ goes to 1. Therefore, it can be seen that on the light cone there are severe singularities for the Pauli-Jordan function. Passing through the light cone $\mathcal{D}(x)$ experience singularities of two types. One of them is just delta-function singularity and the second type of singularity is finite discontinuity is given by the function $\theta(x^2)$. So, finite discontinuity is something which is usually understood as function jump, because of the θ-function.

Now the question may be asked: how to derive the gotten result? In particular, how to construct the function $\mathcal{D}$?

Let's explain how it can be done and let's construct a scheme.

1) It is possible to use an expression for fields in terms of oscillators. Let's start from the expression for the commutator

$$\left[\phi\left(\vec{x}\,,t\right),\phi\left(\vec{x}\,',t'\right)\right] \tag{7.38}$$

given in terms of creation and annihilation operators. Up to a prefactor, which contains c, 2π and $\hbar$, the following will be obtained:

$$\begin{aligned} 7.38 \sim \int \tfrac{\mathrm{d}\vec{k}\,\mathrm{d}\vec{k}\,'}{\sqrt{4\omega\omega'}}\Big(\left[a^{+}\left(\vec{k}\right),a\left(\vec{k}\,'\right)\right] e^{i(\omega t-\omega' t')-i\left(\vec{k}\,\vec{x}-\vec{k}\,'\vec{x}\,'\right)}+ \\ +\left[a\left(\vec{k}\right),a^{+}\left(\vec{k}\,'\right)\right] e^{-i(\omega t-\omega' t')+i\left(\vec{k}\,\vec{x}-\vec{k}\,'\vec{x}\,'\right)}\Big)\ . \end{aligned} \tag{7.39}$$

So, the Fourier decomposition of ϕ was used and then formula (7.39) was recieved.

2) Commutation relations between quantum oscillators are used and it is replaced by delta-functions. Then it is possible to integrate the corresponding delta-function. So, the first and the second commutators in the (7.39) will be replaced by

$$\delta\left(\vec{k}\,-\vec{k}\,'\right)\ \text{and}\ -\delta\left(\vec{k}\,-\vec{k}\,'\right) \tag{7.40}$$

respectively. Then it is possible to integrate over $\vec{k}\,'$ by using these delta functions, simplify a bit and as a result the following expression will be obtained:

$$7.38 = i\hbar c\frac{-2c}{(2\pi)^3}\int\frac{\mathrm{d}\vec{k}}{2\omega}e^{i\vec{k}\,(\vec{x}-\vec{x}\,')}\sin\omega\,(t-t')\ . \tag{7.41}$$

where the coefficient was added before expression (7.38). If the gotten expression is compared with the definition of the Pauli-Jordan function, it will be seen that this prefactor $i\hbar c$ is also in the (7.31). So, it is recieved that

$$\mathrm{D}\,(x) = -\frac{2c}{(2\pi)^3}\int\frac{\mathrm{d}\vec{k}}{2\omega}e^{i\vec{k}\,\vec{x}}\sin\omega t\ . \tag{7.42}$$

So, a subject of our further computation is obtained.

3) Then it is necessary to continue the computation and it is needed to evaluate the three-dimensional integral or evaluate it in terms of known functions like, for instance, a Bessel function. It is convenient here to come to spherical coordinates. So, it is necessary to introduce a quantity r which is

$$r = \sqrt{\vec{x}^{\,2}} \tag{7.43}$$

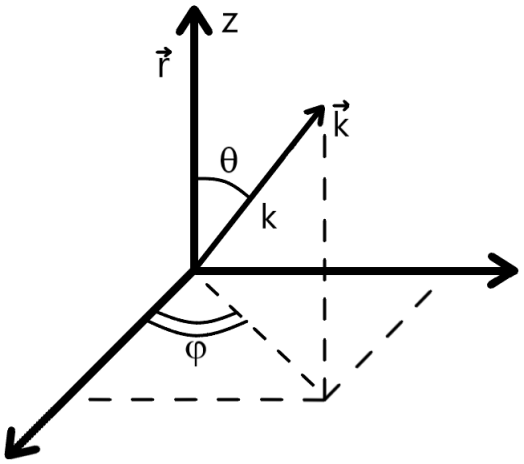

Fig. 7.1. Change of coordinates to r and k variables

and k, which is

$$k = \sqrt{\vec{k}^{\,2}} \ . \tag{7.44}$$

Our coordinate system is oriented along direction of $\vec{r}$ (fig. 7.1).

Therefore, the $\mathcal{D}(x)$ will be obtained in spherical coordinates in the following form:

$$\mathcal{D}(x) = -\frac{2c}{(2\pi)^3}\int_0^\infty \frac{k^2 \mathrm{d}k}{2\omega}\sin\omega t \int_0^{2\pi}\mathrm{d}\varphi \int_0^\pi \sin\theta \mathrm{d}\theta e^{ikr\cos\theta} \ , \tag{7.45}$$

where it was used that

$$e^{i\vec{k}\,\vec{x}} = e^{irk\cos\theta} \ . \tag{7.46}$$

The integral in the (7.45) can be elementary computed. It can be seen that the measure $\sin\theta \mathrm{d}\theta$ can be replaced by

$$\sin\theta \mathrm{d}\theta = -\mathrm{d}(\cos\theta) \ . \tag{7.47}$$

Such a way, integral with exponent can be elementary integrated and then integral over $\mathrm{d}\phi$ is also elementary taken and, therefore, $\mathcal{D}(x)$ upon this integration becomes

$$\mathcal{D}(x) = -\frac{c}{2\pi^2 r}\int_0^\infty \frac{k\mathrm{d}k}{\omega}\sin\omega t \sin kr \ . \tag{7.48}$$

In the (7.48) evaluation was reduced to the one-dimensional integral and this can be further written conveniently in the following way:

$$\frac{c}{2\pi^2}\frac{1}{r}\frac{\partial}{\partial r}\int_0^\infty \frac{\mathrm{d}k}{\omega(k)}\sin\omega t\cos kr \ . \tag{7.49}$$

where ω is the function of k. When we will be integrating back over r, the cosine will turn into a sine, but it also produce k that was in the numerator of the (7.48).

Let's remember what ω is. Now it can be seen that the integral (7.49) is a bit non-trivial, it's not easy to take, because $\omega(k)$ is an explicit function of k, which is not pleasant function, it's equals to

$$\omega(k) = c\sqrt{\vec{k}^{\,2} + \vec{\mu}^{\,2}} = c\sqrt{k^2 + \mu^2}\ , \tag{7.50}$$

where μ is

$$\mu = \frac{mc}{\hbar}\ . \tag{7.51}$$

So, explicitly $\mathcal{D}(x)$ then if we substitute the expression for ω take the following form:

$$\mathcal{D}(x) = \frac{1}{4\pi}\frac{1}{r}\frac{\partial}{\partial r}\int_{-\infty}^{\infty}\frac{\mathrm{d}k}{\sqrt{k^2+\mu^2}}\cos kr\sin x_0\sqrt{k^2+\mu^2}\ , \tag{7.52}$$

where integral limits were changed to from $-\infty$ to $+\infty$. The integration was extended for negative values of k, because the function, that is integrated is even. Then the integral is divided by two and that's why $2\pi^2$ turns into $4\pi^2$ in denominator. An argument of sine was also changed according to the formula

$$t = \frac{x_0}{c}\ . \tag{7.53}$$

Such a way an one dimensional integral (7.52) is obtained, which will be taken. It can be seen that it is quite non-trivial, because it involves the parameter k under the square roots.

It is convenient to introduce the function $F(r,t)$ and simply define it as

$$F(r,t) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{\mathrm{d}k}{\sqrt{k^2+\mu^2}}\cos kr\sin x_0\sqrt{k^2+\mu^2}\ . \tag{7.54}$$

To proceed, of course, it is needed to find a clever choice of variables, which somehow should efficiently deal with the $\sqrt{k^2+\mu^2}$. The way to do it is to introduce the following change of variables:

$$k = \mu\sinh\varphi\,,\ -\infty \le \varphi \le +\infty\ , \tag{7.55}$$

where φ is not an angle, but just a variable. That's efficient change, because if $\sqrt{k^2+\mu^2}$ is now computed, it will be seen that this is

$$\sqrt{k^2+\mu^2} = \sqrt{\mu^2\sinh^2\varphi+\mu^2} = \mu\cosh\varphi > 0\ . \tag{7.56}$$

The gotten in the (7.56) expression is always positive. Moreover, it can be seen that if measure decay is taken it will give us

$$\mathrm{d}k = \mu \cosh\varphi \mathrm{d}\varphi \tag{7.57}$$

and cancels exactly with $\mu \cosh\phi$, which is in the denominator .

Under such change of variables function F will take the following form

$$F(r,t) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \mathrm{d}\varphi \cos\left(\mu r \sinh\varphi\right) \times \sin\left(\mu x^0 \cosh\varphi\right) \ . \tag{7.58}$$

The integral (7.58) is obtained, but the formula can be used further for the product of two trigonometric functions cosine multiplied by sine. So let's proceed and write (7.58) as

$$\begin{aligned} F(r,t) = \tfrac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\varphi \Big[& \sin\left(\mu r \sinh\varphi + \mu x^0 \cosh\varphi\right) - \\ & - \sin\left(\mu r \sinh\varphi - \mu x^0 \cosh\varphi\right) \Big] \ . \end{aligned} \tag{7.59}$$

The first step in our next considerations is to start distinguishing three different cases depending on inequalities between x^0 and r. For definiteness in further computation the case where x^0, which is the same as ct is bigger than r and r is bigger than 0, will be considered. That is needed, because every time a certain inequality is picked up between the ct and r, it is necessary to make its own change of variables in order to proceed. Thus, for the case that is being considered here, it is necessary to perform appropriate further variable modification, namely

$$\frac{x^0}{\sqrt{(x^0)^2 - r^2}} = \cosh\varphi_0 \frac{r}{\sqrt{(x^0)^2 - r^2}} = \sinh\varphi \ . \tag{7.60}$$

Once again, a new change of variables is needed and this is needed in order to turn the sum of sinh and cosh into something simple. Indeed, if a change of variables is done, then the expression for $F(r,t)$ takes the following form

$$\begin{aligned} F(r,t) = \tfrac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}\varphi & \sin\left[\mu\sqrt{(x^0)^2 - r^2}\cosh\left(\varphi + \varphi_0\right)\right] + \\ & + \sin\left[\mu\sqrt{(x^0)^2 - r^2}\cosh\left(\varphi - \varphi_0\right)\right] \ . \end{aligned} \tag{7.61}$$

An interesting point is that since the integration is over all ϕ in the first integral it is possible to make a shift of variables $\phi - \phi_0$ to remove the ϕ_0 from integration and

in the second integral it is possible to make a change $\phi + \phi_0$ to remove this shift. Then integrals are equal to each other and the answer doubles and gives us

$$F(r,t) = \frac{1}{\pi}\int_{-\infty}^{+\infty} \mathrm{d}\varphi \sin\left[\mu\sqrt{(x^0)^2 - r^2}\cosh(\varphi)\right] \ . \tag{7.62}$$

Now finally if we look at the books looking for an explicit integral representation for the Bessel function, then it will be seen that what is written down in the (7.62) is exactly the integral representation for the Bessel function:

$$F(r,t) = J_0\left(\mu\sqrt{(x^0)^2 - r^2}\right) \ . \tag{7.63}$$

So our goal was achieved and the function $F(r,t)$ was reduced to just a Bessel function of index zero. If this calculation is niw done for other inequalities between ct and r, the following will be found:

$$F(r,t) = \begin{cases} +J_0\left(\mu\sqrt{(x^0)^2 - r^2}\right) \ , \ ct > r \\ 0 \quad -r < ct < r \\ -J_0\left(\mu\sqrt{(x^0)^2 - r^2}\right) \ , \ ct < -r \end{cases} \tag{7.64}$$

The three gotten results for $F(r,t)$ can be combined in one expression by saying that $F(r,t)$ is given by

$$F(r,t) = \epsilon(t)\,\theta\left((ct)^2 - r^2\right) J_0\left(\frac{mc}{\hbar}\sqrt{(ct)^2 - r^2}\right) \ . \tag{7.65}$$

4) The final step consists in evaluating $\mathcal{D}(x)$, which is derivative of the expression for $F(r,t)$ with respect to one over r. So, the following is recieved

$$\mathcal{D}(x) = \frac{\epsilon(t)}{4\pi r}\frac{\partial}{\partial r}\left[\theta\left((ct)^2 - r^2\right) J_0\left(\frac{mc}{\hbar}\sqrt{(ct)^2 - r^2}\right)\right] \ . \tag{7.66}$$

When the (7.66) is differentiated over ∂r, differentiating of terms in the brakets will give us delta-function, which has been seen in the expression for $\delta(x)$ and then differentiating the Bessel function with zero index will produce the Bessel function with index one. So, finally the expression for the Pauli-Jordan function will be found:

$$\mathcal{D}(x) = -\frac{\epsilon(t)}{2\pi}\left[\delta\left(x^2\right) - \frac{1}{2}\left(\frac{mc}{\hbar}\right)^2\theta\left(x^2\right)\frac{J_1\left(\frac{mc}{\hbar}\sqrt{x^2}\right)}{\frac{mc}{\hbar}\sqrt{x^2}}\right] \ . \tag{7.67}$$

That's what concerns important commutation function, which is called Pauli-Jordan function and which is a commutator between quantum fields at different space-time points.

In fact, what can be shown further is that it is clear that the commutator has been calculated

$$[\phi(x,t),\phi(x',t')] = i\hbar c\mathcal{D}(x-x') \ . \tag{7.68}$$

It is possible then to show that since x' and t' are different from x and t, it is possible to act on both sides of the relation with the Klein-Gordon's wave operator, which is

$$\Box_{x,t} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_i^2} + \left(\frac{mc}{\hbar}\right)^2 \ . \tag{7.69}$$

It is possible to act with (7.69) on both sides of the (7.68). So, then the next expression will be obtained:

$$[\Box_{x,t}\phi(x,t),\phi(x',t)] = i\hbar c\Box_{x,t}\mathcal{D}(x-x') \ . \tag{7.70}$$

Since ϕ is a solution of the Klein-Gordon equation, the first term in the commutator of the (7.70) is zero and, therefore, on the right hand side it is necessary to find that the result of applying box with respect to coordinates x and t to Pauli-Jordan function also gives us 0. Thus, in fact, what was obtained also means that $\mathrm{D}(x-x')$ is an explicit solution of the Klein-Gordon equation:

$$\Box_{x,t}\mathcal{D}(x-x') \ . \tag{7.71}$$

Lecture 8. Retarded and Advanced Green's Functions. Feynman Propagator. Yukawa Force

In the last lecture we have introduced the Pauli-Jordan function. The Pauli-Jordan function is a commutator of two field operators taken at different space-time points.

Now we will define two more interesting and important functions, namely the retarded Green's function and the advanced Green's function. We will look how they are explicitly constructed and then we will come to the very important notion in quantum field theory, the notion of Feynman propagator.

Let's start with the definition of the retarded Green's function.

Retarded Green's function

So, the retarded Green's function is a Δ_{ret}, which is defined in the following way

$$\Delta_{ret}\left(x-x'\right)=\theta\left(t-t'\right)\left[\phi\left(\vec{x}\,,t\right),\phi\left(\vec{x}\,',t'\right)\right]\ . \tag{8.1}$$

We see that if we compare this to the definition of the Pauli-Jordan function, then what is written down in the (8.1) is simply

$$\Delta_{ret}\left(x-x'\right)=i\hbar c\theta\left(t-t'\right)\mathcal{D}\left(x-x'\right)\ . \tag{8.2}$$

Advanced Green's function

Analogously, the advanced Green's function is defined as

$$\Delta_{adv}\left(x-x'\right)=-\theta\left(t'-t\right)\left[\phi\left(\vec{x}\,,t\right),\phi\left(\vec{x}\,',t'\right)\right]\ . \tag{8.3}$$

Again, we need to compare the (8.3) with the definition of the Pauli-Jordan function:

$$\Delta_{adv}\left(x-x'\right)=-i\hbar c\theta\left(t'-t\right)\mathcal{D}\left(x-x'\right)\ . \tag{8.4}$$

We have to show that the functions introduced in such a way as in (8.2) and (8.4) are really Green's functions. In other words, they solve the Klein-Gordon equation with a delta-function source on the right hand side. This showing can be done in different ways. The simplest way, probably, to take a Δ_{ret} and act on it with the Klein-Gordon propagator.

Of course, we need also to use then the fact that the field $\phi(\vec{x},t)$ solves the Klein-Gordon equation. Let's see how it works.

So, we take the Klein-Gordon operator, which is essentially

$$\left(c^{-2}\partial_t^2 - \partial_i^2 + \mu^2\right) \ , \tag{8.5}$$

where we denoted μ as

$$\mu = \frac{mc}{\hbar} \ . \tag{8.6}$$

And then we apply (8.5) to retarded Green function:

$$\left(-c^2\partial_t^2 - \partial_i^2 + \mu^2\right)\left(\theta(t-t')\left[\phi(\vec{x},t),\phi(\vec{x}\,',t')\right]\right) \ . \tag{8.7}$$

We need first to apply derivative with respect to time and applying it, we will get the following:

$$-c^2\partial_t^2 : \ \partial_t\left(-c^2\delta(t-t')\left[\phi(\vec{x},t),\phi(\vec{x}\,',t')\right] + \theta(t-t')\left[\pi(x,t),\phi(\vec{x}\,',t')\right]\right) \ . \tag{8.8}$$

Then we have the part of the Klein-Gordon operator, which straightforwardly gets inside the bracket and acts on the field ϕ. So, these are derivatives with respect to spatial directions. What we get here is

$$-\partial_i^2 + \mu^2 : \ \theta(t-t')\left[\left(-\partial_i^2 + \mu^2\right)\phi(\vec{x},t),\phi(x',t')\right] \ . \tag{8.9}$$

Now we have to continue and act with the remaining time derivative on the result inside the brackets of the (8.8), but before doing that, we know that, because of the delta-function we would have a non-trivial result only when $t = t'$. So, we go to the support of the delta-function, but in this case, when $t = t'$, we have an equal time commutator of two quantum fields ϕ and we know that from their canonical commutation relations for $t = t'$, this commutator vanishes. So, in fact, the whole term with delta-function produce a vanishing contribution and we can just drop it.

Then we need to apply the time derivative for the second term in the brackets of the (8.7) and the resulting expression for the (8.7) will have the next form:

$$8.7 = \delta(t-t')\left[\pi(\vec{x},t),\phi(\vec{x}\,',t')\right] + \theta(t-t')\left[\left(c^2\partial_t^2 - \partial_i^2 + \mu^2\right)\phi(\vec{x},t),\phi(\vec{x}\,',t')\right] \ . \tag{8.10}$$

What do we see from formula (8.10)? The first thing that we see is that since $\phi(\vec{x},t)$ is a solution of the Klein-Gordon equation, then the first term of the commutator $\left(c^2\partial_t^2 - \partial_i^2 + \mu^2\right)\phi(\vec{x},t)$ in the (8.10) vanishes, because ϕ is a solution.

Such a way the whole result turns to

$$8.7 = \delta\left(t - t'\right)\left[\pi\left(\vec{x}\,, t\right), \phi\left(\vec{x}\,', t'\right)\right] \ . \tag{8.11}$$

We got something, which we can easily compute, because due to the delta-function in the (8.11), the time argument t' can be replaced for t. As a result we got an equal time commutator between the momentum of the field and the field itself. Because of the canonical commutation relation, we know that what is written down in the commutator brackets is in fact

$$\left[\pi\left(\vec{x}\,, t\right), \phi\left(\vec{x}\,', t'\right)\right] = -i\hbar\delta\left(t - t'\right)\delta\left(\vec{x} - \vec{x}\,'\right) \ . \tag{8.12}$$

In other words, we restore on the right hand side of the (8.12) the four dimensional delta-function and the final expression can be written as:

$$8.7 = -i\hbar\delta^{(4)}\left(x - x'\right) \ , \tag{8.13}$$

where

$$x^0 = ct \ . \tag{8.14}$$

Let's act the bracket $(\Box + \mu^2)$ on the retarded Green's function $\Delta_{ret}\left(x - x'\right)$. Then the result of this will be

$$\left(\Box + \mu^2\right)\Delta_{ret}\left(x - x'\right) = -i\hbar c\delta^{(4)}\left(x - x'\right) \ , \tag{8.15}$$

where $\Delta_{ret}\left(x - x'\right)$ is a Green's function for the Klein-Gordon operator.

Why the $\Delta_{ret}\left(x - x'\right)$ is called as retarded? It's called retarded, because due to the presence of θ function in the definition, it can be seen that the Green's function has a distinguished property that $\Delta_{ret}\left(x - x'\right)$ vanishes for $t < t'$. The Green's function with such a property is called retarded.

For $\Delta_{adv}\left(x - x'\right)$ computation is absolutely the same. It can be found that $\Delta_{adv}\left(x - x'\right)$ is also a fundamental solution of the Klein-Gordon equation. The fundamental means that it's a solution with a right hand side represented in the form of a point-like source.

There is another way to show how the $\Delta_{ret}\left(x - x'\right)$ allows to obtain a useful representation for the Green's function.

Let's look at the following representation, which in fact uses the integral representation for the Pauli-Jordan function

$$\Delta_{ret}\left(x\right) = -\hbar c\frac{\theta\left(t\right)}{\left(2\pi\right)^3}\int\frac{\mathrm{d}\vec{k}}{2k^0}\left[e^{i\omega t} - e^{-i\omega t}\right]e^{+i\vec{k}\,\vec{x}} \ . \tag{8.16}$$

Expression (8.16) is integral representation for the function $\Delta_{ret}(x)$, where k_0 is a component of the wave vector

$$k_0 = \sqrt{\vec{k}^{\,2} + \mu^2} = \frac{\omega}{c} \ . \tag{8.17}$$

According to the (8.17), the integral representation can be modified. Let's start from some auxiliary or additional integral, which apparently at first sight has nothing to do with the Δ_{ret}. Then the complex analysis should be used, in particular, as a residue theorem to compute the following integral:

$$\int_{-\infty}^{+\infty} \frac{e^{-ick^0 t}\mathrm{d}k^0}{\left(k^0 + i\epsilon\right)^2 - \vec{k}^{\,2} - \mu^2} \ . \tag{8.18}$$

The integral (8.18) is taken along the real line, but it is more convenient to use a complex analysis. In the (8.18) the complex plane of the variable k^0 was considered for computation of the integral over variable t^0.

Let's look at the denominator of the (8.18) and forget about the little piece $i\epsilon$, which was added to k^0. An ϵ is a positive and small and, therefore, it is a little shift in the imaginary direction. This little term should be used, because without it the denominator in formula (8.18) looks like

$$\left(k^0\right)^2 - \vec{k}^{\,2} - \mu^2 \tag{8.19}$$

and then, when k^0 will be running through the whole real line, eventually it will meet the problem, because at certain values of k^0 the expression in the denominator may become equal to zero. This happens, when

$$k^0 = \pm\sqrt{\vec{k}^{\,2} + \mu^2} \ . \tag{8.20}$$

Therefore, without $i\epsilon$ term singularities along the real line will exist and then the integral will not be well defined and will diverge.

Exactly, to overcome this problem and make sense of the integral one adds up a little term $i\epsilon$ for shift the pole in the lower half plane. It can be seen that now poles are shifted from the real line a little bit down (fig. 8.1).

That's because a different from the (8.19) equation should be solved:

$$\left(k^0 + i\epsilon\right)^2 - \vec{k}^{\,2} - \mu^2 = 0 \ . \tag{8.21}$$

The square root of the (8.21) will give:

$$k^0 + i\epsilon = \pm\sqrt{\vec{k}^{\,2} + \mu^2} \ . \tag{8.22}$$

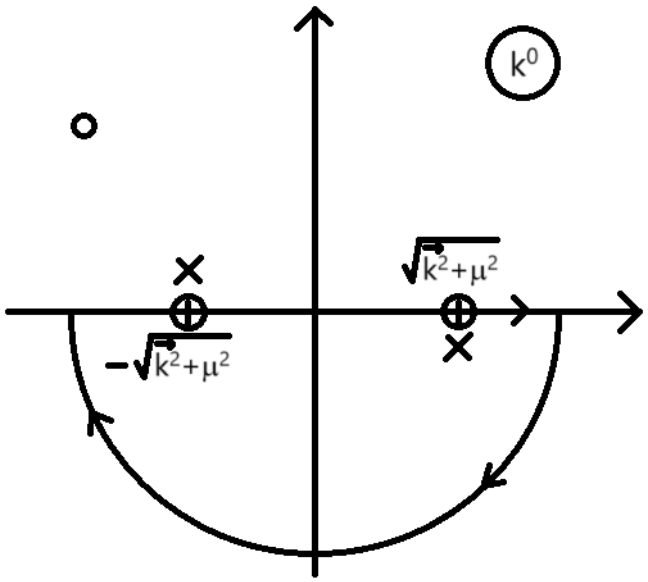

Fig. 8.1. The poles of the integral (8.18) at k^0 plane

Therefore,

$$k^0 = -i\epsilon \pm \sqrt{\vec{k}^{\,2} + \mu^2}\ . \tag{8.23}$$

Because of the $-i\epsilon$ poles are shifted in the lower half plane. This means there are no problems, if integral over the real line will be taken. It also should be noticed that the Cauchy's theorem can be applied to compute the integral by closing the correspondent contour in the lower half plane. Integral goes through infinity and close the contour in the lower half plane like presented at the (fig. 8.1).

Then the integral can be computed by Cauchy's residue theorem in the following way: the two poles are exist and residues at these poles should be taken. So, the integral will be transformed to

$$8.18 = \int_{-\infty}^{+\infty} \frac{\mathrm{d}k^0}{2\sqrt{\vec{k}^{\,2} + \mu^2}} \left[\frac{e^{-ick^0 t}}{k^0 - \sqrt{\vec{k}^{\,2} + \mu^2} + i\epsilon} - \frac{e^{-ick^0 t}}{k^0 + \sqrt{\vec{k}^{\,2} + \mu^2} + i\epsilon} \right] , \tag{8.24}$$

where the expression was transformed into 2 decomposed terms, which are residues at the first and the second pole. According to the Cauchy's residue theorem, the direction of the integration should be taken into account and the integration contour should be enclosed in such a way, that the region enclosed by this contour must remain on the left. The way, which was presented in the (fig. 8.1) stays the integration contour on the right and that is why the sign should be changed.

One more point to discuss is why the contour was enclosed below the real line. That is because for $t > 0$ in the exponential

$$e^{-ick^0 t} \tag{8.25}$$

k^0 should have a negative imaginary part. It should behave like $-i\left|k^0\right|$, in order to produce minus in front of the argument of the exponential. It should be so to provide the dumping in the spatial infinity, which can be achieved by increasing of the radius of a circle. For t bigger than zero the integral was closed below the real line.

Then, of course, two poles should be encountered. One pole is at

$$k^0 = \sqrt{\vec{k}^{\,2} + \mu^2} - i\epsilon \tag{8.26}$$

and the other pole is at

$$k^0 = -\sqrt{\vec{k}^{\,2} + \mu^2} - i\epsilon \ . \tag{8.27}$$

By Cauchy's theorem, the result will equal to

$$8.18 = \frac{2\pi i}{2\sqrt{\vec{k}^{\,2} + \mu^2}} \left[e^{ic\sqrt{\vec{k}^{\,2}+\mu^2}t} - e^{-ic\sqrt{\vec{k}^{\,2}+\mu^2}t} \right] \ . \tag{8.28}$$

Expression (8.28) is the answer for this auxiliary integral. In the original integral t can be anything. It can be bigger than zero and it can be less than zero. For $t > 0$ the contour should be closed in the lower half plane. For $t < 0$ the situation is actually the opposite in order to guarantee the exponential damping of the function at large values of k^0. Therefore, for $t < 0$ the contour should be closed in the upper half plane, but in the upper half plane there are no poles and the function has its poles only in the lower half plane. That is why the integral will be equal to zero.

It can be obtained that that the integral has interesting property: for positive t, it gives a non-trivial answer given by expression (8.28), for negative t it actually vanishes. On the other hand, it can be obtained that the integral, in fact, is the next quantity from (8.16):

$$\frac{1}{2k^0} \left[e^{i\omega t} - e^{-i\omega t} \right] \ , \tag{8.29}$$

where ω and k^0 should be replaced by ck^0 and $\sqrt{\vec{k}^{\,2} + \mu^2}$ respectively.

Instead of expression (8.28), which was gotten before, the original integral representation can be substituted for the retarded Green's function:

$$\Delta_{ret}\left(k\right) = \hbar c \int \frac{\mathrm{d}^4 k}{\left(2\pi\right)^4} \frac{i}{\left(k^0 + i\epsilon\right)^2 - \vec{k}^{\,2} - \mu^2} e^{-ikx} \ , \tag{8.30}$$

where integration was performed not only for spacial directions of k, but also for the k^0 variable. The scalar product kx in the (8.30) is a relativistic scalar product:

$$kx = k_\mu x^\mu \ . \tag{8.31}$$

Let's write one important property of the Δ_{ret}. From the courses on mathematical physics, it can be known that the Green's function is not unique, because of the fact that the following equation should be solved:

$$\left(\Box + \mu^2\right) G = \delta^{(4)}\left(x - x'\right) \ . \tag{8.32}$$

It can be clearly seen that if one particular solution of the equation was found, then a solution of the homogeneous equation can always be added to this solution and the gotten expression will be again a Green's function:

$$G = G_0 + G^{hom} \tag{8.33}$$

Solution of the homogeneous equation means that

$$\left(\Box + \mu^2\right) G^{hom} = 0 \ . \tag{8.34}$$

Playing with the solution of the homogeneous equation, the properties of the Green's function can be adjusted. For instance, we can have retarded solution or advanced solution.

