Introduction to Quantum Field Theory

Rainer F. Hauser
rainer.hauser@gmail.com

November 18, 2020

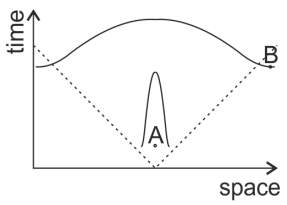
Abstract

Quantum Field Theory combines Special Relativity and Quantum Mechanics and was initially developed to quantize the electrodynamic field. It later turned out that all fundamental interactions are necessarily described by Quantum Field Theory. Markus Luty from the University of California, Davis, held several introductory lectures in spring 2013 covering this topic but only the first seven have been published as videos. His lectures were available on YouTube at the time this transcript has been assembled and may as "Quantum Field Theory" or "QFT" still be available today.

1 Overview of Quantum Field Theory

1.1 Combining Special Relativity and Quantum Mechanics

Quantum Field Theory combines Special Relativity and Quantum Mechanics. However, there is tension between Special Relativity with causality $v \leq c$ and Quantum Mechanics with Heisenberg's uncertainty principle $\Delta p \, \Delta x \geq \hbar$ because an initially well-localized wave function spreads with time and becomes non-zero in spacetime regions forbidden by causality as shown in the figure on the right side. If there is a non-zero probability for measuring the particle at point A and detecting it later at point B, this would mean that the particle has traveled faster than light.



The theory of a single non-interacting particle is very simple. A quantum mechanical theory is defined by a space of states $\{|\psi\rangle\}$ and a set of Hermitian operators $\{\boldsymbol{O}\}$ that act on states. Among the operators is a special operator called the Hamiltonian $\boldsymbol{H}$ which gives the time evolution

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \boldsymbol{H} |\psi\rangle$$

of the state $|\psi\rangle$.

A complete set of states for a single non-interacting particle in Quantum Mechanics is $|\vec{p}\rangle$ consisting of the eigenstates of the three momentum operators, and all possible states of this particle are linear combinations of them. For each of them one needs to know how the Hamiltonian acts on it, and it is assumed that these states are eigenstates of the Hamiltonian

$$\boldsymbol{H} |\vec{p}\rangle = E_{\vec{p}} |\vec{p}\rangle$$

such that the eigenvalues $E_{\vec{p}}$ are the respective energies of the states. The energy is

$$E_{\vec{p}} = +\sqrt{\vec{p}^2 c^2 + m^2 c^4}$$

for a single relativistic and non-interacting particle with the 3-momentum $\vec{p}$. This defines the theory completely, and it is obvious that this is the only way the theory can possibly be defined.

The development of the wave function $\tilde{\psi}(\vec{p},t)$ with respect to time t in momentum space is

$$\tilde{\psi}(\vec{p},t) = e^{-iE_{\vec{p}}t}\,\tilde{\psi}(\vec{p},0)$$

using the usual conventions with units $\hbar = c = 1$. Time development in position space becomes

$$\psi(\vec{x},t) = \int d^3\vec{y}\, G(\vec{x}-\vec{y},t)\,\psi(\vec{y},0)$$

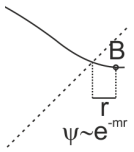
when Fourier-transformed where the Green's function is

$$G(\vec{x}-\vec{y},t) = \int \frac{d^3\vec{p}}{(2\pi)^3}\, e^{-i(E_{\vec{p}}t-\vec{p}\cdot\vec{x})} \tag{1.1}$$

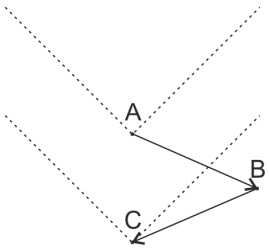
which describes how the wave package spreads. The Green's function evaluated in momentum space is

$$G(\vec{x},t) = \frac{1}{(2\pi)^3}\int_0^\infty p^2dp\int_{-1}^1 d\,\cos\theta\int_0^{2\pi} d\varphi\; e^{-i(E_pt-px\cos\theta)} = \frac{1}{2\pi^2}\int_0^\infty p\,dp\,e^{-iE_pt}\,\sin(px)$$

using the spherical coordinates (p,θ,φ) with $x = |\vec{x}|$ and $p = |\vec{p}|$. The interesting question is whether this integral is non-zero in a region with $x > t$. This would mean that there is a non-zero probability for the particle to go faster than light. This is indeed the case because the energy E_p is positive such that the exponential factor e^{-iE_pt}, extended to complex t, becomes a damped exponential if $\mathrm{Im}(t) < 0$. The whole expression is therefore an analytic function in t for fixed x. An analytic function of t has the property that it vanishes identically if it vanishes for any finite interval of t. Thus, in order to be zero in the entire region $x > t$ it has to vanish identically in a finite strip of t.



This result comes from the basic fact that the energy E_p is positive. If one evaluates this integral one finds as shown in the portion of the above figure on the right side that the wave function in the area of point B is non-zero and it has the form $\psi \sim e^{-mr}$ as one would expect from tunneling into a classically forbidden region with the distance r from the point where the forbidden region begins. This means that there is a possibility to send a signal faster than the speed of light although it is not a very efficient method because of e^{-mr}.



If the straight line from point A to point B is spacelike, then there is another reference frame where B looks like being in the past as seen from A. Thus, as above there is the possibility to send a message from A to B, and similarly there is also the possibility to send a message from B to a point C in the true past of A leading to all kinds of paradoxes. The situation is depicted in the figure on the left side. A man could, for example, send a message from A to C to convince himself to commit suicide, but then he might no longer live at A to be able to send this message.

Processes faster than the speed of light such as this communication vanish identically because of complete destructive interference with processes involving antiparticles as Richard Feynman found out. Assuming an electron e^- comes in to A where an experiment localizes it, there is a non-zero amplitude for this electron to go to B where it is detected and from where it continues. This process going forward in time for one observer as shown on the left side of figure 1 is not vanishing.

The direction in time depends on the reference frame. Another observer may see it as going backward in time. There is, as Feynman pointed out, another process that contributes to the same set of observations

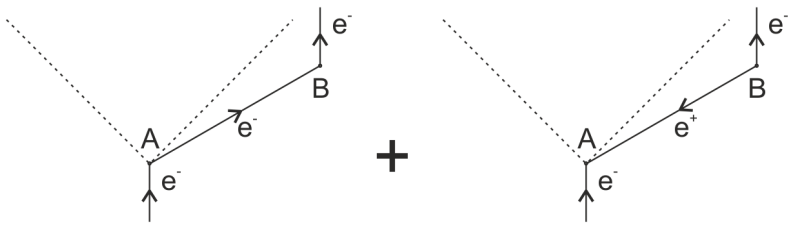


Figure 1: Two processes contributing to the same set of observations

as shown on the right side of figure 1. In this process an electron-positron pair gets created at B where the positron e^+ annihilates with the incoming electron at A. In another reference frame the electron may go backward in time and the positron forward, but important is only that both processes exist, and that they destructively interfere such that their sum is actually zero.

In order to interfere this way the two processes have to be completely indistinguishable. In both processes a negative charge is transferred from A to B. On the left side of the figure an electron is moved from A to B and on the right side an positive charge is moved from B to A. More fundamentally, there has to be an exact symmetry relating electrons and positrons, and there is indeed one. In Relativistic Quantum Field Theory there is invariance under CPT where C is charge conjugation, P is parity with the reversal of all spatial directions, and T is the direction of time. This symmetry interchanges particles and antiparticles. Therefore, electron and positron must have exactly the same mass and the same spin, but they have opposite charge.

1.2 Constructing a Relativistic Quantum Theory of Particles

This looks like a total miracle, but it is not. The way it works is that causality is actually guaranteed if one formulates the theory in terms of observables where observables are Hermitian operators $\boldsymbol{O}(X)$ associated with points of spacetime. They have the crucial property that they commute $[\boldsymbol{O}_i(X), \boldsymbol{O}_j(Y)] = 0$ at two different points X and Y if these two points are spacelike separated and therefore fulfill $(X-Y)^2 < 0$. This means that X and Y can be made living at equal time $t = 0$ in some reference frame and that operators living at different points in space at the same time can be thought of as independent. These commuting operators can therefore be simultaneously measured.

When constructing a quantum field theory there are many operators without this property, but if it is required that the theory is written in terms of operators with this property then one automatically gets causality. The quality which ensures that the operators have this property is locality. If the theory is defined in terms of quantum fields that obey local equations then this property is guaranteed. A non-interacting scalar field, for example, is described by a quantum operator $\varphi(X)$ defined for every point X in spacetime and obeys

$$\Box\varphi(X) + m^2\varphi(X) = 0 \tag{1.2}$$

which is basically the Klein-Gordon equation. The most general solutions $\varphi(X)$ to this equation are observables that do commute at spacelike points. These operators $\varphi(X)$ when acting on particle states relate the creation and the annihilation of particles in such a way as to give the cancellation required by Feynman. Thus, the task is the quantization of a classical theory of fields through the application of the rules of quantum mechanics. This is considered the only way one can get a relativistic quantum theory of particles.

If an experimentalist decides to localize a particle in a tinier and tinier region l the particle get a large momentum p and a large energy E

$$p \sim \frac{\hbar}{l} \qquad\qquad E \sim \sqrt{p^2c^2 + m^2c^4}$$

putting $\hbar$ and c temporarily back. If momentum p becomes large enough, the term m^2c^4 in E does not matter anymore, and the energy becomes much larger than the rest energy associated with any particle. The claim is that particles and antiparticles get copiously created in this situation. This happens as soon as the length l gets smaller than or of order

$$l \lesssim \frac{\hbar}{mc}$$

called the Compton wavelength. This tells that whenever one tries to probe the theory on these distance scales or shorter one has to deal with relativistic particles and particle creation.

This has practical implications and shows how one can create particles in the laboratory because it is the basic mechanism behind particle colliders such as the Large Hadron Collider currently running at CERN. The collision between two beams of particles actually localizes a large amount of energy in a small region and that gives rise to particle-antiparticle pairs. This is the reason why one can study fancy particles, and this is very universal. If there is a particle X with all sorts of weird charges one can always create

a particle-antiparticle pair $X\bar{X}$ because of CPT where the antiparticle $\bar{X}$ has all the opposite quantum numbers. By colliding protons one can produce pairs $X\bar{X}$ usually together with some other particles provided that the mass of the X particles is sufficiently small such that one can produce them, because there is only a limited amount of energy in the colliding beams. Given enough energy it is simply a question of how frequently or infrequently such a process happens. These facts explain why one tries to go to higher and higher energies.

1.3 Some General Features of Quantum Field Theory

Invariance under CPT following from Quantum Mechanics and Special Relativity has already been mentioned. Another feature is the relation between spin and statistics. Particles of integer spin are bosons, and particles of half integer spin are fermions.

Interactions of relativistic particles of higher spin are extremely constrained. For spin 0 and $\frac{1}{2}$ there are no constraints beyond locality and Lorentz invariance. Higher spin values are more constraint such that there are no elementary particles with spin greater than 2. Obviously there are objects with spin (angular momentum) greater than 2 such as spinning basketballs, but they are not elementary particles. The spacing ΔE of energy levels is so tiny for a basketball with mass M that $\Delta E \ll Mc^2$. For a relativistic basketball one does not get one particle for this relativistic limit $M \to 0$ but one gets a lot of states. In this sense the basketball is not just one particle, and this is the key difference. An elementary (or fundamental) particle is one that has only a small number of degrees of freedom for $m \to 0$.

For spin 1 the only theories possible are so-called gauge theories. The spin 1 massless particle exists but its structure is highly constrained. For spin 2 there is a unique theory for relativistic spin 2 particles, and this is Einstein's theory of General Relativity as shown by Feynman and others. For spin $\frac{3}{2}$ and spin > 2 one can write interactions for all these massless theories, but they are so-called irrelevant interactions which become arbitrarily weak as one goes to low energies. These theories can be interesting, and also General Relativity is a theory whose interactions are irrelevant in the same sense. If one believes that gravity is fundamental, the theories with spin $\frac{3}{2}$ and spin > 2 need to be coupled to gravity. For theories with spin 0, $\frac{1}{2}$ and 1 this is completely straight forward. For theories with spin $\frac{3}{2}$ the unique theory one gets is supergravity which is the supersymmetric version of Einstein's General Relativity, and for spin > 2 one cannot find anything at all. Thus, if one includes gravity the highest spin possible is 2.

1.4 Effective Field Theory

Effective field theory is both a set of techniques and a point of view that allows to organize the thinking and to make progress in many areas. At the most basic level effective field theory is nothing more than dimensional analysis. In units with $\hbar = c = 1$, length and time are related $t \sim l$, energy and momentum are related $E \sim p$, and length/time and energy/momentum are inversely related $t \sim l \sim \frac{1}{E} \sim \frac{1}{p}$. In those units the uncertainty principle states that high energy and momentum are equivalent to small length and time. If one wants to uncover the fundamental structure of interactions at short distances or short times, one has to go to high energy and momentum. This is what particle accelerators are made for. They are basically very powerful microscopes that allow to see short distances and times.

Given the fact that there are effects one cannot see in experiments with a finite resolution or a finite amount of energy, there might be a particle whose mass is much bigger than the energy of experiments possible with current accelerators. One can assume that there are such particles which can therefore not be produced by the accelerator but their quantum fluctuations are still there and give non-zero effects. One can tell a lot about them just by dimensional analysis.

The fields φ describing particles with spin 0 and 1 have dimension of mass $[\varphi] = M$, and the fields ψ describing particles with spin $\frac{1}{2}$ have dimension $[\psi] = M^{3/2}$. Also the derivative ∂_μ needed to write down interactions has dimension one over length which is mass $[\partial_\mu] = M$. All of these ingredients used to write interaction terms have a dimension of a positive power of mass. If one writes interactions with many fields and derivatives they have a very large dimension in units of mass. Therefore if there is this new unknown particle its effects will be suppressed by some positive power of E over M where E is the characteristic energy scale of the experiments one can do.

As an example the Hamiltonian $\boldsymbol{H}$ might be

$$\boldsymbol{H} = \int d^3x \left[\lambda\varphi^4 + \frac{c}{M^2}\varphi^6 + ...\right]$$

with $[\boldsymbol{H}] = M$. Thus with $[d^3x] = l^3 = \frac{1}{M^3}$ and $[\varphi] = M$ the coupling constant λ must be dimensionless, and that is called marginal. The electromagnetic coupling, for example, is marginal. They are just numbers and do not have any scale dependence. That is the reason why they are called marginal. The term φ^6 on the other hand is supposed to be suppressed by $(E/M)^2$. Its coupling is called irrelevant because its strength falls off with a power of the energy.

Thus, one sees that up to these irrelevant effects which could be interesting and important but are generally small the theory is actually determined just by the marginal couplings. Because everything that is used to write a Lagrangian has positive mass dimensions, there is always a finite small number of marginal interactions.

For example, in Quantum Electrodynamics, the theory of the electromagnetic field interacting with electrons and positrons, there are only two parameters which are not irrelevant: the electromagnetic coupling strength and the mass of the electron. These two parameters completely specify the whole theory. Contrasting that to the case of non-relativistic Quantum Mechanics, even in simple one-dimensional Quantum Mechanics one has to give a potential to specify the theory, and this is an entire function of x corresponding to an infinite number of parameters.

The fact that low-energy physics is determined by a small number of parameters in Quantum Field Theory has another side. It also limits what one can learn from low-energy measurements about the fundamental theory that predicts everything. The most things one can observe at low energies are the values it predicts for all the marginal couplings. There is only a handful of these, and there is really only a handful of predictions one can directly test. Thus, many fundamental theories may look the same. That is called universality. It means that one can often make predictions without knowing the details of the fundamental theory because many fundamental theories have the same set of low-energy particles and symmetries and therefore also the same set of allowed marginal interactions.

Sometimes irrelevant interactions are not irrelevant in the colloquial sense because one might actually be able to measure them. Because they are small, this happens when they give the leading contribution. In the Standard Model of Particle Physics it turned out that the leading irrelevant interactions would generate neutrino masses and mixing.

1.5 Mass Terms in the Lagrangian and Supersymmetry

Mass means terms in the Lagrangian with positive mass dimension. If these were given by dimensional analysis one would have a kind of absurd conclusion that every mass should be of order of the largest scale that is relevant for physics. The funny thing about this kind of prediction is that the largest scale gives the biggest contribution. Then one would not see any particles, and one would not see the electron which is obviously lighter than the largest scales of physics which probably is the Planck scale $M_P \sim 10^{19}\,\mathrm{GeV}$. This is the scale at which quantum mechanical effects of gravity become important.

The electron mass is $m_e \sim 10^{-3}\,\mathrm{GeV}$, and this is a pretty big discrepancy. The reason for this is that when the electron mass goes to zero the theory would have an additional symmetry called chiral symmetry because it acts differently on left- and right-handed polarized states. If one wants to explain the small electron mass, this symmetry is required to be almost perfect and can only be broken a little bit. Then the electron mass would always be proportional to this little bit of breaking and can be small. The small mass is therefore protected by a symmetry. For spin 1 there is a similar story. It is a little bit more subtle but there is also a symmetry involved.

However, for spin 0 particles there is a genuine problem. There is no additional symmetry when a spin 0 particle has vanishing mass, and what is actually forbidden is a spin 0 particle with marginal interactions. A particle has been detected which is at least approximately a Higgs boson. This means that there is an elementary particle with spin 0. There is no inconsistency in having a light spin 0 particle. If there are particles at the Planck scale M_P then the mass of the Higgs boson squared m_h^2 as it appears in the Lagrangian would get some heavy contributions from the Planck scale in the order of M_P^2. However,

this is not the whole answer. There are other pieces one cannot compute, and they may also be in the order of M_P^2 but with the opposite sign such that the two terms in the order of M_P^2 could cancel. The cancellation is roughly the size of the observed Higgs mass squared over the Planck scale squared

$$\frac{m_h^2}{M_P^2} \sim 10^{-34}$$

such that a cancellation is needed for 34 decimal places which seems rather absurd.

There is a way to make a natural theory of spin 0 interacting particles and that is supersymmetry usually called SUSY. It is an expansion of Lorentz invariance of spacetime and turns fermions into bosons and vice versa. The idea is since it relates fermions and bosons that a spin 0 particle such as the Higgs boson would have a spin $\frac{1}{2}$ partner with the same mass because they are related by supersymmetry. However, it is known that there is an additional symmetry when a spin $\frac{1}{2}$ particle gets massless, and consequently there must be an additional symmetry when both particles get massless.

Supersymmetry gives a beautiful solution to this problem and allows the existence of a light Higgs boson. It also addresses many other conceptual problems in physics. However, it predicts that every standard model particle has a superpartner with a different spin, and there is currently no experimental evidence for this despite many searches. There are more complicated versions of this theory but physicists are feeling a little bit nervous that one has to resort to more and more complicated models just to explain that nothing has been seen.

There is another known way to explain the observation of the Higgs boson and that is that perhaps the Higgs boson is composite or at least partially composite. In this case the extra states of this compositeness have to show up near the TeV scale.

This is the subject of a big debate going on within Particle Physics. Is the Higgs boson elementary? Is supersymmetry there? The hope for finding a simple, beautiful supersymmetric model have not been fulfilled. Some people are questioning despite the discovery of the Higgs boson whether it is worth going to higher energies. Physics goes into unknown territory, and other physicists think therefore that there are many things that point to the TeV scale as an interesting scale of physics, although nobody can guarantee that there will be again spectacular discoveries like the Higgs boson.

1.6 Combining Quantum Field Theory and General Relativity

Combining Special Relativity and Quantum Mechanics resulted in Quantum Field Theory as discussed above. However, at this point there are new tensions coming from trying to combine Quantum Field Theory with Einstein's theory of General Relativity. It is not known what the resulting theory of quantum gravity is, and this combination results in a new set of deep puzzles.

General Relativity states basically that all of spacetime structure is defined by a spacetime metric that is itself a dynamical object. One of the puzzles this leads to is the following. In order to guarantee causality in Quantum Field Theory one has to write the theory in terms of observables that commute at spacelike separation. However, if the metric is dynamical and even a quantum-fluctuating object, the distinction between spacelike and timelike is not well-defined a priori. It is not just a technical problem because the distinction between commuting and non-commuting operators is a basic question that has to do with the degrees of freedom. If one takes the x-component and the y-component of a particle then these operators commute because they are completely independent degrees of freedom. Thus, the fact that now even the degrees of freedom have this quantum fuzziness is very puzzling.

To make progress on this kind of questions, situations are needed where the contradictions are sharp. Einstein used though experiments both in Special and General Relativity and was led very far. Feynman used them as well to resolve the paradoxes with causality mentioned above. Thus, one can ask what kinds of thought experiments one can use to understand quantum gravity. One very promising direction is the physics of black holes. Steven Hawking discovered in the seventies that black holes radiate in the presence of Quantum Mechanics and therefore eventually evaporate. The phenomenon of black-hole radiation is very closely tied to basic features of General Relativity and Quantum Field Theory. Black holes are very simple objects (and play in quantum gravity a similar role as the hydrogen atom plays in Quantum Mechanics) such that they are understood very well.

Radiation from black holes is nearly perfectly thermal as comes out of Hawking's calculations. It is so perfectly thermal that it is hard to understand how the formation and the evaporation of a black hole could be a unitary quantum mechanical process. It may be very complicated but the underlying physics should still be unitary Quantum Mechanics. In Quantum Mechanics everything is probabilistic but the wave function is completely deterministic. Therefore, if one knows the initial wave function one knows the final wave function. Unitarity means that one could, at least in principle, reconstruct the initial state knowing the final state. But Hawking's radiation is so perfectly thermal that it cannot contain the information needed to reconstruct the initial state. This is called the black hole information paradox. Hawking originally believed that this paradox is so strong that unitary must be violated. Now physicists believe that this process is unitary because of another major advance in understanding of quantum gravity. This is the correspondence between quantum gravity on anti-de Sitter space and conformal field theories without gravity in one lower spacetime dimension called AdS/CFT correspondence. It has not been derived from first principles but it is a sharp statement and there is so much evidence for it that the feeling is that it has to be right. One of the important things one has learned from it is that black-hole evaporation information is a unitary process as seen from outside the black hole and that is because AdS/CFT relates it to a process that could be described in an ordinary quantum field theory without gravity.