Another type of solution is also exist, which is more interesting to observe and called Feynman propagator. But before that let's denote what is retarded means. Retarded means just a specific property that Green's function appears to be developed after the acting of the impulse. For instance, if a delta-function source is exist, which switches at $t = t'$, then the signal propagates at later times and that is what retarded Green's function is taken care of. In a more convenient way, a retarded Green's function has a retarded property, which means that at the later values of time, we expect to get consequences of the signal. This consequences is observed, because

$$\Delta_{ret}\left(t - t' > 0\right) \neq 0 \ . \tag{8.35}$$

On the other hand, before the source switched on

$$\Delta_{ret}\left(t - t' < 0\right) = 0 \ . \tag{8.36}$$

An advanced Green's function has an opposite property.

With the integral representation (8.30) it can be also elementary checked that Klein-Gordon operator acting on Δ doesn't produce a delta-function. And if the integral representation is used, then it will be seen that:

$$\left(\Box + \mu^2\right) \Delta_{ret} = i\hbar c \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{(k^0)^2 - \vec{k}^{\,2} - \mu^2}{(k^0 + i\epsilon)^2 - \vec{k}^{\,2} - \mu^2} e^{-ikx} \ . \tag{8.37}$$

As can be seen from the (8.37), the numerator and the denominator have the same expression and there is only different that a little $i\epsilon$ in the denominator exist. Then the limit $\epsilon \to 0$ can be taken and the resulting expression will be:

$$\lim_{\epsilon \to 0} \left(\Box + \mu^2\right) \Delta_{ret} = -i\hbar c \int \frac{\mathrm{d}^4 k}{(2\pi)^4} e^{-ikx} \ . \tag{8.38}$$

The integral in the (8.38) is a Fourier image of the four dimensional delta-function:

$$\lim_{\epsilon \to 0} \left(\Box + \mu^2\right) \Delta_{ret} = -i\hbar c \delta^{(4)}(x) \ . \tag{8.39}$$

Finally, according to (8.39), Klein-Gordon operator acting on delta retarded produce the delta-function source.

The most interesting function to consider, which is essentially used as an element of quantum field theory, is a so-called causal Green's function. This is the same as Feynman propagator.

Feynman propagator

Feynman propagator can be denoted simply by symbol Δ. It also depends on $x - x'$ and the motivation for its introduction can be given in the following way. Let's look at vacuum expectation value of two quantum fields:

$$\langle 0 | \phi(x) \, \phi(x') | 0 \rangle \ , \tag{8.40}$$

where $\phi(x) \, \phi(x')$ is a two point function. Two point means that it depends on x and on x'. It is also important to notice that ϕ is a quantum field.

Expression (8.40) can be written in terms of creation and annihilation operators. After that the product of operators should be put between two vacuum states. Often this is called as vacuum expectation value for the product of two operators. The result of this evaluation will have the following form:

$$\langle 0 | \phi(x) \, \phi(x') | 0 \rangle = \left\langle 0 \middle| \phi^-(x) \, \phi^+(x') \middle| 0 \right\rangle = \left\langle 0 \middle| \left[\phi^-(x) \, \phi^+(x')\right] \middle| 0 \right\rangle \ . \tag{8.41}$$

where the fact that field $\phi(x)$ can be split into positive and negative frequency parts was used. It should be noticed that the contribution of the $\phi^-(x')$ will produce zero, because this part depends on the annihilation operator and the annihilation operator annihilates vacuum. Analogously, if $\phi^+(x)$ is used, than due to the fact that $\phi^+(x)$ depends on the creation operator, zero will be produced after acting on the left vacuum.

The last expression in the (8.41) with commutator was written, because

$$\left\langle 0\middle|\phi^+(x')\,\phi^-(x)\middle|0\right\rangle = 0\ . \tag{8.42}$$

Then the commutator from the (8.41) can be evaluated and, therefore, the result will have the following form:

$$\langle 0|\phi(x)\,\phi(x')|0\rangle = i\hbar c\mathcal{D}^-(x-x')\ . \tag{8.43}$$

where $\mathcal{D}^-$ is the negative frequency part of the Pauli-Jordan function.

On the other hand, the expression on the left hand side of the (8.43) can be physically interpreted as an amplitude for the process, when at the moment $t > t'$ a particle at a point x' is created and a particle at a point x is destroyed (fig. 8.2).

Fig. 8.2. Creation of a particle at a point x' and destroying a particle at a point x

The process presented at (fig. 8.2) can be described in a bit different way. A similar amplitude may be considered, when a particle is created at x, but then it destroyed at x'.

To combine this processes in one go, one introduces the notion of the Feynman propagator $\Delta(x-x')$, which is constructed in the following way:

$$\Delta(x-x') = \theta(t-t')\,\langle 0|\phi(x)\,\phi(x')|0\rangle + \theta(t'-t)\,\langle 0|\phi(x')\,\phi(x)|0\rangle\ , \tag{8.44}$$

where θ function prescribes the way how creation and annihilation happens in time. For instance, $\theta(t-t')$ means that particle is created at a point x' and a time t' and then annihilated at a point x and a time t.

The Feynman propagator can be written in a different way. The two terms presented in the (8.44) can be encoded into one formula:

$$\Delta(x-x') = \langle 0|T(\phi(x)\,\phi(x'))|0\rangle\ , \tag{8.45}$$

where the operation of T ordering explicitly means that the T checks the time arguments of fields $\phi(x)$ and $\phi(x')$ and then it orders them according to the time dependence of fields. The field with a less value of T always go to the right and this rule can be remembered by slogan that „Youth is always right".

The Feynman propagator also can be expressed in terms of a Pauli-Jordan function and, in fact, half the work to obtain this representation has already been done. Additionally, a Pauli-Jordan function should be supplied with θ-function:

$$\Delta(x-x') = i\hbar c\left[\theta(t-t')\,\mathcal{D}^-(x-x') + \theta(t'-t)\,\mathcal{D}^-(x'-x)\right] \ . \tag{8.46}$$

The Pauli-Jordan function have a positive and negative frequency parts, which are connected with each other according to the formula:

$$\mathcal{D}^+(x) = -\mathcal{D}^-(-x) \ . \tag{8.47}$$

Let's consider the Feynman propagator in more detail. It is useful to repeat the same exercise as was done for the retarded Green's function. First of all, the integral expression for $\mathcal{D}^-$ should be written:

$$\mathcal{D}^- = -\frac{i}{(2\pi)^3}\int \frac{\mathrm{d}\vec{k}}{2k^0} e^{-ick_0t + i\vec{k}\,\vec{x}} \ . \tag{8.48}$$

Then, expression (8.48) need to be put in the Feynman propagator. The result will have the following form:

$$\Delta(x-x') = \hbar c\int \frac{\mathrm{d}\vec{k}}{(2\pi)^3} e^{i\vec{k}\,\vec{x}} \times \left[\frac{\theta(t)}{2k_0}e^{-ick_0t} + \frac{\theta(-t)}{2k_0}e^{ick_0t}\right] \ , \tag{8.49}$$

where k_0 as it was before equal to

$$k_0 = \sqrt{\vec{k}^{\,2} + \mu^2} \ . \tag{8.50}$$

The integral can be evaluated by the same trick as before. So, the integral taken along the real line in the complex k_0 plane should be found or, in other words, an auxiliary integral should be found:

$$\int \frac{\mathrm{d}k_0 e^{-ick_0t}}{k^2 - \mu^2 + i\epsilon} \ , \tag{8.51}$$

where the term $i\epsilon$ was used in a different way in comparison to what was done before, when $i\epsilon$ was put close to k_0:

$$(k_0 + i\epsilon) \ . \tag{8.52}$$

It also should be noticed that

$$k^2 = k_0^2 - \vec{k}^{\,2} \tag{8.53}$$

and, therefore (8.51) can be written as

$$\int \frac{\mathrm{d}k_0 e^{-ick_0t}}{k_0^2 - \vec{k}^{\,2} - \mu^2 + i\epsilon} \ . \tag{8.54}$$

Poles of the integral are now located as shown at the (fig. 8.3). Due to the (8.54), the left pole will be shifted up and the right pole a little bit down.

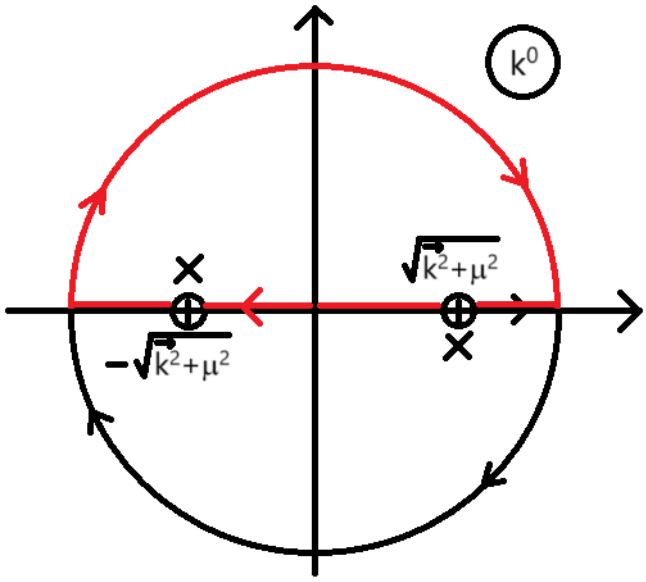

Fig. 8.3. The poles of the integral (8.54) at k^0 plane

What is the purpose of such a different shift? The purpose of the shift is to actually realize the properties of the θ-function. One function should be equal to 1, when t is positive and another should be equal to one when t is negative, where t is a difference between t and t'.

Similarly to the previous contour integral, a contour should be closed in the lower half plane. The difference is that now only one pole in this integration contour exists. And so the non-trivial contribution in this case will be found.

The denominator of the integral should be split into simple poles and the expression with simple poles will be equal to:

$$8.54 = \int \frac{\mathrm{d}k_0}{2\sqrt{\vec{k}^{\,2} + \mu^2}} \times \left[\frac{e^{-ick_0t}}{k_0 - \sqrt{\vec{k}^{\,2} + \mu^2} + i\epsilon} - \frac{e^{-ick_0t}}{k_0 + \sqrt{\vec{k}^{\,2} + \mu^2} - i\epsilon} \right] \ . \tag{8.55}$$

Condition for k_0, which defines the poles, should be written in the following way:

$$k_0 = \pm\sqrt{\vec{k}^{\,2} + \mu^2 - i\epsilon}\ , \tag{8.56}$$

where the term $i\epsilon$ is under the square root. But ϵ is small and for small ϵ the following replacement is right:

$$\pm\sqrt{\vec{k}^{\,2} + \mu^2 - i\epsilon} \to \pm\sqrt{\vec{k}^{\,2} + \mu^2} \mp i\epsilon\ . \tag{8.57}$$

According to the (8.55) the next statement can be formed: for t bigger than 0, the contour should be closed in the lower half plane and only one pole gives the contribution; for t less than 0 the contour should be closed in the upper half plane. For instance, for the retarded Green's function, the integral's result was 0, because there was no pole in the upper half plane.

Now the integral can be evaluated by using the Cauchy residue theorem:

$$8.55 = 2\pi i \frac{1}{2\sqrt{\vec{k}^{\,2} + \mu^2}} \left[-\theta\left(t\right) e^{-ic\sqrt{\vec{k}^{\,2}+\mu^2}t} - \theta\left(-t\right) e^{ic\sqrt{\vec{k}^{\,2}+\mu^2}t} \right]\ . \tag{8.58}$$

The result of evaluation of the integral is that now the integral representation for the Feynman propagator can be written:

$$\Delta\left(x\right) = \hbar c \int \frac{\mathrm{d}^4 k}{\left(2\pi\right)^4} \frac{i}{k^2 - \mu^2 + i\epsilon} e^{-ikx}\ . \tag{8.59}$$

The causal properties are now encoded exactly in the little shift $i\epsilon$ in the denominator. From expression (8.59) it also can be seen that the propagator in the Fourier space will be simply:

$$\Delta\left(k\right) = \frac{i}{k^2 - \mu^2 + i\epsilon}\ . \tag{8.60}$$

It also can be explained where the term $i\epsilon$ comes from. It comes from encoding the causality acts of creation and annihilation particles with the help of the T-ordering product. So, T-ordering seats in the $i\epsilon$ shift in the denominator of the Feynman propagator.

The integral can be actually computed explicitly and as the result Δ in the x-space will be gotten in terms of known special functions:

$$\Delta\left(x\right) = \frac{\hbar c}{4\pi^2} \left(\frac{mc}{\hbar}\right)^2 \frac{K_1\left(\frac{mc}{\hbar}\sqrt{-x^2 + i\epsilon}\right)}{\frac{mc}{\hbar}\sqrt{-x^2 + i\epsilon}}\ , \tag{8.61}$$

where K_1 is the Macdonald function, x^2 is a relativistic interval, which looks like:

$$x^2 = x_0^2 - \vec{x}^{\,2}\ . \tag{8.62}$$

The last thing that needed to be discussed concerning the scalar field is what is called Yukawa force.

Yukawa force

The Yukawa force topic is linked with the action of a scalar field. The expression for the action was already introduced and can be simply written as:

$$S\left[\phi\right]=\frac{1}{c}\int \mathrm{d}^4x\left[\frac{1}{2}\partial_\mu\partial^\mu\phi-\frac{\mu^2}{2}\phi^2+J\left(x\right)\phi\left(x\right)\right]\ , \tag{8.63}$$

where $J\left(x\right)$ is called as source term. This term can be considered as a source for a field ϕ and the simultaneously what it does, it changes the Hamiltonian of the field ϕ itself. In other words, the Hamiltonian and the energy receives an extra contribution from interaction of the field ϕ with the source. The source here is supposed to be in a way external.

It is interesting here to study the response of the field ϕ on the presence of the source $J\left(x\right)$. In fact, the term with J is a part of the theory, which is called linear response theory.

It is important that in the presence of the source the Hamiltonian H receives an extra contribution, which can be denoted as H_J. This is given simply by

$$H_J=-\int \mathrm{d}\vec{x}\,J\left(x\right)\phi\left(x\right)\ . \tag{8.64}$$

H_J comes with a minus sign, because if the Hamiltonian will be derived from the action S, the canonical procedure should be applied, where the Hamiltonian is given by:

$$H=\pi\phi-L\ . \tag{8.65}$$

When taking the minus Lagrangian, then the term comes with a minus sign and this will produce an extra contribution to the Hamiltonian of the field.

On the other hand, equations of motion for the field will be changed. It will not just the Klein-Gordon equation, but it will be inhomogeneous Klein-Gordon equation, where on the right hand side a source term will exist:

$$\left(\partial_\mu\partial^\mu+\mu^2\right)\phi\left(x\right)=J\left(x\right)\ . \tag{8.66}$$

When the action will be varied to find equations of motion, the field ϕ will be also varied and a contribution from the source term will be gotten.

Then all the knowledge about the Green's function can be used. So, Green's function becomes important because it is known that if an homogeneous equation exists, it can always been solved by using the method of Green's functions. The right solution is:

$$\phi\left(x\right)=\frac{i}{\hbar c}\int \mathrm{d}^4yG\left(x-y\right)J\left(y\right)\ . \tag{8.67}$$

If a Klein-Gordon operator is applied to the field $\phi(x)$, then the Klein-Gordon operator will hit the Green's function under the integral:

$$\left(\partial_\mu \partial^\mu + \mu^2\right) \phi(x) = \frac{i}{\hbar c} \int \mathrm{d}^4 y \left(\partial_\mu \partial^\mu + \mu^2\right) G(x-y) J(y) \ . \tag{8.68}$$

When Klein-Gordon operator hits Green's function, it will produce by definition of Green's function a term:

$$\left(\partial_\mu \partial^\mu + \mu^2\right) G(x-y) = -i\hbar c \delta^{(4)}(x-y) \ . \tag{8.69}$$

Then the integral should be taken and due to the delta-function presence, the result will be:

$$\left(\partial_\mu \partial^\mu + \mu^2\right) \phi(x) = J(x) \ . \tag{8.70}$$

Lecture 9. Yukawa Force. Dirac Equation

Yukawa force

At the previous lecture the action for a scalar field was introduced in the presence of the linear source and this action is given by the following expression:

$$S\left[\phi\right]=\frac{1}{c}\int \mathrm{d}^4x\left[\frac{1}{2}\partial_\mu\phi\partial^\mu\phi-\frac{1}{2}\mu^2\phi^2+J\phi\right]\ , \tag{9.1}$$

where J is a linear source.

If the dimension of the field ϕ is recalled, the dimension of the source will be easily deduced. Recalling that the dimension of the action is the same as the dimension of $\hbar$ (the Plank constant) and dimension of $\hbar$ is the same as a dimension of angular momentum, it can be found out that the linear source has the following physical dimension:

$$[J]=\frac{\sqrt{\hbar c}}{l^3}\ . \tag{9.2}$$

An interesting thing is that the quantity $\sqrt{\hbar c}$ has the same dimension as an electric charge. Therefore, the source J has a dimension of the density of electric charge.

As soon as the source is added up, the energy is changed and the Hamiltonian receives an extra contribution from the source term. This extra contribution to the Hamiltonian can be denoted as H_J and it can be given by

$$H_J=-\int \mathrm{d}\vec{x}\,J\left(x\right)\phi\left(x\right)\ , \tag{9.3}$$

where $\phi\left(x\right)$ is a real scalar field.

It is also can be seen that equations of motion for ϕ in the presence of the source gets change and new equations of motion will have the following form:

$$\left(\partial_\mu\partial^\mu+\mu^2\right)\phi\left(x\right)=J\left(x\right)\ . \tag{9.4}$$

The solution of equation (9.4) can be written as

$$\phi\left(x\right)=\frac{i}{\hbar c}\int \mathrm{d}^4yG\left(x-y\right)J\left(y\right)\ , \tag{9.5}$$

where $G\left(x-y\right)$ is the Green's function, which is a solution of the fundamental Klein-Gordon equation, where on the right hand side of this equation a delta-source or point like source:

$$\left(\partial_\mu\partial^\mu+\mu^2\right)G\left(x-x'\right)=-i\hbar c\delta^{(4)}\left(x-x'\right)\ . \tag{9.6}$$

It is well known that there are many solutions of equation (9.6), because to any solution of the inhomogeneous equation the homogeneous part can be added up. That is why, there are such Green's functions as retarded Green's function, advanced Green's function and Feynman propagator exist. But for the present discussion it does not matter which Green's function will be picked up. The (9.4) can be easily gotten if $\phi(x)$ in the (9.4) will be replaced by the (9.5) and if Klein-Gordon operator will be moved into the integral for acting on the Green's function:

$$\left(\partial_\mu \partial^\mu + \mu^2\right) \phi(x) = \frac{i}{\hbar c} \int \mathrm{d}^4 y \left(\partial_\mu \partial^\mu + \mu^2\right) G(x-y) J(y) \ . \tag{9.7}$$

According to the (9.6), expression (9.7) can be modified:

$$9.7 = \int \mathrm{d}^4 y \delta^{(4)}(x-y) J(y) = J(y) \ . \tag{9.8}$$

The goal now is to calculate the interaction energy. This is quantity, which is mediated by a scalar field between two equal charge static sources, represented by point-like particles, sitting at positions x_1 and x_2 (fig. 9.1). So, these particles are static and this means that their positions are fixed and they are not changing with the time. But due to the interaction mediated by field ϕ around, these particles will interact with each other and there will be a force, which this particles will exert on each other, in particular, because a particle at x_2 will be interpreted as a source of the field. Another particle will find itself under the influence of the scalar field and will be considered as a test particle.

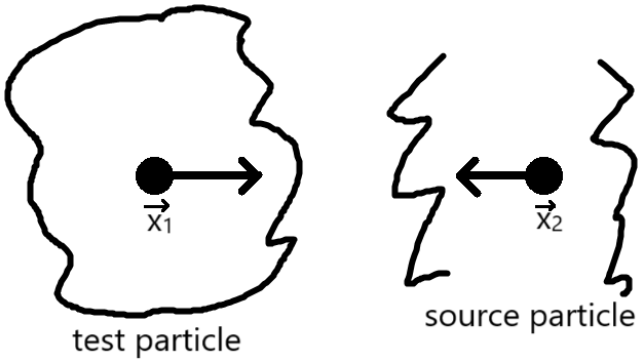

Fig. 9.1. Interaction of the point-like particles, sitting at positions x_1 and x_2

In terms of mathematical formulas this means that:

$$J_2(x) = \sqrt{\hbar c}\delta(\vec{x} - \vec{x}_2) \ , \tag{9.9}$$

where J_2 is a source of a scalar field $\phi(x)$.

It can be seen that, when x is equal to x_2, then the delta function will be different from zero and the magnitude of the source is proportional to the quantity with a dimension of electric charge.

The source function J_2 can be used in the right hand side of equation (9.4) and then the response to the presence of the source will be a scalar field $\phi(x)$ arising in the surrounding space due to the fact that the source was put at a position x_2.

The first particle, on the other hand, will be represented in a similar manner. A function $J_1(x)$ will be associated for this source with the same properties:

$$J_1(x) = \sqrt{\hbar c}\delta(\vec{x} - \vec{x}_1) \ , \tag{9.10}$$

where J_1 is a function of the test particle.

The energy H_J can be interpreted as an additional contribution V:

$$V := H_J \ . \tag{9.11}$$

The energy of field ϕ due to the interaction of the test particle with a field ϕ will be given by:

$$V = -\int \mathrm{d}\vec{x}\, J_1(x)\,\phi(x) \ . \tag{9.12}$$

Formula (9.12) is similar to formula (9.3), except the fact that now a particle concentrated at the position x_1.

On the other hand, since $\phi(x)$ is a scalar field produced by the source J_2, an explicit solution for ϕ can be written in terms of the Green's function:

$$V = -\frac{i}{\hbar c}\int \mathrm{d}\vec{x}\,\mathrm{d}^4y J_1(x)\,G(x-y)\,J_2(y) \ , \tag{9.13}$$

where in the (9.13) $\phi(x)$ was replaced by the integral with the Green's function according to the (9.5). $J_2(y)$ is the source, which generates a field and this a field interacts with a test particle, which can be also considered as another source. Therefore, the source of the field is interchangeable. It is also can be seen actually that formula (9.13) is completely symmetric with respect to x and y. There is only little difference for the moment that $\mathrm{d}\vec{x}$ is a three dimensional integration and $\mathrm{d}y$ is the four dimensional integration, but it can be reduced also to three dimensional integration.

Integrals in the (9.13) can be computed by using the knowledge of the Green's function. First thing to do is substitute the expression for the source. For simplicity a source with a unit value of charge will be introduced, but if an arbitrary value is needed, then the

expression may be multiplied with a number. So, let's multiply each of these sources with a number q, which is just number, and then

$$q\sqrt{\hbar c} \tag{9.14}$$

will have the dimension of electric charge.

And the q value will be chosen in the same way and, therefore, particles will have the same charge. Upon substituting the point like sources J_1 and J_2, the following expression will be gotten:

$$V = -iq^2 \int \mathrm{d}\vec{x}\, \mathrm{d}^4 y G\left(x-y\right) \delta\left(\vec{x}-\vec{x}_1\right) \delta\left(\vec{y}-\vec{x}_2\right) \ . \tag{9.15}$$

Then, the transformation to the Fourier's representation of the Green's function can be done. As a Green's function the Feynman propagator will be used, but it does not really matter which one to take. The result of the transformation will have the following form:

$$V = \frac{q^2\hbar c}{\left(2\pi\right)^4} \int \mathrm{d}\vec{x}\, \mathrm{d}^4 y \mathrm{d}\vec{k}\, \mathrm{d}k^0 \frac{e^{-ik^0\left(x^0-y^0\right)} e^{i\vec{k}\left(\vec{x}-\vec{y}\right)}}{\left(k^0\right)^2 - \vec{k}^{\,2} - \mu^2 + i\epsilon} \delta\left(\vec{x}-\vec{x}_1\right) \delta\left(\vec{y}-\vec{x}_2\right) \ . \tag{9.16}$$

First of all, expression (9.16) can be integrated over $\vec{x}$ and $\vec{y}$. This is trivial, because $\vec{x}$ and $\vec{y}$ in arguments of delta-functions will be replaced: $\vec{x}$ by $\vec{x}_1$ and $\vec{y}$ by $\vec{x}_2$. Another thing, which can be done immediately is the integration over y^0, because

$$\mathrm{d}^4 y = \mathrm{d}\vec{y}\, \mathrm{d}y^0 \tag{9.17}$$

and then the integration over y^0 can be performed, because y^0 comes only with the exponential

$$e^{-ik^0\left(x^0-t^0\right)} = e^{-ik^0 x^0} e^{ik^0 y^0} \ . \tag{9.18}$$

When the integration over y^0 will be completed, the proportionality coefficient will be a delta-function on the variable k^0, because an integral with exponential is just proportional to the delta-function:

$$\frac{1}{2\pi} \int \mathrm{d}y^0 e^{ik^0 y^0} = \delta\left(k^0\right) \ . \tag{9.19}$$

According to the facts, which was discussed above, the simplified expression will have the following form:

$$V = -\frac{q^2\hbar c}{\left(2\pi\right)^3} \int \mathrm{d}\vec{k}\, \frac{e^{i\vec{k}\left(\vec{x}_1-\vec{x}_2\right)}}{\vec{k}^{\,2} + \mu^2 - i\epsilon} \ , \tag{9.20}$$

where in the denominator since the integration over k^0 was completed, k^0 can be replaced by 0.

The term $i\epsilon$ in the denominator of the (9.20) can be removed by taking a limit $\epsilon \to 0$, because the integral over $\vec{k}$ has no problems with the denominator, because $\vec{k}^{\,2} + \mu^2$ is non negative. That's actually shows that any Green's function can be chosen, because all Green's functions differ by the ϵ prescription.

The simplified integral for the V value will have the next form:

$$V = -\frac{q^2\hbar c}{(2\pi)^3}\int \mathrm{d}\vec{k}\,\frac{e^{i\vec{k}\,(\vec{x}_1-\vec{x}_2)}}{\vec{k}^{\,2}+\mu^2}\ . \tag{9.21}$$

The integral (9.21) can be computed explicitly. This can be done by passing to spherical coordinates. Let's introduce the modulus of vector $\vec{k}$, which can be denoted by k without vector symbol on it

$$k = \left|\vec{k}\,\right| \tag{9.22}$$

and the distance between points x_1 and x_2 should be also introduced, which is:

$$r = |\vec{x}_1 - \vec{x}_2|\ . \tag{9.23}$$

Then the spherical coordinate system can be chosen by directing the axis z along $\vec{x}_1 - \vec{x}_2$ and then the integral will be written in the following way:

$$V = -\frac{q^2\hbar c}{(2\pi)^3}\int_0^\infty k^2\mathrm{d}k\int_0^\pi \sin\theta\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\varphi\frac{e^{ikr\cos\theta}}{k^2+\mu^2}\ . \tag{9.24}$$

The spherical coordinate system can be described in the following way: the direction of the z axis is the same as the direction of the vector $\vec{x}_1 - \vec{x}_2$ and x and y axes are also existed. Then the vector $\vec{k}$ is also defined, which is a running vector in the integral, and then angles θ and φ should be introduced as it shown on the (fig. 9.2).

The standard measure of the $\mathrm{d}\vec{k}$ is:

$$\mathrm{d}\vec{k}\ = k^2\sin\theta\mathrm{d}k\ . \tag{9.25}$$

Nothing depends on angle φ in the integral (9.24) and it can be immediately integrated. The result will give just 2π.

Then the integral over angle θ can also be taken, because:

$$\sin\theta\mathrm{d}\theta = -\mathrm{d}\left(\cos\theta\right)\ , \tag{9.26}$$

which is very convenient, because $\cos\theta$ is standing in the exponential $e^{ikr\cos\theta}$. Then, the integral will have the following form:

$$V = -\frac{q^2\hbar c}{(2\pi)^3}\int_0^\infty k^2\mathrm{d}k\frac{4\pi}{kr}\cdot\frac{\sin\left(kr\right)}{k^2+\mu^2}\ , \tag{9.27}$$

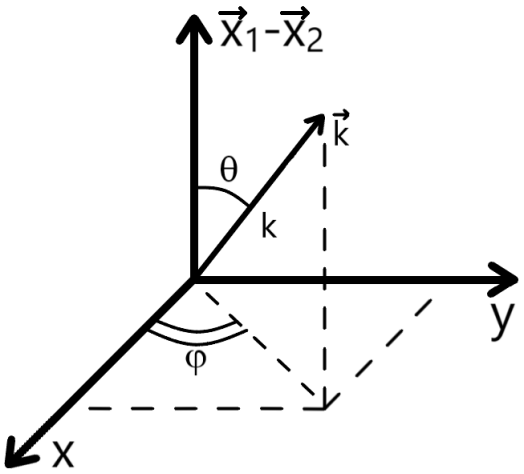

Fig. 9.2. The spherical coordinate system

where the expression

$$\frac{\sin(kr)}{kr} \tag{9.28}$$

come from integrating of the exponential in the (9.24).

Expression (9.27) can be simplified to the next form:

$$V = -\frac{q^2\hbar c}{2\pi^2 r}\int_0^\infty \mathrm{d}k \frac{k\sin(kr)}{k^2+\mu^2}\ . \tag{9.29}$$

The integral from the (9.29) will be denoted as $I(r)$, where:

$$I(r) = \int_0^\infty \mathrm{d}k \frac{k\sin(kr)}{k^2+\mu^2}\ . \tag{9.30}$$

This integral can be computed by using the Cauchy residue theorem in the theory of complex variables.