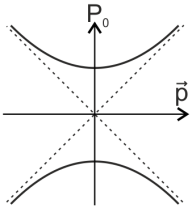
2 From Particles to Fields

2.1 States of a Single Non-Interacting Relativistic Quantum Particle

A complete set of states $|\vec{p}\rangle$ for a single non-interacting relativistic quantum particle is given by the momentum eigenstates. The conventional normalization is $\langle\vec{q}\,|\,\vec{p}\rangle = \delta^3(\vec{p}-\vec{q})$ as in non-relativistic Quantum Mechanics. The overlap of two momentum states is just a delta-function. The question is how these states transform under a Lorentz transformation. A Lorentz transformation on any state $|\psi\rangle$ gives a new state $|\psi'\rangle$ such that $|\psi\rangle \to |\psi'\rangle = \boldsymbol{U}(\Lambda)\,|\psi\rangle$ where $\boldsymbol{U}(\Lambda)$ is a unitary operator depending on the Lorentz transformation $\Lambda^\mu_{\ \nu}$ and acting on the original state $|\psi\rangle$. The transformation has to be unitary and satisfies therefore $\boldsymbol{U}^\dagger(\Lambda)\,\boldsymbol{U}(\Lambda) = \mathbf{1}$ so that it preserves the overlap between wave functions.

This unitary operation acting on $|\vec{p}\rangle$ gives $\boldsymbol{U}(\Lambda)\,|\vec{p}\rangle = |\vec{p}\,'\rangle$ because if $|\vec{p}\rangle$ describes a particle with momentum $\vec{p}$ the transformed state must describe a particle with the Lorentz-transformed momentum $\vec{p}\,'$. This is not quite right because $\vec{p}$ is a 3-momentum but can be turned into a 4-momentum $P^\mu = (E_{\vec{p}}, \vec{p})$ using $E_{\vec{p}} = +\sqrt{\vec{p}^{\,2} + m^2}$. If a Lorentz transformation acts on this 4-vector, then another 4-vector results with the same mass m but a different 3-momentum $\vec{p}$.

The transformation $\boldsymbol{U}(\Lambda)\,|\vec{p}\rangle = |\vec{p}\,'\rangle$ with 3-vectors is actually not unitary. Calculating $\langle\vec{q}|\boldsymbol{U}^\dagger(\Lambda)\,\boldsymbol{U}(\Lambda)|\vec{p}\rangle$ gives $\langle\vec{q}\,'|\,\vec{p}\,'\rangle = \delta^3(\vec{p}\,'-\vec{q}\,')$ because $\langle\vec{q}|\boldsymbol{U}^\dagger(\Lambda)| = \langle\vec{q}\,'|$ and $|\boldsymbol{U}(\Lambda)|\vec{p}\rangle = |\vec{p}\rangle$, but it should give $\langle\vec{q}|\mathbf{1}|\,\vec{p}\rangle$ and this would be $\langle\vec{q}\,|\,\vec{p}\rangle = \delta^3(\vec{p}-\vec{q})$.



The standard completeness relation $\int d^3\vec{p}\,|\vec{p}\rangle\,\langle\vec{p}| = \mathbf{1}$ is not Lorentz-invariant because it integrates only over the spatial components. Going to 4-momentum P^μ leads to $P^2 = m^2$ where $P^2 = P^\mu P_\mu = P_0^2 - \vec{p}^{\,2}$. The equation $P^2 = m^2$ is the Lorentz-invariant way of saying that the particle has mass m and corresponds to the hyperboloid shown in the figure on the right side. The Lorentz-invariant integration of the form

$$\int d^4P\,\delta(P^2-m^2)\,\theta(P^0)\,f(P) = \int d^3\vec{p}\,\frac{1}{2E_{\vec{p}}}\,f(E_{\vec{p}},\vec{p}) \tag{2.1}$$

only covers the upper part with $P_0 > 0$ because of $\theta(P^0)$. The factor $\theta(P^0)$ does not look manifestly Lorentz-invariant but Lorentz transformations preserve the sign of the time component of a 4-vector. The extra factor $1/(2E_{\vec{p}})$ comes from the fact that the argument of the delta-function is not just P^0. There is a general formula

$$\int dx\,\delta(g(x))\,f(x) = \sum_{g(x_n)=0} f(x_n)\,\frac{1}{|g'(x_n)|}$$

because the delta-function is only zero if $g(x_n)$ vanishes and where the extra factor $1/|g'(x_n)|$ comes from the Jacobian. Equation (2.1) allows replacing an integral $\int d^4P$ with an integral $\int d^3\vec{p}$.

The new states $|P\rangle$ with 4-vectors satisfy the completeness relation

$$\int \frac{d^3\vec{p}}{(2\pi)^3}\,\frac{1}{2E_{\vec{p}}}\,|P\rangle\,\langle P| = \mathbf{1} \tag{2.2}$$

where $(2\pi)^3$ is a purely conventional factor. Thus, the new states $|P\rangle$ can be expressed in terms of the conventional states $|\vec{p}\rangle$ as $|P\rangle = (2\pi)^{3/2}\sqrt{2E_{\vec{p}}}\,|\vec{p}\rangle$ such that $\langle Q|P\rangle = (2\pi)^3 2E_{\vec{p}}\,\delta^3(\vec{p}-\vec{q})$, for example, where the right side is basically a Lorentz-invariant delta-function. With this definitions one gets $\boldsymbol{U}(\Lambda)\,|P\rangle = |P'\rangle$ and $\langle Q|P\rangle = \langle Q'|P'\rangle$ where the primed states are the Lorentz-transformed states.

The Hamiltonian is not a covariant object either because it tells what the time evolution of the system is but time is not a priori well defined. What is needed is a 4-vector operator $\boldsymbol{P}^\mu$ acting on momentum eigenstates $|P\rangle$ such that $\boldsymbol{P}^\mu\,|P\rangle = P^\mu\,|P\rangle$ with the c-number eigenvalue P^μ. The Hamiltonian is the time component of the 4-momentum operator $\boldsymbol{P}^\mu$.

Lorentz transformations can act either on operators or on states but not on both. The matrix element $\langle Q|\boldsymbol{P}^\mu|P\rangle$ is supposed to transform like a 4-vector. If the states transform as $|\psi\rangle \to \boldsymbol{U}(\Lambda)\,|\psi\rangle$ and the operators stay the same $\boldsymbol{O} \to \boldsymbol{O}$ then $\langle Q|\boldsymbol{P}^\mu|P\rangle \to \langle Q'|\boldsymbol{P}^\mu|P'\rangle$ with $\langle Q'|\boldsymbol{P}^\mu|P'\rangle = P'^\mu\,\langle Q'|P'\rangle$ and with $\langle Q'|P'\rangle = \langle Q|P\rangle$. This is like the Schrödinger picture for time evolution. The states evolve with time and the operators do not. For Quantum Field Theory it is generally more useful to consider the Heisenberg picture where the states stay the same $|\psi\rangle \to |\psi\rangle$ and the operators transform $\boldsymbol{O} \to \boldsymbol{O'} = \boldsymbol{U}^\dagger(\Lambda)\boldsymbol{O}\boldsymbol{U}(\Lambda)$. In this case $\langle Q|\boldsymbol{P}^\mu|P\rangle \to \langle Q|\boldsymbol{U}^\dagger(\Lambda)\,\boldsymbol{P}^\mu\,\boldsymbol{U}(\Lambda)|P\rangle = \langle Q'|\boldsymbol{P}^\mu|P'\rangle$, and therefore $\langle Q|\boldsymbol{P}^\mu|P\rangle$ transforms in the Heisenberg picture as a 4-vector as it also does in the Schrödinger picture.

$\boldsymbol{U}(\Lambda)$ is called a unitary representation of the Lorentz group. The thinking is that Λ is the actual Lorentz transformation and $\boldsymbol{U}(\Lambda)$ is an operator which depends on it and implements the Lorentz transformations on either the states or the operators. This representation must have certain properties. If $\Lambda_1 \cdot \Lambda_2 = \Lambda_3$ then $\boldsymbol{U}(\Lambda_1)\cdot\boldsymbol{U}(\Lambda_2) = \boldsymbol{U}(\Lambda_3)$, and consequently also $\boldsymbol{U}(\Lambda^{-1}) = \boldsymbol{U}^{-1}(\Lambda)$ is required.

The particle states actually form a unitary representation of the Lorentz group. One can mathematically classify these unitary representations of the Lorentz group, and this was done by Eugene Wigner. These unitary representations which are all the possible particle states one can have are labeled by their mass and spin. One can understand in this way what all possible relativistic particles look like.

2.2 States of Multiple Non-Interacting Relativistic Quantum Particles

Restricted to the state of a single particle there is no way to have causality as shown above. The solution is to allow multiple particles with particle creation and annihilation. Thus, the first question is how to construct states with more than one particle.

The state $|P_1, P_2\rangle$ of two particles is supposed to be an eigenstate of the total momentum operator where the total momentum is $\boldsymbol{P}^\mu\,|P_1, P_2\rangle = (\boldsymbol{P}_1 + \boldsymbol{P}_2)^\mu\,|P_1, P_2\rangle$. Because two particles of the same kind such as two electrons are indistinguishable and have the same mass $P_1^2 = P_2^2 = m^2$, $|P_1, P_2\rangle$ and $|P_2, P_1\rangle$ are physically the same state. There are only the two possibilities

$$|P_1, P_2\rangle = \begin{cases} +\,|P_1, P_2\rangle & \text{for bosons} \\ -\,|P_1, P_2\rangle & \text{for fermions} \end{cases}$$

although one could imagine other possibilities such as that there are phases which might depend on the momenta or other things. However, one can prove that these two cases are the only possibilities in $3+1$ dimensions.

Restricting the discussion to bosons for the moment, the order of the momenta P_i in a multiparticle state $|P_1, ..., P_n\rangle$ labeled by n does not matter. The normalization for two states $\langle Q_1, Q_2|P_1, P_2\rangle$ is obviously $\langle Q_1, Q_2|P_1, P_2\rangle = \langle Q_1|P_1\rangle\,\langle Q_2|P_2\rangle + \langle Q_1|P_2\rangle\,\langle Q_2|P_1\rangle$. In other words, the two states $|Q_1, Q_2\rangle$ and $|P_1, P_2\rangle$ are orthogonal unless they have exactly the same pairs of 4-momenta such that either $|Q_1\rangle = |P_1\rangle$ and $|Q_2\rangle = |P_2\rangle$ or $|Q_1\rangle = |P_2\rangle$ and $|Q_2\rangle = |P_1\rangle$. One particle overlap is just the Lorentz-invariant delta-function. Thus, the normalization is Lorentz-invariant. For the normalization of n particle states $\langle Q_1, ..., Q_n|P_1, ..., P_n\rangle$ one has to sum over all possible permutations for which a state $|Q_i\rangle$ could be equal to a state $|P_j\rangle$ starting with $\langle Q_1|P_1\rangle\,...\,\langle Q_n|P_n\rangle$. This gives $n!$ terms of this form in the sum.

It is useful to find a bigger space of states which includes all the states with any number of particles. It includes the one-particle state $|P_1\rangle$, the two-particle states $|P_1, P_2\rangle$, the three-particle states $|P_1, P_2, P_3\rangle$ and so on in one big space called a Fock space. It also includes the no-particle state denoted by $|0\rangle$ which is defined by the property that it is normalized $\langle 0|0\rangle = 1$ and is orthogonal to all the other states in the Fock space. The zero-particle state is also called the vacuum state, and it will turn out as the ground state of Quantum Field Theory. Any n-particle state is orthogonal to any m-particle state for $n \neq m$.

2.3 Creation and Annihilation Operators

Given the Fock space with all possible states, operators are needed to act on these states. The most general set of operators can be written in terms of creation and annihilation operators. They are very analogous to the raising and lowering operator solution for the simple harmonic oscillator. The reason for this is that the theory for free quantum particles is actually equivalent to an infinite number of harmonic oscillators.

The creation and annihilation operators are defined as

$$|P_1, ..., P_n\rangle = \boldsymbol{\alpha}^\dagger(P_1) \, ... \, \boldsymbol{\alpha}^\dagger(P_n) \, |0\rangle \tag{2.3}$$

through the creation operator $\boldsymbol{\alpha}^\dagger(P)$. For example, $\boldsymbol{\alpha}^\dagger(K) \, |P_1, ..., P_n\rangle = |K, P_1, ..., P_n\rangle$ adds just one more, and the action on an n-particle state is define by (2.3). The definition of the creation operator $\boldsymbol{\alpha}^\dagger$ also determines the annihilation operator $\boldsymbol{\alpha}(P)$ completely because

$$\begin{aligned} \langle Q_1, ..., Q_m| \, \boldsymbol{\alpha}(K) \, |P_1, ..., P_n\rangle &= \langle P_1, ..., P_n| \, \boldsymbol{\alpha}^\dagger(K) \, |Q_1, ..., Q_m\rangle^* \\ &= \langle P_1, ..., P_n|K, Q_1, ..., Q_m\rangle^* \end{aligned}$$

given that $m = n - 1$. The explicit formula for the annihilation operator

$$\boldsymbol{\alpha}(K) \, |P_1, ..., P_n\rangle = \sum_{i=1}^{n} \langle K|P_i\rangle \; |P_1, ..., \widehat{P}_i, ..., P_n\rangle \tag{2.4}$$

follows where the first term in the sum means that K must be one of the P_i and $\widehat{P}_i$ indicates that this P_i has to be omitted. As a special case, applied to the zero-particle state $\boldsymbol{\alpha}(K) \, |0\rangle = 0$ the annihilation operator gives the zero state.

The matrix elements of the creation and annihilation operators are completely determined by the formula (2.3), and if one knows the matrix elements of an operator one knows everything about it. Therefore, one should be able to compute the commutators. Because the order of the P_i in (2.3) does not matter due to the Bose statistics, the order of the $\boldsymbol{\alpha}^\dagger(P_i)$ also does not matter, and the same is true for the $\boldsymbol{\alpha}(P_i)$. Thus, $[\boldsymbol{\alpha}^\dagger(P), \boldsymbol{\alpha}^\dagger(Q)] = [\boldsymbol{\alpha}(P), \boldsymbol{\alpha}(Q)] = 0$. However, $[\boldsymbol{\alpha}(P), \boldsymbol{\alpha}^\dagger(Q)] = \langle Q|P\rangle$ because

$$\boldsymbol{\alpha}(P)\boldsymbol{\alpha}^\dagger(Q) \, |P_1, ..., P_n\rangle = \boldsymbol{\alpha}(P) \, |Q, P_1, ..., P_n\rangle = \langle P|Q\rangle \, |P_1, ..., P_n\rangle + \sum_{i=1}^{n} \langle P|P_i\rangle \; |Q, P_1, ..., \widehat{P}_i, ..., P_n\rangle$$

$$\boldsymbol{\alpha}^\dagger(Q)\boldsymbol{\alpha}(P) \, |P_1, ..., P_n\rangle = \boldsymbol{\alpha}^\dagger(Q) \sum_{i=1}^{n} \langle P|P_i\rangle \; |P_1, ..., \widehat{P}_i, ..., P_n\rangle = \sum_{i=1}^{n} \langle P|P_i\rangle \; |Q, P_1, ..., \widehat{P}_i, ..., P_n\rangle$$

show nearly the same terms except for $\langle P|Q\rangle \, |P_1, ..., P_n\rangle$.

It is possible to write any operator in terms of the creation and annihilation operators because any operator is defined by its matrix elements and one could always put together combinations of the raising and lowering operators to get any matrix elements wanted. The creation and annihilation operators are therefore the building blocks for all possible operators.

States with different numbers of particles and operators to create and destroy particles have been introduced and it may seem to be inevitable if one follows the formalism that it forces the creation and destruction of particles, but this is not the case. Operators such as $(\boldsymbol{\alpha}^\dagger)^n(\boldsymbol{\alpha})^n \, |P_1, ..., P_m\rangle$ called n-body operators applied to an m-particle state will only have matrix elements between states with the same number of particles. In condensed matter physics this sort of formalism with creation and annihilation

operators and n-body operators is actually useful even in situations where one does not create and destroy particles.

As an example of this kind of operator one can look at the 4-momentum operator which has been introduced above as $\boldsymbol{P}^\mu |P\rangle = P^\mu |P\rangle$. With the obvious eigenvalue $(P_1 + ... + P_n)^\mu$ which is the sum of the 4-momenta of the individual particles, it can through $\boldsymbol{P}^\mu |P_1, ..., P_n\rangle = (P_1 + ... + P_n)^\mu |P_1, ..., P_n\rangle$ be extended to multiparticle states. One could define it this way, but one can also see what the operator $\boldsymbol{P}^\mu$ is in terms of creation and annihilation operators. It turns out to be

$$\boldsymbol{P}^\mu = \int \frac{d^4P}{(2\pi)^4}\, 2\pi\, \delta(P^2 - m^2)\, \theta(P^0)\, P^\mu \boldsymbol{\alpha}^\dagger(P)\boldsymbol{\alpha}(Q) = \int (dP)\, P^\mu \boldsymbol{\alpha}^\dagger(P)\boldsymbol{\alpha}(Q) \tag{2.5}$$

where the Lorentz-invariant measure $\frac{d^4P}{(2\pi)}{}^4 2\pi\delta(P^2 - m^2)\theta(P^0)$ will be abbreviated to (dP) in the following and can be replaced by the three-dimensional measure as in (2.1). The calculation

$$\boldsymbol{P}^\mu |P\rangle = \int (dK) K^\mu \boldsymbol{\alpha}^\dagger(K)\boldsymbol{\alpha}(K) |P\rangle = \int (dK) K^\mu \boldsymbol{\alpha}^\dagger(K) \langle K|P\rangle |0\rangle = P^\mu \boldsymbol{\alpha}^\dagger(P) |0\rangle = P^\mu |P\rangle$$

proves (2.5) for one-particle states and can be similarly shown for states with more than one particle.

2.4 Field Operators in Position Space

The fields in position space are

$$\varphi^+(X) = \int (dP)\, \boldsymbol{\alpha}(P)\, e^{-iP\cdot X} \qquad\qquad \varphi^-(X) = \int (dP)\, \boldsymbol{\alpha}^\dagger(P)\, e^{+iP\cdot X} \tag{2.6}$$

by using the Lorentz-invariant measure (dP) and by Fourier-transforming $\boldsymbol{\alpha}(P)$. The notation $\varphi^+(x)$ has possibly been introduced by Pauli a long time ago and the plus sign is related to the fact that $\varphi^+(x)$ is a positive frequency due to $e^{-iP\cdot X} = e^{-iE_{\vec{p}}t + \cdots}$ where the negative sign is the "correct" sign for positive propagation in Quantum Mechanics. Obviously, $\varphi^-(X) = [\varphi^+(X)]^\dagger$.

Since $\boldsymbol{\alpha}^\dagger(P)$ and $\boldsymbol{\alpha}(P)$ are operators $\varphi^+(X)$ and $\varphi^-(X)$ are operators as well. Thus, the fields in position space are time-dependent operators and one works in the Heisenberg picture where the operators change under Lorentz transforms and not the states. But because the Schrödinger picture is more familiar, this picture where the states change $|P\rangle \to |\Lambda P\rangle$ under Lorentz transformations while the operators $\boldsymbol{\alpha}(P)$ stay the same is used to show that the annihilation operator transform as $\boldsymbol{\alpha}(P) \to \boldsymbol{\alpha}(\Lambda^{-1}P)$ in the Heisenberg picture. The matrix elements

$$\langle 0|\boldsymbol{\alpha}(P)|Q\rangle = \langle P|Q\rangle \langle 0|0\rangle = \langle P|Q\rangle$$

transform in the Schrödinger picture (SP) and in the Heisenberg picture (HP) as

$$\begin{aligned} \langle 0|\boldsymbol{\alpha}(P)|Q\rangle &\xrightarrow{\text{SP}} \langle 0|\boldsymbol{\alpha}(P)|\Lambda Q\rangle = \langle P|\Lambda Q\rangle \\ \langle 0|\boldsymbol{\alpha}(P)|Q\rangle &\xrightarrow{\text{HP}} \langle 0|\boldsymbol{\alpha}(\Lambda^{-1}P)|Q\rangle = \langle \Lambda^{-1}P|Q\rangle \end{aligned}$$

and this is the same because firstly the zero-particle state transforms into itself and secondly $P = \Lambda Q$ following from $\langle P|\Lambda Q\rangle$ and $\Lambda^{-1}P = Q$ following from $\langle \Lambda^{-1}P|Q\rangle$ are the same condition. It follows that $\langle P|Q\rangle = \langle P'|Q'\rangle$ where the primed states are the Lorentz-transformed states.

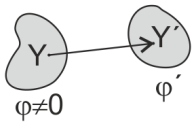
Sticking to the Heisenberg picture the operators $\varphi^+(X)$ in position space (2.6) transform as

$$\varphi^+(X) = \int (dP)\, e^{-iP\cdot X}\, \boldsymbol{\alpha}(P) \to \int (dP)\, e^{-iP\cdot X}\, \boldsymbol{\alpha}(P') = \int (dP') e^{-iP'\cdot(\Lambda^{-1}X)}\, \boldsymbol{\alpha}(P') = \varphi^+(\Lambda^{-1}X)$$

using $P' = \Lambda^{-1}P$, $P = \Lambda P'$, $P \cdot X = (\Lambda P') \cdot X = P' \cdot (\Lambda^{-1}X)$ and renaming the integration variable in the last step from P' to P. The step $(\Lambda P) \cdot X = P \cdot (\Lambda^{-1}X)$ in matrix notation is

$$(\Lambda P) \cdot X = (\Lambda P)^\mu \eta_{\mu\nu} X^\nu = (\Lambda P)^T \eta X = P^T \Lambda^T (\Lambda^{-1T} \eta \Lambda^{-1}) X = P^T (\Lambda^T \Lambda^{-1T}) \eta \Lambda^{-1} X = P^T \eta \Lambda^{-1} X$$

using the identity $\eta_{\mu\nu} = (\Lambda^{-1})^\rho{}_\mu\, (\Lambda^{-1})^\sigma{}_\nu\, \eta_{\rho\sigma}$ or $\eta = \Lambda^{-1T}\, \eta\, \Lambda^{-1}$ due to the Lorentz invariance of η.



Thus, transformation property $\varphi^+(X) \to \varphi^+(\Lambda^{-1}X)$ and the fact that there is $\Lambda^{-1}X$ and not ΛX can be seen also for a classical field φ. Given a special point Y in spacetime where the field φ is not zero in a region around it, a Lorentz transformation Λ will take this point Y to some new point $Y' = \Lambda Y$ and the field configuration φ localized around Y to a new field configuration φ' localized around Y' as shown in the figure on the right side. The new field configuration φ' evaluated at Y' is supposed to be the same as the old field configuration φ evaluated at Y such that $\varphi'(Y') = \varphi(Y)$, but this has to be true for every point X showing that $\varphi'(X') = \varphi(X)$ for $X' = \Lambda X$ and therefore $\varphi'(X) = \varphi(\Lambda^{-1}X)$.

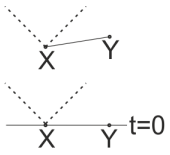
To summarize, fields have been defined by $\varphi^+(X)$ and $\varphi^-(X)$ which are functions of spacetime and they transform by Lorentz transformations according to

$$\varphi^+(X) \to \varphi^+(\Lambda^{-1}X) \qquad\qquad \varphi^-(X) \to \varphi^-(\Lambda^{-1}X) \tag{2.7}$$

which is actually the way that any scalar field transforms in spacetime. Since these fields are nothing but the Fourier-transformed $\boldsymbol{\alpha}(P)$ and $\boldsymbol{\alpha}^\dagger(P)$, any operator in this Fock space can be written in terms of the operators $\varphi^+(X)$ and $\varphi^-(X)$.

2.5 Resolution of the Causality Problem

The basic idea is that the theory places restrictions on what can be measured. Any measurement is an interaction, and the experimentalist together with the apparatus is part of the physical world. Thus, what can be measured is determined by the interactions possible, and these interactions are determined by what terms appear in the Hamiltonian. The operators that can appear in the Hamiltonian are called observables and are Hermitian operators $\boldsymbol{O}_i(X)$ in general associated with a spacetime point X.



The crucial property that these observables must have in order to avoid problems with causality is $[\boldsymbol{O}_i(X), \boldsymbol{O}_j(Y)] = 0$ for spacelike separated points X and Y. Spacelike separated points X and Y can be transformed into a reference frame such that they lie on the same time surface as shown in the figure on the right side. The fact that two operators $\boldsymbol{O}_i(X)$ and $\boldsymbol{O}_j(Y)$ commute means that they are independent operators.

If the two operators $\boldsymbol{O}_i(X)$ and $\boldsymbol{O}_j(Y)$ do not commute means that measurements at X necessarily influence measurements at Y in a way that violates causality. This point is a little bit subtile because ordinary Quantum Mechanics does allow measurements at spacelike separated points to influence each other but not in a way that violates causality.

An example is the classic setup by Einstein, Podolsky, Rosen (EPR). They considered a situation of two different particles with spin s_1 and s_2 separated spatially but prepared in a state where the two spins are correlated. The spin state of the system could be, for example,

$$|\psi\rangle = \frac{1}{\sqrt{2}}\big(\,|\uparrow\rangle\,|\downarrow\rangle - |\downarrow\rangle\,|\uparrow\rangle\,\big)$$

forming a spin singlet. If the experimentalists Alice and Bob measures these two spins at prearranged times (and therefore at spacelike separated spacetime points) then Alice's measurement of s_1 seems to influence Bob's measurement of s_2 in the sense that if Alice measures spin up then Bob measures spin down or vice versa. However, there is no possibility to communicate faster than light in this way because whatever Alice does, Bob will always see fifty percent of the time up and fifty percent of the time down. Bob cannot tell what Alice has done.

Only when the measurements have been done, Alice and Bob can come together or exchange notes and compare their measurements to find that there is a hundred percent anticorrelation between them. Thus, there is no violation of causality. The claim here is that the basic reason why there is no causality problem is $[s_1, s_2] = 0$, or in other words, that s_1 and s_2 commute. These two operators are spins of two different particles and act on different parts of the space.

To understand this point better, one can try to imagine a situation where Alice and Bob can measure (at spacelike separated points in spacetime) two operators that do not commute with each other. This cannot happen, but one can still try to imagine such a situation. For example, if Alice can measure the

spin s_z and Bob can measure the spin s_x which are supposed to be different components of the spin of the same particle. The spin is assumed to be prepared in the state $|\psi\rangle = |s_z = +1/2\rangle$. If Bob tries to communicate with Alice then there are two situations. If Bob does nothing then Alice will a hundred percent of the time measure $s_z = +1/2$. If Bob turns on his apparatus and measures then he will measure fifty percent of the time $s_x = +1/2$ and fifty percent of the time $s_x = -1/2$ and therefore the state will change to either $s_x = +1/2$ or $s_x = -1/2$ such that Alice will see fifty percent of the time $s_z = +1/2$ and fifty percent of the time $s_z = -1/2$. In the first situation Alice always gets $s_z = +1/2$ and in the second situation she gets half of the times $+1/2$ and half the times $-1/2$, and this is a direct consequence of $[s_x, s_z] \neq 0$.

The question is whether the field operators $\varphi^{\pm}(X)$ defined in (2.6) are observables. It is easy to see that $[\varphi^+(X), \varphi^+(Y)] = 0$. They always commute whether they are spacelike separated or not. The basic reason is that the annihilation operators commute with each other. The same is true for $\varphi^-(X)$ such that $[\varphi^-(X), \varphi^-(Y)] = 0$. The remaining case

$$\begin{aligned}
[\varphi^+(X), \varphi^-(Y)] &= \int (dP) e^{-iP\cdot X} \int (dP) e^{+iQ\cdot Y} [\boldsymbol{\alpha}(P), \boldsymbol{\alpha}^\dagger(Q)] \\
&= \int (dP) e^{-iP\cdot X} \int (dP) e^{+iQ\cdot Y} \langle Q|P\rangle \\
&= \int (dP) e^{-iP\cdot(X-Y)} = \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} e^{-iP\cdot(X-Y)}
\end{aligned}$$

shows that $\varphi^+(X)$ and $\varphi^-(Y)$ do not commute with each other because the Lorentz-invariant delta-function $\langle Q|P\rangle$ can be used to calculate the integral over (dQ). The resulting integral that is called $\Delta_+(X-Y)$ looks very similar to the Green's function (1.1) when using (2.1), and one can apply the same argument as had been used above to show that it is not zero. This gives

$$\begin{aligned}
\Delta_+(X) &= \frac{1}{(2\pi)^3} \int_0^\infty p^2\, dp \int_{-1}^{+1} d\, \cos\theta\, 2\pi\, \frac{1}{2E_{\vec{p}}}\, e^{-i(E_{\vec{p}}t - pr\, \cos\theta)} \\
&= \frac{1}{4\pi^2} \int_0^\infty p^2\, dp\, \frac{1}{2E_{\vec{p}}}\, \frac{\sin(pr)}{pr}\, e^{-i\, E_{\vec{p}}t}
\end{aligned}$$

with $X^\mu = (t, \vec{r})$. Because $e^{-i\, E_{\vec{p}}t}$ is an analytic function with $\mathrm{Im}(t) < 0$ becomes a damped exponential, it cannot vanish over an integral of t without vanishing identically. Thus, $\Delta_+(X)$ does not vanish outside the light cone.

2.6 Building Observables from the Field Operators

The two field operators $\varphi^+(X)$ and $\varphi^-(X)$ are not observables because they are not Hermitian but one can build observables out of them. The simplest example is $\varphi^+(X) + \varphi^-(X)$. It is Hermitian because it is $\varphi^+(X)$ plus its Hermitian conjugate $\varphi^-(X)$. One can generalize this to $\varphi(X) = e^{+i\theta}\, \varphi^+(X) + e^{-i\theta}\, \varphi^-(X)$, and one could also multiply with a real number. Because $\Delta_+(X-Y)$ is clearly an even function

$$\Delta_+(X-Y) \propto \int_0^\infty p^2\, dp\, \frac{\sin(pr)}{pr}$$

when choosing the spacelike points X and Y such that $(X-Y)^\mu = (0,0,0,r)$, the commutator

$$[\varphi(X), \varphi(Y)] = \Delta_+(X-Y) - \Delta_+(Y-X)$$

vanishes for spacelike separated X and Y, and the operator $\varphi(X)$ is indeed an observable. (This fact could have been understood without any calculations because Δ_+ is Lorentz-invariant and Lorentz transformations take $X-Y$ to minus itself.) However, the commutator does not vanish for timelike separated X and Y because

$$\Delta_+(X-Y) \propto \int_0^\infty p^2\, dp\, \frac{e^{-i\, E_{\vec{p}}t}}{E_{\vec{p}}} \neq \Delta_+(Y-X)$$

when choosing the timelike points X and Y such that $(X-Y)^\mu = (t,0,0,0)$. (This is due to the fact that there is no proper Lorentz transformation that takes t to $-t$.)

Thus, the particular combination $\varphi(X) = e^{+i\theta}\,\varphi^+(X) + e^{-i\theta}\,\varphi^-(X)$ of the field operators is an observable. This works for any value of θ, but the phases $e^{\pm i\theta}$ can always be absorbed into the $\boldsymbol{\alpha}$ by $\boldsymbol{\alpha} \to e^{i\theta}\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^\dagger \to e^{-i\theta}\boldsymbol{\alpha}^\dagger$. This amounts to rephasing all of the states and does not change anything. The operator $\varphi(X)$ is therefore defined as $\varphi(X) = \varphi^+(X) + \varphi^-(X)$.

One can use the observable $\varphi(X)$ to create an infinite number of new observables by taking local functions of $\varphi(X)$ and its derivatives such as $\varphi^2(X)$, $\varphi^3(X)$, $\partial_\mu\varphi(X)$ and so on. They all contain creation and annihilation operators because $\varphi(X)$ contains both. This allows to construct Hamiltonians and interacting theories.

The combination $\varphi(X) = \varphi^+(X) + \varphi^-(X)$ is the simplest combination of $\varphi^+(X)$ and $\varphi^-(X)$. Thus, the question arises whether it is possible to create other combinations which are observables. One may try, for example, $\boldsymbol{O}(X) = \varphi^-(X)\,\varphi^+(X)$, but the commutator

$$[\boldsymbol{O}(X), \boldsymbol{O}(Y)] = \Delta_+(X-Y)\varphi^-(X)\,\varphi^+(Y) - \Delta_+(Y-X)\varphi^-(Y)\,\varphi^+(X)$$

would not be zero. It is assumed that it can be proven that all possible observables can be made out of $\varphi(X)$ by taking local functions of it.

To summarize, $\varphi(X)$ defined as

$$\varphi(X) = \int (dP)\big[\boldsymbol{\alpha}(P)\,e^{-iP\cdot X} + \boldsymbol{\alpha}^\dagger(P)\,e^{+iP\cdot X}\big] \tag{2.8}$$

according to (2.6) is the most general solution of the Klein-Gordon equation (1.2) $(\Box + m^2)\varphi(X) = 0$. Because $\Box$ only acts on X in the exponentials, the Klein-Gordon equation is satisfied by $\varphi(X)$. It is the most general solution because of the arbitrary operator coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^\dagger$. On the other hand, also $\varphi^+(X)$ and $\varphi^-(X)$ satisfy the Klein-Gordon equation, but $\varphi^+(X)$, for example, is not the most general solution because it contains only the positive frequency part. Thus, $\varphi(X)$ is a quantum field.

3 From Fields to Particles

3.1 Solutions of the Klein-Gordon Equation

It is possible to reverse the process and start with a quantum field and show how that describes particles at least for a non-interacting (or a linear) field equation. In the previous chapter non-interacting spin 0 bosons have been described with the quantized field operator $\varphi(X)$ satisfying the Klein-Gordon equation (1.2) $(\Box + m^2)\varphi(X) = 0$ where $\varphi(X)$ is a Heisenberg operator that acts on Fock space.

Ehrenfest's theorem states that the expectation values in Quantum Mechanics obey their classical equation of motion. Because the only time-dependence in the Heisenberg picture is in the operators and the states do not depend on time, the fact that the quantum field obeys the classical equation of motion here can be thought of as Ehrenfest's theorem.

The Klein-Gordon equation is a linear differential equation of a simple type and can therefore be solved. With

$$\int (dP)\,\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X}$$

plus its Hermitian conjugate, the general solution is an arbitrary linear combination of the exponentials representing plane waves with operator coefficients $\boldsymbol{\alpha}(P)$. Here the $\boldsymbol{\alpha}(P)$ are arbitrary but if $\varphi(X)$ is supposed to act on Fock space then the commutator relation $[\boldsymbol{\alpha}(P), \boldsymbol{\alpha}^\dagger(Q)] = \langle P|Q\rangle$ written in a Lorentz-invariant form is required.

From the point of view of the field theory the Klein-Gordon equation is a classical field equation, one promotes $\varphi(X)$ to an operator and then $\varphi(X)$ with operator coefficients $\boldsymbol{\alpha}(P)$ is the most general solution. Thus, the question from the field point of view is, where this commutator relation comes from. The claim is that this is equivalent to imposing the canonical commutation relations on the field.

Starting from the general solution $\varphi(X)$ of the Klein-Gordon equation and the commutation relations $[\boldsymbol{\alpha}(P), \boldsymbol{\alpha}^\dagger(Q)] = \langle P|Q\rangle$, one can compute the commutation relation of $\varphi(X)$ and derivatives of $\varphi(X)$ to

get

$$[\partial_\mu \varphi(X), \varphi(Y)] = -2i \int (dP)\, P_\mu \cos\left(P \cdot (X - Y)\right)$$

by using the commutation relations of the operator coefficients $\boldsymbol{\alpha}(P)$. Evaluating this for spacelike separated X and Y, one can choose $X^0 = Y^0 = t$ and compute ∂_0. This gives

$$[\partial_0 \varphi(\vec{x}, t), \varphi(\vec{y}, t)] = -2i \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} E_{\vec{p}} \cos(\vec{p} \cdot (\vec{x} - \vec{y})) = -i \int \frac{d^3\vec{p}}{(2\pi)^3} \cos(\vec{p} \cdot (\vec{x} - \vec{y})) = -i\delta^3(\vec{x} - \vec{y})$$

and the result looks very similar to a canonical commutation relation. It is the commutator of a momentum and a position evaluated at equal time. The same commutator still at equal time vanishes

$$[\partial_i \varphi(\vec{x}, t), \varphi(\vec{y}, t)] = -2i \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{P^i}{2E_{\vec{p}}} E_{\vec{p}} \cos(\vec{p} \cdot (\vec{x} - \vec{y})) = 0$$

when evaluated for a spatial derivative $\mu = i = 1, 2, 3$ because cos is an even function and P^i is an odd function. Note that $[\partial_0 \varphi(\vec{x}, t), \varphi(\vec{y}, t)] = -i\delta^3(\vec{x} - \vec{y})$ is local because this commutator vanishes for spacelike separation except at coincident points.

3.2 Creation and Annihilation Operator for a Free Field

The development here has started from commutation relations of the operators $\boldsymbol{\alpha}$ and derived the commutation relations of operators φ. One can turn this around and show that imposing the commutation relations of the operators φ gives the commutation relations of the operators $\boldsymbol{\alpha}$.

The Klein-Gordon inner product (f, g) which is not really an inner product despite its name takes two complex scalar functions f and g of space and time and gives a number. It is defined as

$$(f, g) = i \int d^3\vec{x}\, f^*(\vec{x}, t) \overset{\leftrightarrow}{\partial_0} g(\vec{x}, t) \tag{3.1}$$

where the derivative $\overset{\leftrightarrow}{\partial_0}$ means $f^* \overset{\leftrightarrow}{\partial_0} g = f^*\dot{g} - \dot{f}^* g$. It is an integral over space only, and the functions f and g are evaluated at the same time t.

The Klein-Gordon inner product (3.1) has the property that it is time-independent if both f and g are solutions of the Klein-Gordon equation (1.2). In order to prove this it is shown that its time-derivative vanishes

$$\begin{aligned}
\partial_t (f, g) &= i \int d^3\vec{x} \left[\dot{f}^*\dot{g} + f^*\ddot{g} - \ddot{f}^* g - \dot{f}^*\dot{g}\right] = i \int d^3\vec{x} \left[f^*\ddot{g} - \ddot{f}^* g\right] \\
&= i \int d^3\vec{x} \left[f^*(\vec{\nabla}^2 - m^2) g - (\vec{\nabla}^2 - m^2) f^* g\right] = 0
\end{aligned}$$

using the Klein-Gordon equation in the form $\ddot{f} = (\vec{\nabla}^2 - m^2) f$. The terms with m^2 obviously vanish, and the terms with $\vec{\nabla}^2$ also vanish because one can integrate by parts twice.

The Klein-Gordon inner product (3.1) also has another useful property

$$\begin{aligned}
&(e^{-i\,P\cdot X}, e^{-i\,Q\cdot X}) = \\
&(P_0 + Q_0)\,(2\pi)^3\,\delta^3(\vec{p} - \vec{q})\, e^{-i(P_0 - Q_0)t} \begin{cases} \overset{P_0 = Q_0}{=} \pm 2E_{\vec{p}}\,(2\pi)^3\,\delta^3(\vec{p} - \vec{q}) \begin{cases} \overset{P_0 \geq 0}{=} \langle P|Q\rangle \\ \overset{P_0 \leq 0}{=} -\langle P|Q\rangle \end{cases} \\ \overset{P_0 = -Q_0}{=} 0 \end{cases}
\end{aligned}$$

when applied to plane-wave solutions. (The minus sign in $e^{-i\,P\cdot X}$ has been chosen to get $P_0 > 0$ corresponding to positive frequency.) This is not really an inner product because it is not positive definite but it shows that positive frequency and negative frequency solutions are orthogonal to each other. Positive frequency has positive norm and negative frequency has negative norm. The result $\langle P|Q\rangle$

is Lorentz-invariant even thought the definition of the Klein-Gordon inner product does not look like but can be shown to be Lorentz-invariant.

One can use this result to project out the creation and the annihilation operators from the free field. Taking this inner product gives

$$\left(e^{-i\,P\cdot X},\varphi(X)\right)=\boldsymbol{\alpha}(P)\qquad\qquad\left(e^{+i\,P\cdot X},\varphi(X)\right)=-\boldsymbol{\alpha}^\dagger(P)$$

assuming $P_0>0$, and the commutation relations become

$$\begin{aligned}[\boldsymbol{\alpha}(P),\boldsymbol{\alpha}^\dagger(Q)]&=\int d^3\vec{x}\int d^3\vec{y}\,[e^{+i\,P\cdot X}\overset{\leftrightarrow}{\partial}_0\varphi(X),e^{-i\,Q\cdot Y}\overset{\leftrightarrow}{\partial}_0\varphi(Y)]\\&=\int d^3\vec{x}\int d^3\vec{y}\,\{-iP_0\,[\varphi(X),\dot\varphi(Y)]+iQ_0\,[\dot\varphi(X),\varphi(Y)]\}\;e^{+i\,P\cdot X}\,e^{-i\,Q\cdot Y}\\&=2E_{\vec{p}}\,(2\pi)^3\,\delta^3(\vec{p}-\vec{q})=\langle P|Q\rangle\end{aligned}$$

with $X^0=Y^0=t$ such that X and Y are spacelike separated. This shows that the commutation relations $[\varphi(\vec{x},t),\dot\varphi(\vec{y},t)]=i\delta(\vec{x}-\vec{y})$ imply the commutation relations $[\boldsymbol{\alpha}(P),\boldsymbol{\alpha}^\dagger(Q)]$.

3.3 Quantization as a Procedure

The commutation relations between the field operators $\varphi(X)$ follow from applying the rules of what is usually called quantization. Actually quantization does not make sense if it is seen as the process that starts from a classical theory and somehow derives from it a quantum theory. This does not make sense because Quantum Mechanics is more fundamental than Classical Mechanics which only appears as a limit of Quantum Mechanics. Certain quantum states behave approximately classical. The theory of quantum theory is fundamental and one should be deriving the classical theory from it and not the other way around. Thus, quantization is a procedure to define a quantum theory that has a given classical theory as its classical limit.

Even that is not quite right. What is actually more important is that the quantum theory to be defined has the same symmetries as the classical theory. Most notably Lorentz invariance is the key symmetry. The theory developed so far covers bosons. When looking at fermions then there is no classical limit for a fermion field. There is no classical electron field the way there is a classical electromagnetic field. Quantizing a fermion field uses anticommutators, and one has to make up a new rule of quantization. The only point of this procedure is that it gives a theory that has the same symmetries.

Reviewing the standard procedure for quantization for bosons shows that one starts with a classical action S which is defined in terms of a Lagrangian L and derives a Hamiltonian H from it which is in a first step a classical Hamiltonian. Finally, one derives a quantum Hamiltonian from the classical Hamiltonian by promoting the quantities that appear in the classical Hamiltonian to quantum operators. There is another way of formulating a quantum field theory by path integrals but it is much less familiar. Thus, this is still the best way to make contact with what is known and to set up a general quantum field theory.

Starting from the classical variables $q^a(t)$ with a going from 1 to the number N of classical degrees of freedom one assumes an action S given as

$$S=\int_{t_i}^{t_f}dt\,L(q,\dot{(}q))$$

of a Lagrangian L. The classical equations of motion arise from extremizing S subject to the conditions $q(t_i)=q_i$ and $q(t_f)=q_f$ for fixed quantities q_i and q_f.

If one thinks of the quantities q as points in space, extremizing S means looking at all the possible paths between q_i and q_f in order to find the one that extremises the action. This path is the one that satisfies the classical equations of motion. To extremize the integral one takes an arbitrary trajectory $q^a(t)$ and replaces it by the trajectory plus a small perturbation $\delta q^a(t)$. This perturbation has to vanish on both ends of the paths such that $\delta q^a(t_i)=\delta q^a(t_f)=0$ because the

endpoints are fixed. Extremizing the action means taking δS as the linear term in $\delta q^a(t)$, and $\delta S = 0$ gives the classical equations of motion. Doing so results in

$$\delta S = \int_{t_i}^{t_f} dt \, \left[\frac{\partial L}{\partial q^a} \delta q^a + \frac{\partial L}{\partial \dot{q}^a} \partial_t \delta q^a \right] = \int_{t_i}^{t_f} dt \, \left[\frac{\partial L}{\partial q^a} - \partial_t \frac{\partial L}{\partial \dot{q}^a} \right] \delta q^a$$

with integration of $\partial_t \delta q^a$ by parts. (This step produces a surface term evaluated at the endpoints that vanishes because the variation $\delta q^a(t)$ vanishes at the endpoints.) In order to extremize S for general δq^a the condition

$$\frac{\partial L}{\partial q^a} - \partial_t \frac{\partial L}{\partial \dot{q}^a} = 0$$

must hold gives the equations of motion.