The $I(r)$ can be also rewritten as:

$$I(r) = -\frac{\partial}{\partial r}\int_0^\infty \mathrm{d}k \frac{\cos(kr)}{k^2+\mu^2}\ . \tag{9.31}$$

Then the cosine can be represented as a difference of two exponentials and, in fact, the denominator should be also expanded into simple fractions. One more step to complete is to make integration also from $-\infty$ to $+\infty$ using the fact that the function under the integral is even. Therefore,

$$I(r) = -\frac{1}{4}\frac{\partial}{\partial r}\int_{-\infty}^{+\infty}\left(\frac{e^{ikr}}{(k+ia)(k-ia)} + \frac{e^{-ikr}}{(k+ia)(k-ia)}\right)\ , \tag{9.32}$$

where, the variable a is equal to:

$$a = \mu = \frac{mc}{\hbar} = \frac{1}{\lambda} \ . \tag{9.33}$$

In fact, in physical dimensions the quantity $\frac{mc}{\hbar}$ is known as an inverse lens $\frac{1}{\lambda}$ and this is something which is called as Compton wavelengths of a particle with mass m.

As can be seen the integral (9.32) is always real line, but the contour can be closed in the complex plane and, in particular, for the first integral exponential e^{ikr} it should be closed in the upper half plane. For the second integral the situation is opposite and it should be closed in the lower half plane. Then it can be seen, which poles ends up in the lower half plane and the residue theorem should be used to compute the integral. The corresponding pole in the up half plane will be equal to $k = -ia$. In the lower half plane it will be equal to $k = -ia$.

So, the integral reduces to:

$$I\left(r\right) = -\frac{1}{4}\frac{\partial}{\partial r}\left(2\pi i\frac{e^{-ar}}{2ia} - 2\pi i\frac{e^{-ar}}{-2ia}\right) \ . \tag{9.34}$$

Then it can be simplified to the following expression:

$$I\left(r\right) = -\frac{\pi}{2a}\frac{\partial}{\partial r}e^{-ar} \ , \tag{9.35}$$

where after differentiating the final result will be equal to:

$$I\left(r\right) = \frac{\pi}{2}e^{-ar} \ . \tag{9.36}$$

In terms of Compton wavelengths expression (9.36) can be written as:

$$\frac{\pi}{2}e^{-r/\lambda} \ . \tag{9.37}$$

Now expression (9.29) can be simplified and the result will be equal to:

$$\boxed{V = -\frac{1}{4\pi}\frac{q^2\hbar c}{r}e^{-r/\lambda}} \ . \tag{9.38}$$

Thanks to the (9.38) the potential between two equally charged static sources was gotten and this potential is called as Yukawa potential.

In particular, from expression (9.38), several facts can be highlighted. First of all, a very interesting thing is that this potential appears to be negative and changing $q \to -q$ will produce the same expression, because of the existence of q^2. And this potential with

a minus sign actually means that the force is an attractive, because having a smaller distance between the particles will decrease the potential and it will take more and more negative values.

It should be also noticed that it's preferable to approach each other to diminish the energy. Diminishing the energy corresponds, therefore, in this case, to the desire of particles to be closer to each other.

The second thing which can be seen is that there is the exponential, which dumps the potential, when r becomes sufficiently large, in particular, when r becomes bigger than the wavelengths of the particles potential shows exponential decay. So, potential goes to 0 very fast, actually exponentially with a distance increasing.

It is also should be noticed that the interaction range set up by the exponential does depend on the Compton wavelengths. It's governed by the effective length, which is turns out to be Compton wavelengths.

This additionally shows, why people call the Greens function as a propagator. The propagator propagates interaction from the one source to the other source or from a source to a test particle. And it propagates interaction, represented by the field ϕ.

Generally, a very physical calculation was performed, which showed explicitly what happens if two particles are existed inside the space field with a scalar field ϕ. And because of interacting with this field they start to attract.

One more important fact is that when $\lambda \to \infty$, which essentially means that $m \to 0$ according to

$$\frac{1}{\lambda} = \frac{mc}{\hbar} \ , \tag{9.39}$$

the field will be massless scalar field, instead of massive scalar field and then the potential between two static sources actually becomes a Coulomb potential:

$$V(r) \to -\frac{1}{4\pi}\frac{q^2\hbar c}{r} \ . \tag{9.40}$$

When mass is non trivial, the Coulomb potential gets screened and becomes short range because of presence of the exponential.

In the mid of 1930s Yukawa tried to understand the origin of the strong force, which holds together protons and neutrons in the nucleus. So, the nucleus consists of protons and neutrons. It was also known that the force acting between these protons and neutrons inside nuclear is a very short range.

Yukawa made a conjecture that the strong force was mediated by particle similar to photon, which like photon in electrodynamics mediates interactions between electrons and

positrons. He assumed that something similar creates place inside nucleus and there is a hypothetical particle, which mediates interactions between this protons and neutrons.

Assuming that this mediating particle is massive and following absolutely the same analogy with electromagnetism, one sees that one needs to add mass to the wave equation and this produces as a result, the short range mediating force between two static sources.

Yukawa actually predicted the existence of a novel type of particle and from the known data on the range of nuclear forces he was actually able to predict with quite a good accuracy, the mass of this particle. So, if the range of the strong force is known, which can be actually done experimentally, then λ can be deduced or equivalently the mass of the particle can be found. And indeed later, this particle called π meson or sometimes people call it simply pion, pion was discovered in cosmic rays. It also appeared later in this cyclotron experiments, where it was predicted by Yukawa.

Although, today, there are several types of mesons are known and it is also known that mesons are not elementary themselves, but rather they are composite particles made of the quark and it's anti quark. Nevertheless, Yukawa theory made a very right qualitative picture and it also, in a way, paved the development of what is called QCD or quantum chroma dynamics. Nevertheless, that was very important step towards development of QCD by assuming the nature of strong force as a rising from a massive particle with zero spin.

As a general computation the field of any spin can be applied to a mediator and similar computation takes place also for electromagnetism, where instead of minus sign plus sign can be seen in the potential, and, therefore, there will be repulsive force instead of attraction. Then for particles, it is more preferable to take a large distance between them, in order to decrease the energy, because then energy is positive and it will be decreasing when the distance will tend to infinity. So, sign is very important and sign eventually comes from the equations of motion that induced by the type of action. For instance, if a particle of a given speed was chosen to describe, as a result it leads to the physical consequences of attraction or repulsion.

Then, a new chapter in the quantum field theory, namely the Dirac equation should be introduced.

Dirac equation

It's a new and very interesting story called the Dirac equation. In 1928, the Dirac discovered his relativistic equation trying to overcome difficulties with negative probability densities of the Klein-Gordon equation.

As it was already shown Schrodinger made this attempt before to write down relativistic quantum mechanics by taking dispersion relation between energy and momentum of relativistic particle and transformed this condition in the language of operators by replacing momentum and energy with the corresponding operators, shifts in a space and shifts in time. And also he wrote down the Klein-Gordon equation, which was later rediscovered many times by Klein, Gordon, Fock and many others.

But it was the problem to treat this equation as a relativistic equation for a wave function ϕ, because it has difficulties with probabilistic interpretation, because equation have the second order. That's allows to fix initial conditions and the initial velocities and leads to the appearance of negative probability densities.

The reason, which led Dirac create his equation was essentially the following. First of all, the target is to prevent the occurrence of negative probability densities. This means that time derivatives in the formula for the probability density should be avoided. So, the equation must, therefore, not contain time derivatives higher than the first order. For instance, the Schrodinger equation in the usual quantum mechanics has the first order of time:

$$i\hbar\frac{\partial}{\partial t}\psi = H\psi \ . \tag{9.41}$$

To achieve a good probabilistic interpretation, a wave function must satisfy the equation, which contains only first derivative.

On the other hand, if relativistic covariance is wanted, then it requires that the spacial and time components, in particular, spacial derivatives and the time derivative must be treated on equal footing. For example, in the standard Schrodinger equation of quantum mechanics, this is not the case, because, typically, the Hamiltonian contains second spacial derivative. That's immediately shows that the equation is non-relativistic, because relativism requires that t can be converted to x and x can be converted to t by means of linear Lawrence transformations. It is impossible to have this, if derivatives in the equation have a different order.

Now also, one needs linearity to have a superposition principle like in quantum mechanics. So, equation must be linear.

Like in the usual quantum mechanics a superposition principle exists, because linearity means that if there are two solutions of the equation, then they sum will be also a solution.

The last important point is, that since the relativistic particle is described, it must still satisfy the standard on-shell condition for relativistic particles. So,

$$E^2 = p^2c^2 + m^2c^4 \ . \tag{9.42}$$

For ψ this would mean that ψ must satisfy the second order Klein-Gordon equation:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial^2 t} - \frac{\partial^2}{\partial x_i^2} + \left(\frac{mc}{\hbar}\right)^2\right)\psi = 0 \ , \tag{9.43}$$

because this equation will produce the relativistic dispersion relation.

Dirac said the following: let's try to write down something, which would be linear for ψ and contains derivative of the first order:

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{\hbar c}{i}\left(\alpha_1\frac{\partial\psi}{\partial x^1} + \alpha_2\frac{\partial\psi}{\partial x^2} + \alpha_3\frac{\partial\psi}{\partial x^3}\right) + \beta mc^2\psi \ , \tag{9.44}$$

where the left hand side looks like non-stationary Schrodinger equation, but on the right hand side, there are terms, which are the linear combination of first order derivatives function ψ.

Coefficients α_i cannot be just usual numbers, because if they're usual numbers, which have a certain values, then the equation is not even invariant with respect to usual three dimensional rotations, because three dimensional rotations will act on coordinates x_1, x_2 and x_3 and they will be just then produce changes of coefficients α_1, α_2, α_3 and will immediately destroy the equation.

So, Dirac said that it's not possible for these α_1, α_2, α_3 to be numbers and supposed them to be matrices. If they are matrices, then ψ cannot be just scalar and it should be some multi component object. ψ should be something, which would contain many components: ψ_1 and so on:

$$\psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} \ . \tag{9.45}$$

ψ also cannot be a scalar also for another reason. From relativistic covariance it is known that probability density, which would be

$$\rho = \psi^*\psi \ , \tag{9.46}$$

it must be relativistic invariant or to be a zero component of a vector, because for ρ the continuity equation exists, which is written as

$$\frac{\partial \rho}{\partial t} + \div \vec{j} = 0 \ . \tag{9.47}$$

Expression (9.47) describing a vector probability, where a vector, probability density and probability current are defined. And ρ with $\vec{j}$ forms a four vector $\left(\rho, \vec{j}\right)$. That is why, ρ must be a component of a 4-vector. It means that ψ cannot be just scalar in this case.

Indeed, the suggestion is to treat α_i as matrices and ψ as a column with a number of components $\psi_1, \ldots, \psi_n$. Dirac called the object ψ, presented in the (9.45) as spinor.

By the way, if α are some matrices, then β is also a matrix and ψ is a column.

Then, equation (9.44) can be written in the following form:

$$\left(i\hbar\frac{\partial}{\partial t} + i\hbar c\alpha^i \frac{\partial}{\partial x^i} - \beta mc^2\right)\psi = 0 \ . \tag{9.48}$$

Let's get the relativistic dispersion relation for ψ. This can be seen as follows. If such an equation for a n-component object exists, it can be hit by the following operator:

$$i\hbar\frac{\partial}{\partial t} - i\hbar c\alpha^i \frac{\partial}{\partial x^i} + \beta mc^2 \ , \tag{9.49}$$

which is a matrix differential operator. Firstly, it's differential operator, because it contains derivatives. Secondly, it's a matrix.

When the (9.48) will be hit by the operator (9.49) from the left, the expression will have the following form:

$$-\hbar^2\frac{\partial^2}{\partial t^2} + \hbar^2 c^2 \frac{\alpha^i\alpha^j + \alpha^j\alpha^i}{2}\frac{\partial^2\psi}{\partial x^i \partial x^j} + i\hbar mc^3\left(\alpha^i\beta + \beta\alpha^i\right)\frac{\partial\psi}{\partial x^i} - \left(mc^2\right)^2\beta^2\psi = 0 \ , \tag{9.50}$$

where initially there are nine terms to proceed, but four terms will cancel out. The second derivative of ψ with respect to x^i and x^j . This two derivatives acting on ψ commute and the derivative can be written in different ways, because it's symmetric.

Obtained expression (9.50) is a consequence of equation (9.48). It was assumed that ψ satisfy the linear equation and as a consequence it might satisfy the second order equation (9.50), because it contains second order derivatives. It also contains a first derivative and the term without derivative.

If the equation is multiplied by $-\frac{1}{\hbar^2 c^2}$, then the equation will take the following form:

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \frac{\alpha^i\alpha^j + \alpha^j\alpha^i}{2}\frac{\partial^2}{\partial x^i \partial x^j} - i\frac{mc}{\hbar}\left(\alpha^i\beta + \beta\alpha^i\right)\frac{\partial\psi}{\partial x^i} + \left(\frac{mc}{h}\right)^2\beta^2\psi = 0 \ . \tag{9.51}$$

Now it's pretty clear what conditions on matrices α^i and β should be imposed to turn equation (9.51) into the Klein-Gordon equation.

It is wanted to make exactly the Klein-Gordon equation for each component of ψ and each component of ψ must satisfy the Klein-Gordon equation, because ψ is supposed to describe relativistic particle with a standard dispersion relation. It can be seen that:

$$\alpha^i\alpha^j + \alpha^j\alpha^i = 2\delta^{ij} \ , \tag{9.52}$$

where i and j run from 1 to 3. Then the second term of (9.51) will be turned simply into second derivative $\frac{\partial^2\psi}{\partial x^{i2}}$.

Then the following should be required:

$$\alpha^i\beta + \beta\alpha^i = 0 \ . \tag{9.53}$$

Such a way the term with the first derivative will be removed.

Finally, to get the standard Klein-Gordon equation for a particle of mass m the next condition should be completed:

$$\beta^2 = 1 \ . \tag{9.54}$$

So, the equation, which will be gotten for ψ is the following:

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \frac{\partial^2\psi}{\partial x^{i2}} + \left(\frac{mc}{\hbar}\right)^2\psi = 0 \ . \tag{9.55}$$

This equation is valid for each component of ψ, because there are no matrices. They all now proportional to identity matrix. It also can be written in the next form:

$$\left(\partial_\mu\partial^\mu + \mu^2\right)\psi = \begin{pmatrix} \left(\partial_\mu\partial^\mu + \mu^2\right)\psi_1 \\ \vdots \\ \left(\partial_\mu\partial^\mu + \mu^2\right)\psi_n \end{pmatrix} = 0 \ . \tag{9.56}$$

From relations (9.52) - (9.54) it is immediately can be seen that if indexes i and j will be equal to each other, than expression (9.52) will have the following form:

$$\alpha^i\alpha^i + \alpha^i\alpha^i = 2 \tag{9.57}$$

and then

$$\left(\alpha^i\right)^2 = 1 \ . \tag{9.58}$$

Including the condition (9.54), it was gotten that all squares of matrices equal to 1, where 1 here is the identity matrix, because α and β are matrices.

Conditions (9.54) and (9.58) means all values of this matrices can be only equal to $+1$ or -1.

Second important observation is that all these matrices have zero trace and it can be seen from the condition (9.53). β is invertible, because of the (9.54) and, therefore, it can be written that:

$$\alpha^i = -\beta\alpha^i\beta \ . \tag{9.59}$$

Now a trace of α^i can be taken and due to cyclic property of a trace it will be equal to:

$$\mathrm{Tr}\alpha^i = -\mathrm{Tr}\left(\beta\alpha^i\beta\right) = -\mathrm{Tr}\left(\alpha^i\beta^2\right) = -\mathrm{Tr}\alpha^i = 0 \ . \tag{9.60}$$

Analogously trace for matrix β will be equal to:

$$\mathrm{Tr}\beta = \mathrm{Tr}\left(\left(\alpha^i\right)^2\beta\right) = \mathrm{Tr}\left(\alpha^i\beta\alpha^i\right) = -\mathrm{Tr}\beta \ , \tag{9.61}$$

where the following property was used:

$$\alpha^i\beta + \beta\alpha^i = 0 \ \rightarrow \ \alpha^i\beta\alpha^i = -\beta \ . \tag{9.62}$$

Therefore,

$$\begin{cases} \mathrm{Tr}\alpha^i = 0 \\ \mathrm{Tr}\beta = 0 \end{cases} \tag{9.63}$$

It means that the number of positive eigenvalues must be the same as the number of negative eigenvalues. And as it was gotten eigenvalues can be equal only to $+1$ or -1.

One more important conclusion is that the dimension of matrices α^i and β should be even or, in other words, n must be even.

Lecture 10. The Dirac Equation and Lorentz Transformations

In the last lecture the Dirac's equation was introduced. The physical principles, which are standing behind the derivation of this equation was basically discussed and an object which is called spinor was introduced.

The main consequence of Dirac equation as it was computed is that this equation also satisfies the Klein-Gordon equation. In other words, every component of our spinor satisfies the Klein-Gordon equation, which means that every component describes a relativistic particle, that means a dispersion relation for this particle is relativistic.

The equation also involves a set of matrices which satisfy the following properties: there are three matrices α^i, which satisfy the following algebraic relation:

$$\alpha^i\alpha^j + \alpha^j\alpha^i = 2\delta^{ij} \ . \tag{10.1}$$

And there is one matrix β, for which the following conditions are right:

$$\begin{cases} \alpha^i\beta + \beta\alpha^i = 0 \\ \beta^2 = 1 \end{cases} \tag{10.2}$$

It also was shown that traces of all these matrices are equal to 0. So,

$$\begin{cases} \mathrm{Tr}\alpha^i = 0 \\ \mathrm{Tr}\beta = 0 \end{cases} \tag{10.3}$$

According to (10.1) and (10.2) eigenvalues of matrices α^i and β are equal to +1 and -1. Due to the condition (10.3) matrices can be realized only in even dimension.

Then the concrete realization of these matrices should be gotten. The first idea is that since matrices must exist only in even dimensions, let's try the minimal even dimension, where such matrices have a chance to exist. The dimension, which satisfies this condition is a dimension with $n = 2$. Therefore, matrices α^i and β will be 2x2 matrices.

In two dimensions there are three distinguished matrices, which are Pauli matrices

$$\begin{cases} \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{cases} \tag{10.4}$$

They actually obey the desired relation, namely the anti-commutator of these matrices

$$\{\sigma^i, \sigma^j\} = 2\delta^{ij} \ , \tag{10.5}$$

which is similar to the anti-commutator of matrices α^i and $alpha^j$. However in two dimensions the fourth independent matrix would be just the identity matrix, but, unfortunately, this cannot be identified with β, because identity matrix would commute with Pauli matrices, while we should have a non-trivial relation between σ-s and matrix β, which is not just a simple commutation relation.

This means that in dimension with $n = 2$ such realization of Dirac matrices cannot be found. The next trial would be to ask for such matrices in dimension with $n = 4$. Indeed, it appears that the minimal dimension, where one can construct the four matrices α^i and β with all properties above is 4. Concretely, the following construction for α^i can be taken:

$$\alpha^i = \begin{pmatrix} \mathbb{0} & \sigma^i \\ -\sigma^i & \mathbb{0} \end{pmatrix} \ . \tag{10.6}$$

At the same time, for β the next matrix can be taken

$$\beta = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & -\mathbb{1} \end{pmatrix} \ , \tag{10.7}$$

It can be checked that matrices, indeed, satisfy all the relations above and they are traceless. So, that's concrete and explicit realization of the matrices, which should feature in the Dirac equation.

In the last lecture it was found that the Dirac Hamiltonian and the Dirac equation in the Hamiltonian form can be written in the following way:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \ , \tag{10.8}$$

where H is the Dirac Hamiltonian, which can be called as first quantized Hamiltonian meaning, that here the pedestrian approach to the theory was used. ψ can be treated as a multi component wave function for the Dirac equation, which has a form of the standard Schrodinger equation.

The Dirac Hamiltonian is a matrix, which can be written in the following form:

$$H = c\vec{a}\,\vec{p} + \beta mc^2 \ . \tag{10.9}$$

In a more explicit form, expression (10.9) can be presented as:

$$H = ca^i \left(-i\hbar \frac{\partial}{\partial x^i}\right) + \beta mc^2 \ , \tag{10.10}$$

where $\vec{p}$ was replaced according to the definition of the momentum operator. H is a differential matrix acting on four component wave function ψ or simply on column

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \ . \tag{10.11}$$

Now, the algebra of the matrices, which enter the Dirac's equation, should be investigated and this is needed in order to discuss covariant properties of the Dirac's equation, in particular, to understand how this equation transforms under Lawrence group, because the Lawrence group is very important and it's a symmetry group of relativistic quantum field theory.

In order to to do this the structure of the matrices should be better understood. And to complete it the Dirac's equation should be multiplied by the ratio $\frac{\beta}{c}$. Also, the new notation can be introduced:

$$\gamma^0 = \beta \, , \ \gamma^i = \beta\alpha^i \ , \tag{10.12}$$

where i is running from 1 to 3.

Let's also identify x^0 with:

$$x^0 = ct \tag{10.13}$$

and, therefore,

$$\frac{\partial}{\partial x^0} = \frac{\partial}{c\partial t} \ . \tag{10.14}$$

Then the Dirac's equation can be written in a more symmetric form:

$$\left[i\hbar\left(\gamma^0 \frac{\partial}{\partial x^0} + \gamma^i \frac{\partial}{\partial x^i}\right) - mc\right]\psi = 0 \ , \tag{10.15}$$

where in the last term the variable β is disappeared due to the property (10.2).

Now, the relativistic definition of a relativistic scalar product can be used

$$\left(i\gamma^\mu\partial_\mu - \frac{mc}{\hbar}\right)\psi = 0\ . \tag{10.16}$$

With a natural choice of units

$$\hbar = c = 1 \tag{10.17}$$

equation (10.17) takes the form:

$$\boxed{(i\gamma^\mu\partial_\mu - m)\,\psi = 0}\ . \tag{10.18}$$

And the (10.18) is the final form of the Dirac's equation.

Is also can be seen that with the new matrices γ^μ the algebraic relations, which was introduced for α and β can be also uniformized in one relation for matrices γ^μ, namely

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\cdot\mathbb{1}\ . \tag{10.19}$$

The free algebra generated by symbols γ^μ modulo the relation (10.19) or subject to this relation is called Clifford algebra.

The name goes back to the Clifford and this algebra has been known to mathematicians before the discovery of the Dirac's equation, but miraculously turns out that this mathematical background, on which the Dirac's equation is based on actually already existed in mathematics in the form of the Clifford algebra.

From the definition of matrices γ^μ it is clear that matrix γ^0 is Hermitian. So, it satisfies the following relation:

$$\left(\gamma^0\right)^+ = \gamma^0\ , \tag{10.20}$$

while the matrices γ^i are anti-Hermitian:

$$\left(\gamma^i\right)^+ = -\gamma^i\ . \tag{10.21}$$

(10.21) are right, because γ^i are given by the product of β with α^i. β and α^i are Hermitian, but under Hermitian conjugation the order of matrices should be changed. Therefore,

$$\left(\gamma^i\right)^+ = \left(\alpha^i\right)^+\beta^+ = -\beta\alpha^i = -\gamma^i\ . \tag{10.22}$$

The presented in (10.20) and (10.21) properties can be combined in one relation, which can be written as

$$(\gamma^\mu)^+ = \gamma^0\gamma^\mu\gamma^0\ . \tag{10.23}$$

For instance, when μ is specified to zero,

$$\left(\gamma^0\right)^+ = \gamma^0\gamma^0\gamma^0 = \gamma^0 \tag{10.24}$$

according to the fact that $\left(\gamma^0\right)^2 = 1$.

One more important fact is that $\eta^{\mu\nu}$ in the (10.19) is Minkowski metric. Another important observation is that actually one concrete realization of γ matrices, which satisfies the Clifford algebra relations, is found. In fact, there exist an infinite number of possible realizations and this can be recognized by noting that actually if the following similarity transformation is performed, γ^μ can be taken and transformed as:

$$\gamma^\mu \to U\gamma^\mu U^{-1} \ , \tag{10.25}$$

where U is an arbitrary unitary matrix. Then γ matrices obtained after this transformation will satisfy the same Clifford algebra relations and also the same Hermiticity properties. The fact that Clifford algebra relations are satisfied is trivial, because

$$U\left(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu\right)U^{-1} = U\gamma^\mu\gamma^\nu U^{-1} + U\gamma^\nu\gamma^\mu U^{-1} \ . \tag{10.26}$$

Then the following property can be used

$$U^{-1}U = UU^{-1} = \mathbb{1} \ . \tag{10.27}$$

to simplify the (10.26):

$$10.26 = U\gamma^\mu U^{-1}U\gamma^\nu U^{-1} + U\gamma^\nu U^{-1}U\gamma^\mu U^{-1} \ . \tag{10.28}$$

The (10.28) can be written as:

$$10.28 = \gamma'^\mu\gamma'^\nu + \gamma'^\nu\gamma'^\mu \ . \tag{10.29}$$

As a result

$$10.29 = 2\eta^{\mu\nu}\mathbb{1} \ . \tag{10.30}$$

So the new γ matrices obtained by the unitary transformation satisfy the same Clifford algebra relations and it can be also checked in an analogous way that due to unitarity of U, hermeticity property of newly introduced γ matrices are the same as for the old ones.

The concrete representation, which was introduced for α^i and β in terms of γ matrices, can be read in the following way:

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & -\mathbb{1} \end{pmatrix} \ , \ \gamma^i = \begin{pmatrix} \mathbb{0} & \sigma^i \\ -\sigma^i & \mathbb{0} \end{pmatrix} \ . \tag{10.31}$$

The concrete representation of γ matrices introduced above is called Dirac representation. In this representation γ^0 is diagonal and has the form presented in the (10.31).

It also should be noticed, that other representations can be obtained by applying, for instance, to the Dirac representation the unitary transformation with arbitrary U.

Let's investigate properties of γ matrices in more detail. From matrices γ^μ 16 linearly independent Hermitian matrices was constructed in the following way: first of all, an identity matrix can be built if

$$\mathbb{1} = \eta^{\mu\mu}\gamma^\mu\gamma^\mu \tag{10.32}$$

will be constructed, where there is no summation over index μ and it should be either 0 or 1, 2 and 3. It can be seen that from Clifford algebra the product of γ matrices with the same index and multiplied with Minkowski component of the Minkowski metric will obtain identity.

Then four Hermitian matrices γ^0 exist and then γ^i, which will be also Hermitian matrices if they will be multiplied by i.

Six matrices, which can be denoted as $\sigma^{\mu\nu}$, also can be formed and they will be given by the following formulas:

$$\sigma^{\mu\nu} = \begin{cases} \sigma^{ij} = i\frac{\gamma^i\gamma^j - \gamma^j\gamma^i}{2} = i\gamma^i\gamma^j \;\; i < j \;\; i,j = 1,2,3 \\ \sigma^{0j} = \frac{\gamma^0\gamma^j - \gamma^j\gamma^0}{2} = \gamma^0\gamma^j \;, \;\; j = 1,2,3 \end{cases} \tag{10.33}$$

The number of matrices in the (10.33) is equal to 6. It also should be noticed that these matrices are skew symmetric.

Then, there is one important matrix, which has a special name of γ^5, where 5 stands for historical reasons. These matrix can be introduced as:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \;, \tag{10.34}$$

where i is added to provide uh the hermiticity property for γ^5.

Finally, there are also 4 matrices τ^μ introduced in the following way:

$$\tau^\mu = \begin{cases} \tau^i = \gamma^i\gamma^5 \;, \;\; i = 1,2,3 \\ \tau^0 = i\gamma^0\gamma^5 \end{cases} \tag{10.35}$$

So, in total 16 hermitian γ matrices was constructed with the help of γ-s.

It is also should be noticed that γ^5 matrix introduced in the (10.34) satisfies the following properties:

$$\left(\gamma^5\right)^2 = \mathbb{1} \,, \;\; \left(\gamma^5\right)^+ = \gamma^5 \;. \tag{10.36}$$

In the Dirac representation due to the definition of γ^5 (10.34) the explicit form of this matrix is:

$$\gamma^5 = \begin{pmatrix} \mathbb{0} & \mathbb{1} \\ \mathbb{1} & \mathbb{0} \end{pmatrix} . \tag{10.37}$$

Then γ^5 is off diagonal in Dirac representation as soon as γ^0 is diagonal.

It can be checked that all other possible products of γ matrices and their linear combinations can be expressed by 16 hermitian matrices that was introduced. This means that these 16 matrices provide the basis in the space of 4x4 hermitian matrices.

Indeed, they form a basis it is needed to be shown that they are actually linearly independent. This can be done in the following way. To show linear independence it may be noticed that all these matrices except identity matrix have vanishing trace. For instance, let's do it for γ^μ and use the fact that

$$\gamma_\nu \gamma^\nu = \mathbb{1} , \tag{10.38}$$

where one matrix γ have low index ν and another is upper index ν and there is no summation over index ν. The result is gotten due to the Clifford algebra.

Indeed, if the upper index μ in the (10.19) is replaced by the lower index μ, then the expression will have the following form:

$$\gamma_\mu \gamma^\nu + \gamma^\nu \gamma_m u = 2\delta^\nu_\mu \cdot \mathbb{1} , \tag{10.39}$$

where Minkowski metric was replaced by the Kronecker's delta $\delta^n_\mu u$. Now, if index μ is equal to index ν, then formula (10.39) will be transformed into:

$$2\gamma_\nu \gamma^\nu = 2\delta^\nu_\nu \cdot \mathbb{1} = 2 \cdot \mathbb{1} . \tag{10.40}$$

That is why, (10.38) is right.