One can apply this to a scalar field theory. The only thing that changes is the notation. Instead of $q^a(t)$ depending on time as a degree of freedom there is $\varphi(\vec{x}, t)$ as a scalar depending on space and time. It is one degree of freedom for each spatial point $\vec{x}$. One possible action is

$$S = C \int dt \int d\vec{x} \, \left[\frac{1}{2} \, \partial^\mu \varphi \partial_\mu \varphi - V(\varphi) \right] \tag{3.2}$$

with an arbitrary constant C. The two integrals do not have boundary conditions, and nobody is carefully about that. It is assumed here that the boundary conditions are such that the fields fall off fast enough at infinity (space and time) to allow to freely integrate by parts.

Because $\partial^\mu \varphi \partial_\mu \varphi = \eta^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi = \dot{\varphi}^2 - (\vec{\nabla} \varphi)^2$ and the two integrals can be written as one integral $\int d^4X$, this looks Lorentz-invariant. Due to (2.7) a Lorentz transformation Λ acts as $\varphi(X) \to \varphi'(X) = \varphi(\Lambda^{-1} X)$ on $\varphi(X)$. Inserting this into equation (3.2) with $Y = \Lambda^{-1} X$ and S' split into $S' = S'_1 - S'_2$ gives

$$\begin{aligned}
S'_1 &= C \int d^4X \, \frac{1}{2} \, \partial^\mu \varphi \, \partial_\mu \varphi = C \int d^4X \, \frac{1}{2} \, \eta^{\mu\nu} \frac{\partial}{\partial X^\mu} \varphi'(X) \, \frac{\partial}{\partial X^\nu} \varphi'(X) \\
&= C \int d^4X \, \frac{1}{2} \, \eta^{\mu\nu} \frac{\partial}{\partial X^\mu} \varphi(Y) \, \frac{\partial}{\partial X^\nu} \varphi(Y) = C \int d^4X \, \frac{1}{2} \, \eta^{\mu\nu} \frac{\partial Y^\rho}{\partial X^\mu} \frac{\partial}{\partial Y^\rho} \varphi(Y) \, \frac{\partial Y^\sigma}{\partial X^\nu} \frac{\partial}{\partial Y^\sigma} \varphi(Y) \\
&= C \int d^4Y \, \frac{1}{2} \, \eta^{\mu\nu} (\Lambda^{-1})^\rho{}_\mu \, \partial_\rho \, \varphi(Y) \, (\Lambda^{-1})^\sigma{}_\nu \, \partial_\sigma \, \varphi(Y) = C \int d^4Y \, \frac{1}{2} \, \eta^{\mu\nu} (\Lambda^{-1})^\rho{}_\mu \, (\Lambda^{-1})^\sigma{}_\nu \, \partial_\rho \, \varphi(Y) \, \partial_\sigma \, \varphi(Y) \\
&= C \int d^4Y \, \frac{1}{2} \, \eta^{\rho\sigma} \, \partial_\rho \, \varphi(Y) \, \partial_\sigma \, \varphi(Y) \\
S'_2 &= C \int d^4Y \, V(\varphi(Y))
\end{aligned}$$

and therefore $S = S'$ because Y, ρ, σ are just bound variables. The action S is truly Lorentz-invariant.

To get the equations of motion for the action S in (3.2) one takes $\varphi(X) \to \varphi(X) + \delta\varphi(X)$. The linear term is

$$\delta S = C \int d^4X \, \left[\frac{1}{2} \, 2 \, \partial^\mu \varphi \, \partial_\mu (\delta\varphi) - V'(\varphi) \delta\varphi \right] = C \int d^4X \big[- \Box\varphi - V'(\varphi) \big] \, \delta\varphi$$

where $\Box = \eta^{\mu\nu} \partial_\mu \partial_\nu$. Thus, the field equation is

$$\Box\varphi + V'(\varphi) = 0 \tag{3.3}$$

similarly to the classical equations of motion. For the special case that $V = \frac{1}{2} m^2 \varphi^2$, the field equation is the Klein-Gordon equation (1.2), but V can be different and also contain higher powers of φ. (Note that the equations of motion do not depend on C.)

3.4 From the Lagrangian Formulation to the Hamiltonian Formulation

The classical Hamiltonian formalism defines for every generalized coordinate $q^a(t)$ a generalized momentum $p_a = \frac{\partial L}{\partial \dot{q}^a}$ to allow to express $\dot{q}^a = f^a(p, q)$. The Hamiltonian is defined as

$$H = \big[p_a \, \dot{q}^a - L(q, \dot{q}) \big]_{\dot{q} = f(p,q)}$$

in terms of the Lagrangian L and using the summation convention. The Hamiltonian has its own equations of motion not shown here.

Applying this to the scalar field theory defined by the action S (3.2) with using the momenta in the form $p_a(t) = \delta S/\delta q^a(t)$ gives the conjugate momentum

$$\pi(X) = \frac{\delta}{\delta\dot\varphi(X)}\, C \int d^4Y \left[\frac{1}{2}\,\dot\varphi^2(Y) - \frac{1}{2}\,(\vec\nabla\varphi(Y))^2 - V(\varphi(Y))\right] = C\,\dot\varphi(X)$$

and the Hamiltonian

$$\mathcal{H} = \int d^3\vec{x}\left\{\frac{1}{2C}\pi^2(X) + C\left[\frac{1}{2}(\vec\nabla\varphi)^2 + V(\varphi)\right]\right\}$$

which contrary to the action does not look at all Lorentz-invariant. However, the action and the Hamiltonian contain the same information and lead to the same equations of motion. When starting with the Hamiltonian one defines momenta which have to do with time derivatives and therefore one must have picked a direction of time. That is why the Lorentz invariance gets lost. The constant C appears in a non-trivial way and would disappear when setting $C = 1$, but there is no reason so far to do so.

In a quantum theory the coordinates q^a and the momenta p_a have to be turned into operators. These operators are supposed to satisfy the canonical commutation relations $[q^a, p_b] = -i\delta^a_{\ b}$. In the Schrödinger picture the states evolve with time $|\psi(t)\rangle_S = \boldsymbol{U}(t)\,|\psi(0)\rangle_S$ with a unitary time-evolution operator $\boldsymbol{U}(t)$ while the operators $\boldsymbol{O}_S$ are independent of time. In the Heisenberg pictures the states are independent of time and can be written as $|\psi\rangle_H = \boldsymbol{U}^\dagger(t)\,|\psi(t)\rangle_S = |\psi(0)\rangle_S$. The matrix elements of operators $\boldsymbol{O}_H$ must be defined to agree with the Schrödinger picture because both pictures try to describe the same physics. Thus, they are $\boldsymbol{O}_H(t) = \boldsymbol{U}^\dagger(t)\,\boldsymbol{O}_S\,\boldsymbol{U}(t)$. The fact that the two pictures have been chosen to agree at time $t = 0$ is arbitrary, of course.

The commutation relations $[q^a, p_b] = -i\delta^a_{\ b}$ are represented in the Schrödinger picture. In the Heisenberg picture they are the same $[q^a(t), p_b(t)] = -i\delta^a_{\ b}$ but the time-dependent operators have to be imposed at equal time t.

In Quantum Field Theory one is nearly always working in the Heisenberg picture. Thus, the commutation relations $[\pi(\vec{x},t), \varphi(\vec{y},t)] = i\delta^3(\vec{x}-\vec{y})$ at equal time become

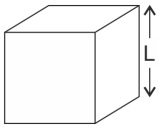
$$[\dot\varphi(\vec{x},t), \varphi(\vec{y},t)] = \frac{i}{C}\,\delta^3(\vec{x}-\vec{y}) \tag{3.4}$$

because $\pi(\vec{x},t) = C\,\dot\varphi(\vec{x},t)$. The arbitrary constant C normalizes the action, and it did not matter from the classical equations of motion because it factors out. Choosing different values C here just means rescaling the states because these commutation relations are for a free field theory and are related to the creation and annihilation operators. Thus, choosing different values C give different normalizations of the states, and the so-called canonical norm $C = 1$ will be used in the following.

3.5 Relation Between Free Fields and Harmonic Oscillators

Using the simple action

$$S = \int d^4X \left[\frac{1}{2}\,\dot\varphi^2 - \frac{1}{2}\,(\vec\nabla\varphi)^2 - \frac{1}{2}\,m^2\,\varphi^2\right]\omega$$



one can see the connection to the harmonic oscillator from the fact that the action is quadratic in the fields. To determine that explicitly the system is put into a box of size L where the field has periodicity L such that $\varphi(\vec{x},t) = \varphi(\vec{x}+L\hat{e},t)$ with the unit direction $\hat{e} = \hat{x},\hat{y},\hat{z}$. The field φ can be Fourier-transformed

$$\varphi(\vec{x},t) = \frac{1}{V}\sum_{\vec{p}} \varphi_{\vec{p}}(t)\,e^{i\vec{p}\cdot\vec{x}}$$

into a sum over all momenta $\vec{p}$. The factor $\frac{1}{V}$ with the volume $V = L^3$ is convenient. The sum is really a discrete sum because the allowed momenta are given by $\vec{p} = \frac{2\pi}{L}(n_1, n_2, n_3)$ where the three values n_i

are integers. The sum becomes an integral

$$\frac{1}{V}\sum_{\vec{p}} \to \int \frac{d^3\vec{p}}{(2\pi)^3}$$

for $L \to \infty$. If one integrates over the whole volume V

$$\int_V d^3\vec{x}\, e^{i(\vec{p}-\vec{q})\cdot\vec{x}} = V\,\delta_{\vec{p},\vec{q}}$$

the result is 0 for $\vec{p} \neq \vec{q}$ and V for $\vec{p} = \vec{q}$, and the infinite volume limit gives $(2\pi)^3\,\delta^3(\vec{p}-\vec{q})$.

Thus, the action S becomes

$$S = \frac{1}{2}\int dt\,\frac{1}{V}\sum_{\vec{p}}\left[\dot{\varphi}_{\vec{p}}\,\dot{\varphi}_{-\vec{p}} - (\vec{p}^{\,2}+m^2)\,\varphi_{\vec{p}}\,\varphi_{-\vec{p}}\right]$$

where the factors $\varphi_{\vec{p}}$ and $\varphi_{-\vec{p}}$ come from the fact that the integral over $\vec{x}$ has been resolved. This almost looks like a simple harmonic oscillator except for the $\vec{p}$ and $-\vec{p}$, but this makes sense because the original $\varphi(X)$ is real (and therefore Hermitian) such that $(\varphi_{\vec{p}})^\dagger = \varphi_{-\vec{p}}$ due to the complex Fourier transform. This makes the action S a real number as it should be.

There is a simple way to fix that because one can use even and odd combinations

$$\varphi^{\text{even}}_{\vec{p}} = \frac{1}{2}\,(\varphi_{\vec{p}} + \varphi_{-\vec{p}}) \qquad\qquad \varphi^{\text{odd}}_{\vec{p}} = \frac{i}{2}\,(\varphi_{\vec{p}} - \varphi_{-\vec{p}})$$

where both are Hermitian. The action S

$$S = \frac{1}{2}\int dt\,\frac{1}{V}\sum_{\vec{p}}\sum_{i=\text{even,odd}}\left[\dot{\varphi}^i_{\vec{p}}\,\dot{\varphi}^i_{\vec{p}} - E^2_{\vec{p}}\,\varphi^i_{\vec{p}}\,\varphi^i_{\vec{p}}\right] = \frac{1}{2}\int dt\,\frac{1}{V}\sum_{\vec{p}}\sum_{i=\text{even,odd}}\left[(\dot{\varphi}^i_{\vec{p}})^2 - E^2_{\vec{p}}\,(\varphi^i_{\vec{p}})^2\right]$$

shows in this form that it is a sum over harmonic oscillators. Note that the first summation is only over half the $\vec{p}$ in order not to double count. (This is anyway not the notation used later on).

Recalling some notions from Quantum Mechanics, the one-dimensional harmonic oscillator in the Heisenberg picture, the relation between operators in the Heisenberg picture $\boldsymbol{O}_H = \boldsymbol{O}(t)$ and the Schrödinger picture $\boldsymbol{O}_S = \boldsymbol{O}(0)$, and the equation of motion in the Heisenberg picture are

$$\boldsymbol{H} = \frac{\boldsymbol{p}^2}{2M} + \frac{1}{2}\,k\,\boldsymbol{p}^2 \qquad\qquad \boldsymbol{O}(t) = e^{i\boldsymbol{H}t}\,\boldsymbol{O}(0)\,e^{-i\boldsymbol{H}t} \qquad\qquad \frac{d}{dt}\boldsymbol{O}(t) = i[\boldsymbol{H},\boldsymbol{O}]$$

where k is the spring constant, while the mass M is not the mass m of a particle here. The application of the third equation to the harmonic oscillator gives

$$\dot{\boldsymbol{x}} = i\left[\frac{\boldsymbol{p}^2}{2\,M},\boldsymbol{x}\right] = \frac{\boldsymbol{p}}{M} \qquad\qquad \dot{\boldsymbol{p}} = i\left[\frac{1}{2}\,k\,\boldsymbol{x}^2,\boldsymbol{p}\right] = -k\,\boldsymbol{x} \qquad\qquad \ddot{\boldsymbol{x}} = -\frac{k}{M}\,\boldsymbol{x} = -\omega^2\,\boldsymbol{x}$$

for the derivatives and the equation of motion with the frequency $\omega = \sqrt{k/M}$. The general solution $\boldsymbol{x}(t)$ must be of the form $\boldsymbol{A}\,e^{-i\omega t}$ plus Hermitian conjugate with the commutation relations $[\boldsymbol{A},\boldsymbol{A}^\dagger] = \frac{1}{2\,M\,\omega}$. Defining $\boldsymbol{a} = \sqrt{2\,M\,\omega}\,\boldsymbol{A}$ gives

$$\left[\boldsymbol{a},\boldsymbol{a}^\dagger\right] = 1 \qquad\qquad \boldsymbol{H} = \frac{\omega}{2}\left(\boldsymbol{a}^\dagger\,\boldsymbol{a} + \boldsymbol{a}\,\boldsymbol{a}^\dagger\right) = \omega\left(\boldsymbol{a}^\dagger\,\boldsymbol{a} + \frac{1}{2}\right) \tag{3.5}$$

as the commutation relations and the corresponding Hamiltonian.

In order to calculate the energy eigenstates $\boldsymbol{H}\,|E\rangle = E\,|E\rangle$ one determines

$$\begin{aligned}\boldsymbol{H}(\boldsymbol{a}\,|E\rangle) &= \omega\left(\boldsymbol{a}^\dagger\,\boldsymbol{a}\,\boldsymbol{a} + \frac{1}{2}\,\boldsymbol{a}\right) = \omega\left((\boldsymbol{a}\,\boldsymbol{a}^\dagger - 1)\,\boldsymbol{a} + \frac{1}{2}\,\boldsymbol{a}\right) = \boldsymbol{a}(\boldsymbol{H}\,|E\rangle) - \omega\,\boldsymbol{a}\,|E\rangle \\ &= E\,\boldsymbol{a}\,|E\rangle - \omega\,\boldsymbol{a}\,|E\rangle = (E-\omega)\,\boldsymbol{a}\,|E\rangle\end{aligned}$$

and observes that $\boldsymbol{a}\,|E\rangle$ is an energy eigenstate with the energy reduced by ω. In other words, $\boldsymbol{a}$ is a lowering operator. Since there has to be a lowest energy to make the system stable there must be a state $|0\rangle$ with $\boldsymbol{a}\,|0\rangle = 0$. It is called the ground state. The other states $|n\rangle = \mathcal{N}_n(\boldsymbol{a}^\dagger)^n\,|0\rangle$ which have to be there come from the application of $\boldsymbol{a}^\dagger$ to state $|0\rangle$ several times and are normalized with $\mathcal{N}_n$. Calculations

$$\begin{aligned}\langle m|n\rangle &= \mathcal{N}_m^*\,\mathcal{N}_n\ \langle 0|\boldsymbol{a}^m(\boldsymbol{a}^\dagger)^n|0\rangle = \mathcal{N}_m^*\,\mathcal{N}_n\ \langle 0|\boldsymbol{a}^{m-1}[\boldsymbol{a},(\boldsymbol{a}^\dagger)^n]|0\rangle \\ &= \mathcal{N}_m^*\,\mathcal{N}_n\ \langle 0|\boldsymbol{a}^{m-1}n(\boldsymbol{a}^\dagger)^{n-1})|0\rangle = \mathcal{N}_m^*\,\mathcal{N}_n\,n\ \langle 0|\boldsymbol{a}^{m-1}(\boldsymbol{a}^\dagger)^{n-1}|0\rangle \\ &= \mathcal{N}_m^*\,\mathcal{N}_n\,\delta_{mn}\,n!\end{aligned}$$

using the commutation relations $[\boldsymbol{a},(\boldsymbol{a}^\dagger)^n] = n(\boldsymbol{a}^\dagger)^{n-1}$ iteratively show that $\langle m|n\rangle$ is only non-zero for $m = n$ because otherwise there are no remaining $\boldsymbol{a}$ acting on the left side or no remaining $\boldsymbol{a}^\dagger$ acting on the right side. Thus, a convenient definition of the normalization factors $\mathcal{N}_n$ is

$$\mathcal{N}_n = \frac{1}{\sqrt{n!}} \qquad\Rightarrow\qquad \langle m|n\rangle = \delta_{mn}$$

leading to

$$\boldsymbol{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle \qquad\qquad \boldsymbol{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \tag{3.6}$$

for the raising and lowering operators.

The energy is therefore

$$\boldsymbol{H}\,|n\rangle = \omega\left(\boldsymbol{a}^\dagger\boldsymbol{a} + \frac{1}{2}\right)|n\rangle = \left(n + \frac{1}{2}\right)\omega\,|n\rangle \tag{3.7}$$

because of (3.5) and (3.6). (Note that there is a factor $\hbar$ if conventional units are used.)

Coming back to quantum fields, each $\varphi^i_{\vec{p}}$ in the action S

$$S = \frac{1}{2}\int dt\,\frac{1}{V}\sum_{\vec{p}}\sum_{i=\text{even,odd}}\left[(\dot{\varphi}^i_{\vec{p}})^2 - E_{\vec{p}}^{\,2}\,(\varphi^i_{\vec{p}})^2\right]$$

for the free scalar field is a harmonic oscillator with its own raising and lowering operator $\boldsymbol{a}^i_{\vec{p}}$. The state of the field is $|n_{\vec{p}_1}, n_{\vec{p}_2}, ...\rangle$ for all the possible operators. This looks slightly different from Fock space introduced above where the assumption was that all occupation numbers $n_{\vec{p}_i}$ are either 0 or 1. There the momenta were continuous and one could pretend that all the momenta were occupied only once, while here the momenta are discrete such that some of the momenta can be identical.

3.6 Measuring Quantum Fields

The Hamiltonian of a scalar quantum field φ is

$$\boldsymbol{H} = \frac{1}{2}\,\frac{1}{V}\sum_{\vec{p}}\sum_{i=\text{even,odd}}\left[(\dot{\varphi}^i_{\vec{p}})^2 + E_{\vec{p}}^{\,2}\,(\varphi^i_{\vec{p}})^2\right]$$

where the simple harmonic oscillators $\varphi^i_{\vec{p}}$ have frequency $\omega = E_{\vec{p}}$ and mass $M = V^{-1}$. Thus, the vacuum expectation value $\langle 0|\varphi^i_{\vec{p}}|0\rangle = 0$ is zero because there is a symmetry for φ going to $-\varphi$. More interesting in order to measure the field φ as an observable is

$$\langle 0|\varphi^i_{\vec{p}}\,\varphi^j_{\vec{q}}\,|0\rangle = \delta_{\vec{p},\vec{q}}\,\delta_{ij}\,\frac{V}{2E_{\vec{p}}}$$

because $\langle 0|\varphi^2(X)|0\rangle = \frac{1}{2M\omega}$ for a simple harmonic oscillator. The field φ is

$$\varphi(\vec{x},t) = \frac{1}{V}\sum_{\vec{p}}\left[\cos(\vec{p}\cdot\vec{x})\,\varphi^{\text{even}}_{\vec{p}}(t) + \sin(\vec{p}\cdot\vec{x})\,\varphi^{\text{odd}}_{\vec{p}}(t)\right]$$

and the fluctuations of the field become

$$\langle 0|\varphi^2(\vec{x},t)|0\rangle = \frac{1}{V^2}\sum_{\vec{p}}\left[\cos^2(\vec{p}\cdot\vec{x})\,\frac{V}{2E_{\vec{p}}} + \sin^2(\vec{p}\cdot\vec{x})\,\frac{V}{2E_{\vec{p}}}\right]$$
$$= \frac{1}{V}\sum_{\vec{p}}\frac{1}{2E_{\vec{p}}} \to \int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2E_{\vec{p}}} = \infty$$

with the infinite volume limit which diverges quadratically with large momenta. Thus, the result is that the fluctuations of the field are infinite because there are modes with arbitrarily large momenta. This comes from the fundamental property of locality such that there is a mode at every single position. This infinity is telling that such quantities are very sensitive to the structure of the theory at very short distances. There are ways to deal with this problem but in the following it will be mostly ignored.

If one defines a "smeared" field φ_R as

$$\varphi_R(t) = \frac{1}{V_R}\int_R d^3\vec{x}\,\varphi(\vec{x},t)$$

where R stands for some region in space with volume V_R. This averages the field over a finite size region. The square of the smeared field is

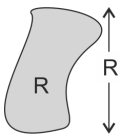
$$\langle 0|\varphi_R^2(t)|0\rangle = \frac{1}{V_R^2}\int_R d^3\vec{x}\int_R d^3\vec{y}\,\langle 0|\varphi(\vec{x},t)\,\varphi(\vec{y},t)|0\rangle$$

where one can insert

$$\langle 0|\varphi(\vec{x},t)\,\varphi(\vec{y},t)|0\rangle = \frac{1}{V^2}\sum_{\vec{p}}\left[\cos(\vec{p}\cdot\vec{x})\,\cos(\vec{p}\cdot\vec{y})\,\frac{V}{2E_{\vec{p}}} + \sin(\vec{p}\cdot\vec{x})\,\sin(\vec{p}\cdot\vec{y})\,\frac{V}{2E_{\vec{p}}}\right]$$
$$= \frac{1}{V}\sum_{\vec{p}}\cos\left(\vec{p}\cdot(\vec{x}-\vec{y})\right)\frac{V}{2E_{\vec{p}}} \to \int\frac{d^3\vec{p}}{(2\pi)^3}\,\cos\left(\vec{p}\cdot(\vec{x}-\vec{y})\right)\frac{1}{2E_{\vec{p}}}$$

with the infinite volume limit which converges.

One can calculate this integral, but given the region R and calling also its characteristic spacial size R one can do order of magnitude estimates for R smaller than the Compton wavelength corresponding to $R \ll m^{-1}$ for the particle of mass m. This means $m \to 0$ and $E_{\vec{p}} \simeq |\vec{p}|$. Since there is no other scale one gets



$$\langle 0|\varphi(\vec{x},t)\,\varphi(\vec{y},t)|0\rangle \propto \frac{1}{|\vec{x}-\vec{y}|^2}$$

from dimensional analysis. Thus, the fluctuations for the smeared field are in the order

$$\langle 0|\varphi_R^2(t)|0\rangle \sim \frac{1}{R^2}$$

that goes to infinity for $R \to 0$.

This is basically just dimensional analysis, and one can do the same calculations for any dimension D. The action is of the form

$$S = \int d^Dx\left[\frac{1}{2}\,\partial^\mu\varphi\,\partial_\mu\varphi + ...\right]$$

and must be dimensionless due to the canonical normalization of the fields such that the average $\langle\varphi_R^2\rangle$ satisfies $\langle\varphi_R^2\rangle \sim \frac{1}{R^{D-2}}$ because the dimension of the field must be $[\varphi] = \frac{D-2}{2}$. For lower dimensions this divergence $R \to 0$ is less rapid, and becomes $\langle\varphi_R^2\rangle \sim \frac{1}{\ln R}$ in two dimensions. But the flipside of this is that for large R it dies off slowly. Thus, fluctuations in two dimensions averaged over large regions only die of as a logarithm and are very important. One of the implications of this is that these fluctuations are large enough to prevent the formation of long-range order, or in the language of particle physics, prevent spontaneous symmetry breaking. This is the Coleman-Mermin-Wagner theorem for which there is a version that holds in Quantum Field Theory and a version that holds in Statistical Mechanics. This shows that simple dimensional analysis can tell something very basic about the field as an observable.

3.7 Vacuum Energy

The expectation value $\langle 0|\boldsymbol{H}|0\rangle$ of the Hamiltonian $\boldsymbol{H}$ in the ground state is

$$\langle 0|\boldsymbol{H}|0\rangle = \sum_{\vec{p}} \frac{1}{2} E_{\vec{p}}$$

because for every momentum there is an oscillator. There is no $\frac{1}{V}$ because the correct quantity to think about is the density ρ_{vac} defined as

$$\rho_{\text{vac}} = \frac{\langle 0|\boldsymbol{H}|0\rangle}{V} = \frac{1}{V} \sum_{\vec{p}} \frac{1}{2} E_{\vec{p}} \to \frac{1}{2} \int \frac{d^3\vec{p}}{(2\pi)^3} \sqrt{\vec{p}^2 + m^2} = \infty \tag{3.8}$$

in the infinite volume limit.

If one does not believe that the theory is valid for arbitrarily large momenta because it has not been tested there, one can introduce a cut-off and limit the integral to values $|\vec{p}| \leq \Lambda$ where this mass scale Λ is chosen much larger than anything ever measured. This is an artificial mutilation and the only justification for it is that it mutilates it at short distances. Thus, for $\Lambda \gg m$ the density ρ_{vac} becomes

$$\rho_{\text{vac}} = \frac{1}{16\pi^2} \left[\Lambda^4 + \Lambda^2 m^2 - \frac{1}{2} m^4 \ln \frac{2\Lambda}{m} + O\left(\frac{m^6}{\Lambda^2}\right) \right]$$

where the terms that have been left out are small for $\Lambda \to \infty$ instead of being big. The artificial cut-off is certainly nonsense, but it is believed that something such as a new kind of physics will come in to make this theory finite. (There are actually models like that.) If one evaluates the integral (3.8) in the fundamental theory the integrand will be as shown here for $|\vec{p}| \ll \Lambda$ and for larger $|\vec{p}|$ it will be something different. The question is just what exactly it is. The believe is that the answer has to look like ρ_{vac} particularly with the huge term Λ^4 and the huge m-dependent piece $\Lambda^2 m^2$.

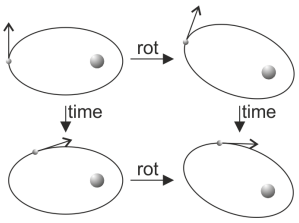
One might say that this is fine but who cares because one cannot observe the energy of the vacuum. The vacuum is the lowest energy state, and the only thing that one can measure is the difference of energies. This sounds like a reasonable argument but is not correct. One reason is that some of these divergent terms when Λ goes to infinity depend on the mass m. In the Standard Model of Particle Physics the mass of all the elementary particles is proportional to the Higgs field, and it turns out that the m-dependent contribution to the vacuum energy $\Lambda^2 m^2$ corresponds to a field-dependent contribution to the energy of the system. This is something that can be measured because if one changes the field one changes m and one changes the vacuum energy by this divergent amount. Precisely this kind of contribution $\Lambda^2 m^2$ where the m comes from a Higgs gives the Higgs in the Standard Model a huge contribution to its mass from the fluctuations of other fields.

Another reason for why not only energy differences can be measured comes from gravity where energy differences can be observed because a constant term in the action still depends on the metric and this is the so-called cosmological constant. Thus, the huge divergent term Λ^4 contributes directly to the expansion of the universe. There are several solutions for the problem with the Higgs field from the $\Lambda^2 m^2$ term and nobody knows whether one is right. However, for the problem with the cosmological constant from the Λ^4 term there is no solution so far, and it seems to be a really deep problem.

4 Symmetry

4.1 Symmetries Leaving the Laws of Physics Invariant

Much of what is known about Quantum Field Theory comes from the studies of symmetries. Lorentz invariance, for example, completely changes the nature of Quantum Mechanics. Symmetry usually means some set of transformations that leave something such as a geometrical figure invariant but one is interested here in symmetries that leave the laws of physics invariant. This means, in other words, symmetries of the laws are studied and not symmetries of the solutions.



To make this distinction clear, classical orbits in a spherically symmetric potential like a planet going around a star are used as an example. The spherically symmetric potential makes sure that the equations of motion are unchanged if one rotates the system about the center of the potential. However, the solutions are certainly not invariant. A star with a planet on an elliptical orbit as in the figure on the right side is certainly not the same when rotated. However, it does not matter whether the system is rotated first and then evolved in time or time evolved first and then rotated. In this sense the laws are the same. The fact that the laws are the same for the unrotated and the rotated system makes the figure commute, and rotations take solutions to solutions.

Since it does not matter whether the symmetry transformation is done before the time evolution or the other way around, this means in Quantum Mechanics that the symmetry operation $|\psi\rangle \to \boldsymbol{U}|\psi\rangle$ performed by a unitary operator $\boldsymbol{U}$ in the Schrödinger picture commutes with the Hamilton operator $\boldsymbol{H}$ responsible for the time evolution giving $\boldsymbol{HU} = \boldsymbol{UH}$ respectively $[\boldsymbol{H}, \boldsymbol{U}] = 0$. In the Heisenberg picture with $|\psi\rangle \to |\psi\rangle$ and $\boldsymbol{O} \to \boldsymbol{U}^\dagger \boldsymbol{OU}$ this translates to $\boldsymbol{H} \to \boldsymbol{U}^\dagger \boldsymbol{HU} = \boldsymbol{H}$.

4.2 Internal and Spacetime Symmetries of a Physical System

If a system is defined by an action $S[\varphi]$ depending on some variables φ, if there is a transformation $\varphi(X) \to \varphi'(X)$ and if the action is invariant under this transformation of φ such that $S[\varphi] = S[\varphi']$ then this is a symmetry of the action. The claim is that any symmetry of the action is a symmetry of the laws of motion because the action determines the laws of motion. The classical equations of motion come from extremizing the action $\delta S = 0$. If the action is invariant under a transformation then every stationary point is mapped to a new stationary point and solutions of the equations of motion to other solutions of the equations of motion.

The free scalar field, for example, with $S[\varphi] = \int d^4X \left[\frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m^2\varphi^2\right]$ where $(\partial\varphi)^2 = \partial^\mu\varphi\,\partial_\mu\varphi$ is used as common notation has a symmetry $\varphi \to \varphi' = -\varphi$ because the action is quadratic and the symmetry leaves the action invariant. Thus, if $\varphi(X)$ is a solution then also $-\varphi(X)$ is a solution. This is an example of a discrete symmetry.

More interesting are continuous symmetries such as rotations defined by angles as continuous parameters. Symmetries that act on spacetime are a little bit more complicated than other symmetries such as the example of two scalar fields φ_1 and φ_2 with the action

$$S = \frac{1}{2}\int d^4X \left[(\partial\varphi_1)^2 + (\partial\varphi_2)^2 + m^2(\varphi_1^2 + \varphi_2^2)\right]$$

chosen such that both fields have the same mass m. They do not need to have the same mass, but here it is assumed that they do. Therefore, one obvious symmetry is $\varphi_1 \leftrightarrow \varphi_2$ because the two fields are interchangeable. This symmetry can be generalized to

$$\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} \to \begin{pmatrix}\varphi_1'\\ \varphi_2'\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix}$$

where the transformation matrix belongs to SO(2). A well-known property of matrices in SO(2) is that $\varphi_1^2 + \varphi_2^2$ is invariant. Also $(\partial\varphi_1)^2 + (\partial\varphi_2)^2$ is invariant because the derivatives do not act on the $\sin\theta$ and $\cos\theta$ factors. This shows that the given action is indeed invariant. This is a continuous transformation labeled by the angle θ.

Another example is translation invariance. Translation in spacetime is $\varphi(X) \to \varphi'(X) = \varphi(X - A)$ and has the four parameters A^μ. The action $S[\varphi]$ is invariant. For the term in the action S with φ^2 this follows from

$$\int d^4X\varphi^2(X) \to \int d^4X\varphi^2(X) = \int d^4X\varphi^2(X - A) = \int d^4Y\varphi^2(Y)$$

by changing the integration variables X^μ to $Y^\mu = X^\mu - A^\mu$ because the integration variable does not

change such that $d^4X = d^4Y$. This work for any function of φ. For the kinetic term in S with $(\partial\varphi)^2$

$$\begin{aligned}\int d^4X(\partial\varphi(X))^2 &= \int d^4X\,\eta^{\mu\nu}\frac{\partial}{\partial X^\mu}\varphi(X)\,\frac{\partial}{\partial X^\nu}\varphi(X) \to \int d^4X\,\eta^{\mu\nu}\frac{\partial}{\partial X^\mu}\varphi'(X)\,\frac{\partial}{\partial X^\nu}\varphi'(X)\\ &= \int d^4Y\,\eta^{\mu\nu}\frac{\partial Y^\rho}{\partial X^\mu}\frac{\partial}{\partial Y^\rho}\varphi(Y)\,\frac{\partial Y^\sigma}{\partial X^\nu}\frac{\partial}{\partial Y^\sigma}\varphi(Y) = \int d^4Y\,\eta^{\mu\nu}\delta^\rho{}_\mu\frac{\partial}{\partial Y^\rho}\varphi(Y)\,\delta^\sigma{}_\nu\frac{\partial}{\partial Y^\sigma}\varphi(Y)\\ &= \int d^4Y\,\eta^{\mu\nu}\frac{\partial}{\partial Y^\mu}\varphi(Y)\,\frac{\partial}{\partial Y^\nu}\varphi(Y) = \int d^4Y(\partial\varphi(Y))^2\end{aligned}$$

shows its invariance again setting $Y^\mu = X^\mu - A^\mu$ and using the chain rule. This shows $S[\varphi] = S[\varphi']$.

The two examples present two different kinds of symmetries. The first example shows a transformation that changes the values of the fields leading to so-called internal symmetries and in the second example the transformation acts on the arguments of the fields leading to so-called spacetime symmetries.

4.3 Noether's Theorem

Whenever there is a symmetry with a continuous parameter Noether's theorem applies. Calling the continuous parameter θ and supposing that $\theta = 0$ is the identity transformation Noether's theorem assumes that θ can be different at different spacetime points X such that one can write $\theta(X)$. This is no longer a symmetry of the action and $S[\varphi] \neq S[\varphi']$. Expanding the difference between $S[\varphi]$ and $S[\varphi']$ gives

$$\delta S = -\int d^4X\,\partial_\mu\theta\,J^\mu$$

to linear order. If θ would be independent of X then $S[\varphi] = S[\varphi']$ and therefore there has to be a derivative acting on θ because this is the only way to make $\delta S = 0$ for θ independent of the spacetime location X. The minus sign is conventional and the coefficients are called J^μ.

However, there might be more than one derivative acting on θ. In this case one can always integrate by parts to move all derivatives but one to J^μ. For example, $\partial_\mu\partial_\nu\theta K^{\mu\nu} = -\partial_\mu\theta\partial_\nu K^{\mu\nu}$ would allow to define J^μ to be $J^\mu = \partial_\nu K^{\mu\nu}$. Using integration by parts, there is no loss of generality by writing δS in the form above.

The question remains why one should stop here and not eliminate $\partial_\mu\theta$ also using integration by parts to give

$$\delta S = -\int d^4X\,\partial_\mu\theta\,J^\mu = +\int d^4X\,\theta\,\partial_\mu J^\mu$$

where $\partial_\mu J^\mu$ is is not an operator but a 4-divergence. Whenever one takes something that is a symmetry in case θ goes to a constant the variation under that transformation has to look like this.

The coefficients J^μ are functions of φ. If there is a solution to the equations of motion then any variation δS around a solution is zero. Thus, also $\delta S = \int d^4X\theta(X)\,\partial_\mu J^\mu(\varphi)$ is zero, and the divergence vanishes such that $\partial_\mu J^\mu = 0$ on solutions of the equations of motion. The equation $\partial_\mu J^\mu = 0$ is called a current conservation condition because it has the same form as charge conservation in Electrodynamics where

$$\rho = J^0 \qquad\qquad \frac{\partial}{\partial t}\rho - \vec{\nabla}\cdot\vec{J} = 0$$

for the charge density ρ and the current density $\vec{J}$. In Quantum Field Theory it is conventional to call the whole J^μ a current, and

$$\partial_\mu J^\mu(X) = 0 \tag{4.1}$$

is the current conservation condition. It is a local condition because it holds at any point X in spacetime. In Electrodynamics, $\frac{\partial}{\partial t}\rho$ states that the charge density is changing precisely due to the presence of a current flowing out, and this is true for every point in spacetime. One can also make a global statement using $Q(t) = \int d^3\vec{x}\,\rho(t,\vec{x})$ depending on time t, but since there is this local current conservation condition it actually does not depend on time t because

$$\dot{Q}(t) = \int d^3\vec{x}\,\partial_0 J^0 = \int d^3\vec{x}\,\vec{\nabla}\cdot\vec{J} = 0$$

following from $\rho = J^0$ and $\frac{\partial}{\partial t}\rho - \vec{\nabla}\cdot\vec{J} = 0$ is guaranteed by the local conservation condition and integration by parts.

The values J^μ are called a current and Q is called a charge in Quantum Field Theory as a generalized terminology used outside of the context of Electromagnetism. Noether's theorem is not just abstract mathematics but it is usually also the best way to calculate the current.

In order to show this one can use the above example of an internal symmetry

$$\begin{pmatrix}\varphi_1'\\ \varphi_2'\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} \qquad \text{and} \qquad \begin{aligned}\delta\varphi_1 &= -\theta\,\varphi_2\\ \delta\varphi_2 &= +\theta\,\varphi_1\end{aligned}$$

with the transformation for any θ and the infinitesimal transformations for small values of θ in linear order. The variation of the action

$$S = \frac{1}{2}\int d^4X\,\left[(\partial\varphi_1)^2 + (\partial\varphi_2)^2 + m^2(\varphi_1^2+\varphi_2^2)\right]$$

gives

$$2(\varphi_1\,\delta\varphi_1 + \varphi_2\,\delta\varphi_2) = 2\big[\varphi_1(-\theta\varphi_2) + \varphi_2(+\theta\varphi_1)\big] = 0$$

for the term $\varphi_1^2+\varphi_2^2$. The term $(\partial\varphi_1)^2 + (\partial\varphi_2)^2$ leads to the same kind of calculations if the derivatives never act on θ and also these two terms cancel. However, if $\theta(X)$ depends on X since Noether's theorem requires to replace θ by $\theta(X)$ and the derivatives do act on them, the only terms of δS one has to keep for finding the Noether current are those where the derivatives act on θ. This gives

$$\delta S = \frac{1}{2}\int d^4X\big[2\partial^\mu\varphi_1\,\partial_\mu(-\theta\varphi_2) + 2\partial^\mu\varphi_2\,\partial_\mu(+\theta\varphi_1) + ...\big] = \int d^4X\big[\,-\partial^\mu\varphi_1\,\partial_\mu\theta\varphi_2 + \partial^\mu\varphi_2\,\partial_\mu\theta\varphi_1\,\big]$$

and the Noether current is $J_\mu = -\varphi_1(\partial_\mu\varphi_2) + (\partial_\mu\varphi_1)\varphi_2 = -\varphi_1\overset{\leftrightarrow}{\partial_\mu}\varphi_2$ as the coefficients of $\partial_\mu\theta$ with a minus sign. If one derives a Noether current one should always check it. Taking the divergence of the current just derived gives

$$\begin{aligned}\partial_\mu J^\mu(X) &= \partial_\mu\big[\,-\varphi_1(\partial_\mu\varphi_2) + (\partial_\mu\varphi_1)\varphi_2\,\big] = -\partial_\mu\varphi_1\partial^\mu\varphi_2 - \varphi_1\Box\varphi_2 + \Box\varphi_1\varphi_2 + \partial^\mu\varphi_1\partial_\mu\varphi_2\\ &= -\varphi_1\Box\varphi_2 + \Box\varphi_1\varphi_2 = -\varphi_1(-m^2\varphi_2^2) + (-m^2\varphi_1^2)\varphi_2 = 0\end{aligned}$$

by using the equations of motion in the step from $\Box\varphi_i$ to $-m^2\varphi_i^2$. The result is zero as long as the mass m is the same for both fields making the action symmetric.

For the spacetime symmetry $\varphi'(X) = \varphi(X-A)$ as a second example the infinitesimal transformation can be obtained using the Taylor expansion. The result is

$$\varphi'(X) = \varphi(X-A) = \varphi'(X) = \varphi(X) - A^\mu\partial_\mu\varphi(X) + O(A^2) \qquad \delta\varphi(X) = -A^\mu\partial_\mu\varphi(X)$$

with the nice property that both sides of the equation for the infinitesimal transformation depend on the same spacetime point X. The price for this is the occurrence of a derivative of $\varphi(X)$. To apply Noether's theorem A^μ has to be replaced by $A^\mu(X)$. The action is the usual action for a scalar field

$$S = \int d^4X\,\left[\frac{1}{2}(\partial\varphi)^2 - V(\varphi)\right]$$

but for a general potential function $V(\varphi)$. As shown above this is invariant under translation and is therefore also invariant under infinitesimal translation, but it is no longer invariant after the replacement $A^\mu \to A^\mu(X)$. The form of δS is

$$\delta S = -\int d^4X\,\partial_\mu A_\nu\,\Theta^{\mu\nu}$$

with a general tensor $\Theta^{\mu\nu}$, and Noether's argument states $\partial_\mu\Theta^{\mu\nu} = 0$. This gives four conservation conditions, one for each ν. Notice that there is no particular symmetry between μ and ν and therefore there is no requirement for any symmetry between μ and ν in $\Theta^{\mu\nu}$. Thus, in general $\Theta^{\mu\nu} \neq \Theta^{\nu\mu}$, although

in this case it will turn out below that it is symmetric such that $\Theta^{\mu\nu} = \Theta^{\nu\mu}$. The variation of the term with $V(\varphi)$ in S is

$$\begin{aligned}\delta \int d^4X\, V(\varphi) &= \int d^4X\, \delta V(\varphi) = \int d^4X\, V'(\varphi)\, \delta\varphi = \int d^4X\, V'(\varphi)\, \big[- A^\mu \partial_\mu \varphi \,\big] \\ &= \int d^4X\, \big\{ \,\partial_\mu \big[- A^\mu V(\varphi) \,\big] + \partial_\mu A^\mu\, V(\varphi) \big\} \\ &= \int d^4X\, \partial_\mu \big[- A^\mu V(\varphi) \,\big] + \int d^4X\, \partial_\mu A^\mu\, V(\varphi) = 0 + \int d^4X\, \partial_\mu A_\nu\, \eta^{\mu\nu}\, V(\varphi)\end{aligned}$$

using $\partial_\mu V = V'(\varphi)\, \partial_\mu \varphi$ and the fact that the first integral in the third line is an integral of a total derivative and therefore vanishes. The second integral has been written such that it is obviously in the form for Noether's theorem and has a term $\partial_\mu A_\nu$. The other term in the variation of S is

$$\begin{aligned}\delta \int d^4X\, \frac{1}{2}(\partial\varphi)^2 &= \int d^4X\, \frac{1}{2}\, 2\, \partial^\mu\varphi\, \partial_\mu(\delta\varphi) = \int d^4X\, \partial^\mu\varphi\, \partial_\mu \big[- A^\nu \partial_\nu \varphi \,\big] \\ &= \int d^4X\, \big\{ - \partial_\mu A^\nu\, \partial^\mu\varphi\, \partial_\nu\varphi - A^\nu\, \partial^\mu\varphi\, \partial_\mu\partial_\nu\varphi \,\big\} \\ &= \int d^4X\, \left\{ -\partial_\mu A^\nu\, \partial^\mu\varphi\, \partial_\nu\varphi - \partial_\nu \left[\frac{1}{2} A^\nu\, (\partial\varphi)^2 \right] + \partial_\nu A^\nu \cdot \frac{1}{2}(\partial\varphi)^2 \right\}\end{aligned}$$

and collecting both terms from δS gives

$$\Theta_{\mu\nu} = \partial_\mu\varphi\partial_\nu\varphi - \eta_{\mu\nu}\, \mathcal{L} \qquad\qquad \mathcal{L} = \frac{1}{2}\, (\partial\varphi)^2 - V(\varphi)$$

with the Lagrangian density $\mathcal{L}$. Note that the tensor $\Theta_{\mu\nu}$ in this case turns out to be symmetric. To check that $\Theta_{\mu\nu}$ is indeed a conserved quantity the calculation using

$$\partial_\nu \mathcal{L} = \partial^\mu \left[\frac{1}{2}\, (\partial\varphi)^2 - V(\varphi) \right] = \partial^\mu\varphi\, \partial_\nu\partial_\mu\varphi - V'(\varphi)\, \partial_\nu\varphi$$

gives

$$\begin{aligned}\partial^\mu \Theta_{\mu\nu} &= \partial^\mu \big[\partial_\mu\varphi\, \partial_\nu\varphi - \eta_{\mu\nu}\, \mathcal{L} \big] = \Box\varphi\, \partial_\nu\varphi + \partial_\mu\varphi\, \partial^\mu\partial_\nu\varphi - \partial_\nu\mathcal{L} \\ &= \Box\varphi\, \partial_\nu\varphi + \partial_\mu\varphi\, \partial^\mu\partial_\nu\varphi - \big\{ \partial^\mu\varphi\, \partial_\nu\partial_\mu\varphi - V'(\varphi)\, \partial_\nu\varphi \big\} \\ &= \Box\varphi\, \partial_\nu\varphi + V'(\varphi)\, \partial_\nu\varphi = \big[\Box\varphi + V'(\varphi) \big] \partial_\nu\varphi\end{aligned}$$

because two terms cancel when resolved with equal upper and lower indices. Since $\Box\varphi + V'(\varphi)$ vanishes given the equation of motions also the divergence of $\Theta_{\mu\nu}$ vanishes as supposed.