Coming back to the trace of γ^μ, the following standard trick to prove that trace vanishes can be used:

$$\mathrm{Tr}\gamma^\mu = \mathrm{Tr}\mathbb{1}\gamma^\mu = \mathrm{Tr}\left(\gamma_\nu \gamma^\nu \gamma^\mu\right)\Big|_{\nu \neq \mu} = \frac{1}{2}\mathrm{Tr}\left(\gamma_\nu \left(\gamma^\nu \gamma^\mu + \gamma^\mu \gamma^\nu\right)\right) , \tag{10.41}$$

where the property (10.38) and cyclic property was used. This is the way how to manipulate the product $\gamma_\nu \gamma^\nu$ under the trace and then due to the Clifford algebra relation since index ν was chosen not to coincide with index μ, the sum in the brackets of the (10.41) is

actually 0. So, the result of the (10.41) is 0, which shows that all traces of γ^μ is equal to 0:

$$\mathrm{Tr}\gamma^\mu = 0 \ . \tag{10.42}$$

Analogously it can be shown that traces of all other matrices are 0. In particular, trace of γ^5 is 0 and in the Dirac representation it can be explicitly seen from the form of the matrix (10.37):

$$\mathrm{Tr}\gamma^5 = 0 \ . \tag{10.43}$$

The fact that traces of all matrices excepting the identity matrix vanish can be used to show that these 16 matrices are linearly independent. This can be shown by assuming the opposite: it should be supposed that this matrices are linearly dependent and then if a linear combination of all of them is created, it will be equal to

$$F = a \cdot \mathbb{1} + b^\mu \gamma^\mu + c^{\mu\nu}\sigma^{\mu\nu} + d^\mu \tau^\mu + e\gamma^5 \ , \tag{10.44}$$

where a, b^μ, $c^{\mu\nu}$, d^μ and e are arbitrary coefficients. The statement that this matrices are linearly dependent means that F can be equal to zero with some non-zero coefficients. In other words, the goal is to show that the situation when $F = 0$ and some coefficients of the linear combination are not equal to zero is not possible.

The proof can be done in the following way. First of all, a trace of the left and the right hand side should be taken and since on the left hand side F is zero, trace of F is also zero. So, the expression will have the following form:

$$0 = a\mathrm{Tr}\mathbb{1} = 4 \cdot a \to a = 0 \ , \tag{10.45}$$

where the fact that traces of all matrices excepting the identity matrix are vanished was used.

Since on the left hand side of the (10.45) stands zero, a must be equal to zero. So, the term with a is absent. Then the expression for F should be taken again without identity matrix and it should be multiplied with some γ matrix γ^λ. So,

$$\mathrm{Tr}\left(\gamma^\lambda F\right) = b^\mu \mathrm{Tr}\left(\gamma^\lambda \gamma^\mu\right) + c^{\mu\nu}\mathrm{Tr}\left(\sigma^{\mu\nu}\gamma^\lambda\right) + d^\mu \mathrm{Tr}\left(\gamma^\lambda \tau^\mu\right) + e\mathrm{Tr}\left(\gamma^\lambda\gamma^5\right) \ . \tag{10.46}$$

Now, it can be shown by analyzing all traces, that all of them are equal to zero except the trace with the argument $\gamma^\lambda\gamma^\mu$. For μ equal to λ due to the Clifford algebra the identity matrix will be gained and, therefore,

$$\mathrm{Tr}\left(\gamma^\lambda F\right) = 4b^\mu \ \to \ b^\mu = 0 \ . \tag{10.47}$$

In this case, only coefficients $c^{\mu\nu}$, d^{μ} and e remain. Then multiplying the rest by $\gamma^{\lambda}\gamma^{\beta}$ it can be shown that coefficients $c^{\mu\nu}$ are equal to zero, then the same is for d^{μ} and, finally, for e. In this way proceeding similarly to what was explained before, it can be shown that all coefficients in the linear combination must be equal to zero. Therefore, these matrices are linearly independent. This means that all found matrices form a basis in the space of all complex 4x4 matrices.

Linear combinations of 16 matrices with complex coefficients allow to reconstruct any complex 4x4 complex matrix. This statement is very similar to two dimensional case with the Pauli matrices. If there are three Pauli matrices and an identity matrix, then it can be seen that linear combinations of these matrices with complex coefficients can form any 2x2 complex matrix.

It also should be noticed that transforming γ matrices to another set of γ matrices would be equivalent to transforming the wave function ψ by unitary transformation and unitary transformations are always allowed in quantum field theory and quantum mechanics. This is the same statement as that changing the basis of γ matrices can be compensated by the change of the wave function. And it can be seen from the Dirac equation

$$(i\gamma^{\mu}\partial_{\mu} - m)\,\psi = 0\ . \tag{10.48}$$

Now, this equation should be passed to another basis. So, γ matrices will be changed to the new basis according to:

$$\gamma^{\mu} \to U\gamma^{\mu}U^{-1} = U\gamma^{\mu}U^{+}\ , \tag{10.49}$$

where

$$U^{-1} = U^{+}\ . \tag{10.50}$$

Then in the new basis of γ-s expression (10.48) will be changed in the following way:

$$\left(iU\gamma^{\mu}U^{+}\partial_{\mu} - m\right)\psi = 0\ . \tag{10.51}$$

Since the fact that $UU^{+} = 1$, (10.51) can be transformed to the following expression:

$$\left(iU\gamma^{\mu}U^{+}\partial_{\mu} - mUU^{+}\right)\psi = 0\ . \tag{10.52}$$

Equation (10.52) then should be multiplied from the left by U^{-1} and after this manipulation the expression will look like

$$(i\gamma^{\mu}\partial_{\mu} - m)\,U^{+}\psi = 0\ , \tag{10.53}$$

(10.53) returned to the original form of the Dirac equation, where wave function ψ was modified by a unitary transformation to wave function ψ':

$$\psi' = U^{+}\psi \ . \tag{10.54}$$

The change of the wave function by unitary transformation should not result in the change of the probability.

Now, the discussion of transformation properties of the Dirac equation under Lorentz transformations can be performed.

Let's remember what is spinor ψ. ψ is a function on space-time, which represents a 4-component object of the following form:

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \\ \psi_3(x) \\ \psi_4(x) \end{pmatrix} \ . \tag{10.55}$$

So, it consists of 4 functions on space-time. The question, which should be understood is what happens to ψ under performing the Lorentz transformation.

Lorentz transformations of spinors

It is already known that all Lorentz transformations form a Lorentz group and the Lorentz group has four different components and one component, which contains identity is a subgroup of the Lorentz group. It's a subgroup, which contains proper orthochronous transformations. Other components of the Lorentz group are obtained from this one by means of parity or time reversal operations or combined parity with time reversal.

Under proper Lorentz transformations, it is known how the variable x transforms. So, x is a space-time coordinate and under Lorentz transformations it undergoes the following change:

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu} \ , \tag{10.56}$$

where Λ^{μ}_{ν} is a matrix of Lorentz transformations with the defining property that

$$\eta_{\mu\nu}\Lambda^{\mu}_{\alpha}\Lambda^{\nu}_{\beta} = \eta_{\alpha\beta} \ . \tag{10.57}$$

The (10.57) is the defining property of the Lorentz transformation, which actually tells what Lorentz transformation is. Lorentz transformation is a linear transformation of space-time coordinates, which preserves the Minkowski metric.

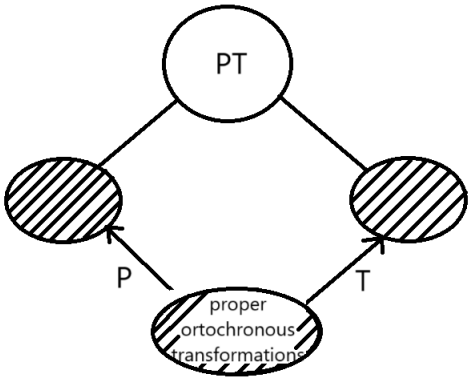

Fig. 10.1. Four components of a Lorentz group

It is very natural to assume that what happens under Lorentz transformation is that wave function $\psi(x)$ goes to $\psi'(x')$, which is related to the old function $\psi(x)$ by means of application of some matrix, which is unknown, but which is needed to be found, and act on $\psi(x)$:

$$\psi(x) \to \psi'(x') = S\psi(x) \ , \tag{10.58}$$

where S is some matrix and this matrix must be a function of the transformation Λ, which is applied to transform space-time points:

$$S = S(\Lambda) \ . \tag{10.59}$$

This matrix S should be found. This can be done by applying the Einstein relativity principle. This is a general principle, which can be applied to all equations of physics saying that in the new Lorentz frame the Dirac equation must look the same as in the original frame.

Namely, after the Lorentz transformation new function $\psi'(x')$ must still satisfy the same Dirac equation:

$$\left(i\gamma^{\mu}\frac{\partial}{\partial x'^{\mu}} - \frac{mc}{\hbar}\right)\psi'(x') = 0 \ . \tag{10.60}$$

Then, the reverse procedure should be completed. It should be turned back to the original form of the Dirac equation, which would allow to relate $\psi'(x')$ with $\psi(x)$. To do this, first of all, the derivative $\frac{\partial}{\partial x^{\mu}}$ should be investigated:

$$\frac{\partial}{\partial x^{\mu}} = \frac{\partial x'^{\nu}}{\partial x^{\mu}}\frac{\partial}{\partial x'^{\nu}} \ , \tag{10.61}$$

where the transformation of derivative under the change of coordinates was written and then using the (10.56) the expression can be simplified:

$$\frac{\partial}{\partial x^{\mu}} = \Lambda^{\nu}_{\mu}\frac{\partial}{\partial x'^{\nu}} \ . \tag{10.62}$$

The relation (10.62) can be inverted and written in the following form:

$$\left(\Lambda^{-1}\right)^{\nu}_{\mu}\frac{\partial}{\partial x^{\nu}}=\frac{\partial}{\partial x'^{\nu}}\ . \tag{10.63}$$

Expression (10.63) can be gotten by the (10.62) by multiplying both sides of the equation with an inverse matrix Λ^{-1}.

Then, the Dirac equation in the new frame (10.60) should be multiplied from the left with the matrix S^{-1}:

$$\left(iS^{-1}\gamma^{\mu}S\frac{\partial}{\partial x'^{\mu}}-\frac{mc}{\hbar}\right)\psi\left(x\right)=0\ , \tag{10.64}$$

where $\psi'\left(x'\right)$ was replaced according to the (10.58) and in the second term the expression $S^{-1}S$ was gotten, which is actually equal to 1.

Let's also substitute the change of the derivative. So,

$$\left(iS^{-1}\gamma^{\nu}S\left(\Lambda^{-1}\right)^{\mu}_{\nu}\frac{\partial}{\partial x^{\mu}}-\frac{mc}{\hbar}\right)\psi\left(x\right)=0\ , \tag{10.65}$$

The old form of the Dirac equation in the original frame will be reproduced if the condition

$$S^{-1}\gamma^{\nu}S\left(\Lambda^{-1}\right)^{\mu}_{\nu}=\gamma^{\mu} \tag{10.66}$$

is right.

Expression (10.66) means that matrix S is a solution of this expression. So, it is needed to understand how to solve this equation and find how S is related to Λ.

The (10.66) can be multiplied by matrix Λ from the right:

$$S^{-1}\gamma^{\mu}S=\Lambda^{\mu}_{\nu}\cdot\gamma^{\nu}\ . \tag{10.67}$$

Another way to write it is to multiply the original equation by S from the left and by S^{-1} from the right and also then it will lead to another form:

$$S\gamma^{\mu}S^{-1}=\left(\Lambda^{-1}\right)^{\mu}_{\nu}\gamma^{\nu}\ . \tag{10.68}$$

It can be seen that the index ν as the index of a vector is transforming with matrix Λ the same as index of a vector should transform under Lorentz transformation, but this should be compensated by acting of on γ^{μ} with the matrix S by similarity transformation. And one needs to find the matrix S, which does the Lorentz rotation on the indices of γ matrices by means of similarity transformation. So, it's very convenient to come to the infinitesimal form of the equation by assuming that a Lorentz transformation is close to

identity. In other words, there exists a small parameter, which can be called as ϵ and Λ^{μ}_{ν} can be considered as a function of this small parameter ϵ and then we expand it in powers of ϵ. Terms of this operation can be written in the form:

$$\Lambda^{\mu}_{\nu}(\epsilon) = \delta^{\mu}_{\nu} + \epsilon\eta^{\mu\rho}\omega_{\rho\nu} + \dots \ , \tag{10.69}$$

where $\omega_{\rho\nu}$ is a infinitesimal generator of Lorentz transformations. $\eta^{\mu\rho}$ is a Minkowski metric, which raises the index of ρ. It was introduced just for convenience, because indices of Λ are located on a different level: μ is up, ν is down. Then, for Lorentz generator one of lower indices up with the help of the Minkowski metric.

Then, it can be actually found what are ω-s in the equation by taking the defining relation for Lorentz transformation:

$$\eta_{\mu\nu}\Lambda^{\mu}_{\alpha}(\epsilon)\,\Lambda^{\nu}_{\beta}(\epsilon) = \eta_{\alpha\beta} \ . \tag{10.70}$$

Λ-s in the (10.70) can be expanded and then according to the (10.69) only terms of linear order in ϵ will be kept:

$$\eta_{\mu\nu}\left(\delta^{\mu}_{\alpha} + \epsilon\eta^{\mu\rho}\omega_{\rho\alpha} + \dots\right)\left(\delta^{\nu}_{\beta} + \epsilon\eta^{\nu\tau}\omega_{\tau\beta} + \dots\right) = \eta_{\alpha\beta} \ . \tag{10.71}$$

Then brackets in the (10.71) should be opened and only the leading term in ϵ should be kept. It will be seen that in the equation on ω the leading term on the left hand side will be equal to the leading term on the right hand side and this is $\eta_{\alpha\beta}$. So, there will be cancelation of terms and at linear order in ϵ it will be found that

$$\omega_{\mu\nu} = -\omega_{\nu\mu} \ , \tag{10.72}$$

which is simply means that the generator of Lorentz transformation with lower indices is simply a skew symmetric, where μ and ν run values from 0 to 3. Matrix has six components from the expression

$$\frac{4\times 3}{2} = 6 \ . \tag{10.73}$$

And these are exactly 6 Lorentz transformations making all 6 parameters of Lorentz transformations: 3 rotations and 3 Lorentz boosts. So, $\omega_{\mu\nu}$ are infinitesimal parameters of arbitrary Lorentz transformation.

Then, a relationship between the Lie algebra and its Lie group should be used. The representation of the one parametric subgroup of the Lorentz group can be written down by simply exponentiating their Lie algebra element representing the Lorentz transformation:

$$\Lambda(\epsilon) = \exp(\epsilon\eta\omega) \ . \tag{10.74}$$

The exponential (10.74) will give a one parametric subgroup of the Lorentz group, where the parameter of one parametric subgroup is exactly the ϵ. $\eta\omega$ is a matrix, which by definition equal to

$$(\eta\omega)^{\mu}_{\nu} = \eta^{\mu\rho}\omega_{\rho\nu} \ . \tag{10.75}$$

Two matrices η and ω was already introduced and $\eta\omega$ is one parametric subgroup of a Lorentz group. This means that the original equation, which is wanted to be solved can be now written in the following way:

$$S^{-1}(\epsilon)\,\gamma^{\mu}S(\epsilon) = \Lambda^{\mu}_{\nu}(\epsilon)\,\gamma^{\nu} \ , \tag{10.76}$$

where the equation is written by picking up as a Lorentz transformation elements $\Lambda^{\mu}_{\nu}(\epsilon)$ of the one parametric subgroup of the Lorentz group.

When ϵ is equal to zero, the identity matrix will be gotten and, therefore, on the right hand side γ^{μ} will be gotten and on the left hand side a matrix S must satisfy the property that when $\epsilon = 0$, it's just an identity matrix:

$$S(0) = \mathbb{1} \ . \tag{10.77}$$

The first strategy will be to find a differential equation for S with respect to the parameter ϵ and then solve this differential equation. In this way, to find what exactly the matrix S is. Therefore, the (10.76) should be differentiated with respect to ϵ:

$$-\underbrace{S^{-1}\frac{\mathrm{d}S}{\mathrm{d}\epsilon}S^{-1}}_{\frac{\mathrm{d}\left(S^{-1}(\epsilon)\right)}{\mathrm{d}\epsilon}}\gamma^{\mu}S + S^{-1}\gamma^{\mu}\frac{\mathrm{d}S}{\mathrm{d}\epsilon} = \frac{\mathrm{d}\Lambda^{\mu}_{\nu}(\epsilon)}{\mathrm{d}\epsilon}\gamma^{\nu} \ . \tag{10.78}$$

This relation further can be written in the following way. If the auxiliary variable X is introduced, which is by definition

$$X := \frac{\mathrm{d}S}{\mathrm{d}\epsilon}S^{-1} \ , \tag{10.79}$$

then equation (10.78) can take the following form

$$[\gamma^{\mu}, X] = \frac{\mathrm{d}\Lambda^{\mu}_{\nu}}{\mathrm{d}\epsilon}S\gamma^{\nu}S^{-1} \ . \tag{10.80}$$

In other words, in the (10.80) equation (10.78) was multiplied by S from the right and by s^{-1} from the left. Then, the left hand side reduces to a commutator of γ^{μ} with X. The expression $S\gamma^{\nu}S^{-1}$ comes from multiplications and then this is actually nothing as:

$$S\gamma^{\nu}S^{-1} = \left(\Lambda^{-1}\right)^{\nu}_{\beta}\gamma^{\beta} \ . \tag{10.81}$$

Therefore,

$$[\gamma^\mu, X] = \left(\frac{\mathrm{d}\Lambda}{\mathrm{d}\epsilon}\Lambda^{-1}\right)^\mu_\nu \gamma^\nu \ . \tag{10.82}$$

If the solution for the one parametric subgroup is substituted into the (10.82), than the result will be equal to:

$$\boxed{[\gamma^\mu, X] = \eta^{\mu\rho}\omega_{\rho\nu}\gamma^\nu} \ . \tag{10.83}$$

Equation (10.83) for X, which is needed to be solved, is purely matrix equation. So, all elements here are matrices. This equation tells that the commutator of two matrices γ^μ and X is another matrix, where $\omega_{\rho\nu}$ are parameters of Lorentz transformations.

Lecture 11. Lorentz Boosts. Parity

Last time, the matrix S was introduced, which performs the transformation of wave function ψ in such a way that the transformed Dirac equation has the same form as an original one. And it was found out that if for a Lorentz transformation a one parametric subgroup given by

$$\Lambda(\epsilon) = \exp(\epsilon\eta\omega) \ , \tag{11.1}$$

where $\eta\omega$ has components $(\eta\omega)^{\mu}_{\nu}$ given by the formula:

$$(\eta\omega)^{\mu}_{\nu} = \eta^{\mu\rho}\omega_{\rho\nu} \ . \tag{11.2}$$

Then for matrix S the corresponding formula for Lorentz transformation turns into the following equation:

$$S^{-1}(\epsilon)\,\gamma^{\mu}S(\epsilon) = \Lambda^{\mu}_{\nu}(\epsilon)\,\gamma^{\nu} \ . \tag{11.3}$$

As the next step equation (11.3) should be differentiated with respect to ϵ and the quantity X given by

$$X = \frac{\mathrm{d}S}{\mathrm{d}\epsilon}S^{-1} \tag{11.4}$$

should be introduced.

Then the equation for X takes the form:

$$\boxed{[\gamma^{\mu}, X] = \eta^{\mu\rho}\omega_{\rho\nu}\gamma^{\nu}} \ . \tag{11.5}$$

And this is an equation, which should be solved for the matrix X.

The convenient way to solve it is to look for a solution in the following form:

$$X = \lambda^{\rho\lambda}\omega_{\rho\lambda} \ . \tag{11.6}$$

So, a solution, which is linear in generators ω and ω are generators of Lorentz transformations, carries indices ρ and λ and this is the way how the form, in which the expression for matrix X contains, looks.

Substituting the (11.6) into formula (11.5) will lead to an equation of the following form:

$$g^{\mu}\lambda^{\rho\lambda} - \lambda^{\rho\lambda}\gamma^{\mu} = \frac{1}{2}\left(\eta^{\mu\rho}\gamma^{\lambda} - \eta^{\mu\lambda}\gamma^{\rho}\right) \ , \tag{11.7}$$

where on both sides ω is canceled. Also the fact that ω is skew symmetric was used:

$$\omega_{\rho\lambda} = -\omega_{\lambda\rho} \ . \tag{11.8}$$

λ is also can be taken to be skew symmetric

$$\lambda^{\rho\lambda} = -\lambda^{\lambda\rho} \ . \tag{11.9}$$

to satisfy the condition of this type.

With this assumption equation for $\lambda^{\rho\lambda}$ can be solved and, directly, the solution will have the following form:

$$\lambda^{\mu\nu} = \frac{1}{8}\left[\gamma^{\mu}, \gamma^{\nu}\right] = \frac{1}{8}\left(\gamma^{\mu}\gamma^{\nu} - \gamma^{\nu}\gamma^{\mu}\right) \ . \tag{11.10}$$

Due to the Clifford algebra expression (11.10) can also be written as

$$\lambda^{\mu\nu} = \frac{1}{4}\gamma^{\mu}\gamma^{\nu} \ , \tag{11.11}$$

where index μ is not equal to ν. This is explicitly taken into account by the anti-symmetrization procedure and, therefore, it can be checked that this is solution by substituting matrix λ into equation (11.7) and using the Clifford algebra relation for γ matrices.

Then, on the other hand, since λ was found, the equation for S can be now specified as follows:

$$\frac{\mathrm{d}S}{\mathrm{d}\epsilon} = XS = \frac{1}{4}\gamma^{\mu}\gamma^{\nu}\omega_{\mu\nu} \cdot S \ . \tag{11.12}$$

The (11.12) is an explicit differential equation, on the right hand side of which is the product of a four by four matrix, because γ^{μ} and γ^{ν} are 4x4 matrices multiplied with matrix S. This equation must be supplemented with the initial condition. If ϵ is equal to zero, the corresponding matrix S must be simply equal to the identity matrix:

$$S\left(\epsilon = 0\right) = \mathbb{1} \ . \tag{11.13}$$

The differential equation together with the initial condition has a unique solution and this unique solution is simply given by:

$$S\left(\epsilon\right) = \exp\left(\frac{1}{4}\epsilon\gamma^{\mu}\gamma^{\nu}\omega_{\mu\nu}\right) \ . \tag{11.14}$$

If expression (11.14) is differentiated with respect to epsilon, then the matrix

$$\frac{1}{4}\gamma^{\mu}\gamma^{\nu}\omega_{\mu\nu} \tag{11.15}$$

comes down and exponential remains and, therefore, expression (11.12) will be right.

So, what is written down in the (11.14) is a solution, where ϵ is a one parametric subgroup and it is equal to any real number. In the described case, the variable ϵ is small in order to consider expansion around ϵ is equal to zero. It can be fixed, for instance, to be equal to one and, in this way, S is taken as:

$$S = \exp\left(\frac{1}{4}\gamma^{\mu}\gamma^{\nu}\omega_{\mu\nu}\right) \ , \tag{11.16}$$

which can be also written due to the skew symmetry of the matrix ω in the following form:

$$\boxed{S = \exp\left(\frac{1}{2}\sum_{\mu<\nu}\gamma^{\mu}\gamma^{\nu}\omega_{\mu\nu}\right)} \ . \tag{11.17}$$

The (11.17) is the final formula for the Lorentz transformation of a spinor. This means that under the Lorenz transformation λ from the proper Lorenz group the Dirac spinner transforms in the following way:

$$\psi'\left(x'\right) = \exp\left(\frac{1}{4}\sum_{\mu,\nu=0}^{3}\gamma^{\mu}\gamma^{\nu}\omega_{\mu\nu}\right)\psi\left(x\right) \ , \tag{11.18}$$

where x' is obtained from x by means of Lorenz transformation with the element Λ:

$$x' = \Lambda x \ . \tag{11.19}$$

The answer that has been found is rather interesting. The first thing that can be noticed is that the matrices Λ^{μ}_{ν}, which feature in Lorenz transformations, satisfy the following algebraic relation, which is the commutator of two matrices:

$$[\lambda^{\mu\nu}, \lambda^{\rho\sigma}] = \eta^{\nu\rho}\lambda^{\mu\sigma} - \eta^{\mu\rho}\lambda^{\nu\sigma} - \eta^{\nu\sigma}\lambda^{\mu\rho} + \eta^{\mu\sigma}\lambda^{\nu\rho} \ . \tag{11.20}$$

In fact, it can be recognized that the (11.20) is nothing else as the commutation relation between generators of the Lie algebra of the Lorenz group. This means that the Lorenz transformations, which were constructed, are transformations realizing the spinor representation of the Lorenz group.

It's important to realize that this representation is different from the vector representation acting on coordinates x^{μ} with the help of Λ^{μ}_{ν}.

Let's investigate this Lorenz transformation that was found before in more details. First of all, the transformation can be specified for the case of spatial rotations.

Spatial rotations

This means that just rotations are considered. As it well known the usual three-dimensional rotations can be specified by given a rotation axis and this can be done by specifying the unit vector $\vec{n}$ with components n_i:

$$\vec{n} = \{n^i\}\Big|_{i=1,2,3} , \tag{11.21}$$

where unit means that if the length of this vector is computed, then the length is equal to one:

$$|\vec{n}\,| = 1 . \tag{11.22}$$

So, $\vec{n}$ is a rotation axis e.g. the axis around which rotations will be performed. A rotation angle should also be specified and it can be denoted by θ. In other words, there is axis, which is specified by the unit vector and an angle theta by means of each rotations will be completed (fig. 11.1).

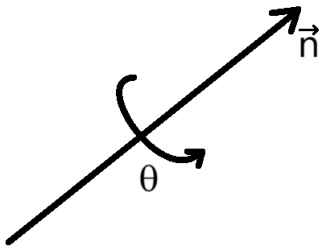

Fig. 11.1. Rotational axis formed by the unit vector $\vec{n}$

Then there are also elements ω_{ij}, which should be expressed and which are equal to

$$\omega_{ij} = -\theta \cdot \epsilon_{ijk} n^k , \tag{11.23}$$

where n^k is the infinitesimal matrix of a transformation.

Then, the following expression can be computed

$$\gamma^i \gamma^j \omega_{ij} = -\theta \gamma^i \gamma^j \epsilon_{ijk} n^k . \tag{11.24}$$

Expression (11.24) sometimes can be written in the following way:

$$\gamma^i \gamma^j \omega_{ij} = 2i\theta \Sigma_k n^k , \tag{11.25}$$

where the following property was used:

$$\gamma^i \gamma^j \epsilon_{ijk} = -2i\Sigma_k . \tag{11.26}$$

where Σ with an upper index due to the Clifford algebra relations of γ-s can be written in the next form:

$$\Sigma^i = \gamma^5\gamma^0\gamma^i \ . \tag{11.27}$$

It can be shown that due to the Clifford algebra relations (11.26) is the same as an expression (11.27). It can be also seen that the object, which was introduced is the matrix and this matrix has 3 components. Σ is the collection of 3 matrices labeled by the index i:

$$\Sigma^i = \{\Sigma^1, \Sigma^2, \Sigma^3\} \ . \tag{11.28}$$

In this way, Σ can be identified with 3 vector:

$$\Sigma^i = \vec{\Sigma} \ . \tag{11.29}$$

In fact, if a variable J_i is equal to

$$J_i = \frac{1}{2}\Sigma_i = -\frac{1}{2}\Sigma^i \ , \tag{11.30}$$

then it satisfies the following algebraic relation:

$$\boxed{[J_i, J_j] = -i\epsilon_{ijk}J_k} \ . \tag{11.31}$$

From expression (11.31) it can be recognized that this algebraic relation is nothing else as the algebraic relations of the generators of the Lie algebra of the rotation group. On the other hand, it is known that the (11.31) is a commutation relation for angular momentum. In this way, operators J_i can naturally interpret as spin operators. So Σ and J can be called as spin operators. These objects are related to the notion of spin and the notion of spin comes from characterizing the transformation properties of an object with respect to three-dimensional rotations. Spin is a property of a wave function to transform under three-dimensional rotations of space.

If not rotation by infinitesimal, but by finite angle θ is performed, then the general formula should be used and for such a transformation S will be equal to

$$S\left(\vec{n}\,, \theta\right) = \exp\left(\frac{i}{2}\vec{\Sigma}\,\vec{n}\,\theta\right) \ . \tag{11.32}$$

where indices ν and μ are taken only values 1, 2 and 3.

Expression (11.32) can be gotten explicitly starting from the formula

$$S\left(\vec{n}\,, \theta\right) = \exp\left(\frac{1}{2}\gamma^i\gamma^j\omega_{ij}\theta\right) \ , \tag{11.33}$$

where the realization of ω_{ij} (11.23) can be used, which involves the Levi-Civita tensor. Then the definition of Σ should be used, which are 4x4 matrices. Such a way, the exponential can be computed explicitly in the usual way by expanding in the Taylor series:

$$S\left(\vec{n}\,,\theta\right) = \mathbb{1}\cdot\cos\frac{\theta}{2} + i\vec{\Sigma}\,\vec{n}\,\sin\frac{\theta}{2}\ . \tag{11.34}$$

Formula (11.34) realizes a transformation by finite angle θ. What is interesting about this formula is that if a rotation by angle 2π is performed, then because in the expression the arguments of half of the angles are involved, this will be nothing else as:

$$S\left(\vec{n}\,,\theta+2\pi\right) = -S\left(\vec{n}\,,\theta\right)\ . \tag{11.35}$$

In other words, when a full rotation is performed, the spinor does not come to itself, but it changes the sign. If the transformation of a spin is considered, then

$$\psi'\left(x'\right) \to -\psi\left(x\right)\ , \tag{11.36}$$

but it does not return to itself. The point x' in the opposite actually returns to itself, because x' transforms with a matrix Λ and matrix Λ is a matrix, which depend on angle θ in such a way that it is periodic, up to upon rotation by 2π and, therefore, after rotation by 2π

$$\psi'\left(x\right) = -\psi\left(x\right)\ . \tag{11.37}$$

Such a representation, which has the property described above is called double valued representation of the rotation group. This is a specific of the spinor representation, because the property occurs at the one and the same space-time point. Under performing a loop in space and returning back to the original position a spinor can change the sign. From this transformation it can be seen that such representations have a specific name in the representation theory, they are called double valued representations of the rotation group. So, the rotation group is $SO(3)$ has a topological property of being not simply connected. If one try to draw it from a topological point of view, it is looks like a circle and it is connected, but it's not simply connected, which means that there are loops in $SO(3)$ that are not contractable to an identity. Such loops, which are not contractable to the to the identity, mean that the corresponding space is not simply connected.