4.4 Energy and Momentum Conservation

One gets the four conserved charges by integrating the time component over space

$$P^\mu = \int d^3\vec{x}\, \Theta^{0\mu} \tag{4.2}$$

and the result is supposed to be the energy-momentum 4-vector P. The facts that $\dot{P}^0 = 0$ and $\dot{\vec{p}} = 0$ as well as the facts that energy is the conserved quantity associated with time translation and momentum is the conserved quantity associated with spatial translation in familiar systems justify this identification.

More fundamentally, equation (4.2) is the definition of energy and momentum because they are the quantities conserved as a result of time translation invariance and spatial translation invariance. (Newtonian physics is wrong twice because Newtonian relativity has to be replaced by Special Relativity and classical mechanics has to be replaced by Quantum Mechanics. Thus, the only things that survive from Newtonian physics are the conservation laws because they are associated with these very deep symmetries. The laws of physics are the same tomorrow as they are today, and they are the same in different locations.)

Calculating the energy from (4.2) gives

$$P^0 = \int d^3\vec{x}\,\Theta^{00} = \int d^3\vec{x}\left[\partial^0\varphi\partial^0\varphi - \eta^{00}\,\mathcal{L}\right] = \int d^3\vec{x}\,\left[\dot{\varphi}^2 - \left(\frac{1}{2}\dot{\varphi}^2 - \frac{1}{2}(\vec{\nabla}\varphi)^2 - V(\varphi)\right)\right]$$
$$= \int d^3\vec{x}\,\left[+\frac{1}{2}\dot{\varphi}^2 + \frac{1}{2}(\vec{\nabla}\varphi)^2 + V(\varphi)\right]$$

with the convention $\eta_{00} = +1$. The integrand is a sum of squares and this shows that $P^0 \geq 0$. The analog calculation for the momenta gives

$$P^i = \int d^3\vec{x}\,\Theta^{0i} = \int d^3\vec{x}\left[\partial^0\varphi\partial^i\varphi - \eta^{0i}\,\mathcal{L}\right] = \int d^3\vec{x}\left[\partial^0\varphi\partial^i\varphi\right] \qquad \vec{p} = -\int d^3\vec{x}\;\dot{\varphi}\,\vec{\nabla}\varphi$$

because $\partial^i\varphi = -\partial_i\varphi = -\frac{\partial}{\partial X^i}\varphi$ is the gradient and $\eta^{0i} = 0$ as the off-diagonal elements of the metric are zero.

4.5 Connection between Noether's Theorem and Quantum Mechanics

Noether's theorem as described above is a classical result but the interest here lies in quantum theories. A continuous symmetry is labeled by θ and each transformation in Quantum Mechanics is implemented by a unitary operator $\boldsymbol{U}(\theta)$ that depends on θ. In the Schrödinger picture where the states transform and operators do not transform this means $|\psi\rangle \to |\psi'\rangle = \boldsymbol{U}(\theta)\,|\psi\rangle$. Because in Noether's theorem infinitesimal transformations near the identity play an important role, $\boldsymbol{U}(\theta)$ can be expanded

$$\boldsymbol{U}(\theta) = \mathbf{1} - i\,\theta\,\boldsymbol{Q} + O(\theta^2)$$

for small θ where $\boldsymbol{Q}$ is some operator with

$$\mathbf{1} = \boldsymbol{U}^\dagger\boldsymbol{U} = \left[\mathbf{1} + i\,\theta\,\boldsymbol{Q} + O(\theta^2)\right] \times \left[\mathbf{1} - i\,\theta\,\boldsymbol{Q} + O(\theta^2)\right] = \mathbf{1} + i\,\theta\,(\boldsymbol{Q}^\dagger - \boldsymbol{Q}) + O(\theta^2)$$

because $\boldsymbol{U}(\theta)$ is unitary. It follows that $\boldsymbol{Q}^\dagger = \boldsymbol{Q}$ and $\boldsymbol{Q}$ is therefore Hermitian.

Spatial translations in the usual position-space representation of Quantum Mechanics with the wave function $\psi(\vec{x})$ transform as

$$\psi(\vec{x}) \to \psi'(\vec{x}) = \psi(\vec{x} - \vec{a}) = \psi(\vec{x}) - \vec{a}\cdot\vec{\nabla}\psi(\vec{x}) + O(\vec{a}^2)$$

expanded to linear order. The generator $\vec{\nabla}$ in this case is just the derivative, and this is the momentum such that $\vec{a}\cdot\vec{\nabla}$ can be replaced by $-i\,\vec{a}\cdot\vec{p}$. This can be written as $\delta\,|\psi\rangle = -i\,\vec{a}\cdot\vec{p}\,|\psi\rangle$. The term $-i\,\vec{a}\cdot\vec{p}$ is sort of $-i\,\theta\,\boldsymbol{Q}$ in quotes, and the object $\boldsymbol{Q}$ which is the generator of the transformation is the momentum. The fact that $\boldsymbol{Q}$ uses the same letter as the charge is not a coincidence because this generator is a conserved charge. If $\boldsymbol{U}(\theta)$ is a symmetry then $\boldsymbol{U}(\theta)\boldsymbol{H} = \boldsymbol{H}\boldsymbol{U}(\theta)$ and looking at the linear term tells that $[\boldsymbol{H}\boldsymbol{U}(\theta)] = 0$ and that $\boldsymbol{Q}$ is therefore a conserved charge in Quantum Mechanics. In general, the generator of a unitary symmetry transformation is always a Hermitian operator and commutes with the Hamiltonian.

4.6 From Charged Free Particles to Scalar Fields

A charged particle $|P,\pm\rangle$ is not only characterized by a 4-momentum P with $P^2 = m^2$ but also by a charge that can take one of the two states $+q$ or $-q$. This is a one-particle state and the most general state with any number of particles is $|\{P\}_+,\{Q\}_-\rangle$ where $\{P\} = \{P_1,...,P_n\}$ and $\{Q\} = \{Q_1,...,Q_m\}$ are two lists of momenta for the positively and negatively charged particles, respectively. The charge operator $\boldsymbol{Q}$ acting on this space of states is defined as

$$\boldsymbol{Q}\,|P,\pm\rangle = \pm q\,|P,\pm\rangle \qquad\qquad \boldsymbol{Q}\,|\{P\}_+,\{Q\}_-\rangle = (n-m)q\,|\{P\}_+,\{Q\}_-\rangle \tag{4.3}$$

on this space. So far, this is a free theory and not much is happening, but when adding interactions the charge should be conserved in all interactions such that no net charge is created or destroyed.

This charge $\boldsymbol{Q}$ creates a symmetry. In the Heisenberg picture the states do not change but the operators do such that $\boldsymbol{O} \to e^{i\,\theta\,\boldsymbol{Q}}\,\boldsymbol{O}\,e^{-i\,\theta\,\boldsymbol{Q}}$ with the continuous parameter θ. In $e^{i\,\theta\,\boldsymbol{Q}}\,\boldsymbol{O}\,e^{-i\,\theta\,\boldsymbol{Q}} = \boldsymbol{O} + i\,\theta\,[\boldsymbol{Q},\boldsymbol{O}] + O(\theta^2)$ for small θ the term $i\,\theta\,[\boldsymbol{Q},\boldsymbol{O}]$ can be thought of as the action of generator $\boldsymbol{Q}$ on the operator $\boldsymbol{O}$. With the correct factors this action is $-i\,\theta\,\boldsymbol{Q}{*}\boldsymbol{O}$ where $*$ means "acting on" such that one can write $\boldsymbol{Q}{*}\boldsymbol{O} = -[\boldsymbol{Q},\boldsymbol{O}]$. This is the infinitesimal generator.

Any operator can be made out of creation and annihilation operators, but here two creation and two annihilation operators are needed, one for positive and one for negative charge. The two creation operators

$$\boldsymbol{\alpha}^\dagger(P)\,|0\rangle = |P,+\rangle \qquad\qquad \boldsymbol{\beta}^\dagger(P)\,|0\rangle = |P,-\rangle \tag{4.4}$$

allow the creation of the general state

$$|\{P\}_+,\{Q\}_-\rangle = \left(\prod_P \boldsymbol{\alpha}^\dagger(P)\right)\left(\prod_Q \boldsymbol{\beta}^\dagger(Q)\right)|0\rangle$$

from the vacuum state $|0\rangle$. The commutation relations such as $[\boldsymbol{\alpha}^\dagger(P),\boldsymbol{\beta}^\dagger(P)]$ are all zero except for

$$[\boldsymbol{\alpha}(P),\boldsymbol{\alpha}^\dagger(Q)] = \langle P|Q\rangle = [\boldsymbol{\beta}(P),\boldsymbol{\beta}^\dagger(Q)] \tag{4.5}$$

with the Lorentz-invariant delta-function $\langle P|Q\rangle$. The charge operator $\boldsymbol{Q}$ becomes

$$\boldsymbol{Q} = q(\boldsymbol{N}_+ - \boldsymbol{N}_-) = q\int (dP)\Big[\boldsymbol{\alpha}^\dagger(P)\,\boldsymbol{\alpha}(P) - \boldsymbol{\beta}^\dagger(P)\,\boldsymbol{\beta}(P)\Big] \tag{4.6}$$

expressed in terms of these creation and annihilation operators.

The next question is what the observables are. The most general operators have been defined and the charge operator $\boldsymbol{Q}$ has been defined, but the question is what the ingredients are to be used for the interacting theories, for example. One might build operators just out of $\boldsymbol{\alpha}(P)$ and $\boldsymbol{\alpha}^\dagger(P)$ or just out of $\boldsymbol{\beta}(P)$ and $\boldsymbol{\beta}^\dagger(P)$, but it turns out that one would like to have the observables to have a definite charge. As it is convenient to work with states that have a definite charge plus or minus it is also useful to work with fields that have a definite charge.

Thus, the field can be defined as

$$\varphi(X) = \int (dP)\Big[\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\beta}^\dagger(P)\,e^{+i\,P\cdot X}\Big]$$

via a Fourier transform. An obvious property of it is that $\varphi(X) \neq \varphi^\dagger(X)$, and this looks silly because it is not Hermitian. However, it is useful as a building block for observables because one can define linear combinations of φ and $\varphi^\dagger$ such as $\varphi + \varphi^\dagger$ or $i(\varphi - \varphi^\dagger)$ which are both Hermitian. The other crucial property observables are supposed to have is that they commute at spacelike distance.

To show in a first step that $\varphi(X)$ has a definite charge, one gets using Fourier transform twice and (4.5)

$$\begin{aligned}
\boldsymbol{Q}*\varphi(X) = -[\boldsymbol{Q},\varphi(X)] &= -q\int (dP)\int (dQ)\Big[\boldsymbol{\alpha}^\dagger(P)\boldsymbol{\alpha}(P) - \boldsymbol{\beta}^\dagger(P)\boldsymbol{\beta}(P),\boldsymbol{\alpha}(Q)\,e^{-i\,Q\cdot X} + \boldsymbol{\beta}^\dagger(Q)\,e^{+i\,Q\cdot X}\Big] \\
&= -q\int (dP)\int (dQ)\Big\{\Big[\boldsymbol{\alpha}^\dagger(P),\boldsymbol{\alpha}(Q)\Big]\boldsymbol{\alpha}(P)\,e^{-i\,Q\cdot X} - \boldsymbol{\beta}^\dagger(P)\Big[\boldsymbol{\beta}(P),\boldsymbol{\beta}^\dagger(Q)\Big]e^{+i\,Q\cdot X}\Big\} \\
&= -q\int (dP)\int (dQ)\Big\{ -\langle P|Q\rangle\;\boldsymbol{\alpha}(P)\,e^{-i\,Q\cdot X} - \boldsymbol{\beta}^\dagger(P)\,\langle P|Q\rangle\;e^{+i\,Q\cdot X}\Big\} \\
&= +q\int (dP)\Big[\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\beta}^\dagger(P)\,e^{+i\,P\cdot X}\Big] = q\,\varphi(X)
\end{aligned}$$

for $\boldsymbol{Q}$ acting on $\varphi(X)$. The result is $\boldsymbol{Q}*\varphi(X) = q\,\varphi(X)$, and in this sense $\varphi(X)$ has a definite charge. To show in a second step that $[\varphi(X),\varphi(Y)] = 0$ for spacelike separated X and Y no calculations are needed because φ is a sum of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}^\dagger$ and these operators commute with each other in any combination. However, one has also to take $[\varphi(X),\varphi^\dagger(Y)]$ into account because observables are made out of φ and $\varphi^\dagger$. This commutator is not trivial and gives

$$\begin{aligned}
[\varphi(X),\varphi^\dagger(Y)] &= \int (dP)\int (dQ)\Big[\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\beta}^\dagger(P)\,e^{+i\,P\cdot X},\boldsymbol{\alpha}^\dagger(Q)\,e^{+i\,Q\cdot Y} + \boldsymbol{\beta}(Q)\,e^{-i\,Q\cdot Y}\Big] \\
&= \int (dP)\Big\{e^{-i\,P\cdot(X-Y)} - e^{+i\,P\cdot(X-Y)}\Big\} = -2\,i\int (dP)\,\sin\big(P\cdot(X-Y)\big)
\end{aligned}$$

for general spacetime points X and Y. For spacelike separated points chosen as $X = (\vec{x}, 0)$ and $Y = (\vec{y}, 0)$ the commutator vanishes

$$[\varphi(\vec{x},0), \varphi^\dagger(\vec{y},0)] = -2\,i \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \sin\big(-\vec{p}\cdot(\vec{x}-\vec{y})\big) = 0$$

because the integrand is odd under $\vec{p} \to -\vec{p}$.

4.7 From Charged Free Scalar Fields to Particles

The theory of plus and minus charged quantum particles just constructed is the quantum version of the theory with the two real scalar fields φ_1 and φ_2 which are defined by the action

$$S = \int d^4X \left[\frac{1}{2}(\partial\varphi_1)^2 + \frac{1}{2}(\partial\varphi_2)^2 - \frac{1}{2}m^2(\varphi_1^2 + \varphi_2^2)\right]$$

and are invariant under SO(2)-rotations

$$\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} \to \begin{pmatrix}\varphi_1'\\ \varphi_2'\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix}$$

because both fields have the same mass m.

The complex field $\varphi = \frac{1}{\sqrt{2}}(\varphi_1 + i\,\varphi_2)$ can be defined, and its action becomes

$$S = \int d^4X \Big[\partial^\mu\varphi^\dagger\,\partial_\mu\varphi - m^2\,\varphi^\dagger\varphi\Big]$$

making the SO(2)-rotations simple phases $\varphi \to \varphi' = e^{-i\,\theta}\,\varphi$. The charge q does not appear here but can be introduced by replacing θ in the SO(2)-rotations by $q\,\theta$ such that $\varphi \to \varphi' = e^{-i\,q\,\theta}\,\varphi$. This looks just like rescaling θ but it is not quite that because a full rotation is $\theta = 2\pi$ and by choosing different values for q one changes the values of θ for a full rotation.

In this complex theory the transformation is $\varphi \to \varphi' = e^{-i\,q\,\theta}\,\varphi = \varphi - i\,q\,\theta\,\varphi + O(\theta^2)$ for finding the Noether current giving $\delta\varphi = -i\,q\,\theta\,\varphi$ such that δS becomes

$$\begin{aligned}\delta S &= \int d^4X \Big[\partial^\mu\varphi^\dagger\,\partial_\mu(-iq\theta\varphi) + \partial^\mu(+iq\theta\varphi^\dagger)\,\partial_\mu\varphi - m^2\Big[\varphi^\dagger(-iq\theta\varphi) + (+iq\theta\varphi^\dagger)\varphi\Big]\Big]\\ &= \int d^4X \Big[\partial^\mu\varphi^\dagger\,\partial_\mu(-iq\theta\varphi) + \partial^\mu(+iq\theta\varphi^\dagger)\,\partial_\mu\varphi\Big]\end{aligned}$$

and vanishes if θ is independent of X. If $\theta \to \theta(X)$ for Noether's theorem

$$\delta S = \int d^4X \Big[\partial^\mu\varphi^\dagger\,\partial_\mu(-iq\theta\varphi) + \partial^\mu(+iq\theta\varphi^\dagger)\,\partial_\mu\varphi\Big] = \int d^4X\,\partial_\mu\theta\Big[-i\,q\,(\partial^\mu\varphi^\dagger)\varphi + i\,q\,\varphi^\dagger(\partial^\mu\varphi)\Big]$$

and the Noether current is therefore $J^\mu = -i\,q\,\varphi^\dagger \overset{\leftrightarrow}{\partial_\mu}\varphi = -i\,q\,\big(\varphi^\dagger(\partial_\mu\varphi)\,\varphi - (\partial_\mu\varphi^\dagger)\varphi\big)$.

Thus, the conserved charge is $\boldsymbol{Q} = \int d^3\vec{x}\,J^0 = i\,q\int d^3\vec{x}\,\varphi^\dagger \overset{\leftrightarrow}{\partial_\mu}\varphi$ and the question is whether it leads to the same result as the definition (4.6) via creation and annihilation operators. Calculating it gives

$$\begin{aligned}\boldsymbol{Q} &= \int d^3\vec{x}\,J^0 = i\,q\int d^3\vec{x}\,\varphi^\dagger \overset{\leftrightarrow}{\partial_0}\varphi\\
&= i\,q\int d^3\vec{x}\int(dP)\int(dQ)\Big[\boldsymbol{\alpha}^\dagger(P)\,e^{+i\,P\cdot X} + \boldsymbol{\beta}(P)\,e^{-i\,P\cdot X}\Big]\overset{\leftrightarrow}{\partial_0}\Big[\boldsymbol{\alpha}(Q)\,e^{-i\,Q\cdot X} + \boldsymbol{\beta}^\dagger(Q)\,e^{-i\,Q\cdot X}\Big]\\
&= i\,q\int d^3\vec{x}\int(dP)\int(dQ)\Big[\boldsymbol{\alpha}^\dagger(P)\boldsymbol{\alpha}(Q)(-i)(Q_0+P_0)e^{i(P-Q)\cdot X} + \boldsymbol{\alpha}^\dagger(P)\boldsymbol{\beta}^\dagger(Q)\,i\,(Q_0-P_0)e^{i(P+Q)\cdot X}\\
&\qquad + \boldsymbol{\beta}(P)\boldsymbol{\alpha}(Q)(-i)(Q_0-P_0)e^{-i(P+Q)\cdot X} + \boldsymbol{\beta}(P)\boldsymbol{\beta}^\dagger(Q)\,i\,(Q_0+P_0)e^{-i(P-Q)\cdot X}\Big]\\
&= i\,q\int\frac{d^3\vec{p}}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}\frac{1}{2E_{\vec{p}}}\Big[(-i)2E_{\vec{p}}\,\boldsymbol{\alpha}^\dagger(P)\boldsymbol{\alpha}(P) + (+i)\boldsymbol{\beta}(P)\boldsymbol{\beta}^\dagger(P)\,2E_{\vec{p}}\Big]\\
&= q\int(dP)\Big[\boldsymbol{\alpha}^\dagger(P)\boldsymbol{\alpha}(P) - \boldsymbol{\beta}(P)\boldsymbol{\beta}^\dagger(P)\Big]\end{aligned}$$

because the integral over $d^3\vec{x}$ acts only on the exponentials turning them into delta-functions with $\vec{q} = \pm\vec{p}$ and $Q_0 = P_0$. The result is equal to (4.6) except for the order of $\boldsymbol{\beta}(P)$ an $\boldsymbol{\beta}^\dagger(P)$, and these operators do not commute. To change the order with

$$\boldsymbol{\beta}(P)\boldsymbol{\beta}^\dagger(P) = \boldsymbol{\beta}^\dagger(P)\boldsymbol{\beta}(P) + [\boldsymbol{\beta}(P), \boldsymbol{\beta}^\dagger(P)] = \boldsymbol{\beta}^\dagger(P)\boldsymbol{\beta}(P) + \langle P|P\rangle$$

gives an infinite difference of $\langle P|P\rangle = \infty$.

However, Noether's theorem is a classical theorem where the order of factors does not matter, and it is therefore not clear in what order the operators have to be written in a quantum theory. To fix this ambiguity the charge is defined as $\boldsymbol{Q} = \int d^3\vec{x} : J^0 :$ where $: J^0 :$ means normal ordered such that $\boldsymbol{\alpha}^\dagger$ goes to the left of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}^\dagger$ to the left of $\boldsymbol{\beta}$. The most basic property of a normal ordered operator is that its vacuum expectation value $\langle 0| : \boldsymbol{O} : |0\rangle$ is always zero because all the annihilation operators go to the right side where they hit $|0\rangle$ and all the creation operators go to the left side where they hit $\langle 0|$. General operators which are not normal ordered have infinite vacuum expectation values. In some sense part of what normal ordering does is subtracting off infinite constants.

4.8 Discrete Spacetime Symmetries

Lorentz transformations, as a reminder, have the defining property $\Lambda^\mu_{\ \rho}\,\Lambda^\nu_{\ \sigma}\,\eta_{\mu\nu} = \eta_{\rho\sigma}$ and leave the metric $\eta_{\mu\nu}$ in this sense invariant. It follows that $\det(\Lambda) = \pm 1$ and $\Lambda^0_{\ 0} \geq +1$ or $\Lambda^0_{\ 0} \leq -1$. These Λ define more general transformations than rotations and boost, but rotations and boost are the only transformations for which one has experimental evidence because all the classical tests for Special Relativity like myon decay, the existence of time dilatation, the fact that nature seems rotationally invariant are only testing rotations and boost. Rotations and boosts satisfy

$$\det(\Lambda) = 1 \qquad\qquad \Lambda^0_{\ 0} \geq +1$$

and those are the only Lorentz transformations that can be continuously connected to the identity transformation.

To get the more general transformations with $\det(\Lambda) = -1$ or $\Lambda^0_{\ 0} \leq -1$ there are two special Lorentz transformations

$\Lambda^\mu_{\ \nu} = P^\mu_{\ \nu} = \mathrm{diag}(1,-1,-1,-1)$ called parity P

$\Lambda^\mu_{\ \nu} = T^\mu_{\ \nu} = \mathrm{diag}(-1,1,1,1)$ called time reversal T

where the parity transformation reverses the space components and keeps the time component while the time reversal transformation on the other hand reverses the time component and keeps the space components. Both transformations leave the metric unchanged and change the determinant, but only time reversal changes $\Lambda^0_{\ 0}$ as summarized in figure 2.

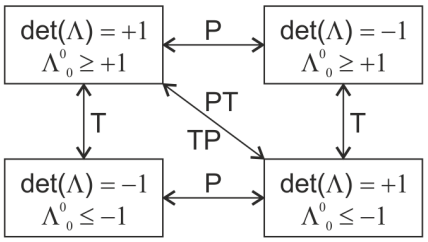


Figure 2: Effects of the parity and time reversal transformations

General transformations satisfying $\Lambda^\mu_{\ \rho}\,\Lambda^\nu_{\ \sigma}\,\eta_{\mu\nu} = \eta_{\rho\sigma}$ can be thought off as generalized Lorentz transformations. They include rotations and boosts as in the upper left rectangle in figure 2 but also parity and time reversal. It is an experimental question whether theories are invariant under parity and time reversal.

The transformation $\varphi(\vec{x},t) \to \varphi'(\vec{x},t) = \mathsf{P}^{-1}\varphi(\vec{x},t)\mathsf{P}$ where φ is a free real scalar quantum field is defined to give $\varphi(-\vec{x},t)$ but it is not clear whether it is a unitary transformation. (Note that this is not a rotation.) An obvious property of P is $\mathsf{P}^2 = 1$. Since any operator can be written in terms of creation

and annihilation operators, $\mathsf{P}^{-1}\boldsymbol{\alpha}(P^0,\vec{p})\mathsf{P} = \boldsymbol{\alpha}(P^0,-\vec{p})$ is to be expected. Using notation $\tilde{P} = (P^0,-\vec{p})$ if $P = (P^0,\vec{p})$ this can be written as $\mathsf{P}^{-1}\boldsymbol{\alpha}(P)\mathsf{P} = \boldsymbol{\alpha}(\tilde{P})$. With this property P is completely defined but the question remains whether it is unitary.

A transformation like this is a symmetry only if it is unitary. Thus, this is a crucial property. To check it one can evaluate the matrix elements of P between an arbitrary state

$$\begin{aligned}\langle Q_1,...,Q_m|\mathsf{P}|P_1,...,P_n\rangle &= \langle Q_1,...,Q_m|\mathsf{P}\,\boldsymbol{\alpha}^\dagger(P_1)...\boldsymbol{\alpha}^\dagger(P_n)|0\rangle\\ &= \langle Q_1,...,Q_m|(\mathsf{P}\,\boldsymbol{\alpha}^\dagger(P_1)\mathsf{P}^{-1})...(\mathsf{P}\,\boldsymbol{\alpha}^\dagger(P_n)\mathsf{P}^{-1})\,\mathsf{P}|0\rangle\\ &= \langle Q_1,...,Q_m|\boldsymbol{\alpha}^\dagger(\tilde{P}_1)...\boldsymbol{\alpha}^\dagger(\tilde{P}_n)|0\rangle\\ &= \langle Q_1,...,Q_m|\tilde{P}_1,...,\tilde{P}_n\rangle\end{aligned}$$

using the fact that $\mathsf{P}^{-1} = \mathsf{P}$ because of $\mathsf{P}^2 = \mathsf{1}$ and the fact $\mathsf{P}\,|0\rangle = |0\rangle$ because the vacuum is invariant under P. The result is just a product of delta-functions. Similar calculations for $\mathsf{P}^\dagger$ give

$$\begin{aligned}\langle Q_1,...,Q_m|\mathsf{P}^\dagger|P_1,...,P_n\rangle &= \langle P_1,...,P_n|\mathsf{P}|Q_1,...,Q_m\rangle^*\\ &= \langle P_1,...,P_n|\tilde{Q}_1,...,\tilde{Q}_m\rangle^*\\ &= \langle \tilde{Q}_1,...,\tilde{Q}_m|P_1,...,P_n\rangle\end{aligned}$$

and this is the same as above because $\langle Q|\tilde{P}\rangle = \langle\tilde{Q}|P\rangle$ is a delta-function. This shows that $\mathsf{P}^\dagger = \mathsf{P} = \mathsf{P}^{-1}$ and therefore that P is unitary. In other words, parity is a symmetry.

To check T, it is much simpler to look at PT because it takes the whole 4-vector X^μ to $-X^\mu$. From

$$(\mathsf{PT})^{-1}\,\varphi(X)\,(\mathsf{PT}) \stackrel{?}{=} \varphi(-X) \qquad \text{and} \qquad \varphi(X) = \int (dP)\Big[\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\alpha}^\dagger(P)\,e^{+i\,P\cdot X}\Big]$$

where one equation is labeled with a question mark follows that replacing X with $-X$ means exchanging $\boldsymbol{\alpha}(P)$ with $\boldsymbol{\alpha}^\dagger(P)$ giving $(\mathsf{PT})^{-1}\,\boldsymbol{\alpha}(P)\,\mathsf{PT} = \boldsymbol{\alpha}^\dagger(P)$. However, this definition leads to

$$\begin{aligned}\langle Q_1,...,Q_m|\mathsf{PT}|P_1,...,P_n\rangle &= \langle Q_1,...,Q_m|\mathsf{PT}\,\boldsymbol{\alpha}^\dagger(P_1)...\boldsymbol{\alpha}^\dagger(P_n)|0\rangle\\ &= \langle Q_1,...,Q_m|\boldsymbol{\alpha}(P_1)...\boldsymbol{\alpha}(P_n)|0\rangle = 0\end{aligned}$$

and this would mean that the matrix elements of PT between an arbitrary state are all zero. That is certainly not a good definition of PT. It is not easy to see what went wrong here because the concept of linearity for operators is built into the notation and the thinking.

The notation will therefore be changed, and a state is now written as ψ instead of $|\psi\rangle$. An operator $\boldsymbol{O}$ acting on a state is written as $\boldsymbol{O}\psi$ instead of $\boldsymbol{O}\,|\psi\rangle$, and the inner product $\langle\chi,\psi\rangle$ that takes two states and returns a complex number is used instead of $\langle\chi|\psi\rangle$. The inner product satisfies

$$\langle\chi, c_1\psi_1 + c_2\psi_2\rangle = c_1\,\langle\chi,\psi_1\rangle + c_2\,\langle\chi,\psi_2\rangle \qquad \langle c_1\chi_1 + c_2\chi_2,\psi\rangle = c_1^*\,\langle\chi_1,\psi\rangle + c_2^*\,\langle\chi_2,\psi\rangle \tag{4.7}$$

as usual. The condition for an operator $\boldsymbol{U}$ to be unitary is $\langle\boldsymbol{U}\chi,\boldsymbol{U}\psi\rangle = \langle\chi,\psi\rangle$ for all states χ and ψ because this simply means $\langle\chi|\boldsymbol{U}^\dagger\,\boldsymbol{U}|\psi\rangle = \langle\chi|\psi\rangle$ in the usual notation.

This new notation allows to define a so-called antiunitary operator $\boldsymbol{V}$ with $\langle\boldsymbol{V}\chi,\boldsymbol{V}\psi\rangle = \langle\chi,\psi\rangle^*$ for all states χ and ψ. An important property of unitary operators is that they preserve the probability amplitudes but if an operator is antiunitary it turns basically χ into χ^* and ψ into ψ^* and therefore also preserves probability amplitudes. An antiunitary operator is, in other words, just as good to define a symmetry transformation as a unitary operator. Eugene Wigner proved that every symmetry transformation can be written either as a unitary operator or as an antiunitary operator, but not both. In practice, time reversal invariance is the only known antiunitary transformation that exists. From

$$\begin{aligned}\langle\chi,\boldsymbol{V}(c_1\psi_1 + c_2\psi_2)\rangle &= \langle\boldsymbol{V}^{-1}\chi, c_1\psi_1 + c_2\psi_2\rangle^* = c_1^*\,\langle\boldsymbol{V}^{-1}\chi,\psi_1\rangle^* + c_2^*\,\langle\boldsymbol{V}^{-1}\chi,\psi_2\rangle^*\\ &= c_1^*\,\langle\chi,\boldsymbol{V}\psi_1\rangle + c_2^*\,\langle\chi,\boldsymbol{V}\psi_2\rangle\end{aligned}$$

follows

$$\boldsymbol{V}(c_1\psi_1 + c_2\psi_2) = c_1^*\boldsymbol{V}\psi_1 + c_2^*\boldsymbol{V}\psi_2 \tag{4.8}$$

and this is the basic property of antilinearity. The way of keeping track of this is to think of it as an operator that acts on c-numbers by complex conjugation and then does to the vectors whatever it is supposed to do to them.

The claim is that PT is an antiunitary operator in this sense. With $(\mathsf{PT})^{-1}\,\boldsymbol{\alpha}(P)\,(\mathsf{PT}) = \boldsymbol{\alpha}(P)$ one gets

$$\begin{aligned}(\mathsf{PT})^{-1}\,\varphi(X)\,(\mathsf{PT}) &= (\mathsf{PT})^{-1}\int (dP)\Big[\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\alpha}^\dagger(P)\,e^{+i\,P\cdot X}\Big]\,(\mathsf{PT})\\ &= \int (dP)\Big[\boldsymbol{\alpha}(P)\,e^{+i\,P\cdot X} + \boldsymbol{\alpha}^\dagger(P)\,e^{-i\,P\cdot X}\Big] = \varphi(-X)\end{aligned}$$

and this shows that PT is antiunitary and therefore a symmetry. This has only been shown for a free theory but amazingly if one allows arbitrary interactions of this free scalar field theory it is still invariant under PT.

Although interactions have not been introduced yet the proof for the PT theorem is outlined here. (There is the CPT theorem, but because the scalar field φ used here is real, charge conjugation C does not do anything.) The proof is based on the beautiful idea to allow the spacetime coordinates X^μ to be complex, and one can analytically continue the whole theory into this complex domain. This leads also to complex Lorentz transformations with the effect that one can now go continuously between the two regions $\Lambda^0{}_0 \geq +1$ and $\Lambda^0{}_0 \leq -1$ which are disconnected in the real domain. A possible transformation is

$$\Lambda^\mu{}_\nu(\theta) = \begin{pmatrix}\cos\theta & i\,\sin\theta & 0 & 0\\ i\,\sin\theta & \cos\theta & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \qquad \Lambda^\mu{}_\nu(0) = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \qquad \Lambda^\mu{}_\nu(\pi) = \begin{pmatrix}-1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

which is not a legitimate Lorentz transformation because of the imaginary i in $i\,\sin\theta$ but it is a complex Lorentz transformation. It keeps the metric invariant in the sense that it satisfies $\Lambda^\mu{}_\rho\,\Lambda^\nu{}_\sigma\,\eta_{\mu\nu} = \eta_{\rho\sigma}$, and it has a continuous parameter. With θ continuously going from 0 to π the element $\Lambda^0{}_0$ goes from $+1$ to -1. One can now ask whether one can also go from det $= +1$ to det $= -1$ but this is not possible because det is always either $+1$ or -1. As shown in figure 2, PT changes $\Lambda^0{}_0$ from $+1$ to -1 but keeps the determinant $+1$. The PT theorem can be proven rigorously by using some basic properties of the analytically continued theory.

When defining P not the most general transformation has been used because it could have been defined as $\mathsf{P}^{-1}\varphi(\vec{x},t)\mathsf{P} = \omega\,\varphi(-\vec{x},t)$ including a complex factor ω with $|\omega| = 1$. However, ω^2 must be 1 such that ω must be ± 1 because of $\mathsf{P}^2 = \mathsf{1}$. This is actually used in Particle Physics where those with $\omega = +1$ are called scalars and those with $\omega = -1$ are called pseudoscalars. The pseudoscalars also flip sign under parity.

The reason why $(\tilde{\mathsf{PT}}^{-1})\varphi(X)(\tilde{\mathsf{PT}}) = -\varphi(-X)$ is not possible is that not all theories would be invariant under $\tilde{\mathsf{PT}}$, and an example is a theory with an interaction term $\mathcal{L}_{\text{int}} = -\lambda\,\varphi^3$ in the Lagrangian or any other odd power of φ. One can interpret $\tilde{\mathsf{PT}} = \mathsf{PT}\cdot\mathsf{S}$ with $\mathsf{S}^{-1}\varphi(X)\mathsf{S} = -\varphi(X)$ changing the sign. The PT theorem states that PT is invariant and not $\tilde{\mathsf{PT}}$. Sometimes P is a symmetry and sometimes $\mathsf{P}\cdot\mathsf{S}$ is a symmetry but PT is always a symmetry.

4.9 Charge Conjugation

Another discrete symmetry is charge conjugation C. If φ is a free complex scalar quantum field with

$$\varphi(X) = \int (dP)\Big[\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\beta}^\dagger(P)\,e^{+i\,P\cdot X}\Big]$$

charge conjugation should act as $\mathsf{C}^{-1}\,\boldsymbol{\alpha}(P)\,\mathsf{C} = \boldsymbol{\beta}(P)$ and $\mathsf{C}^{-1}\,\boldsymbol{\beta}(P)\,\mathsf{C} = \boldsymbol{\alpha}(P)$ because it is supposed to change creation and annihilation operators for positive charged particles into those for negative charged particles and vice versa. Applied to the field gives $\mathsf{C}^{-1}\,\varphi(X)\,\mathsf{C} = \varphi^\dagger(X)$.

To check whether C is unitary one can apply C to a general state

$$|\{P\}_+,\{P'\}_-\rangle = \left(\prod_P \boldsymbol{\alpha}^\dagger(P)\right)\left(\prod_{P'} \boldsymbol{\beta}^\dagger(P')\right)|0\rangle$$

and use $\mathsf{C}\,|\{P\}_+,\{P'\}_-\rangle = |\{P'\}_+,\{P\}_-\rangle$ to show that

$$\langle\{Q\}_+,\{Q'\}_-|\,\mathsf{C}\,|\{P\}_+,\{P'\}_-\rangle \qquad\qquad \langle\{Q\}_+,\{Q'\}_-|\,\mathsf{C}^\dagger\,|\{P\}_+,\{P'\}_-\rangle$$

are the same. This proves that $\mathsf{C} = \mathsf{C}^\dagger = \mathsf{C}^{-1}$ and therefore that C is unitary. Because PT acting on the complex field φ gives also $(\mathsf{PT})^{-1}\,\varphi(X)\,(\mathsf{PT}) = \varphi^\dagger(X)$ it follows that C as well as PT turn φ into $\varphi^\dagger$ and therefore CPT is the identity. This is the CPT symmetry for free complex scalar fields. The same sort of analytic continuation mentioned above tells that it is a symmetry also for interacting complex scalar fields. The same argument can be generalized for arbitrary spins and that is the famous CPT theorem.

5 Interactions

5.1 Using Perturbation for Solving Theories with Interactions

The free theories can be solved, but as soon as there are interactions this is no longer the case. If one has, for example, a scalar field theory and its corresponding equations of motion

$$S = \int (d^4X \left[\frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m^2\varphi^2 - \frac{\lambda}{4!}\varphi^4\right] \qquad\qquad \big(\Box + m^2\big) = \frac{\lambda}{3!}\varphi^3$$

this theory can be solved for $\lambda = 0$ by giving the most general solution because the Heisenberg equations of motion are linear. For $\lambda \neq 0$, the equations become non-linear and cannot be solved. However, interactions are crucial for particles, and almost everything known about the interacting Quantum Field Theory comes from treating the interaction terms perturbatively.

The idea is that for zeroth order, the field is a free field with the most general solution

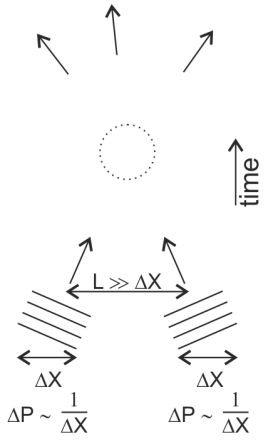
$$\varphi(X) = \int (dP)\Big[\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\alpha}^\dagger(P)\,e^{+i\,P\cdot X}\Big] + O(\lambda)$$

as in (2.8) but with a mistake $O(\lambda)$ coming from ignoring the interaction term. Given

$$\boldsymbol{H}_{\rm int} = \int d^3\vec{x}\,\frac{\lambda}{4!}\varphi^4 \qquad\qquad \sim\lambda(\boldsymbol{\alpha}+\boldsymbol{\alpha}^\dagger)^4 + O(\lambda^2)$$

the interaction Hamiltonian on the left side plugged into (2.8) leads to the correction on the right side where most factors have been dropped. The term $\boldsymbol{\alpha}^2\,(\boldsymbol{\alpha}^\dagger)^2$ in $\sim \lambda(\boldsymbol{\alpha}^4+\boldsymbol{\alpha}^3(\boldsymbol{\alpha}^\dagger)+\boldsymbol{\alpha}^2\,(\boldsymbol{\alpha}^\dagger)^2+\boldsymbol{\alpha}(\boldsymbol{\alpha}^\dagger)^3+(\boldsymbol{\alpha}^\dagger)^4)$ annihilates two particles and creates two particles back again but may not give back the same momenta because of the sum over momenta. In other words, this is scattering where two particles come in and two particles go out. Thus, it is somehow clear what the term $\boldsymbol{\alpha}^2\,(\boldsymbol{\alpha}^\dagger)^2$ does, but it is less clear what the role of the other terms is. To summarize, the basic idea is that the interaction Hamiltonian gives rise to transitions between states and that includes scattering.

5.2 Basics of Scattering



Scattering is experimentally and conceptually the main physical observable available on relativistic particle interactions. The basic setup for scattering is shown in the figure on the right side. In the initial state there are wave packets of single particle states. Thus, single particles are localized in wave packets with some position uncertainty ΔX and some momentum uncertainty $\Delta P \sim \frac{1}{\Delta X}$. It is assumed that all the initial wave packets have the same position space width and the same momentum space width. If an experimentalist wants to specify the momentum very accurately he has to use larger wave packets, but for any given accuracy for momentum one can always make the wave packets initially separated by distance $L \gg \Delta X$. These wave packets have therefore some quasi well-defined position and momentum consistent with the uncertainty principle. The experimentalist aims these wave packets such that they propagate into a

region where they overlap, interact and scatter. A lot of complicated things like the creation of particles may happen but the assumption is that all these particles will eventually stream out such that they are separated at some point in the future and do no longer interact. Thus, the particles in a scattering experiment do not interact at very early times and very late times.

The assumption that the particles do not interact at early and late times means that the Heisenberg fields $\varphi(X)$ become free fields at $t \to \pm\infty$. This property has not been derived but has been asserted. In physics one often does not derive assumptions but makes assumptions and is interested in testing them and their consequences. If things do not work out then the assumptions have to be modified. It will turn out that there are problems with this assumption for massless particles.

The conjecture is more precisely that

$$\varphi(\vec{x},t) \begin{cases} \to \sqrt{z} \cdot \varphi_{\text{in}}(\vec{x},t) & t \to -\infty \\ \to \sqrt{z} \cdot \varphi_{\text{out}}(\vec{x},t) & t \to +\infty \end{cases}$$

with the same factor $\sqrt{z}$ for the incoming and outgoing canonically normalized free fields φ_{in} and φ_{out} where $\varphi_{\text{in}}(X) = \int (dP)[\alpha_{\text{in}}(P)e^{-i\,P\cdot X} + \alpha^\dagger_{\text{in}}(P)e^{+i\,P\cdot X}]$ and $[\alpha_{\text{in}}(P), \alpha^\dagger_{\text{in}}(Q)] = \langle P|Q\rangle$ (and similarly for φ_{out}) is meant by "canonically normalized". The fact that $\varphi_{\text{in}}(X)\,|0\rangle = \int (dP) e^{+i\,P\cdot X}\,|P\rangle$ because only the creation operator matters is the reason for the factor $\sqrt{z}$. Further $\varphi(X)\,|0\rangle$ may not only create one particle but many. The fact that $\varphi(X)\,|0\rangle = \sqrt{z}\int (dP) e^{+i\,P\cdot X}\,|P\rangle + ...$ also creates other states is what makes it an interacting field, and $\sqrt{z} < 1$ because of these other states and as a consequence of unitarity. Actually $\sqrt{z} = 1 + O(\lambda)$ can be computed where $O(\lambda)$ stands for the perturbative corrections.

5.3 Perturbation Expansion in the Interaction Picture

In the Schrödinger picture the states evolve with time such that $|\psi(t)\rangle_S = e^{-i\,\boldsymbol{H}(t-t_0)}\,|\psi(t_0)\rangle_S$ while the operators $\boldsymbol{O}_S$ are time independent. In the Heisenberg picture the states are time independent such that $|\psi\rangle_H = e^{+i\,\boldsymbol{H}(t-t_0)}\,|\psi(t)\rangle_S = |\psi(t_0)\rangle_S$ where time t_0 is the time where both pictures agree, but the operators depend on time such that $\boldsymbol{O}_H(t) = e^{+i\,\boldsymbol{H}(t-t_0)}\,\boldsymbol{O}_S\,e^{-i\,\boldsymbol{H}(t-t_0)}$.

In the interaction picture also called Dirac picture states as well as operators depend on time. It is useful when the full Hamiltonian is $\boldsymbol{H} = \boldsymbol{H}_0 + \boldsymbol{H}_{\text{int}}$ where the leading order Hamiltonian $\boldsymbol{H}_0$ can be solved exactly and $\boldsymbol{H}_{\text{int}}$ contains some interaction terms to be treated perturbatively as a small perturbation in practical applications. Thus, the first term is free field theory and the second is interaction. In the interaction picture one does not take out all time dependence but only the one in $\boldsymbol{H}_0$ such that $|\psi(t)\rangle_I = e^{+i\,\boldsymbol{H}_0(t-t_0)}\,|\psi(t)\rangle_S$ and $\boldsymbol{O}_I(t) = e^{+i\,\boldsymbol{H}_0(t-t_0)}\,\boldsymbol{O}_S\,e^{-i\,\boldsymbol{H}_0(t-t_0)}$.