At each point of $SO(3)$ two values of another group exists, which is called a double cover of $SO(3)$ and it turns out that this double cover of $SO(3)$ is is know as $SU(2)$ group. And $SU(2)$ group is a double cover of $SO(3)$, which is simply connected (fig. 11.2).

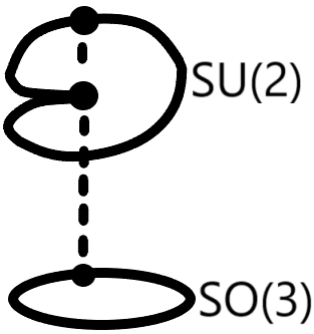

Fig. 11.2. $SO(3)$ and $SU(2)$ groups

This leads to the existence of a special class of representations, which are called as spinor representations or double valued representations of the rotational group. In fact, this also means something else. This means that the spinor cannot be observed itself, because if one is an observer and a transformation of the Lorentz frame by 2π is performed, in fact, one returns back to the original point by making a loop in space. This means that $\psi(x)$ cannot be observed, but what can be observed is the quadratic combinations, for instance, quadratic combinations of spinor components or some quantities like $\overline{\psi}\psi$.

So, these quantities are quadratic in ψ and these quantities remain invariant upon making the rotation in space. Such quantities would be invariant and such quantities then provide observables, but not the components of the spinor itself.

It is also should be noticed that every representation is characterized by a spin. Spin can take half integer values, which are $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ and any representation of a given spin has $2j+1$ components. For instance, for the case of spin zero, only one component exists and this is a scalar. For the case of one half there are two components and although the representation is realized on 4-dimensional spinors, the following is can be seen: if one compute explicitly what Σ is in the Dirac representation, then it will be found out that this is given by:

$$\Sigma^i = \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} . \tag{11.38}$$

Expression (11.38) is a diagonal in terms of Pauli matrices. Then with this explicit form of Σ^i the formula for S can be written in the following way:

$$S(\vec{n}, \theta) = \begin{pmatrix} \mathbb{1} \cdot \cos\frac{\theta}{2} + i\vec{\sigma}\,\vec{n}\,\sin\frac{\theta}{2} & \mathbb{0} \\ \mathbb{0} & \mathbb{1} \cdot \cos\frac{\theta}{2} + i\vec{\sigma}\,\vec{n}\,\sin\frac{\theta}{2} \end{pmatrix} \tag{11.39}$$

If a four component spinor is represented in terms of two component spinors, the

expression for ψ can be written as follows:

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \ , \tag{11.40}$$

where two components ϕ and χ transform independently with respect to three-dimensional rotations and, moreover, they transform in the same way. That is exactly the representation of spin $\frac{1}{2}$, which realized by means of the following 2x2 matrices

$$R\left(\vec{n}\,,\theta\right) = \mathbb{1} \cdot \cos\frac{\theta}{2} + i\vec{\sigma}\,\vec{n}\,\sin\frac{\theta}{2}\ . \tag{11.41}$$

In fact, one can build up representation corresponding to any half integer spin. In particular, vector representation is a three-dimensional representation.

It should be noticed that for the representation of spin $\frac{1}{2}$ the matrix S factorizes into diagonal blocks and a spinor can be reduced on two component spinor, which do not mix with each other. So, they transform completely independent for the reason that ϕ and χ do not mix at all. That means that the four-dimensional representation just breaks into two representations for two component spinors and each of them is just a representation of spin $\frac{1}{2}$ realized by 2x2 matrices. This classification of representations characterized by spin is connected with the term of irreducible representations. In other words, representations, which cannot be decomposed further and do not have inside themselves invariant subspaces, are so-called irreducible representations. The classification of irreducible representations of the rotation group tells that any such representation is uniquely characterized up to unitary equivalence by a half integer number, which is called spin. And if this number is fixed, then the dimension in which the representation is realized is related to spin by the formula $2j + 1$.

Once again, every reducible representation in representation theory sometimes are called as $\mathcal{D}_j$ and parameterized by half integer number, where j is half integer number called spin. The dimension of this representation $\mathcal{D}_j$ for a given spin j is equal to

$$\mathrm{dim}\mathcal{D}_j = 2j + 1\ . \tag{11.42}$$

Lorentz boosts

For Lorentz boosts there is ω_{0i} exists, which is given by by the formula:

$$\omega_{0i} = n^i\vartheta\ , \tag{11.43}$$

where ϑ ia an hyperbolic angle. Then the matrix of Lawrence boosts $S(\vec{n}, \vartheta)$ is given by:

$$S(\vec{n}, \vartheta) = \exp\left(\frac{1}{2}\gamma^0\gamma^i n^i \vartheta\right) , \tag{11.44}$$

which is the same as

$$S(\vec{n}, \vartheta) = \exp\left(\frac{1}{2}\vec{\alpha}\,\vec{n}\,\vartheta\right) , \tag{11.45}$$

where the following replacement was applied:

$$\gamma^0\gamma^i = \alpha^i . \tag{11.46}$$

The (11.45) again can be explicitly computed and the following expression can be found:

$$S(\vec{n}, \vartheta) = \mathbb{1}\cosh\frac{\vartheta}{2} + \vec{\alpha}\,\vec{n}\,\sinh\frac{\vartheta}{2} . \tag{11.47}$$

The fact that we are dealing with Lorentz boosts now shows itself in the appearance of the hyperbolic functions. Hyperbolic functions signaling that now transformations occurs due to Lawrence boosts.

Since in the Dirac representation matrices α^i are block off-diagonal, the spinors ϕ and χ, which was introduced above mix under Lawrence boosts, because explicitly matrices α^i are given by:

$$\vec{\alpha} = \begin{pmatrix} \mathbb{0} & \sigma^i \\ \sigma^i & \mathbb{0} \end{pmatrix} . \tag{11.48}$$

As a result, when the matrix $\vec{\alpha}$ will be computed explicitly, it will be seen that this matrix will mix the ϕ and χ. Therefore, this matrices are not relativistic covariant.

Let's also mention the Hermeticity properties of Lorentz transformations that were found. For space rotations it can be seen that if one take S^+ e.g. compute the Hermitian conjugate of matrix S, then this will be equal to:

$$S^+ = \exp\left(\frac{1}{2}\sum_{i<j}\gamma^{j+}\gamma^{i+}\omega_{ij}\right) , \tag{11.49}$$

where ω_{ij} are real numbers. Using the fact that matrices γ^i are anti-Hermitian, then the (11.49) can be written in the following form:

$$S^+ = \exp\left(\frac{1}{2}\sum_{i<j}\gamma^{j}\gamma^{i}\omega_{ij}\right) . \tag{11.50}$$

Then if the matrices γ is put in the old order by anti-commuting them, it will be gotten that:

$$S^+ = \exp\left(-\frac{1}{2}\sum_{i<j}\gamma^i\gamma^j\omega_{ij}\right) \ . \tag{11.51}$$

This, finally, it was gotten that

$$S^+ = S^{-1} \ . \tag{11.52}$$

In other words, what was found is that

$$S^+S = 1 \ . \tag{11.53}$$

Obviously, such matrices was called as unitary. It is seen that rotations are realized by unitary matrices or, in other words, a unitary representation of the rotation group was realized.

For Lorenz boosts the following is right:

$$S^+ = \exp\left(\frac{1}{2}\gamma^{i+}\gamma^{0+}\omega_{0i}\right) \ . \tag{11.54}$$

It is known that γ^i is anti-hermitian and γ^0 is hermitian and, therefore

$$S^+ = \exp\left(-\frac{1}{2}\gamma^i\gamma^0\omega_{0i}\right) \ . \tag{11.55}$$

Then the anti-commutation relation can be used for matrices γ^i and γ^0 by using the Clifford algebra and the following expression will be gotten:

$$S^+ = \exp\left(\frac{1}{2}\gamma^0\gamma^i\omega_{0i}\right) = S \ . \tag{11.56}$$

This is different from the unitarity condition, rather the condition (11.56) shows that matrix S realizing Lorenz boosts is Hermitian. In fact, both properties under Hermitian conjugation e.g. behavior of rotation matrices and behavior of Lorenz boosts can be combined in one formula, which looks as follows:

$$\boxed{S^+ = \gamma^0 S^{-1}\gamma^0} \, , \tag{11.57}$$

when it can be seen that actually γ^0 commutes through matrix S^{-1} and $(\gamma^0)^2$ is equal to 1. On the other hand, if S is a Lorenz boosts, then γ^0 anti-commutes with γ^i standing in the exponential and for this reason the equality of S^{-1} and S will be gotten.

Both together Hermiticity properties are encoded in one formula, which looks like (11.57). In fact, this formula motivates to introduce the notion of a Dirac conjugate spinor, which is defined as follows:

$$\overline{\psi} = \psi^{*t}\gamma^0 \ . \tag{11.58}$$

So the $\overline{\psi}$ is not a column, but this is a row consisting of 4 numbers, which can be also written as

$$\overline{\psi} = \psi^{+}\gamma^0 \ , \tag{11.59}$$

which means conjugation and transposition multiplied with γ^0. This is because it's spinor $\overline{\psi}$ has a simple transformation property under Lorenz transformation:

$$\overline{\psi'(x')} = (S\psi)^{*\,t}\gamma^0 \ , \tag{11.60}$$

which is according to the discussion the same as

$$\overline{\psi'(x')} = \psi^{+}S^{+}\gamma^0 \ . \tag{11.61}$$

In all cases, S^+ is given by formula (11.57) and then

$$\overline{\psi'(x')} = \psi^{+}\gamma^0 S^{-1}\gamma^0\gamma^0 \ . \tag{11.62}$$

The (11.62) can be simplified and then the final expression for the $\overline{\psi'(x')}$ will look as follows:

$$\overline{\psi'(x')} = \overline{\psi(x)}S^{-1} \ . \tag{11.63}$$

In other words, under Lorenz transformations the Dirac conjugate spinor transforms simply by multiplying it from the right by S^{-1}. The Dirac equation now can be written in the following form:

$$\overline{\psi}\left(i\gamma^\mu\partial_\mu + \frac{mc}{\hbar}\right) = 0 \ , \tag{11.64}$$

where now the derivative acting on the $\overline{\psi}$, which is on the left hand side. The (11.64) represents the Dirac equation for a Dirac conjugate spinner. It is of course follows from the original Dirac equation by taking conjugation of the original Dirac equation and multiplying it with γ^0. Then it can be independently checked that this equation for the Dirac conjugate spinor remains to be Lorentz covariant exactly due to the properties that under Lorentz transformations a Dirac conjugate spinor transforms by multiplying with matrix S^{-1} from the right.

In general, the transformation properties of spinors with respect to the rotation group and with respect to Lorentz boosts, which all together provide the realization of the proper orthochronous Lorentz transformations were introduced and it was done in the Dirac representation of γ matrices. And in this Dirac representation for γ matrices it was shown that a rotation group is realized in the form of a reducible representation, which splits into two irreducible two-dimensional representations, while Lorentz boosts mix these two-dimensional representations between themselves.

Now discrete transformations should be introduced, which are left over. They have to be taken into account if we want to talk about the full Lorentz group. Therefore, it should be realized how spinors transform under parity and under time reversal. Let's start from parity.

Parity

Parity is an operation, which takes three-dimensional $\vec{x}$ and sends it to minus $-\vec{x}$:

$$\mathcal{P}: \ \vec{x} \to -\vec{x} \ , \tag{11.65}$$

where time remains unchanged.

While coordinates x^i are reflected this means that

$$\begin{cases} \vec{x}\,' = -\vec{x} \\ t' = t \end{cases} \tag{11.66}$$

Then, again for $\psi'(x')$ in the new Lorentz frame obtained from the old one by means of the parity transformation as follows:

$$\psi'(x') = \mathcal{P}\psi(x) \ , \tag{11.67}$$

where $\mathcal{P}$ is some matrix is needed to be found and which leaves the Dirac equation to be a covariant. It is very simple to check, because if for the Dirac equation in the old frame

$$(i\gamma^\mu \partial_\mu - m)\,\psi(x) = 0 \ , \tag{11.68}$$

then this must be the same in the new frame

$$\left(i\gamma^\mu \partial'_\mu - m\right)\psi'(x') = 0 \ . \tag{11.69}$$

Then, from the (11.69) it can be seen that the term

$$i\gamma^0 \partial_0 \tag{11.70}$$

will not be changed, because time is not changing under the parity transformation, while all spatial derivative must change a sign, because of the (11.66):

$$\left(i\gamma^0\partial_0 - i\gamma^i\frac{\partial}{\partial x^i} - m\right)\mathcal{P}\psi\left(x\right) = 0\ . \tag{11.71}$$

From the (11.71) it is seen that if this equation is multiplied by $\mathcal{P}^{-1}$ from the left, then to get the original Dirac equation following conditions should be to satisfied:

$$\mathcal{P}^{-1}\gamma^0\mathcal{P} = \gamma^0\ , \tag{11.72}$$

while

$$\mathcal{P}^{-1}\gamma^i\mathcal{P} = -\gamma^i\ . \tag{11.73}$$

If the matrix $\mathcal{P}$, which satisfies conditions (11.72) and (11.73) is found, then the transformation of spinors under parity transformation will be gotten.

Equations (11.72) and (11.73) can be elementary solved if matrix $\mathcal{P}$ will be proportional to γ^0 with some proportionality coefficient, which can be denoted by $\eta_{\mathcal{P}}$:

$$\mathcal{P} = \eta_{\mathcal{P}}\gamma^0\ . \tag{11.74}$$

From the (11.74), $\mathcal{P}$ will commute with γ^0 and the (11.72) is trivially satisfied and then due to the fact that γ^0 is anti-commuting with γ^i the second equation (11.73) will be satisfied.

It is also should be noticed that a number $\eta_{\mathcal{P}}$ is called as internal parity. $\mathcal{P}$ can be required to be a unitary operator

$$\mathcal{P}^{+}\mathcal{P} = 1\ , \tag{11.75}$$

which means that the number $\eta_{\mathcal{P}}$ must be a phase and it's modulus should be equal to

$$|\eta_{\mathcal{P}}| = 1\ . \tag{11.76}$$

Then, if parity is performed twice, the original point must be achieved. Therefore,

$$\mathcal{P}^2 = \mathbb{1}\ . \tag{11.77}$$

On the other hand, it is known that we are dealing with a spinor and actually in this case we can say that after application of parity twice ones returning to itself up to a minus sign. And, therefore, in the second situation the following can be required:

$$\mathcal{P}^2 = -\mathbb{1}\ . \tag{11.78}$$

So, in the first case (11.77) for internal parity

$$\eta_{\mathcal{P}} = \pm 1 \ , \tag{11.79}$$

while in the second case

$$\eta_{\mathcal{P}} = \pm i \ . \tag{11.80}$$

It can also be found that $\mathcal{P}$ anti-commutes with γ^5:

$$\mathcal{P}^{-1}\gamma^5\mathcal{P} = \mathcal{P}^{-1}\left(i\gamma^0\gamma^1\gamma^2\gamma^3\right)\mathcal{P} \ . \tag{11.81}$$

Then $\mathcal{P}$ is proportional to γ^0 and γ^0 should be moved through all γ matrices. As a result it will be found that

$$11.81 = -\gamma^5 \ . \tag{11.82}$$

If a four component spinor written in terms of two components ϕ and χ is taken, then the parity operation act on it as follows:

$$\mathcal{P}\begin{pmatrix}\phi\\ \chi\end{pmatrix} = \eta_{\mathcal{P}}\gamma^0\begin{pmatrix}\phi\\ \chi\end{pmatrix} \ . \tag{11.83}$$

In the Dirac representation this is nothing else as

$$11.83 = \begin{pmatrix}\eta_{\mathcal{P}}\phi\\ -\eta_{\mathcal{P}}\chi\end{pmatrix} \ , \tag{11.84}$$

where parity preserves two component spinors and they do not mix, but from this computation it is seen that components have a different parity.

If the internal parity is chosen, then with respect to the parity it can be seen that the bilinear combinations of spinors can be classified in the following way:

1) $\overline{\psi}\psi$, which is the scalar product of the Dirac conjugate spinor with the spinor itself and, therefore $\overline{\psi}\psi$ is a number. And under Lorentz transformations this object here transforms as a scalar.

2) $\overline{\psi}\gamma^5\psi$. Under Lorentz transformations, because Lorentz transformations anti-commute with matrix γ^5, then this object appears to be a pseudo scalar. It also can be seen that this object goes to minus itself under parity transformation due to the property that under parity

$$\mathcal{P}\gamma^5 = -\gamma^5 \ . \tag{11.85}$$

3) $\overline{\psi}\gamma^{\mu}\psi$, which transforms under Lorentz transformation as a vector. It is very important combination, because it's a current, which has the dependence on the index μ.

4) $\overline{\psi}\gamma^{\mu}\gamma^{5}\psi$, which is a pseudo vector. It's transforms as a usual vector with respect to Lorentz transformations from proper orthochronous Lorentz group, but it does not change a sign under parity.

5) $\overline{\psi}\sigma^{\mu\nu}\psi$ is bilinear combination of spinors, where $\sigma^{\mu\nu}$ is anti-symmetric tensor, which was introduced earlier.

Lecture 12. Time Reversal. Weyl Spinors and Weyl Equations

At the previous lecture the transformations of the Dirac equation with respect to Lorenz group was discussed and it was finished by understanding how a spinor transforms under parity transformations.

Now, it is time to discuss the remaining type of transformations, which wasn't discussed, namely, time reversal.

Time reversal

At this topic it will be observed how a Dirac spinor transforms under transformations of the type

$$t \to -t \ . \tag{12.1}$$

Classically, having such a transformation or having such a symmetry means that all solutions or all trajectories $q(t)$, which are solutions of equations of motion, are time reversible.

In other words, if a solution is passed to another solution or to another function $q_{rev}(t)$, which is

$$q_{rev}(t) = q(-t) \ , \tag{12.2}$$

then this $q_{rev}(t)$ also appears to be a solution of equations of motion.

In this respect, it can be said that this transformation is in fact the symmetry of equations. In a quantum mechanical situation of the Dirac equation, the procedure is absolutely the same as was done, for instance, for parity.

First of all, one assume that time reversal symmetry is some operation, which can be denoted as $\mathcal{T}$. This transformation is assumed to act on a spinor ψ in the following way:

$$\mathcal{T} : \ \psi'(x') = \mathcal{T}\psi(x) \ , \tag{12.3}$$

where an operation of time reversal does not do anything to spatial coordinates, but transforms the time coordinate. So,

$$x'^0 = -x^0 \ , \tag{12.4}$$

while

$$x'^i = x^i \ . \tag{12.5}$$

Let's look at what actually happens in quantum mechanics under time reversal without going even to a field theory to the Dirac equation. First it should be understood how this operator actually acts in usual quantum mechanics. And in quantum mechanics, transformations of operators under symmetries is done in the following way. If some operator A exists, then under time reversal it transforms as:

$$\mathcal{T}A\mathcal{T}^{-1} = A' \ . \tag{12.6}$$

In quantum mechanics, it should be assumed that if an operator of coordinate exists, then this operator does not change under the $\mathcal{T}$ transformation:

$$\mathcal{T}\vec{x}\,\mathcal{T}^{-1} = \vec{x} \ . \tag{12.7}$$

So, coordinates of a particle do not change.

In addition to an operator of coordinate, there is also an operator of momentum and since momentum through its definition involves the time:

$$\vec{p} = \frac{\mathrm{d}\vec{x}}{\mathrm{d}t} \cdot m \ , \tag{12.8}$$

where the (12.8) is a classical definition of the momentum. Therefore, in non relativistic quantum mechanics under such an operation operator of momentum must change the sign. So,

$$\vec{p} \to -\vec{p} \ . \tag{12.9}$$

On the other hand, if the fundamental commutation relation of quantum mechanics

$$\left[p_i, x^j\right] = i\hbar\delta_i^j \tag{12.10}$$

is introduced, which is Heisenberg commutation relation, then this commutation relation will be invariant under such a transformation with $\mathcal{T}$ only if this operation $\mathcal{T}$ acts also on imaginary unit in the following way

$$\mathcal{T}i\mathcal{T}^{-1} = -i \ . \tag{12.11}$$

If the property (12.11) is applied to the (12.10), then

$$\mathcal{T}\left[p_i, x^j\right]\mathcal{T}^{-1} = \left[p_i', x'^j\right] = -\left[p_i, x^j\right] \ , \tag{12.12}$$

where the fact that $\vec{p}$ is change the sign and $\vec{x}$ is not is used. On the right hand side of the (12.10) the following expression will be gotten:

$$-\left[p_i, x^j\right] = \mathcal{T}i\mathcal{T}^{-1}\hbar\delta_i^j = -i\hbar\delta_i^j \ . \tag{12.13}$$

It is clearly seen that the (12.13) is the same as the (12.10). This means that in the quantum mechanics Hamiltonian is invariant and, therefore, Heisenberg equations are invariant with respect to time reversal operation. In general, such a transformation is realized by the so called anti-linear anti-unitary operators. Anti-linear transformation is clear from the definition of what linear transformation is. If a wave function exists and it multiplied with a complex number α and then for this construction $\mathcal{T}$ is applied, an usual linear transformation would mean that α can be taken out, but for anti-linear transformation the complex number α will be conjugated:

$$\mathcal{T}(\alpha\psi) = \alpha^{*}\mathcal{T}\psi \ . \tag{12.14}$$

Anti-unitarity means something else. The notion of unitarity is defined with respect to a chosen scalar product in a Hilbert space and anti-unitarity means that if there is a scalar product of two elements of an space x and y is existed and the operator $\mathcal{T}$ is applied to this scalar product, then the result of this transformation will be equal to

$$\langle \mathcal{T}x|\mathcal{T}y\rangle = \overline{\langle x|y\rangle} = \langle y|x\rangle \ , \tag{12.15}$$

where the conjugated expression was gotten.

Time reversal quantum mechanics is different from what was normal for transformations and symmetries in a sense that symmetries were realized by unitary operators. And the time reversal is realized by anti-linear anti-unitary transformations. Of course, this feature of quantum mechanics goes to the quantum field theory and the operator which will be found in the Dirac theory, which corresponds to this time reversal, has a property of similar types.

Therefore, the following should be assumed: transformation from $\psi(x) \to \psi'(x')$, which is transformation

$$\psi'(x') = \mathcal{T}\psi(x) \tag{12.16}$$

can be explicitly realized in the next way. First of all the complex conjugation should be involved. From a spinor ψ we come to a complex conjugate spinor and then this is not the end of the story, because the equation is a matrix equation and, therefore, the possible multiplication of the conjugate spinor ψ^{*} with some matrix T should be admitted, where T is a normal unitary matrix:

$$12.16 = T\psi^{*}(x) \ . \tag{12.17}$$

Now, to find what is actually the matrix T should be, the Dirac equation should be taken and, as before, we assume that in the transformed Lawrence frame by means of

time reversal, the Dirac equation should look like in the original frame:

$$\left(i\gamma^{\mu}\frac{\partial}{\partial x'^{\mu}}-\frac{mc}{\hbar}\right)\psi'\left(x'\right)=0\ . \tag{12.18}$$

In other words, the (12.18) is the usual Dirac equation, but with x replaced by x'. That's the standard way of implementing the Lawrence covariance of the Dirac equation.

First of all, the usual Dirac equation can be taken and then transformed into the (12.18) by means of complex conjugation and multiplication with the matrix T. The usual Dirac equation has the next form:

$$\left(i\gamma^{\mu}\frac{\partial}{\partial x^{\mu}}-\frac{mc}{\hbar}\right)\psi\left(x\right)=0\ , \tag{12.19}$$

Then the (12.19) should be conjugated:

$$\left(-i\gamma^{*0}\frac{\partial}{\partial x^{0}}-i\gamma^{*i}\frac{\partial}{\partial x^{i}}-\frac{mc}{\hbar}\right)\psi^{*}\left(x\right)=0\ . \tag{12.20}$$

The (12.20) can be simplified and the result will be equal to

$$\left(i\gamma^{*0}\frac{\partial}{\partial x'^{0}}-i\gamma^{*i}\frac{\partial}{\partial x'^{i}}-\frac{mc}{\hbar}\right)\psi^{*}\left(x\right)=0\ . \tag{12.21}$$

The final thing which can be done is a multiplication of the equation from the left by T. Then the equation will have the following form:

$$\left(iT\gamma^{*0}T^{-1}\frac{\partial}{\partial x'^{0}}-iT\gamma^{*i}\frac{\partial}{\partial x'^{i}}-\frac{mc}{\hbar}\right)T\psi^{*}\left(x\right)=0\ . \tag{12.22}$$

Now, expression (12.22) should be compared with the (12.18). From this comparison, it is clearly seen that

$$\begin{cases}\psi'\left(x'\right)=T\psi^{*}\left(x\right)\\ T\gamma^{*0}T^{-1}=\gamma^{0}\\ T\gamma^{*i}T^{-1}=\gamma^{i}\end{cases} \tag{12.23}$$

To solve relations for a matrix T the representation for γ^0 and γ^i should be specified, because we need to explicitly realize how the conjugation acts on this matrices.

For instance, in the Dirac representation for γ matrices

$$g^{*0}=\gamma^{0} \tag{12.24}$$

and

$$\gamma^{*2}=-\gamma^{2}\ . \tag{12.25}$$

That's because γ^2 is constructed with the help of the Pauli matrix σ^2 and this is a diagonal matrix containing i and $-i$.

And matrices γ^1 and γ^3 are real, because they are constructed with the help of Pauli matrices σ^1 and σ^3, which are real matrices. Therefore, eventually for this matrices the following is right:

$$\begin{cases} \gamma^{*1} = \gamma^1 \\ \gamma^{*3} = \gamma^3 \end{cases} \tag{12.26}$$

Finally, for expressions in the (12.23) can be rewritten as follows:

$$\begin{cases} T\gamma^0 T^{-1} = \gamma^0 \\ T\gamma^2 T^{-1} = \gamma^2 \\ T\gamma^1 T^{-1} = -\gamma^1 \\ T\gamma^3 T^{-1} = -\gamma^3 \end{cases} \tag{12.27}$$

It can be easily seen that the solution for T is equal to:

$$T = \gamma^1\gamma^3 \ , \tag{12.28}$$

because γ^0 anti-commutes with γ^1 and anti-commutes with γ^3. Therefore, when $\gamma^1\gamma^3$ will be passed through γ^0 one will get γ^0. In other words, matrix T must commute with γ^0 and γ^2, but anti-commute with γ^1 and γ^3.

In principle T is defined up to an arbitrary phase and very often people just fix this phase, take a convenient choice and usually they pick up T to be equal to

$$T = i\gamma^1\gamma^3 \ . \tag{12.29}$$

This means that the action of the time reversal operation on the spinor is built up and it looks like the T operation apply to $\psi(x)$:

$$\psi'(x') = \mathfrak{T}\psi(x) = i\gamma^1\gamma^3\psi^*(x) \ . \tag{12.30}$$

So, this operation is known and in the presented case it is called as a Wigner time reversal operation and it is represented by anti-linear anti-unitary operator, which can explicitly realized as

$$\mathfrak{T} = T \cdot K \ , \tag{12.31}$$

where K is an operator of complex conjugation.

The same result about the form of time reversal operator can be alternatively derived just by requiring that Dirac Hamiltonian is invariant with respect to this operation of time reversal.

Now, an another important class of spinors can be introduced and the question about reducibility of the proper orthochronous Lawrence transformations can be clarified. And the subject, which will be discussed is the subject about Weyl spinners.

Weyl spinoes

In the last two lectures a representation of the proper orthochronous Lawrence group was built on 4 component spinors. This representation is realized in terms of 4x4 matrices and this matrices were denoted by letter $S(R)$. They depend on the Lawrence transformation, which can be performed on the space-time points or vectors. And it was shown that S is a spinor representation of the Lawrence group. Very interestingly that it turns out that the representation, which was constructed is reducible.

There are two definitions of what can be called a reducible representation and this definitions are equivalent. The definition number one says that representation of a Lie group, which is Lawrence group is irreducible if there are no proper invariant subspaces, where proper means subspaces, which are different from zero and the space representation space itself. If zero subspace is excluded, then the definition of irreducibility means that the representation is irreducible if there are no other invariant subspaces exist.

In other words, if start from any vector is performed in a representation space, then by group transformations any other vector in a space will be reached and it does not happen such a situation that there will be a subspace in a representation space, whose elements transform via themselves and never get out of this subspace.

The second definition is telling that representations of a Lie group is irreducible if only operators, which commute with all group elements are only those, which are proportional to the identity operator $(c \cdot \mathbb{1})$. In other words, these operators, which commute with any group elements are operators of the form constant multiplied with identity operator and no any other non-trivial operators except this one exists.

If the action of the group element in the representation space is exist and depend on the group element g, it's like a matrix $\pi(g)$ acting on the concrete representation space. Then, irreducibility means that there is non-trivial operator, like O such that it commutes

with the $\pi(g)$ for any group element g:

$$\pi(g)\, O = O\pi(g) \ , \tag{12.32}$$

for $\forall g \in G$ where

$$O \neq c \cdot \mathbb{1} \ , \tag{12.33}$$

because $c \cdot \mathbb{1}$ always commute with any $\pi(g)$, but if this case is excluded, then there any such O should not exist.

In fact, in representation theory, a theorem that the first and the second definitions are equivalent exist.

If there is an invariant subspace, then there is also a projector on this invariant subspace and an nontrivial projection operator, which projects on this invariant subspace, can be constructed. But then, because it is invariant, it can be seen that this projection operator, which does appear, is an operator which breaks the irreducibility from the point of view of the second definition.

So, it is then clear how the proof should go. One should rely on the fact that actually the projection of the representation space can be built and it gives a nontrivial operator which breaks irreducibility.

If we look from the point of view of the second definition on the representation, which was constructed, and ask ourselves why a representation that was constructed by $S(R)$ acting on spinor representation of the proper orthochronous Lorentz group is reducible.

It will be reducible if the nontrivial operator, which does commute with all group elements or, in other words, with all $S(R)$ representing elements of the proper orthochronous Lorentz group can be found. The statement is that such an nontrivial operator does exist and it's very easy to see that this operator is simply equal to

$$\gamma^5 \cdot S(R) \ , \tag{12.34}$$

because

$$\left[\gamma^5, S(R)\right] = 0 \ . \tag{12.35}$$

The reason for that is that $S(R)$ realized explicitly as

$$S(R) = \exp\left(\frac{1}{4}\gamma^\mu\gamma^\nu\omega_{\mu\nu}\right) \ , \tag{12.36}$$

where γ^μ is Dirac matrices and an index μ is from 0 to 3. Then, γ^5 is a matrix, which has a property that the anticommutator of γ^5 and γ^μ is equal to

$$\{\gamma^5, \gamma^\mu\} = 0 \ . \tag{12.37}$$

But if γ^5 anticommutes with γ^μ, then, in fact, it commutes with the product of any two γ matrices. More explicitly, it can be written that

$$\gamma_5 S\left(R\right) = \gamma_5 e^{\gamma^\mu \gamma^\nu \omega_{\mu\nu}/4} \gamma_5 \gamma_5 \ , \tag{12.38}$$

where the fact that $\left(\gamma_5\right)^2 = \mathbb{1}$ was used. Then the (12.38) can be rewritten as

$$\gamma_5 S\left(R\right) = e^{\gamma_5 \gamma^\mu \gamma^\nu \gamma_5 \omega_{\mu\nu}/4} \gamma_5 \ . \tag{12.39}$$

γ_5 can be moved through $\gamma^\mu \gamma^\nu$ and γ_5 matrices will cancel each other and, as a result, the following expression will be gotten:

$$\gamma_5 S\left(R\right) = S\left(R\right) \gamma_5 \ . \tag{12.40}$$

According to the logic of the definition of irreducibility the representation of the proper orthochronous Lorentz group is reducible and the matrix γ^5 helps use them to construct invariant subspaces, which remain invariant under proper orthochronous Lorentz transformations. This is done by means of constructing two projectors $P_\pm$, which are given by

$$P_\pm = \frac{1}{2}\left(\mathbb{1} \pm \gamma^5\right) \ . \tag{12.41}$$

$P_\pm$ have properties of being a projector. This means that

$$P_\pm^2 = P_\pm \tag{12.42}$$

and that $P_\pm$ are orthogonal:

$$P_+ P_- = P_- P_+ = 0 \ . \tag{12.43}$$

The third property is that these projectors realize the decomposition of identity:

$$P_+ + P_- = \mathbb{1} \ . \tag{12.44}$$

Other properties follow immediately from the properties of γ^5. For instance,

$$P_+ P_+ = \frac{1}{4}\left(\mathbb{1} + \gamma^5\right)\left(\mathbb{1} + \gamma^5\right) = \frac{1}{4}\left(1 + 2\gamma^5 + 1\right) = \frac{1}{2}\left(1 + \gamma^5\right) = P_+ \ . \tag{12.45}$$

At the same time, $P_+ P_-$ is equal to:

$$P_+ P_- = \frac{1}{4}\left(\mathbb{1} + \gamma^5\right)\left(\mathbb{1} - \gamma^5\right) = \frac{1}{4}\left(\mathbb{1} - \underbrace{\left(\gamma^5\right)^2}_{=1}\right) = 0 \ . \tag{12.46}$$

Projectors provide a decomposition of the identity and this means also that $P_{\pm}$ commute with all proper orthochronous Lorentz transformations. This means that objects, which realize the invariant subspaces, can be defined and these objects are special spinors, which are called Weyl spinors.

A Weyl spinor is defined is defined as a two component complex spinor, which transforms irreducibly under proper orthochronous Lorentz transformations, where this transformations form a group with the name $SO^{+}(1,3)$. Complex means that two components of the Weyl spinor are complex numbers. In other words, a four component Dirac spinor realizes the reducible representation of the proper orthochronous Lorentz group, but it can be split into two component spinors and each of these two component spinors will realize irreducible representation of a proper orthochronous Lorentz group.

Explicitly, these two component spinors are constructed from a four component Dirac spinor by applying the projection $P_{+}\ \psi$ and the projection P_{-} to ψ:

$$\begin{cases} \psi_{+} = P_{+}\psi \\ \psi_{-} = P_{-}\psi \end{cases} \tag{12.47}$$

where ψ is a four component Dirac spinor and $\psi_{\pm}$ are two component Weyl spinors.

In fact, Weyl spinors $\psi_{\pm}$ are called chiral and people say that $\psi_{\pm}$ have different chirality.

To understand more how two component Weyl spinors look like and why they are two component a transformation from the Dirac representation of the Clifford algebra, which explicitly is given by γ matrices of the next form

$$\gamma^{0} = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & \mathbb{1} \end{pmatrix}\ ,\ \gamma^{i} = \begin{pmatrix} \mathbb{0} & \sigma^{i} \\ -\sigma^{i} & \mathbb{0} \end{pmatrix}\ , \tag{12.48}$$

to another representation, which is called as Weyl representation, can be done, where in the Weyl representation matrices are realized in the following way:

$$\gamma^{0}_{c} = \begin{pmatrix} \mathbb{0} & \mathbb{1} \\ \mathbb{1} & \mathbb{0} \end{pmatrix}\ ,\ \gamma^{i}_{c} = \begin{pmatrix} \mathbb{0} & \sigma^{i} \\ -\sigma^{i} & \mathbb{0} \end{pmatrix}\ , \tag{12.49}$$

where γ^{0}_{c} is taken to be off-diagonal and γ^{i}_{c} remains to be the same as γ^{i}. Lower index c means chiral.

It can be seen that γ^{0}_{c} is nothing else as γ^{5} in the Dirac representation and γ^{5}_{c}, which can be construct through chiral matrices in the chiral representation is equal to

$$\gamma^{5}_{c} = -\gamma^{0}\ . \tag{12.50}$$

As it was pointed earlier all representations of the Clifford algebra in 4 dimensions are related to each other by means of unitary transformation and, therefore, there should be a unitary transformation, which relates these two explicit realizations of the Clifford algebra. Indeed, it can be seen that there is a matrix U, which does this job. When it is applied to all matrices γ of the Dirac representation, matrices of the Weyl representation will be gotten:

$$U\gamma^{\mu}U^{-1}=\gamma_{c}^{\mu}\ . \tag{12.51}$$

Explicitly, this matrix U is given by

$$U=\frac{1}{\sqrt{2}}\begin{pmatrix}\mathbb{1} & -\mathbb{1}\\ \mathbb{1} & \mathbb{1}\end{pmatrix} \tag{12.52}$$

and it has the following property:

$$U^{+}U=\mathbb{1}\ . \tag{12.53}$$

What is good about new γ matrices is that if spinor is written

$$\psi=\begin{pmatrix}\phi\\ \chi\end{pmatrix}\ , \tag{12.54}$$

where each of this ϕ and χ is two component, then the following can be seen.

In the Dirac representation the matrix S has the form

$$S\left(\vec{n}\,,\theta\right)=\begin{pmatrix}\mathbb{1}\cos\frac{\theta}{2}+i\vec{\sigma}\,\vec{n}\,\sin\frac{\theta}{2} & \mathbb{0}\\ \mathbb{0} & \mathbb{1}\cos\frac{\theta}{2}+i\vec{\sigma}\,\vec{n}\,\sin\frac{\theta}{2}\end{pmatrix} \tag{12.55}$$

and shows that under rotations ϕ and χ transforms independently. This matrix was constructed from the general formula:

$$S\left(\vec{n}\,,\theta\right)=e^{\frac{1}{4}\gamma^{i}\gamma^{j}\omega_{ij}}\ , \tag{12.56}$$

where the indices of γ matrices are 1, 2 and 3. Therefore, the construction (12.55) does not involve γ^0, but since passing from Dirac basis to Weyl basis of γ matrices there is no change of γ^i and then this result for $S\left(\vec{n}\,,\theta\right)$ for Lorentz transformations corresponding to rotations will hold also in the Weyl representation and the result will be the same.

For Lorentz boosts in the case of the Dirac representation ϕ and χ mix, but it is interesting to see what happens in the new basis, because here the matrix γ^0 in the Weyl basis is off-diagonal and Lorentz transformations corresponding to Lorentz boosts have the form:

$$e^{\gamma^{0}\gamma^{i}\omega_{0i}/2}\ . \tag{12.57}$$

Now, the following can be seen: γ^0 is off-diagonal and $gamma^i$ are off-diagonal and if two such off-diagonal matrices is multiplied, a diagonal matrix will be gotten. If it is diagonal, then it can be seen that the matrix corresponding to Lorentz boosts will be diagonal and it can be computed explicitly:

$$S\left(\vec{n}\,,\vartheta\right)=\begin{pmatrix}\mathbb{1}\cosh\frac{\vartheta}{2}-\vec{s}\,\vec{n}\,\sinh\frac{\vartheta}{2} & \mathbb{0}\\ \mathbb{0} & \mathbb{1}\cosh\frac{\vartheta}{2}+\vec{s}\,\vec{n}\,\sinh\frac{\vartheta}{2}\end{pmatrix}\,. \tag{12.58}$$

The meaning of the (12.58) and (12.55) is that ϕ and χ do not mixed under any proper Lorentz transformations neither under rotations nor under Lorentz boosts, because the corresponding matrices in the Weyl representation are block diagonal. That's, of course, shows that, indeed, ϕ and χ are Weyl spinors, they transform independently under proper orthochronous Lorentz transformations.

The other important remark is the following. It can be seen that the difference between ϕ and χ is also visible from formulas (12.58) and (12.55). Under rotations ϕ and χ transform completely in the same manner, but for Lorentz boosts ϕ and χ looks different and the difference is in the sign of the elements of $S\left(\vec{n}\,,\vartheta\right)$.

Once again, a Lorentz group topologically looks like some set set of transformations. Within this set there is a smaller subset, which has a name of proper orthochronous Lorentz transformations and this subset has a name $SO^+\left(1,3\right)$. Now we established a transformations of the spinor ψ with respect to $SO^+\left(1,3\right)$ and when it was realized that this representation of $SO^+\left(1,3\right)$ on 4 dimensional spinors turns out to be reducible and it splits into two invariant subspaces.

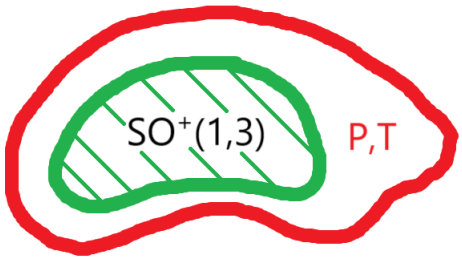

Fig. 12.1. Topological representation of the Lorentz group

The following question can be asked: if components of ψ transform with different matrices realizing Lorenz boosts, then representations can be still equivalent. Two representations are called equivalent if there exists a unitary transformation such as U and matrices of the first representation $\pi\left(g\right)$ transformed with this unitary transformation have the following

form:

$$U\pi\left(g\right)U^{-1}=\rho\left(g\right)\ . \tag{12.59}$$

In this case equivalence is just a statement that one representation is gotten from the other by changing the basis with the help of the unitary transformation. In fact, in representation theory people agree to consider unitary representations of Lie groups up to unitary equivalence.

The question can be formulated as follows: in spite of the fact that in the (12.58) there are different signs, representations may be still equivalent and changing the basis in the representation space say of spinor χ will lead to a representation of spinor ϕ and in this case these two representations will be equivalent. The answer to this question is negative and it can be proofed in the following way. Let's suppose that they are equivalent. It means that there exists a matrix U such that it takes a matrix of the first representation

$$U\left(\mathbb{1}\cosh\frac{\vartheta}{2}-\sigma^{i}n^{i}\sinh\frac{\vartheta}{2}\right)U^{-1} \tag{12.60}$$

and takes out of it the matrix of the second representation

$$11.61=\mathbb{1}\cosh\frac{\vartheta}{2}+\sigma^{i}n^{i}\sinh\frac{\vartheta}{2}\ . \tag{12.61}$$

The (12.61) be valid for any ϑ and any n. If this should be valid for any ϑ and any n, then n can be picked up, for instance, in the form $(1,0,0)$ or $(0,1,0)$ or $(0,0,1)$. Then three equations will be gotten

$$U\left(\mathbb{1}\cosh\frac{\vartheta}{2}-\sigma^{i}\sinh\frac{\vartheta}{2}\right)U^{-1}=\mathbb{1}\cosh\frac{\vartheta}{2}+\sigma^{i}\sinh\frac{\vartheta}{2}\ , \tag{12.62}$$

where now instead of summed up σ^i, there are individual σ^i standing. In the (12.62) U can be passed through the first term and then the gotten expression can be divided by $\sinh\frac{\vartheta}{2}$. As a result the following formula will be exist:

$$U\sigma^{i}U^{-1}=-\sigma^{i}\ , \tag{12.63}$$

which can be rewritten as follows:

$$U\sigma^{i}=-\sigma^{i}U\ . \tag{12.64}$$

The (12.64) is an equation, which should be satisfied. From this expression it can be seen that if a matrix U exists, it must be 2x2 matrix, which anti-commutes with all three Pauli matrices, but such a matrix does not exist. This means that these two representations

of the proper orthochronous group acting on spinors ϕ and χ are inequivalent. Therefore, these two representations of spin one half people agreed to denote differently. So, representation for ϕ is denoted as $\left(\frac{1}{2},0\right)$ and the representation realized by χ is called $\left(0,\frac{1}{2}\right)$. The spinors realizing this representations for ϕ and χ can be denoted as ψ_α, which is a two component spinor with index α, where α takes values 1 and 2. But to distinguish them on writing we will denote spinor of the type $\left(\frac{1}{2},0\right)$ by putting dot on α.

The two spinors ψ_α and $\psi_{\dot\alpha}$ are two Weyl spinors, which realize different representations of the Lorenz group. Earlier, when the construction of the Lorenz group was discussed, it was mentioned that, in fact, the proper orthochronous Lorenz group, which is $SO^+(1,3)$ is not simply connected and it looks like a circle topologically, but it has a simply connected cover. So it can be covered by another group, which is simply connected, and the name of this group is $SL(2,\mathcal{C})$. $SO^+(1,3)$ and $SL(2,\mathcal{C})$ have the same Lee algebra, where the same means that Lee algebras are isomorphic, but these are not isomorphic as groups and one is the universal cover of the other. In fact $SL(2,\mathcal{C})$ is a double cover, which covers twice the group $SO^+(1,3)$ and, in fact, $SL(2,\mathcal{C})$ are 2x2 complex matrices. And spinors ψ_α and $\psi_{\dot\alpha}$ realize two inequivalent two-dimensional representations of $SL(2,\mathcal{C})$. In fact, this is a kind of a general statement. So, in general irreducible representations of a proper orthochronous Lorentz are characterized by two half-integer numbers (j_1,j_2). In one case, representation $\psi_{\dot\alpha}$ is a representation, which corresponds to the choice j_1 to be $1/2$ and j_2 to be 0, while the undotted spinor is characterized by choice of 0 and $1/2$.

From the point of view of rotation group the group $SO^+(1,3)$ looks like a multi-valued representation and it's related to the fact that when the rotation by 2π is performed a spinor is not returning to itself, but it changes the sign. From the point of view of universal covers $SL(2,\mathcal{C})$ there is nothing strange happens and we deal with two-dimensional spinors, which are usual two-dimensional objects transforming by multiplying with the 2x2 complex matrices.

Another remark is the following. Let's return to the (fig. 12.1), which was drawn earlier. It is understood that the representation of $SO^+(1,3)$ is reducible and it can be reduced to Weyl spinors, which form irreducible representation of $SO^+(1,3)$. But what about full Lorentz group, which includes all the other elements, which can be obtained from a $SO^+(1,3)$ by applying parity and time reversal operations? For instance, under parity operation components transforms not independently and if one look at parity operation, which was discussed at the previous lecture, it will found that the parity operation is

realized by means of

$$\mathcal{P}\psi = \eta_{\mathcal{P}}\gamma^0\psi \ . \tag{12.65}$$

In the Dirac representation γ^0 was diagonal and thanks to that (12.65) can be written as

$$\mathcal{P}\psi = \eta_{\mathcal{P}} \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & -\mathbb{1} \end{pmatrix} \psi \ . \tag{12.66}$$

If ψ is written in the matrix form, than the expression for the (12.66) will be simplified to:

$$\mathcal{P}\psi = \begin{pmatrix} \eta_{\mathcal{P}}\phi \\ -\eta_{\mathcal{P}}\chi \end{pmatrix} \ . \tag{12.67}$$

Therefore, in the Dirac representation components of ψ do not mix. However, when pass to the Weyl representation will be completed, parity will be still given by the same formula, but γ^0_c is off-diagonal and

$$\mathcal{P}\psi = \eta_{\mathcal{P}} \begin{pmatrix} \mathbb{0} & \mathbb{1} \\ \mathbb{1} & \mathbb{0} \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} \ . \tag{12.68}$$

Therefore,

$$\mathcal{P} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = \begin{pmatrix} \eta_{\mathcal{P}}\chi \\ \eta_{\mathcal{P}}\phi \end{pmatrix} \ . \tag{12.69}$$

According to the (12.69) to describe the full Lorentz group both spinors are needed, because with respect to the full Lorentz group the parity operation should be realized.

If the Dirac equation is observed from the point of view of ϕ and χ, it will be seen the following. The Dirac equation in the Weyl representation will take the following form

$$\begin{cases} \frac{1}{c}\frac{\partial\phi}{\partial t} - \sigma^i\frac{\partial\phi}{x^i} + i\frac{mc}{\hbar}\chi = 0 \\ \frac{1}{c}\frac{\partial\chi}{\partial t} + \sigma^i\frac{\partial\chi}{x^i} + i\frac{mc}{\hbar}\phi = 0 \end{cases} \tag{12.70}$$

In fact equations (12.70) are coupled and equation for ϕ contains χ and equation for χ contains ϕ. But there is one case, when they actually decoupled from each other and become independent and this is the case, when mass of a particle is equal to 0, because when mass is 0 the last term disappears. The equation becomes independent and this equation for massless particle are called Weyl equations

$$\begin{cases} \frac{1}{c}\frac{\partial\phi}{\partial t} - \sigma^i\frac{\partial\phi}{x^i} = 0 \\ \frac{1}{c}\frac{\partial\chi}{\partial t} + \sigma^i\frac{\partial\chi}{x^i} = 0 \end{cases} \tag{12.71}$$

Weyl equation is covariant with respect to the proper orthochronous Lorentz transformation, but it's is not invariant under parity.

The time reversal operation can be also included. So, time reversal operation in the Wiley representation can be written as

$$\mathcal{T}\psi\left(x\right)=T\psi^{*}=i\gamma_{c}^{1}\gamma_{c}^{3}\psi^{*}=i\gamma^{1}\gamma^{3}\psi^{*}\ , \tag{12.72}$$

where γ^1 and γ^3 are the same in Weyl and in Dirac representations. They are both off-diagonal and their product is diagonal. So, under time reversal nothing bad happens and spinors ϕ and χ continue to transform independently.

Lecture 13. Solution of the Dirac Equation

In the last lecture the concept of Weyl spinors, the concept of irreducibility were discussed and it was found that representation of the proper orthochronous Lorentz group on four-dimensional spinors is reducible and it is decomposable into two irreducible representations, which are realized by means of two-dimensional Weyl spinors, which are called as dotted and undotted spinors. Also it was found that under parity transformation one species of Weyl spinors, say, dotted spinors go to undotted ones and vice versa. From this point of view these four-component spinors realize irreducible representation of the full Lawrence group, which includes in addition to proper orthochronous Lorentz transformations also parity and time reversal. Especially, parity acts non-trivial by transforming into each other Weyl spinors.

Let's write the Dirac equation again. This can be written as

$$\left(i\gamma^{\mu}\partial_{\mu} - \frac{mc}{\hbar}\right)\psi = 0\ . \tag{13.1}$$

It is interesting to look for solutions of equation (13.1), which are given by plane waves or which are described by plane waves. An answer for $\psi\left(\vec{x}\,,t\right)$ given by the following formula

$$\psi\left(\vec{x}\,,t\right) = u\left(p\right)e^{-i(Et-\vec{p}\cdot\vec{x})/\hbar}\ , \tag{13.2}$$

where $u\left(\vec{p}\right)$ is an amplitude, which depends on components of the momentum p only.

So, the (13.2) is a usual quantum mechanical plane wave and this wave is called plane, because the value of the amplitude $u\left(p\right)$ depends only on momentum, but it does not depend on x and the phase of this wave is one and the same follow points on the same plane. This is a plane for which the scalar product

$$\vec{p}\cdot\vec{x}\ = const\ . \tag{13.3}$$

In fact, this plane is perpendicular to vector $\vec{p}$ and $\vec{p}$ points to the direction of propagation of this wave. In other words, if there is momentum $\vec{p}$ pointing in a certain direction, then a plane, which is orthogonal to $\vec{p}$ to every point on this plane the phase $\vec{p}$ multiplied by x is one and the same. That is why the front of this wave is just a plane. It can be seen geometrically (fig. 13.1), because every plane is described by the following equation

$$\vec{x}\ = \vec{x}_{\,0} + s\vec{a}\ + t\vec{b}\ , \tag{13.4}$$

where s and t belongs to $\mathcal{R}$.

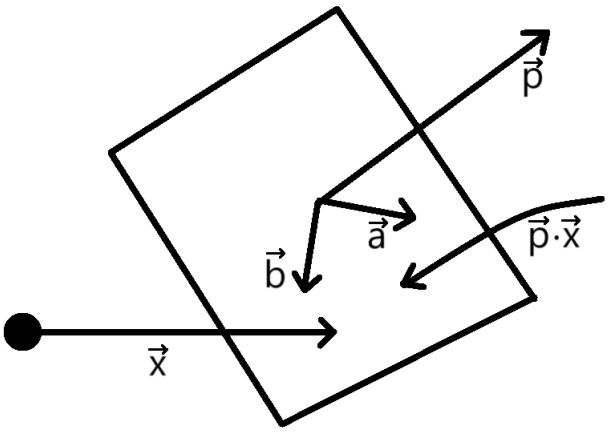

Fig. 13.1. Plane, for which (13.3) is satisfied

$\vec{a}$ and $\vec{b}$ will be two independent vectors on the plane. Then (13.4) is how any point on the plane is described. Then it can be seen that product (13.3) will be equal according to the (13.4) to

$$\vec{p} \cdot \vec{x} = \vec{p} \cdot \vec{x}_0 = const \ , \tag{13.5}$$

where $\vec{p}$ is a vector orthogonal to the plane and, therefore, it's orthogonal to vectors $\vec{a}$ and $\vec{b}$.

Now the ansatz (13.2) should be taken and pluged into the Dirac equation and in this way an equation for the amplitude $u(p)$ takes the following form:

$$\left(E\gamma^0 - c\gamma^i p^i - mc^2\right) u(p) = 0 \ . \tag{13.6}$$

Now, it should be noticed that what is written down in the (13.6) is nothing else as a eigenvalue problem for the Dirac Hamiltonian H, for which the following expression is right:

$$H\psi = \underbrace{\left(c\alpha^i p^i + mc^2\beta\right)}_{\text{Dirac Hamiltonian}} u = Eu \ , \tag{13.7}$$

where E is an eigenvalue of the (13.2) which coincides with the energy and to come from equation (13.7). And to transform the (13.6) into (13.7) the first equation should be multiplied with γ^0 from the left and γ^0 is the same as:

$$\gamma^0 = \beta \ . \tag{13.8}$$

In this way, γ^i will turn into

$$\gamma^i = \alpha^i \ . \tag{13.9}$$

And then we should get the Dirac Hamiltonian and put the energy to the other side. Then, it will be seen that what is gotten is simply eigenvalue for the energy. This is very similar to the way of solving Schrodinger equation by means of the plane wave. Plane wave should be plug into the Schrodinger equation and then a eigenvalue problem for the Schrodinger Hamiltonian will be absolutely the same.

It is convenient to take u and represent it in terms of two component spinors. So, the u is a four-component, but it can be split into ϕ and χ:

$$u = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{13.10}$$

and plug into equation (13.6) and then the expression for α and β in terms of Pauli matrices can be used. Therefore, the equation will take the following form:

$$\begin{cases} (E - mc^2)\,\phi - c\vec{\sigma}\cdot\vec{p}\cdot\chi = 0 \\ -c\vec{\sigma}\cdot\vec{p}\cdot\phi + (E + mc^2)\,\chi = 0 \end{cases} \tag{13.11}$$

The (13.11) is simply a system of 4 homogeneous equations for the components of spinors ϕ and χ. In terms of the 2x2 blocks the corresponding matrix, which acting on ϕ and χ can be written as follows:

$$\begin{pmatrix} E - mc^2 & -c\vec{s}\,\vec{p} \\ -c\vec{\sigma}\,\vec{p} & E + mc^2 \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = 0\ . \tag{13.12}$$

This is a homogeneous system of of four algebraic equations for components of this spinors ϕ and χ and it is known that in order for this system to have a non-trivial solution it is necessary that the determinant of the matrix, which defines this homogeneous system must be equal to zero. Otherwise there is no non-trivial solution of this system at all. This means that the condition to have a non-trivial solution should be imposed and determinant of the matrix from the (13.12) must be equal to zero:

$$\begin{vmatrix} E - mc^2 & -c\vec{s}\,\vec{p} \\ -c\vec{\sigma}\,\vec{p} & E + mc^2 \end{vmatrix} = 0\ . \tag{13.13}$$

The determinant (13.13) is a

$$13.13 = E^2 - \left(mc^2\right)^2 - c^2\vec{p}^{\,2} = 0\ , \tag{13.14}$$

From the (13.14), it can be seen that the condition of vanishing of determinant of the matrix is nothing else as a relativistic dispersion relation or dispersion relation for

relativistic particle, which tells that the energy of the plane wave solution is related to the momentum of the plane wave by the following condition:

$$E = \pm\sqrt{\vec{p}^{\,2}c^2 + m^2c^4}\ . \tag{13.15}$$

The relation (13.15) can be also called in the context of Klein-Gordon equation as the on-shell condition.

It is interesting that the classical Dirac theory contains solutions with both signs of energy. It has positive solutions and it has negative solutions. Therefore, from classical point of view, the Dirac theory is not well defined, because solutions with negative energy should not exist. Energy typically in usual dynamical systems is quantity which must be restricted from below. But it can be seen that solutions with any value of negative energy are possible. It will be shown that this kind of unwanted feature of the classical Dirac theory will turn into advantage of this theory, when a transition to the quantum case will be completed.

Meanwhile, the sign in front of the expression for the energy can be chosen and an explicit solution can be produced. If the positive sign is fixed, then the following solution can be gotten:

$$u_+(p) = \begin{pmatrix} \phi \\ \frac{c\vec{\sigma}\,\vec{p}}{E(\vec{p})+mc^2}\phi \end{pmatrix}\ , \tag{13.16}$$

where u_+ means the solution with positive sign of the energy. So, solution is given in terms of two components spinor ϕ, which remains to be unrestricted. Once forever $E(\vec{p})$ will be fixed to be positive:

$$E(\vec{p}) = \sqrt{\vec{p}^{\,2}c^2 + m^2c^4}\ . \tag{13.17}$$

If the negative sign is fixed in front of the square root, then the following solution will be gotten:

$$u_-(p) = \begin{pmatrix} \phi \\ \frac{c\vec{\sigma}\,\vec{p}}{-E(\vec{p})+mc^2}\phi \end{pmatrix}\ , \tag{13.18}$$

where $E(\vec{p})$ again given by the same positive solution (13.17).

Since ϕ remains arbitrary there are two solutions with positive energy, where two solutions means two components of ϕ, and two solutions with negative energy. The fact that we have two solutions hints that there exists another operator, which commutes with the Dirac Hamiltonian and which distinguishes between themselves two components of ϕ. So, the fact that ϕ remains to be arbitrary and unspecified and has two components

means that there is another operator, which commutes with the Dirac Hamiltonian and which distinguishes between these two components of ϕ and this operator indeed can be constructed in an explicit way and it has a name Helicity operator.

Helicity operator

The Helicity operator is constructed as follows:

$$\mathcal{J} = \frac{1}{|\vec{p}|}\gamma^5\gamma^0\gamma^i p^i \ . \tag{13.19}$$

If γ matrices are presented in the Dirac representation, then the following answer will be found:

$$\mathcal{J} = \frac{1}{|\vec{p}|}\begin{pmatrix} \vec{\sigma}\,\vec{p} & 0 \\ 0 & \vec{\sigma}\,\vec{p} \end{pmatrix} \ . \tag{13.20}$$

Then from the discussion of the law of matrices transformation, which realize a representation of the Lorenz group, the (13.20) can be written as:

$$\mathcal{J} = \left(\vec{\Sigma}\cdot\vec{n}\right) \ , \tag{13.21}$$

where $\vec{n}$ is a unit vector in the direction of momentum and it's equal to

$$\vec{n} = \frac{\vec{p}}{|\vec{p}|} \ . \tag{13.22}$$

Σ is understood as a 4x4 matrix on the one hand. On the other hand, $\vec{\Sigma}$ is a vector built up from three Pauli matrices

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & \mathbb{0} \\ \mathbb{0} & \vec{\sigma} \end{pmatrix} \ . \tag{13.23}$$

and the operator Σ can be called as the spin operator, because 3 matrices $\vec{\Sigma}$ satisfy the Lie algebra relations for the angular momentum and this is an operator, which can be identified with the spin of a particle.

It can be seen that the operator, which was constructed in (12.19)-(12.21) is nothing else as a projection of spin on the direction of propagation or the direction of motion.

Now, it can be checked that an operator $\mathcal{J}$ commutes with the Dirac Hamiltonian:

$$[H, \mathcal{J}] = 0 \ . \tag{13.24}$$

From the course on linear algebra it is known that if there are two commuting operators, then they can be simultaneously diagonalized and this means that these operators will

have the same basis in the Hilbert space made from common eigenstates of these two commuting operators.

Properties of this operator are the following. The square of the Helicity operator is equal to

$$\mathcal{J}^2 = \mathbb{1} \ , \tag{13.25}$$

which means that the eigenvalues of this operator are simply ± 1. Thus a state with positive energy can also be an eigenstate of the helicity operator with either positive or negative helicity and the same, of course, true for states with negative energy. If we now take the general solution for states with positive energy and the same general solution for states with negative energy and further use these states to diagonalize the Helicity operator on these states, which are already made to be eigenstates of the Hamiltonian, then the following explicit solution for this the amplitude u will be found, which is simultaneously an eigenstate of both operators of H and of $\mathcal{J}$. Somehow to distinguish states with positive and negative Helicity indices 1 and 2 will be used. For instance, for $u_+(p)$ a solution with positive energy and positive Helicity can be written in the following way:

$$u^1_+(\vec{p}) = \eta(p) \begin{pmatrix} p^3 + |\vec{p}| \\ p^1 + ip^2 \\ \frac{c|\vec{p}|\left(p^3+|\vec{p}|\right)}{mc^2+E(\vec{p})} \\ \frac{c|\vec{p}|\left(p^1+ip^2\right)}{mc^2+E(\vec{p})} \end{pmatrix} \ , \tag{13.26}$$

where $\eta(p)$ is a normalization, which can be used for normalization of the spinor in a convenient way. This is how the solution with positive Helicity and the positive energy looks like.

Then, it will also be an eigenstate of the Helicity operator with negative sign and the solution for positive energy and this eigenstate will look as follows:

$$u^2_+(p) = \eta(p) \begin{pmatrix} -\left(p^1 - ip^2\right) \\ p^3 + |\vec{p}| \\ \frac{c|\vec{p}|\left(p^1-ip^2\right)}{mc^2+E(\vec{p})} \\ -\frac{c|\vec{p}|\left(p^3+|\vec{p}|\right)}{mc^2+E(\vec{p})} \end{pmatrix} \ . \tag{13.27}$$

Analogously, solutions with negative energy and different helicities can be written as

$$u_{-}^{1}\left(\vec{p}\right)=\eta\left(p\right)\begin{pmatrix}-\frac{c|\vec{p}|\left(p^{3}+|\vec{p}|\right)}{mc^{2}+E(\vec{p})}\\-\frac{c|\vec{p}|\left(p^{1}+ip^{2}\right)}{mc^{2}+E(\vec{p})}\\p^{3}+|\vec{p}|\\p^{1}+ip^{2}\end{pmatrix} \tag{13.28}$$

and

$$u_{-}^{2}\left(p\right)=\eta\left(p\right)\begin{pmatrix}-\frac{c|\vec{p}|\left(p^{1}-ip^{2}\right)}{mc^{2}+E(\vec{p})}\\\frac{c|\vec{p}|\left(p^{3}+|\vec{p}|\right)}{mc^{2}+E(\vec{p})}\\-\left(p^{1}-ip^{2}\right)\\p^{3}+|\vec{p}|\end{pmatrix}\ . \tag{13.29}$$

Overall, normalization can be chosen in a convenient way to provide a convenient orthogonality relations for the four different solutions presented above. It turns out then if the normalization is chosen in the following way:

$$\eta\left(p\right)=\frac{1}{2}\sqrt{\frac{mc^{2}+E\left(\vec{p}\right)}{mc^{2}\left(\vec{p}^{\,2}+p^{3}\left|\vec{p}\right|\right)}}\ , \tag{13.30}$$

then solutions are conveniently normalized in the following way:

$$\overline{u}_{+}^{r}\left(\vec{p}\right)u_{+}^{s}\left(\vec{p}\right)=\delta^{rs}\ , \tag{13.31}$$

where $\bar{u}$ is Dirac conjugate spinor. In other words, the scalar product of Dirac conjugate spinor with positive energy spinor of different helicity gives 0. For negative energy solutions with different helicity the story is similar with a little exception that on the right hand side it will be -1:

$$\overline{u}_{-}^{r}\left(\vec{p}\right)u_{-}^{s}\left(\vec{p}\right)=-\delta^{rs}\ . \tag{13.32}$$

In fact, these relations also can be found by using the projector operations or projectors on solutions with a definite helicity simply by using the fact (13.25). So helicity plus and minus projection operators can be introduced, which are simply equal to

$$\mathcal{J}_{\pm}=\frac{\mathbb{1}\pm\mathcal{J}}{2}\ . \tag{13.33}$$

Because of the fact that $\mathcal{J}$ squared is equal to 1, $\mathcal{J}_{\pm}$ have all the properties of projectors and, in particular, it can be also checked that they satisfy the following properties. If $\mathcal{J}_{\pm}$ is an operator, which depends on momentum, then it can be seen that

$$\mathcal{J}_{\pm}\left(\vec{p}\right)=\mathcal{J}_{\mp}\left(-\vec{p}\right)\ . \tag{13.34}$$

This is one of the properties that can be verified explicitly from the form of the operator. In fact, solutions that was found for an amplitude u look pretty complicated, because there is a certain amount of symmetry broken, because of the fact that components of momentum entering differently in different elements of the corresponding spinors. But, in fact, these formulas immensely simplify if we go to the rest frame of a massive particle. So, if particle is massive, a transition to the rest frame can be done, where

$$\vec{p} = 0 \ . \tag{13.35}$$

Then it can be checked that components with positive and negative energy and definite helicity are simply turn into

$$u^1_+ = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \ , \ u^2_+ = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \ , \ u^1_- = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \ , \ u^2_- = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \ , \tag{13.36}$$

where matrices are turned to the standard basis in the space $\mathcal{C}^4$.

In fact, this simple expression for components of the spinors can be used to perform a Lorentz transformation in the arbitrary Lorentz frame. And then the result, which already was found, will be obtained. The only point is that solutions, which will be obtained in this way do not have a definite helicity and an extra rotation of the spinors should be performed to diagonalize the helicity operator by taking linear combinations of two positive energy solutions and then separately two negative energy solutions.

If solutions with positive energy are considered, then solutions for positive energy, but for different helicities are orthogonal. The same is true for solutions with negative energies.

It is interesting to discuss what happens if now solutions with positive and negative energy are taken and a scalar product is needed to be evaluated. There is a little problem, because helicity operator commutes with the matrix γ^0 that can be understood from the explicit form of this operator.

$$\left[\mathcal{J}, \gamma^0\right] = 0 \tag{13.37}$$

Indeed helicity operator does commute with γ^0 and γ^0 is important, because it is used to define the Dirac conjugate spinor, which can be recalled as

$$\overline{u}\left(\vec{p}\right) = u^+\left(\vec{p}\right)\gamma^0 \ . \tag{13.38}$$

It should be noticed that the Dirac conjugate spinors, but not Hermitian conjugate spinors, are interested, because Dirac conjugate spinors transform under Lorentz transformation with the inverse matrix of Lorentz transformation, in comparison to the matrix of under which the spinor u transforms. That is useful, because then the Lorentz covariant combinations can be easily built up. For instance,

$$\overline{u}\left(p\right)u\left(p\right) \tag{13.39}$$

is a scalar,

$$\overline{u}\left(p\right)\gamma^{\mu}u\left(p\right) \tag{13.40}$$

is a vector and so on. So, that is very useful for constructing objects with well-defined transformation properties under Lorentz transformations.

With respect to Helicity everything is fine, Helicity goes through γ^0 and act on the u itself, because Helicity commutes with γ^0, but this is not the case for the Hamiltonian H, for which such a nice property does not exist. It does not commute with γ^0

$$\left[H,\gamma^{0}\right]\neq 0\ . \tag{13.41}$$

This leads to the fact that if two spinners of different energy are taken, then the scalar product will not be equal to zero. γ^0 enters into the definition of the Dirac conjugate spinner rather the exact relation between H and γ^0 is as follows:

$$\gamma^{0}H\left(\vec{p}\right)\gamma^{0}=H\left(-\vec{p}\right)\ . \tag{13.42}$$

That is clear from the definition of H:

$$H\left(\vec{p}\right)=c\alpha^{i}p^{i}+mc^{2}\beta\ , \tag{13.43}$$

where α^i are simply given by

$$\alpha^{i}=\gamma^{0}\gamma^{i} \tag{13.44}$$

and β is again equal to:

$$\beta=\gamma^{0}\ . \tag{13.45}$$

In fact, the property (13.42) can be used in the following way. First of all, it can be seen that if a solution with a positive energy is taken and the Hamiltonian is applied to it, then a positive eigenvalue will be gotten:

$$H\left(\vec{p}\right)u_{+}\left(p\right)=E\left(\vec{p}\right)u_{+}\left(p\right)\ . \tag{13.46}$$

Then from the (13.46) the next expression follows. Since the Hamiltonian is Hermitian, then:

$$u_+^\dagger(p)\, H(\vec{p}) = E(\vec{p})\, u_+^\dagger(p) \ . \tag{13.47}$$

$u_+^\dagger(p)$ contains γ^0 and now the γ^0 can be compensated by:

$$\overline{u}_+(p)\,\gamma^0 H(\vec{p})\,\gamma^0 \ . \tag{13.48}$$

Then the (13.48) is simply the following thing:

$$13.48 = E(\vec{p})\,\overline{u}_+(p) \ , \tag{13.49}$$

because what is written down here is simply the same relation as (13.47), but with plug between $u_+^\dagger(p)$ and $H(\vec{p})\,\gamma^0$ squared matrix and multiplied by γ^0 from the right.

On the other hand, it is known that

$$\gamma^0 H(\vec{p})\,\gamma^0 = H(-\vec{p}) \tag{13.50}$$

and, therefore, the (13.49) can be written as:

$$\overline{u}_+(p)\, H(-\vec{p}) = E(\vec{p})\,\overline{u}_+(p) \ . \tag{13.51}$$

The relation (13.51) can be multiplied by a spinor $u_-(-p)$ from the right:

$$\overline{u}_+(p)\, H(-\vec{p})\, u_-(-p) = E(\vec{p})\,\overline{u}_+(p)\, u_-(-p) \ . \tag{13.52}$$

The (13.52) can be simplified further if the following property will be used:

$$H(-\vec{p})\, u_-(-p) = -E(\vec{p})\, u_-(-p) \ , \tag{13.53}$$

where minus sign appeared due to the fact that $-E(\vec{p})$ is eigenvalue for the negative energy solution.

According to the (13.53), the (13.52) can be simplified and the following equation can be gotten:

$$-E(\vec{p})\,\overline{u}_+(p)\, u_-(-p) = E(\vec{p})\,\overline{u}_+(p)\, u_-(-p) \ . \tag{13.54}$$

From the (13.54) it can be concluded that

$$\overline{u}_+(p)\, u_-(-p) = 0 \ . \tag{13.55}$$

This means that solutions $\overline{u}_+(p)$ and $u_-(-p)$ carry a different sign of energy: one energy is positive, another energy is negative. Moreover they are orthogonal to each other and

this orthogonality is proved by using the consideration that is done. This is what should be taken as an orthogonality condition of positive-negative solution, because if the pairing of solutions with the same momentum is done, then it will be found that it is not-zero for the reason, which was mentioned that γ^0 does not commute with the Hamiltonian. Although, if the sign of solution with negative energy is changed, then the result is zero. This consideration suggests that it is convenient to introduce a spinor v^s_-, which should be defined or as

$$v^s_- (\vec{p}) = u^s_- (-\vec{p}) \ . \tag{13.56}$$

Then the full set of the orthogonality conditions, which can be deduced from relations above is the following:

$$\overline{u}^r_+ (\vec{p}) \, u^s_+ (\vec{p}) = \delta^{rs} \ . \tag{13.57}$$

Then it also can be checked that with minuses or with v-s, which correspond to negative energy solutions, these solutions are orthogonal for different helicities:

$$\overline{v}^r_- (\vec{p}) \, v^s_- (\vec{p}) = -\delta^{rs} \ . \tag{13.58}$$

Finally, the orthogonality condition for different energy solutions is

$$\overline{v}^r_- (\vec{p}) \, u^s_+ (\vec{p}) = 0 \tag{13.59}$$

for any r and s. Analogously,

$$\overline{u}^r_+ (\vec{p}) \, v^s_- (\vec{p}) = 0 \ . \tag{13.60}$$

For helicity two projections, which project on states with positive and negative helicity, were introduced, where in fact the same also can be done for projections on solutions with positive and negative energy. This is convenient if an arbitrary spinor exist, it can be decomposed into solutions of positive and negative energy by using the projectors. One projector projects on solution with positive energy and the other will project on solutions with negative energy and these projections are explicitly constructed as follows:

$$\begin{cases} \Lambda_+ = \frac{mc \cdot \mathbb{1} + \not{p}}{2mc} \\ \Lambda_- = \frac{mc \cdot \mathbb{1} - \not{p}}{2mc} \end{cases} \tag{13.61}$$

This is how projectors on positive and negative energy are constructed and if $\not{p}$ is written down explicitly, then the (13.61) will look as follows:

$$\begin{cases} \Lambda_+ = \frac{\gamma^0 E(\vec{p}) - c\gamma^i p^i + mc^2 \cdot \mathbb{1}}{2mc} \\ \Lambda_- = -\frac{\gamma^0 E(\vec{p}) - c\gamma^i p^i - mc^2 \cdot \mathbb{1}}{2mc} \end{cases} \tag{13.62}$$

Projectors Λ_+ and Λ_- are projectors, which means that the following relations are satisfied:

$$\begin{cases} \Lambda_+^2 = \Lambda_+ \\ \Lambda_-^2 = \Lambda_- \end{cases} \tag{13.63}$$

Then, these projectors are orthogonal, which means that

$$\Lambda_+ \Lambda_- = 0 \ . \tag{13.64}$$

Finally, it can be seen that

$$\Lambda_+ + \Lambda_- = \mathbb{1} \ / \tag{13.65}$$

It cab be also checked that Λ_+ applied to spinors will give the following expression

$$\Lambda_+ u_+^{1,2} = u_+^{1,2} \ . \tag{13.66}$$

Analogously, for L_- the following expression can be gotten:

$$\Lambda_- v_-^{1,2} = u_-^{1,2} \ . \tag{13.67}$$

So, u_+ and v_- form a basis in the full space of solutions and these are separately eigenstates in the subspaces, which are defined by means of projectors Λ_+ and Λ_-. There are also important formulas, which people often use, namely, Λ_+ and Λ_- can be expressed in the components of spinors u_+ and v_-.

In particular, Λ_+ is a 4x4 matrix and its elements i and j can be written in the following way:

$$(\Lambda_+)_{ij} = \sum_{r=1}^{2} \left(u_+^r\right)_i \left(\overline{u}_+^r\right)_j = \sum_{r=1}^{2} \left(u_+^r\right)_i \left(u_+^{r\,+}\gamma^0\right)_j \ . \tag{13.68}$$

Also the similar expression can be written for Λ_-:

$$(\Lambda_-)_{ij} = -\sum_{r=1}^{2} \left(v_-^r\right)_i \left(\overline{v}_-^r\right)_j = -\sum_{r=1}^{2} \left(v_-^r\right)_i \left(v_-^{r\,+}\gamma^0\right)_j \ . \tag{13.69}$$

That is basically all, which allows now to write down the general solution of the Dirac equation or basically a decomposition of the Dirac spinor via the basis of plane waves. If a superposition of solutions is taken, then it will be gotten a general solution of the Dirac equation. So, plane waves actually provide the basis over which an arbitrary solution can be expanded and this means that a superposition principle can be implemented, because Dirac equation is the linear equation. If there are two solutions, then the sum of these

two solutions is another solution of this equation. If we want to construct an arbitrary solution, then a sum of all solutions should be taken, which means that an integral with a certain integration measure implied by the normalization of spinors should be taken:

$$\psi(x) = c^{\frac{1}{2}} \int \frac{\mathrm{d}\vec{p}}{(2\pi\hbar)^{3/2}} \left(\frac{mc^2}{E(\vec{p})}\right)^{1/2} \times \sum_{r=1}^{2} \left(b_r(\vec{p})\, u_+^r(\vec{p})\, e^{-ip_\mu x^\mu/\hbar} + d_r^*(\vec{p})\, v_-^r(\vec{p})\, e^{ip_\mu x^\mu/\hbar}\right) , \tag{13.70}$$

where $b_r(\vec{p})$ and $d_r^*(\vec{p})$ are amplitudes. The written above integral is a general solution of the Dirac equation, which is written as a superposition of plane wave solutions that was found. And the fact that the $u_+^r(\vec{p})$ and $v_-^r(\vec{p})$ form a basis in the space of solutions was used. Then for each momentum $\vec{p}$ the corresponding spinors are exist.

By direct calculation it can be checked that the (13.70) is a solution of the Dirac equation. All what is needed to do is just to take the Dirac operator

$$\left(i\gamma^0 \frac{\partial}{c\partial t} + i\gamma^i \frac{\partial}{\partial x^i} - \frac{mc}{\hbar}\right) \tag{13.71}$$

and act with it on $\psi(x)$ and when the Dirac operator acts on $\psi(x)$ this essentially means that it can be passed under the integral and act on the exponent, because it's an exponent, which depends on axis over which differentiation is performed. Therefore,

$$\left(i\gamma^0 \frac{\partial}{c\partial t} + i\gamma^i \frac{\partial}{\partial x^i} - \frac{mc}{\hbar}\right) e^{-i(Et-\vec{p}\vec{x})/\hbar} u_+^r = \frac{\gamma^0}{c\hbar}\left(E(\vec{p}) - H(\vec{p})\right) u_+^r e^{-i(Et-\vec{p}\vec{x})/\hbar} . \tag{13.72}$$

u_+^r is a positive energy solution of the Dirac equation with eigenvalue $E(\vec{p})$ and, therefore

$$E(\vec{p}) - H(\vec{p}) = 0 . \tag{13.73}$$

and

$$13.72 = 0 . \tag{13.74}$$

Analogously, it can be shown for the negative energy solutions that they satisfy Dirac equation.

In the quantum theory the operator $b_r(\vec{p})$ will become an operator of annihilation of electron, while $d_r^*(\vec{p})$ will turn into the operator of creation of positron. This is how amplitudes of classical theory will be reinterpreted in the quantum theory.

So, it's very similar to what was done when a general solution of the Klein-Gordon was discussed. The amplitudes a and a^* there were only a and a^*, because in that case a real scalar field was implied. If a complex scalar field is implied than there will be found two independent amplitudes a and b or a^* and b^* star. And here not just complex

scalar field was introduced, but a spinor field and as it can be seen the spinor field must have four complex components and, indeed, it has two components, which are described by b_r like independent coefficients, which are taken values 1 and 2, and the other two independent complex components are related to d_r^*, because $\psi(x)$ has arbitrary four complex components.

One more moment to discuss here is the following. For a massive particle the helicity operator does not commute with Lorentz transformations. Besides helicity and the Hamiltonian, in principle, what is also in the game are Lorentz transformations. Under arbitrary Lorentz transformations neither Hamiltonian nor helicity operators are preserved even with respect to their usual rotations. The following transformation properties are exist. So, the Dirac Hamiltonian transforms as follows. If a Lorentz transformation is taken as

$$S^{-1}(R) H(p) S(R) \ , \tag{13.75}$$

then the result of this will be

$$13.75 = H(R\vec{p}) \ . \tag{13.76}$$

This is related to the fact that H is a matrix and $S(R)$ is also a 4x4 matrix. The (13.76) means that, when a Lorentz frame is transformed by means of Lorentz transformation, in general, the Hamiltonian undergoes a transformation particular, it is the Hamiltonian evaluated at the transformed value of momentum $\vec{p}$. Similarly, if the helicity is taken, it transforms also under this Lorentz transformations in a similar way:

$$S^{-1}(R) \mathcal{J}(p) S(R) = \mathcal{J}(R\vec{p}) \ . \tag{13.77}$$

In the case of the Klein-Gordon field there was a difference. So, the Hamiltonian was invariant with respect to such transformations, because the Hamiltonian in that case depends on $\vec{p}^2$ and $\vec{p}$ is a vector, which length is invariant under rotations. So, if a vector is taken and rotated, by the definition of the rotation group the length of this vector must be preserved. The Dirac Hamiltonian is different, because it involves momentum not like a $\vec{p}^2$, but involves it as

$$\gamma^i p^i \tag{13.78}$$

and it's a big difference. Now under rotations by Lorentz transformations result is that we have Hamiltonian at rotated momentum $\vec{p}$ and that is a different Hamiltonian. In general, it can be seen the only thing, which preserves actually the Hamiltonian and also helicity are rotations, which happen in the plane orthogonal to momentum. If the momentum is

looking in one direction of particles, then if a orthogonal plane to this momentum is taken and rotated around the direction of $\vec{p}$ in this plane, then this will be invariant momentum $\vec{p}$. In this case under the restricted rotations neither Hamiltonian nor helicity will be changed (fig. 13.2).

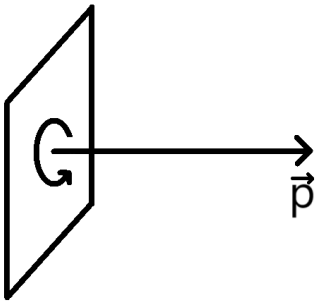

Fig. 13.2. The plane orthogonal to momentum $\vec{p}$

The situation with helicity is actually dramatic, because helicity depends on the Lorenz frame. In other words, a state with a positive helicity in a different Lorenz frame can become a state with a negative helicity. So, helicity is not a Lorenz invariant quantity for a massive particle. If particle is massive, then helicity is not a good observable, because it depends on the choice of the Lorenz frame. But for massless particle the situation is different. This is due to the massless Dirac equation, because if the Dirac equation for the massless particle is written, then there is no the term with mc^2 and the massless Dirac equation takes the following form:

$$\gamma^i p^i = \frac{E}{c}\gamma^0 \ . \tag{13.79}$$

So, in the momentum space for a plane wave this is how the massless Dirac equation looks like. For massless equation it is known that energy is given by

$$E\left(\vec{p}\right) = \pm c\left|\vec{p}\right| \ . \tag{13.80}$$

This is a formula, which can be gotten from the general formula

$$E = \sqrt{p^2c^2 + \cancel{m^2c^4}} \ , \tag{13.81}$$

if the mass turns to zero. This means that the operator of helicity, in general, is equal to

$$\mathcal{J} = \frac{1}{|\vec{p}|}\gamma^5\gamma^0\gamma^i p^i \tag{13.82}$$

and because of the relation (13.59) this can be simplified to the following form:

$$\mathcal{J} = \pm\gamma^5 \ . \tag{13.83}$$

So, for massless particle helicity operator is the same as chirality operator, which was introduced at the previous lecture. Helicity and chirality becomes the same and they become an absolute thing, which does not depend on the Lorentz frame and this is also something obvious, because in the Lorentz frame for a massless particle a rest frame is not exist. Basically a particle is moving with the speed of light and since we're a massive persons, we can never reach it.

For a massive particle if a momentum is high enough, the observer can be placed in the particle's frame and in this system the sign of momentum will change. So the helicity in this case can be changed, because the sign of $\vec{p}$ can be changed in the frame of observer if it, for instance, will go faster than a particle.

Lecture 14. Charge Conjugation and Anti-Particles

At this lecture the operation of quantum field theory, which acts on the Dirac equation, will be observed. This operation of charge conjugation, which actually allows to understand the notion of anti-particles. To do this, firstly it is needed to understand, how a Dirac particle interacts with an electromagnetic field. And this is needed in order to be able to introduce the notion of electric charge in the context of a Dirac particle, because the Dirac equation didn't involve electric charge at all. On the other hand, the properties of the electron concerning its electric charge only shown up, when an electron is put in the an external electromagnetic field. There should be electromagnetic field around, which then influences the motion of an electrically charged particle.

Indeed, there is a coupling of the particle to surrounding electromagnetic field and tracing the change of the particle trajectory. Therefore, to detect the presence of electromagnetic field around on the one hand side and on the other hand side it is needed to say that a particle possess a property, which can be called as electric charge. In other words, the ability of a particle to interact with electromagnetic field is called as electric charge.

The first step would be to introduce the coupling of the Dirac spinor or a Dirac particle to an external electromagnetic field. In this case, the four electromagnetic field has been given. And as it known from classical electrodynamics already, the electromagnetic field is defined by means of an electromagnetic potential A_μ, which can be introduced in the Dirac equation in the following way:

$$\left[\gamma^\mu \left(i\hbar\partial_\mu - \frac{e}{c}A_\mu\right) - mc\right]\psi = 0 \ , \tag{14.1}$$

where e is a charge of a Dirac particle.

In other words, the introduction of the potential amounts to a procedure of, as people sometimes say, making the usual derivative to become with long derivative, which means that the electromagnetic potential A_μ was extended. In fact, the quantity in the brackets of the (14.1) is a covariant derivative or derivative where A_μ is considered to be a connection on a certain bundle. A_μ is an electromagnetic potential given in the whole space-time. So, it's a function of a space-time point x.

In natural units electric charge has a physical dimension of

$$[e] = \left[\sqrt{\hbar c}\right] \ . \tag{14.2}$$

This is useful to have in mind that electric charges have the same dimension as $\sqrt{\hbar c}$ in terms of fundamental constants. And this fact can be deduced from the Coulomb force

between two equal charges, because, as it is known, the Coulomb force between two such charges simply equal to

$$F = \frac{e^2}{r^2} \ , \tag{14.3}$$

where r is a distance between charges and charges are equal to each other.

From the (14.3) it is clearly seen that

$$[e]^2 = \left[F \cdot l^2\right] = \left[F \cdot l \underbrace{\cdot t \cdot c}_{l}\right] \ , \tag{14.4}$$

where F multiplied by length is a work, which is the same as energy. And it is known that an energy multiplied by time is a Plank constant $\hbar$. Therefore,

$$[e]^2 = [\hbar \cdot c] \ . \tag{14.5}$$

So, electromagnetic field is not a matrix, it's just the full component vector $A_\mu(x)$ and the coupling is coming through the product $\gamma^\mu A_\mu$. Then there is a constant, which regulates the strengths of this coupling and this is what is called as an electric charge.

It should be noticed, that an introduced electromagnetic field is external, because the dynamical equations for electromagnetic field were not added, because the dynamical equations for an electromagnetic field would be Maxwell equations. And if electromagnetic field interacts with a spinor particle, then this particle would serve as a source for electromagnetic field. So, there will be nontrivial back reaction of a spinor field on an electromagnetic field, but this back reaction for the moment is completely neglected. That is because the given field is strong enough not to be influenced by the motion of a Dirac particle itself.

But on the other hand, the presence of this electromagnetic field, of course, influences the dynamics of the Dirac particle itself and that's the reason why it is said that a particle couples to the field.

The new Dirac equation, which was written in the (14.1) turns out that it has a new fundamental discrete symmetry, which is a symmetry over the theory with respect to the change of the sign of the electric charge and this is the symmetry, which is also considered as symmetry replacing the particle by its antiparticle. To understand this a name for it should be introduced. This symmetry is called charge conjugation. The generator of this symmetry can be denoted as $\mathbb{C}$, which is a charge conjugation. It is a new fundamental symmetry of the Dirac equation.

Now, existence of this symmetry implies that there is a one to one correspondence between solutions, which was denoted as ψ, of the Dirac equation with a negative energy and positive energy (a particle have the same mass). Wave functions for this cases are connected by the following expression:

$$\psi^c = \mathbb{C}\psi \ , \tag{14.6}$$

where ψ^c is a wave function of a particle with positive charge and ψ is a wave function of a particle with negative charge.

If there is a particle, which is an electron and it is described by solution with positive energy, as it should be, then there will be another solution, which is constructed from side with negative energy as ψ^c and this solution will carry opposite charge. It has to be treated as a positron. So, electrons with E less than zero should be treated as positrons with normal situation, where energy is positive. That's a fundamental thing in this existence of the theory actually, in a way, returns to the Dirac equation the status of being physical, because normally dynamical systems, which have solutions with negative energy should be neglected. Energy must always be restricted either from above or from below and in the Dirac equation this is not the case, because there are solutions with negative energy and with positive energy and they are not restricted or bounded, neither from above nor from below.

From this point of view we might say that this equation should not be interpreted as physical. On the other hand, due to the new fundamental symmetry a new interpretation for the solution with negative energy can be given. Namely, they can be treated as solutions with positive energy, but opposite sign of electric charge.

In fact, the real understanding of what's happening with this fundamental symmetry comes only upon performing the quantization of the Dirac theory.

If second quantization in this case is performed, there will be no any more solutions with negative energy at all, but there will be electrons and positrons solutions with opposite sign of electric charge, but only with positive energy.

Now, it is needed to be found how the symmetry acts on ψ. The equation for the charge conjugated ψ should be understood in the following way. The equation for ψ^c have the next form:

$$\left[\gamma^\mu \left(i\hbar\partial_\mu - \frac{e}{c}A^c_\mu\right) - mc\right]\psi^c = 0 \ , \tag{14.7}$$

where A^c_μ is a charge conjugated potential and ψ^c is exactly the wave function, which is obtained from side by application of the charge conjugation operation. The A^c_μ is simply

amounts to change the sign of the electromagnetic potential. So, charge conjugation is an operation, which acts on fields and the index. Since now two kinds of fields are exist: ψ and A_μ, and the charge conjugation is operation, which acts on all fields, which are presented in a theory, it does not touches constants.

And charge conjugation acts on the potential A_μ in a very simple way. In a sense, it can be seen that changing the sign of electromagnetic potential is exactly equivalent to changing the sign of the electric charge, because they come together:

$$A^c_\mu(x) = -A_\mu(x) \ . \tag{14.8}$$

If the (14.7) is compared with the original Dirac equation, then it has exactly the same form except that it acts on charge conjugate ψ^c not on ψ.

Once fields was transformed by symmetry transformation, the Dirac equation should be covariant and, in other words, in terms of new transformed quantities, it should look the same as in terms of the old or the original quantities.

Now, the following can be seen that the implementation of charge conjugation on the electromagnetic field amounts to the change of sign minus into plus. So, if A_μ will be replaced by A^c_μ in the (14.1), then it will lead to appearance of plus sign:

$$\left[\gamma^\mu\left(i\hbar\partial_\mu + \frac{e}{c}A^c_\mu\right) - mc\right]\psi = 0 \ , \tag{14.9}$$

but this change of sign can be easily done by the operation of conjugation, because the electromagnetic field is real.

In other words, what we should start from equation (14.1) and perform an operation of complex conjugation. In this way, what will be gotten can be then written as follows:

$$\left[\gamma^{*\mu}\left(-i\hbar\partial_\mu - \frac{e}{c}A^c_\mu\right) - mc\right]\psi^* = 0 \ . \tag{14.10}$$

Then minus in the brackets of the (14.10) can be taken out and the equation can be written in the following form:

$$\left[\gamma^{*\mu}\left(i\hbar\partial_\mu + \frac{e}{c}A^c_\mu\right) + mc\right]\psi^* = 0 \ . \tag{14.11}$$

Then, to reach a goal, it should be assumed that the operation of charge conjugation can be explicitly realized by applying a certain matrix to ψ^c, where ψ^c is a spinor, which is charge conjugation of the original spinor ψ:

$$\psi^* = \left(\mathbb{C}\gamma^0\right)^{-1}\psi^c \ , \tag{14.12}$$

where $\mathbb{C}$ is a 4x4 matrix and $\mathbb{C}\gamma^0$ is a matrix of transformation from the ψ^c to ψ^*

In other words,

$$\psi^c = \mathbb{C}\gamma^0 \cdot \psi^* \ . \tag{14.13}$$

Now, the ψ^* in the (14.11) can be replaced by ψ^c and, therefore the following expression will be gotten:

$$\left[\left(\mathbb{C}\gamma^0\right)\gamma^{*\mu}\left(\mathbb{C}\gamma^0\right)^{-1}\left(i\hbar\partial_\mu - \frac{e}{c}A^c_\mu\right) + mc\right]\psi^c = 0 \tag{14.14}$$

If we want to restore equation (14.7), then the following should be required

$$\left(\mathbb{C}\gamma^0\right)\gamma^{*\mu}\left(\mathbb{C}\gamma^0\right)^{-1} = -\gamma^\mu \ . \tag{14.15}$$

The (14.15) also can be written in the following way:

$$\left(\mathbb{C}\gamma^0\right)^{-1}\gamma^{\mu}\left(\mathbb{C}\gamma^0\right) = -\gamma^{*\mu} \ . \tag{14.16}$$

The first question, which arises here is if such a matrix $\mathbb{C}$, which has the property (14.15), exists, because this transformation formula should be valid for all μ, which can be equal to 0, 1, 2 and 3.

As can be see easily seen easily an transformation

$$\gamma^\mu \to -\gamma^{*\mu} \tag{14.17}$$

is automorphism of the Clifford algebra, which means that under this transformation a Clifford algebra is not changing, it turns to itself. But, on the other hand, it is known from the power of theorems that any of the morphism of the Clifford algebra in four dimensions is internal. In other words there should exist a unitary matrix U, such that

$$-U\gamma^{*\mu}U^{-1} = \gamma^\mu \tag{14.18}$$

So transformation from a set $\gamma^{*\mu}$ to γ^μ is done by unitary transformation, where U is equal to:

$$U = \mathbb{C}\gamma^0 \ . \tag{14.19}$$

It's form depends on the representation of γ matrices that is used. It will have one form in the Dirac representation and it will have other forms in other representations.

It is known that a spinor transforms under the group of proper orthochronous Lorentz transformation or, in other words under the $SO^+(1,3)$. And an interesting point is that, in fact, the charge conjugate spinor transforms under the same transformation, under which ψ transforms.

If ψ undergoes under Lawrence transformations, it can be written in the following way:

$$\psi \to S\psi \ , \tag{14.20}$$

where S as it known is given by

$$S = \exp\left(\frac{1}{4}\gamma^{\mu}\gamma^{\nu}\omega_{\mu\nu}\right) \ . \tag{14.21}$$

Then for ψ^c the same formula will be found:

$$\psi^c \to S\psi^c \ . \tag{14.22}$$

And this is a consequence of the relation (14.15) and (14.13):

$$\mathbb{C}\gamma^0\psi^* \to \mathbb{C}\gamma^0 S^*\psi^* \ . \tag{14.23}$$

Then, ψ^* can be replaced by ψ^c:

$$14.23 = \left(\mathbb{C}\gamma^0\right) S^* \left(\mathbb{C}\gamma^0\right)^{-1} \psi^c \ . \tag{14.24}$$

Then, we can look in more detail at the term $\left(\mathbb{C}\gamma^0\right) S^* \left(\mathbb{C}\gamma^0\right)^{-1}$:

$$\left(\mathbb{C}\gamma^0\right) S^* \left(\mathbb{C}\gamma^0\right)^{-1} = \mathbb{C}\gamma^0 \exp\left(\frac{1}{4}\gamma^{*\mu}\gamma^{*\nu}\omega_{\mu\nu}\right) \left(\mathbb{C}\gamma^0\right)^{-1} \ . \tag{14.25}$$

Using the property of exponential term $\mathbb{C}\gamma^0$ can be raised into the exponent and, therefore,

$$14.25 = \exp\frac{1}{4}\left(\mathbb{C}\gamma^0\right)\gamma^{*\mu}\left(\mathbb{C}\gamma^0\right)^{-1}\left(\mathbb{C}\gamma^0\right)\gamma^{*\nu}\left(\mathbb{C}\gamma^0\right)^{-1}\omega_{\mu\nu} \ . \tag{14.26}$$

Then according to the (14.15), formula (14.26) can be written as:

$$\left(\mathbb{C}\gamma^0\right) S^* \left(\mathbb{C}\gamma^0\right)^{-1} = \exp\left(\frac{1}{4}\gamma^{\mu}\gamma^{\nu}\omega_{\mu\nu}\right) = S \ . \tag{14.27}$$

In other words, the (14.26) can be transformed into

$$\mathbb{C}\gamma^0\psi^* \to S\psi^c \ . \tag{14.28}$$

Moreover, if U is introduced as $\mathbb{C}\gamma^0$, then it's natural to require that this transformation U is unitary, because the equivalence of γ matrices γ^{μ} and $-\gamma^{*\mu}$ will be then unitary equivalence of two representations of the Clifford algebra. And this requirement as can be seen amounts to the following.

$$UU^+ = \mathbb{C}\gamma^0\gamma^{+0}\mathbb{C}^+ = \mathbb{C}\mathbb{C}^+ = 1 \ . \tag{14.29}$$

So, matrix $\mathbb{C}$ itself must be unitary.

Then, for transformation of the Lawrence representation S the following can be seen:

$$\boxed{US^*U^+ = S}\ . \tag{14.30}$$

So, this is a property of the Lawrence transformation from $SO^+(1,3)$, which can be derived in this way. It shows in fact that complex conjugate representation Lorentz representation S^* is related to S by means of the formula:

$$S^* = U^+SU\ . \tag{14.31}$$

Mathematically this fact means that the original representation $SO^+(1,3)$ on four-dimensional spinors and it's complex conjugate representation are simply a unitary equivalent.

On the other hand, now a new property of this representation was found, namely, if he complex conjugated that in reality we don't get new representation, but the unitary equivalent representation is gotten and representations, for which complex conjugate are equivalent to the original representation are called pseudoreal.

Therefore, from here the representation of $SO^+(1,3)$ on four-component Dirac spinors is reducible and it is pseudoreal.

If the unitary matrix U would be one, then such representations would be called in the representation theory as real representations. Representations for each complex conjugate are unitary equivalent to the original representation are called pseudoreal. And finally, if, in general, S^* is not related to S by unitary transformation, then such representations are called complex. If two group elements A and B are taken, then this is a homomorphism, which satisfy the relation:

$$\rho(AB) = \rho(A)\,\rho(B)\ , \tag{14.32}$$

where A and B are any two elements of the group.

If the representation is complex conjugated than this representation still will be a representation:

$$\rho^*(AB) = \rho^*(A)\,\rho^*(B)\ . \tag{14.33}$$

And then the questions, which can be asked is how ρ^* is related to ρ. Is it unitary equivalent to the original representation? Is it's absolutely the same as original representation?

The answer to this questions is to give a classification of representations as real, pseudoreal and complex. And for the case of the Lorentz representation of $SO^+(1,3)$ and four-component Dirac spinors will learn that such a representation is pseudoreal.

Then it is interesting to observe what will happen if electromagnetic field is absent or if there is no electromagnetic field at all. So, in this case, if there is no electromagnetic field or this field is switched off, then it can be seen that the transformation $\psi \to \psi^c$ is just a novel discrete symmetry of the Dirac equations. This is because it transforms solutions into solutions of the Dirac equation. So, there is one solution, which can be called as ψ and, then, if the operation of charge conjugation is applied, another solution of the same Dirac equation will be gotten. Moreover, this discrete symmetry is very similar to, for instance, parity or time reversal. So, it acts in a similar way. It turns out that this symmetry maps positive energy solutions into negative energy solutions and vice versa. And it can be seen from the following. If the symmetry of charge conjugation is taken and acts on positive energy solution, which is obtained by acting on the Dirac spinor, then this can be written in the following way:

$$\mathbb{C}\left(\Lambda_+ u\right) = \mathbb{C}\gamma^0\left(\Lambda_+^* u^*\right) \ . \tag{14.34}$$

The matrix Λ_+^* can be replaced by definition and the following expression can be gotten:

$$\mathbb{C}\left(\Lambda_+ u\right) = \mathbb{C}\gamma^0 \underbrace{\frac{\gamma^{*0}E - c\gamma^{*i}p^i + mc^2\cdot\mathbb{1}}{2mc^2}}_{\Lambda_+^*}\left(\mathbb{C}\gamma^0\right)^{-1}\left(\mathbb{C}\gamma^0\right)u^* \ . \tag{14.35}$$

Expression (14.35) can be simplified and as a result the following expression will be gotten:

$$\mathbb{C}\left(\Lambda_+ u\right) = \underbrace{\frac{-\gamma^0 E + c\gamma^i p^i + mc^2\cdot\mathbb{1}}{2mc^2}}_{\Lambda_-} c\gamma^0 u^* = \Lambda_-\left(\mathbb{C}u\right) \ . \tag{14.36}$$

In other words, the following commutation relations exist:

$$\boxed{\mathbb{C}\Lambda_+ = \Lambda_-\mathbb{C}} \ . \tag{14.37}$$

That's why a positive energy solution and the charge conjugation go to the negative energy solutions and vice versa. So it's a symmetry, which maps positive solutions into negative solutions.

Now the same trick with helicity can be done, because for helicity projectors on solution with positive helicity and negative helicity are exist.

$$\mathbb{C}\left(\mathcal{J}_\pm\left(\vec{p}\right)u\right) = \mathbb{C}\gamma^0\left(1 \pm \frac{\gamma^{*5}\gamma^{*0}\gamma^{*i}p^i}{|\vec{p}|}\right)\left(\mathbb{C}\gamma^0\right)^{-1}\left(\mathbb{C}\gamma^0\right)u^* \ . \tag{14.38}$$

Then it can be seen, analogously to the previous case, that what expression (14.38) will be simplified into:

$$\mathbb{C}\left(\mathcal{J}_{\pm}\left(\vec{p}\right)u\right)=\frac{1}{2}\left(1\mp\frac{\gamma^{5}\gamma^{0}\gamma^{i}p^{i}}{|\vec{p}|}\right)\left(\mathbb{C}u\right)=\mathcal{J}_{\mp}\left(\vec{p}\right)\mathbb{C}u\ . \tag{14.39}$$

So, intertwining of helicity happens in the following way:

$$\mathbb{C}\mathcal{J}_{\pm}=\mathcal{J}_{\mp}\left(\vec{p}\right)\mathbb{C}\ . \tag{14.40}$$

In other words, solutions with positive helicity go to solutions with negative helicity and vice versa.

In general, now it can be concluded that this operation acts like positive energy solution and the definite helicity goes to negative energy solution and opposite helicity and vice versa. So, this is how this symmetry acts on solutions of the Dirac equation even in the absence of electromagnetic field, because if there is an electromagnetic fields and charge conjugation also implies that electromagnetic field changes the sign. But if electromagnetic field is absent, there is nothing to charge conjugate from the point of view of electromagnetic field and, therefore, what then is gotten is just a statement about the charge conjugation been in U discrete symmetry of the Dirac equation, which transforms solutions into solutions.

Further determination of the concrete matrix C relies on the representation chosen. And, for instance, for the Dirac representation of γ matrices, the matrix C can be chosen to be equal to:

$$\mathbb{C}=\gamma^{2}\gamma^{0}\ . \tag{14.41}$$

Then it can be seen that this matrix exactly does the job, which is needed. So, for instance,

$$\mathbb{C}\gamma^{0}=\gamma^{2}\ . \tag{14.42}$$

Then, it can be seen that the following expression is right

$$\left(\gamma^{2}\right)^{-1}\gamma^{\mu}\gamma^{2}=-\gamma^{*\mu}\ . \tag{14.43}$$

According to the property

$$\left(\gamma^{2}\right)^{2}=-1\ , \tag{14.44}$$

the (14.43) can be written as follows:

$$\gamma^{2}\gamma^{\mu}\gamma^{2}=\gamma^{*\mu}\ . \tag{14.45}$$

For instance, it can be seen that in the Dirac representation a matrix γ^0 is a diagonal matrix and it's given by

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & -\mathbb{1} \end{pmatrix} . \tag{14.46}$$

So, from this point of view, γ^0 is real. Then,

$$\gamma^2\gamma^0\gamma^2 = -\gamma^0\left(\gamma^2\right)^2 = \gamma^0 . \tag{14.47}$$

Now, for other matrices the relation (14.45) can be checked. For this purpose, γ^i matrices should be recalled:

$$\gamma^i = \begin{pmatrix} \mathbb{0} & \sigma^i \\ -\sigma^i & \mathbb{0} \end{pmatrix} . \tag{14.48}$$

And it should be noticed that σ^1 and σ^3 are real. Therefore, under complex conjugation γ^1 and γ^3 matrices will be unchanged

$$\left(\gamma^1\right)^* = \gamma^1 , \ \left(\gamma^3\right)^* = \gamma^3 , \tag{14.49}$$

while for γ^2 the following expression will be gotten:

$$\left(\gamma^2\right)^* = -\gamma^2 . \tag{14.50}$$

It is exactly seen that γ^1 and γ^3 matrices will behave themselves exactly like γ^0:

$$\begin{cases} \gamma^2\gamma^1\gamma^2 = -\gamma^1\left(\gamma^2\right)^2 = \gamma^1 \\ \gamma^2\gamma^3\gamma^2 = -\gamma^3\left(\gamma^2\right)^2 = \gamma^3 \end{cases} \tag{14.51}$$

For γ^2 the situation is different. So, we will get

$$\gamma^2\underbrace{\gamma^2\gamma^2}_{=-1} = -\gamma^2 . \tag{14.52}$$

Therefore, in the (14.52) the required expression is gotten.

Indeed, for the Dirac representation $\mathbb{C} = \gamma^2\gamma^0$ is a matrix, which satisfies the required property. Then it was also seen that unitarity conditions are satisfied.

In addition the matrix $\mathbb{C}$ has the following properties:

$$\mathbb{C}^2 = 1 , \ \rightarrow \gamma^2\gamma^0\gamma^2\gamma^0 = -\underbrace{\gamma^2\gamma^2}_{-\mathbb{1}}\underbrace{\gamma^0\gamma^0}_{\mathbb{1}} = \mathbb{1} , \tag{14.53}$$

$$\mathbb{C}^* = -\mathbb{C} , \ \left(\gamma^2\gamma^0\right)^* = \gamma^{*2}\gamma^0 = -\gamma^2\gamma^0 = -\mathbb{C} , \tag{14.54}$$

$$\mathbb{C}\mathbb{C}^{+} = \mathbb{C}\mathbb{C}^{*t} = -\mathbb{C}\mathbb{C}^{t} = -\gamma^{2}\underbrace{\gamma^{0}\gamma^{t0}}_{\mathbb{1}}\gamma^{t2} = -\gamma^{2}\underbrace{\gamma^{t2}}_{\gamma^{2}} = \mathbb{1}\ . \tag{14.55}$$

Then the charge conjugate spinor in the Dirac representation explicitly can be realized as:

$$\psi^{c} = \mathbb{C}\gamma^{0}\psi^{*} = \gamma^{2}\psi^{*}\ . \tag{14.56}$$

The ψ^{c} can also be written in the following way. The $\overline{\psi}$ should be introduced, which is explicitly

$$\overline{\psi} = \psi^{*t}\gamma^{0}\ , \tag{14.57}$$

where ψ^{*t} is a row vector. So, if then $\overline{\psi}$ is transposed, then the following expression will be gotten:

$$\overline{\psi}^{t} = \gamma^{0}\psi^{*}\ . \tag{14.58}$$

According to the (14.58), the (14.56) can be written as:

$$\boxed{\psi^{c} = \mathbb{C}\overline{\psi}^{t}}\ . \tag{14.59}$$

One remark, also can be given here is the following:

$$\mathbb{C}^{-1}\gamma^{\mu}\mathbb{C} = -\gamma^{0}\gamma^{*\mu}\gamma^{0}\ , \tag{14.60}$$

where the following relation was used:

$$\left(\mathbb{C}\gamma^{0}\right)^{-1}\gamma^{\mu}\left(\mathbb{C}\gamma^{0}\right) = -\gamma^{*\mu} \tag{14.61}$$

and the brackets $\left(\mathbb{C}\gamma^{0}\right)^{-1}$ was opened to get:

$$\gamma^{0}\mathbb{C}^{-1}\gamma^{\mu}\mathbb{C}\gamma^{0} = -\gamma^{*\mu} \tag{14.62}$$

where the (14.60) can be gotten by multiplying by γ^{0} from the left and from the right.

But, on the other hand, what is written down in the (14.60) is nothing else in the Dirac representation as

$$\mathbb{C}^{-1}\gamma^{\mu}\mathbb{C} = -\left(\gamma^{\mu}\right)^{t}\ . \tag{14.63}$$

So, complex conjugate matrix multiplied from the left and from the right by γ^{0} is the same as γ^{μ} transposed. Alternatively, in the Dirac representation, the matrix $\mathbb{C}$ can be characterized by the property that conjugating of γ^{μ} with the unitary matrix $\mathbb{C}$ gives the (14.63).

Again, this shows the unitary equivalence of γ^μ with $(\gamma^\mu)^t$ e.g. with the other representation of the Clifford algebra, which is performed with the help of the unitary conjugation by by the matrix $\mathbb{C}$.

Also here another interesting simple consequence can be deduced that if

$$\mathbb{C}^{-1}\gamma^\mu\mathbb{C} = -\left(\gamma^\mu\right)^t \ , \tag{14.64}$$

then it can be checked what happens with γ^5, when it conjugated. And this simply amounts to

$$\mathbb{C}^{-1}\gamma^5\mathbb{C} = \left(\gamma^5\right)^t \ , \tag{14.65}$$

because γ^5 is the product

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{14.66}$$

and when $\mathbb{C}^{-1}$ and $\mathbb{C}$ are applied, then this transposes each of this matrices. So, it becomes

$$\mathbb{C}^{-1}\gamma^0\gamma^1\gamma^2\gamma^3\mathbb{C} = \gamma^{t0}\gamma^{t1}\gamma^{t2}\gamma^{t3} \ . \tag{14.67}$$

On the other hand, it is known that in the Dirac representation the γ^5 matrix is symmetric and it does not change under transposition, but then under the result (14.67) it follows that since

$$\mathbb{C}^{-1}\gamma^5\mathbb{C} = \gamma^5 \ , \tag{14.68}$$

$\mathbb{C}$ actually commutes with γ^5:

$$\left[\mathbb{C}, \gamma^5\right] = 0 \ . \tag{14.69}$$

Sometimes this property is also interesting and should be taken into account in some considerations.

So in any case, this is a consequence of the explicit representation for $\mathbb{C}$ in the case of Dirac representation.

In summary, the Dirac theory is invariant under charge conjugation of a spinor supplied with a simultaneous change of electromagnetic field to minus electromagnetic field. The physical meaning of this charge conjugation operation is that any physically realizable state of electron in the field A_μ corresponds to a physically realizable state of a positron in the field $-A_\mu$. So, operation of charge conjugation changes electrons with negative energy and spin up to positrons with positive energy and spin down. If we have a solution of the Dirac equation, for instance, with a negative energy, which can be called as $\psi^{(-)}$, which is wave function corresponding to a solution with negative energy, and an electric charge,

which can be denoted by e, then we should take as a physically realizable or physically relevant quantity not the wave function $\psi^{(-)}$, because it's a solution with negative energy, but rather it's a charge conjugated solution $\psi^{c(+)}$, which has then positive energy and given explicitly by

$$\psi^{c(+)} = \mathbb{C}\overline{\psi}^{(-)t} \, . \tag{14.70}$$

This $\psi^{c(+)}$ will then describe a particle with a positive energy, which is a physical particle, but with different sign in comparison to the sign of original electron.

Therefore, we can identify the positive energy plane waves ψ for an electron as

$$\psi_e^r \left(\vec{x}\,, t\right) = u_+^r \left(\vec{p}\right) e^{-ip_\mu x^\mu/\hbar} \tag{14.71}$$

and for a positron as

$$\psi_p^r \left(\vec{x}\,, t\right) = \mathbb{C} \left[\overline{u}_-^r \left(-\vec{p}\right)\right]^t e^{-ip_\mu x^\mu/\hbar} \, . \tag{14.72}$$

The (14.72) can be also written in terms of $v_- \left(\vec{p}\right)$, because as it was shown in the previous lecture:

$$v_- \left(\vec{p}\right) = u_- \left(-\vec{p}\right) \tag{14.73}$$

and the (14.73) will be linked with (14.72) by the following expression:

$$v_- \left(\vec{p}\right) \left(\vec{p}\right) e^{ip_\mu x^\mu/\hbar} \underset{\mathbb{C}}{\rightarrow} \mathbb{C} \left[\overline{u}_-^r \left(-\vec{p}\right)\right]^t e^{-ip_\mu x^\mu/\hbar} \, . \tag{14.74}$$

So, then it can be checked that under a charge conjugation the (14.73) will turn into the (14.72).

Let's return back to the general solution of the Dirac equation, which was represented in terms of plane waves and then it would be natural to treat the amplitude, which is given in terms of the coefficient b_r in the general solution of Dirac equation as an annihilation operator of electron, while then d_r^*, which appears as another amplitude standing in front of the $i\hbar p_\mu x^\mu$ as a creation operator of a positron. That's consideration relates to the positive and negative energy solutions in the general expression for the Dirac equation. So, negative energy solutions we would like to interpret from the point of view of the new understanding of the charge conjugation symmetry as those, which correspond to actually positive energy solution, but for another particle, which is called as positron. Correspondingly, this gives an interpretation to the amplitudes, which are existed in general solution of the Dirac equation. There are amplitudes, which introduces b_r and d_r^* as correspondingly annihilation operator of electrons and the creation operator

of positrons. That's something, which can be further understood, when transition to the quantization of the Dirac equation is performed.

The final topic, which linked with the consideration of reality properties and the charge conjugation, concerns with the issue of so-called Majorana spinors.

Majorana spinors

We already learned that there are four-component Dirac spinors. There are also Weyl spinors, which are two-component complex spinors. And now another kind of spinors, which are called Majorana will be introduced. It is known that the Dirac spinor in four-dimensional Minkowski space has four complex components, which is equivalent to 8 real degrees of freedom. It is possible to reduce a number of independent components of this Dirac spinner by imposing certain conditions, which are compatible with Lorentz transformations from $SO^+(1,3)$. One of these conditions is a Weyl condition, which was discussed and that allows to define the Weyl spinor, which has two complex components. And the spinor transforms as it known irreducibly under $SO^+(1,3)$.

Now, another condition can be imposed, which is called Majorana condition. This is condition, which says to take spinor ψ and identify it with its charge conjugation.

$$\psi = \psi^c \ . \tag{14.75}$$

In other words, Majorana condition is a condition that states that particle is its own antiparticle.

The special Dirac spinner, which satisfies the condition (14.75) is called Majorana. This condition is allowed to impose, because it is known that ψ and ψ^c transform in the same way under Lorentz transformations. So, this condition does not break Lorentz symmetry and it is Lorentz invariant.

An explicit form of the Majorana spinner depends on the chosen representation for γ matrices, because the explicit form of the charge conjugate matrix $\mathbb{C}$ depends on the representation chosen. For instance, in the Dirac and Weyl representations, $\mathbb{C}$ is given by

$$\mathbb{C} = \gamma^2 g^0 \tag{14.76}$$

and, therefore, when the Majorana condition is imposed, that means that we have to solve the following equation:

$$\psi = \mathbb{C}\gamma^0 \psi^* = \gamma^2 \psi^* \ . \tag{14.77}$$

If ψ is written in terms of two components spinors ϕ and χ:

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \ , \tag{14.78}$$

then the (14.77) can be written as:

$$\begin{pmatrix} \phi \\ \chi \end{pmatrix} = \begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix} \begin{pmatrix} \phi^* \\ \chi^* \end{pmatrix} \ . \tag{14.79}$$

It can be seen that from the condition (14.79) χ can be expressed and the answer will be that such a Majorana spinor in the Dirac representation of γ matrices will have the following form:

$$\psi = \begin{pmatrix} \phi \\ -\sigma^2\phi^* \end{pmatrix} \ . \tag{14.80}$$

Moreover, there exists another representation of γ matrices, in which the Majorana condition becomes very simple and transparent. This representation of γ matrices is called Majorana representation. Majorana representation is a representation in which all γ matrices are imaginary. So the Dirac operator contains, in fact, only real coefficients and explicitly it can be written down straight away. In this representation:

$$\gamma^0_M = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \ , \ \gamma^1_M = \begin{pmatrix} i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix} \ , \ \gamma^2_M = \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \ , \ \gamma^3_M = \begin{pmatrix} -i\sigma^1 & 0 \\ 0 & -i\sigma^1 \end{pmatrix} \ . \tag{14.81}$$

It can be checked that the Clifford algebra relations for these γ matrices will be satisfied and this means by Pauli theorem that there exists a unitary matrix that transforms Dirac representation into this new Majorana representation by means of conjugation with the unitary matrix U. So,

$$U\gamma^\mu_D U^{-1} = \gamma^\mu_M \ , \tag{14.82}$$

where it can be checked that explicitly the matrix U is given by

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & \sigma^2 \\ \sigma^2 & -\mathbb{1} \end{pmatrix} \tag{14.83}$$

and this matrix is explicitly unitary

$$U^+U = 1 \ . \tag{14.84}$$

Now, the charge conjugation matrix should be determined from the condition

$$\left(\mathbb{C}\gamma^0\right)^{-1} \gamma^\mu \left(\mathbb{C}\gamma^0\right) = -\gamma^{*\mu} \ . \tag{14.85}$$

But in the Majorana representation it is know that the $\gamma^{*\mu}$ is simply

$$\gamma^{*\mu} = -\gamma^{\mu} \ , \tag{14.86}$$

because all γ matrices are purely imaginary and, therefore, the (14.85) can written as

$$\left(\mathbb{C}\gamma^0\right)^{-1} \gamma^{\mu} \left(\mathbb{C}\gamma^0\right) = \gamma^{\mu} \ . \tag{14.87}$$

From here, it follows that

$$\left[\mathbb{C}\gamma^0, \gamma^{\mu}\right] = 0 \ . \tag{14.88}$$

So, this means that $\mathbb{C}\gamma^0$ is the only matrix, which commutes with all γ^{μ} in four dimensions. So, it can be gotten that

$$\mathbb{C}\gamma^0 = \mathbb{1} \tag{14.89}$$

and from here, it can be concluded that $\mathbb{C}$ is simply equals to:

$$\mathbb{C} = \gamma^0 \ . \tag{14.90}$$

Then, charge conjugate spinor in this representation is given by:

$$\psi^c = \mathbb{C}\gamma^0\psi^* = \gamma^0\gamma^0\psi^* = \psi^* \ . \tag{14.91}$$

Therefore, we got that ψ^* in Majorana representation of γ matrices is equal to the charge conjugate spinor. Therefore, in this representation Majorana condition becomes extremely simple and transparent:

$$\psi = \psi^* \ , \tag{14.92}$$

which means that in the Majorana representation spinor ψ is simply real. It has simply four real components and in this representation it can be also seen that S matrix of Lorentz transformations must preserve reality of a spinor. So, spinor ψ is equal to ψ^* and that means that ψ is real and, therefore, S acting on it must transform a real matrix into real matrix. This means that in this case S itself is real and that's can be easily seen from the fact that it's given by

$$S = \exp\left(\frac{1}{4}\gamma^{\mu}_M\gamma^{\nu}_M\omega_{\mu\nu}\right) \ . \tag{14.93}$$

In the Majorana representation each of the γ matrices is real and then the product of the two imaginary γ matrices then gives a real result. So, in this representation Lorentz transformations are realized by real matrices.

ФИЗИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in

ЛЕКЦИИ УЧЕНЫХ МГУ