The time evolution in the interaction picture is defined as $|\psi(t)\rangle_I = \boldsymbol{U}_I(t,t_0)\,|\psi(t_0)\rangle_I$ for the operator

$$\boldsymbol{U}_I(t,t_0) = e^{+i\,\boldsymbol{H}_0(t-t_0)}\,e^{-i\,\boldsymbol{H}(t-t_0)} \tag{5.1}$$

which is unitary since it is the product of unitary operators. This shows why the interaction picture is useful for scattering. Since the time evolution is free at early and late times it is governed by $\boldsymbol{H}_0$. The state $|\psi(t)\rangle_I$ is therefore time independent for $t \to \pm\infty$. The S-matrix is defined as

$$\mathsf{S} = \lim_{\substack{t_i \to -\infty \\ t_f \to +\infty}} \boldsymbol{U}_I(t_f,t_i)$$

and this limit exists. The states stop evolving in time at early and late times. The S-matrix shows the probability to have a transition between different interaction picture states and these interaction picture states are defined to have the free time evolution. This sort of trivial time evolution is taken out of the states.

One can relate this directly to physical observables because the probability to make a transition from some initial state to some final state is $\text{Prob}(i \to f) \propto |\langle f|\mathsf{S}|i\rangle|^2$. Thus, this S-matrix is the basic object one needs to inspect in order to study scattering.

In order to compute the S-matrix a Lorentz-invariant definition of it is needed to replace the above definition (5.1). To get a new expression for this, the strategy is to write a differential equation for this

sought object and then solve it. The differential equation is

$$\begin{aligned}\frac{\partial}{\partial t}\boldsymbol{U}_I(t,t_0) &= e^{+i\,\boldsymbol{H}_0(t-t_0)}\,i\,\boldsymbol{H}_0\,e^{-i\,\boldsymbol{H}(t-t_0)} + e^{+i\,\boldsymbol{H}_0(t-t_0)}\,(-i\,\boldsymbol{H})\,e^{-i\,\boldsymbol{H}(t-t_0)}\\ &= e^{+i\,\boldsymbol{H}_0(t-t_0)}\,(-i\,\boldsymbol{H}_{\text{int}})\,e^{-i\,\boldsymbol{H}(t-t_0)}\\ &= e^{+i\,\boldsymbol{H}_0(t-t_0)}\,(-i\,\boldsymbol{H}_{\text{int}})\,e^{-i\,\boldsymbol{H}_0(t-t_0)}\,e^{+i\,\boldsymbol{H}_0(t-t_0)}\,e^{-i\,\boldsymbol{H}(t-t_0)}\\ &= -i\,\boldsymbol{H}_I(t)\,\boldsymbol{U}_I(t,t_0)\end{aligned}$$

where $\boldsymbol{H}_I(t)$ is the interaction Hamiltonian in the interaction picture. This differential equation is a Schrödinger equation with the initial condition is $\boldsymbol{U}_I(t_0,t_0)=\boldsymbol{1}$ according to (5.1). The solution is

$$\boldsymbol{U}_I(t,t_0) = \mathbf{Texp}\left\{-i\int_{t_0}^{t} dt'\,\boldsymbol{H}_I(t')\right\} = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{t_0}^{t} dt_1 ... \int_{t_0}^{t} dt_n \mathbf{T}\big\{\boldsymbol{H}_I(t_1)\,...\,\boldsymbol{H}_I(t_n)\big\}$$

where **Texp** is the time-ordered exponential and **T** is the time-ordering operator acting as

$$\mathbf{T}\big\{\boldsymbol{O}_1(t_1)\,\boldsymbol{O}_2(t_2)\big\} = \begin{cases}\boldsymbol{O}_1(t_1)\,\boldsymbol{O}_2(t_2) & t_1 > t_2\\ \boldsymbol{O}_2(t_2)\,\boldsymbol{O}_1(t_1) & t_1 < t_2\end{cases}$$

on operators. Time ordering has the obvious property $\mathbf{T}\big\{\boldsymbol{A},\boldsymbol{B}\big\} = \mathbf{T}\big\{\boldsymbol{B},\boldsymbol{A}\big\}$. Another property is that the time-ordered product of Hermitian operators is also Hermitian.

To check this solution one can calculate

$$\begin{aligned}\frac{\partial}{\partial t}\boldsymbol{U}_I(t,t_0) &= \frac{\partial}{\partial t}\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{t_0}^{t} dt_1 ... \int_{t_0}^{t} dt_n \mathbf{T}\big\{\boldsymbol{H}_I(t_1)\,...\,\boldsymbol{H}_I(t_n)\big\}\\ &= \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\left[\int_{t_0}^{t} dt_2 ... \int_{t_0}^{t} dt_n \mathbf{T}\big\{\boldsymbol{H}_I(t)\,\boldsymbol{H}_I(t_2)\,...\,\boldsymbol{H}_I(t_n)\big\} + ...\right.\\ &\qquad\left. + \int_{t_0}^{t} dt_1 ... \int_{t_0}^{t} dt_{n-1} \mathbf{T}\big\{\boldsymbol{H}_I(t)\,\boldsymbol{H}_I(t_1)\,...\,\boldsymbol{H}_I(t_{n-1})\big\}\right]\\ &= \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\,n\,\boldsymbol{H}_I(t)\left[\int_{t_0}^{t} dt_2 ... \int_{t_0}^{t} dt_n \mathbf{T}\big\{\boldsymbol{H}_I(t_2)\,...\,\boldsymbol{H}_I(t_n)\big\} + ...\right.\\ &\qquad\left. + \int_{t_0}^{t} dt_1 ... \int_{t_0}^{t} dt_{n-1} \mathbf{T}\big\{\boldsymbol{H}_I(t_1)\,...\,\boldsymbol{H}_I(t_{n-1})\big\}\right]\\ &= -i\boldsymbol{H}_I(t)\,\boldsymbol{U}_I(t,t_0)\end{aligned}$$

where $\boldsymbol{H}_I(t)$ is in the time ordering always the term with the highest value of t because t is the upper limit of the integrals. The sum becomes a sum from 1 to ∞ and the $n!$ becomes $(n-1)!$ such that the original sum and the resulting sum are the same when adapting the summation variable $n \to n-1$. This is called Dyson's formula.

The S-matrix is now

$$\mathsf{S} = \mathbf{Texp}\left\{-i\int_{-\infty}^{+\infty} dt\,\boldsymbol{H}_I(t)\right\} = \mathbf{Texp}\left\{+i\int d^4X\,\mathcal{L}_{\text{int}}(\varphi_I(X))\right\} \tag{5.2}$$

because $\boldsymbol{H}_I(t) = -\int d^3\vec{x}\,\mathcal{L}_{\text{int}}(\varphi_I(X))$. Formula (5.2) is manifestly Lorentz-invariant. The time-ordered exponential is Lorentz-invariant because fields commute at spacelike separation where any time ordering does not matter and at timelike separation time ordering matters but is Lorentz-invariant. Expanding the time-ordered exponential into time-ordered products of fields gives

$$\mathsf{S} = \boldsymbol{1} + i\int d^4X\left(-\frac{\lambda}{4!}\,\varphi_I^4(X)\right) + \frac{i^2}{2!}\int d^4X\int d^4Y\left(-\frac{\lambda}{4!}\right)\,\mathbf{T}\big\{\varphi_I^4(X)\,\varphi_I^4(Y)\big\} + O(\lambda^3)$$

and the object to be studied are therefore time-ordered products of fields.

5.4 Calculations in a Toy Model

One needs a theory to compute something, and the theory used here is a toy model not describing the real world. It is a simple model that allows to do the first explicit calculations. The real scalar field φ is called a "meson" and its mass is M. There is also a complex scalar field ψ called "nucleon" with mass m. In reality nucleons are fermions but in this toy model they are just scalars. The complex field ψ has a U(1)-symmetry acting as $\varphi \to \varphi$ and $\psi \to e^{-i\,\theta}\,\psi$. In this toy world one can think of this symmetry as a charge and may call it "baryon number". The Lagrangian density is the sum of the free term $\mathcal{L}_0$ and the interaction term $\mathcal{L}_{\rm int}$

$$\mathcal{L}_0 = \frac{1}{2}\,(\partial\varphi)^2 - \frac{1}{2}\,M^2\,\varphi^2 + \partial\psi^\dagger\,\partial\psi - \frac{1}{2}\,m^2\,\psi^\dagger\psi \qquad\qquad \mathcal{L}_{\rm int} = g\,\varphi\,\psi^\dagger\psi$$

where $\partial\psi^\dagger\,\partial\psi$ stands for $\partial^\mu\psi^\dagger\,\partial_\mu\psi$.

With $c = \hbar = 1$ there is only the unit of mass left, and counting units of mass gives $[\boldsymbol{H}] = +1$ and $[x] = [t] = -1$ such that the derivative ∂_t has $[\partial_t] = +1$. Therefore, dimensional analysis reveals from the Hamiltonian $\boldsymbol{H} = \int d^3\vec{x}[\frac{1}{2}(\partial_t\varphi)^2 + ...]$ that $1 = 3(-1) + 2(+1 + [\varphi])$ and consequently that $[\varphi] = +1$. It is more convenient to use the action $S = \int dt\,L = \int d^4X\,\mathcal{L}$ with $[S] = 0$ and therefore $[\mathcal{L}] = +4$ for the calculations. This leads also to $[\varphi] = +1$ because $4 = 2([\varphi] + 1)$. Further follows from $\mathcal{L}_{\rm int} = g\,\varphi\,\psi^\dagger\psi$ that $4 = [g] + 3$ and $[g] = +1$. This is important because the plan is to do a perturbative expansion in orders of g.

Therefore the dimension of g is important, and the corrections have to be powers of g divided by something such as an energy E with dimension mass. For large energies E this is good because going to higher and higher energies in the same scattering process with the same g the expansions $(g/E)^{n>0}$ become better and better. On the other hand, this expansion breaks down for small values of E. Thus, without doing any calculations this dimensional analysis reveals fundamentally important properties. This type of interaction which is very important at small energies but not important at high energies is called relevant.

Operators in the interaction picture $\boldsymbol{O}_I(t) = e^{+i\,\boldsymbol{H}_0(t-t_0)}\,\boldsymbol{O}_S\,e^{-i\,\boldsymbol{H}_0(t-t_0)}$ satisfy the equation of motion $\frac{d}{dt}\boldsymbol{O}_I(t) = i[\boldsymbol{H}_0, \boldsymbol{O}_I(t)]$ as in the Heisenberg picture but now with $\boldsymbol{H}_0$ instead of $\boldsymbol{H}$. In other words, the fields have time evolution in the interaction picture but the only time evolution they have is the free time evolution. This means in field theory that the interaction picture fields $\varphi_I(X)$ are free fields and therefore that one can actually do calculations.

At time t_0 the interaction picture agrees with the Heisenberg picture. Thus, with $t_0 \to -\infty$ the interaction picture fields $\varphi_I(X)$ agree with the Heisenberg picture fields at time $t = -\infty$ and $\varphi_I(X) = \sqrt{z}\,\varphi_{\rm in}(X)$, and one can write the interaction picture fields in terms of free fields. In the following it is assumed $\sqrt{z} = 1$ and $\varphi(X)$ means $\varphi_I(X)$ because the calculations are all done in the interaction picture.

The meson field φ and the nucleon field ψ in the toy model are

$$\varphi(X) = \int (dP)_M \Big[\boldsymbol{\alpha}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\alpha}^\dagger(P)\,e^{+i\,P\cdot X}\Big] \quad \psi(X) = \int (dP)_m \Big[\boldsymbol{\beta}(P)\,e^{-i\,P\cdot X} + \boldsymbol{\gamma}^\dagger(P)\,e^{+i\,P\cdot X}\Big]$$

where $\boldsymbol{\alpha}(P)$ is used for the creation and annihilation operator for φ while for ψ the operators $\boldsymbol{\beta}(P)$ and $\boldsymbol{\gamma}(P)$ are used. The operators $\boldsymbol{\alpha}^\dagger$ and $\boldsymbol{\alpha}$ create or annihilate mesons, $\boldsymbol{\beta}(P)$ annihilates nucleons and therefore $\boldsymbol{\gamma}^\dagger(P)$ creates antinucleons. The distinction between

$$(dP)_M = \frac{d^3\vec{p}}{(2\pi)^3}\,\frac{1}{2\sqrt{\vec{p}^2 + M^2}} \qquad\qquad (dP)_m = \frac{d^3\vec{p}}{(2\pi)^3}\,\frac{1}{2\sqrt{\vec{p}^2 + m^2}}$$

is used to indicate that the two integrals go over different mass shells.

The simplest process one can compute in this toy model is the decay process where a mesons decays into a nucleon and an antinucleon if $M > 2m$. The meson cannot decay into two nucleons because of charge conservation. One can ask how fast the decay process messon $\to$ nucleon+antinucleon happens because the decay rate is a physical observable. The initial state is $|i\rangle = \boldsymbol{\alpha}^\dagger(P)\,|0\rangle$ and makes sure that there is one meson. The final state is $|f\rangle = \boldsymbol{\beta}^\dagger(Q_1)\,\boldsymbol{\gamma}^\dagger(Q_2)\,|0\rangle$ making sure that there is a nucleon and an antinucleon. The value to be computed is $\langle f|\mathsf{S}|i\rangle = \langle 0|\boldsymbol{\beta}(Q_1)\boldsymbol{\gamma}(Q_2)\mathsf{S}\boldsymbol{\alpha}^\dagger|0\rangle$ where the equation for the S-matrix to be used is (5.2) adapted to the toy model using $\mathcal{L}_{\rm int} = g\,\varphi\psi^\dagger\psi$.

This gives

$$\langle f|\mathsf{S}|i\rangle = \langle 0|\boldsymbol{\beta}(Q_1)\,\boldsymbol{\gamma}(Q_2)\left[\mathbf{1} + i\,g\int d^4X\,\psi^\dagger(X)\,\psi(X)\varphi(X) + O(g^2)\right]\boldsymbol{\alpha}^\dagger(P)|0\rangle$$

for the leading terms. This is

$$\langle f|\mathsf{S}|i\rangle = \langle 0|\boldsymbol{\beta}(Q_1)\,\boldsymbol{\gamma}(Q_2)\,\boldsymbol{\alpha}^\dagger(P)|0\rangle + O(g) = \langle 0|\boldsymbol{\alpha}^\dagger(P)\,\boldsymbol{\beta}(Q_1)\,\boldsymbol{\gamma}(Q_2)|0\rangle + O(g) = 0 + O(g)$$

if only the term $\mathbf{1}$ is taken into account but this gives zero because the involved creation and annihilation operators commute with each other. In order to get a non-zero result a $\boldsymbol{\gamma}^\dagger$ must stop the $\boldsymbol{\gamma}$ from moving to the vacuum state on the right. Similarly there must be a $\boldsymbol{\alpha}$ and a $\boldsymbol{\beta}^\dagger$. The next higher order term is

$$\begin{aligned}
\langle f|\mathsf{S}|i\rangle &= i\,g\int d^4X\int (dK_1)_m\int (dK_2)_m\int (dK_3)_M \\
&\qquad \langle 0|\boldsymbol{\beta}(Q_1)\,\boldsymbol{\gamma}(Q_2)\,\boldsymbol{\beta}^\dagger(K_1)\,e^{+i\,K_1\cdot X}\,\boldsymbol{\gamma}^\dagger(K_2)\,e^{+i\,K_2\cdot X}\,\boldsymbol{\alpha}(K_3)\,e^{-i\,K_3\cdot X}\,\boldsymbol{\alpha}^\dagger(P)|0\rangle + O(g^2) \\
&= i\,g\int d^4X\int (dK_1)_m\int (dK_2)_m\int (dK_3)_M\ \langle Q_1|K_1\rangle\,\langle Q_2|K_2\rangle\,\langle P|K_3\rangle\ e^{i(K_1+K_2-K_3)\cdot X} + O(g^2) \\
&= i\,g\int d^4X\,e^{i(Q_1+Q_2-P)\cdot X} + O(g^2) = i\,g\,(2\pi)^4\,\delta^4\big(P-(Q_1+Q_2)\big) + O(g^2)
\end{aligned}$$

where all the integrals can be resolved such that the result is basically a four-dimensional delta function ensuring that the amplitude is only non-zero if energy and momentum is preserved.

It is not clear how to interpret this result physically because the probability is proportional to $|\langle f|\mathsf{S}|i\rangle|^2$, and the question is how to interpret a delta function squared. The delta function squared diverges and this is related to the infinite volume and the infinite time. To understand the result one has to go back to a finite volume and a finite time.

5.5 Fermi's Golden Rule

Putting physics back into a spatial box of size $L\times L\times L$ and limiting the time extent by T leads to the boundary conditions $\varphi(\vec{x}+L\,\hat{e},t+T)=\varphi(\vec{x},t)$ with $\hat{e}=\hat{x},\hat{y},\hat{z}$. Now the states are discrete such that momenta, for example, become $\vec{p}=\frac{2\pi}{L}(n_1,n_2,n_3)$ where the vector components n_i are integers, and delta functions will be replaced by Kronecker deltas such that ${}_V\langle\vec{p}|\vec{q}\rangle_V=\delta_{\vec{p},\vec{q}}$ where ${}_V\langle\vec{p}|$ and $|\vec{p}\rangle_V$ indicate that the states are meant in the box with volume $V=L^3$.

To relate the discrete normalized state $|\vec{p}\rangle_V$ in the box with the continuum normalized state $|P\rangle$ one can make use of

$$\sum_{\vec{p}} \to V\int\frac{d^3\vec{p}}{(2\pi)^3}$$

when turning sums over momenta into integrals, and the continuous limit for the Fourier transform

$$\int d^3\vec{x}\,e^{i(\vec{p}-\vec{q})\cdot\vec{x}} = V\,\delta_{\vec{p},\vec{q}} \to (2\pi)^3\,\delta^3(\vec{p}-\vec{q})$$

shows how to identify a continuous delta function with discrete Kronecker delta. Thus,

$${}_V\langle\vec{p}|\vec{q}\rangle_V = \delta_{\vec{p},\vec{q}} \to \frac{1}{V}(2\pi)^3\,\delta^3(\vec{p}-\vec{q}) = \frac{1}{V}\frac{1}{2E_{\vec{p}}}\,\langle P|Q\rangle$$

with the relativistically normalized state $\langle P|Q\rangle$, and

$$|P\rangle = \sqrt{2E_{\vec{p}}\,V}\ |\vec{p}\rangle_V$$

is the relation between the continuous and the discrete normalized state. This is the basic formula that connects the relativistically normalized state $|P\rangle$ to the discretely normalized state $|\vec{p}\rangle_V$.

The next step is to write the S-matrix elements of these discrete states $|\vec{p}\rangle_V$ because for these discrete states the S-matrix is literally just a transition probability. The calculations give

$$
\begin{aligned}
{}_V\langle f|\mathsf{S}|i\rangle_V &= {}_V\langle \vec{q}_1...\vec{q}_m|\mathsf{S}|\vec{p}_1...\vec{p}_n\rangle_V = \left(\prod_i \frac{1}{\sqrt{2E_i\,V}}\right)\left(\prod_f \frac{1}{\sqrt{2E_f\,V}}\right)\langle Q_1...Q_m|\mathsf{S}|P_1...P_n\rangle \\
&= VT\,\delta_{\mathcal{P}_i,\mathcal{P}_f}\left(\prod_i \frac{1}{\sqrt{2E_i\,V}}\right)\left(\prod_f \frac{1}{\sqrt{2E_f\,V}}\right) i\,\mathcal{M}_{fi}
\end{aligned}
$$

for states normalized in finite volume. This uses

$$
\langle \mathcal{P}_f|\mathsf{S}|\mathcal{P}_i\rangle = (2\pi)^4\,\delta^4(\mathcal{P}_i-\mathcal{P}_f)\,i\,\mathcal{M}_{fi} = VT\,\delta_{\mathcal{P}_i,\mathcal{P}_f}
$$

with the definitions $\mathcal{P}_f = Q_1...Q_m$ and $\mathcal{P}_i = P_1...P_n$. The coefficient $\mathcal{M}_{fi}$ is the Lorentz-invariant amplitude, and a main goal is to calculate it. In the previous example this coefficient has been $\mathcal{M}_{fi} = g$. The probability to go from the initial state to the final state is

$$
\text{Prob}(i\to f) = |{}_V\langle f|\mathsf{S}|i\rangle_V|^2 = V^2T^2\,\delta_{\mathcal{P}_i,\mathcal{P}_f}\left(\prod_i \frac{1}{2E_i\,V}\right)\left(\prod_f \frac{1}{2E_f\,V}\right)|\mathcal{M}_{fi}|^2
$$

because the square of a Kronecker delta is just a Kronecker delta. Back in the continuum

$$
|{}_V\langle f|\mathsf{S}|i\rangle_V|^2 \to VT\,(2\pi)^4\,\delta^4(\mathcal{P}_i-\mathcal{P}_f)\left(\prod_i \frac{1}{2E_i\,V}\right)\left(\prod_f \frac{1}{2E_f\,V}\right)|\mathcal{M}_{fi}|^2
$$

is the probability to go to one particular discrete state but in the continuum this does not make sense because one has to integrate over a range of final states. Summing over final states means

$$
\sum_f \to \int\left(\prod_f V\frac{d^3\vec{q}_f}{(2\pi)^3}\right)
$$

such that the differential rate

$$
\begin{aligned}
d\Gamma &= \left(\prod_f V\frac{d^3\vec{q}_f}{(2\pi)^3}\right)\frac{\text{Prob}(i\to f)}{T} \\
&= \left(\prod_f V\frac{d^3\vec{q}_f}{(2\pi)^3}\right) V\,(2\pi)^4\,\delta^4(\mathcal{P}_i-\mathcal{P}_f)\left(\prod_i \frac{1}{2E_i\,V}\right)\left(\prod_f \frac{1}{2E_f\,V}\right)|\mathcal{M}_{fi}|^2 \\
&= \left(\prod_f (dQ_f)\right)(2\pi)^4\,\delta^4(\mathcal{P}_i-\mathcal{P}_f)\,V\left(\prod_i \frac{1}{2E_i\,V}\right)|\mathcal{M}_{fi}|^2
\end{aligned}
$$

can be defined. The result for the differential rate is

$$
d\Gamma = \left(\prod_f (dQ_f)\right)(2\pi)^4\,\delta^4(\mathcal{P}_i-\mathcal{P}_f)\,V\left(\prod_i \frac{1}{2E_i\,V}\right)|\mathcal{M}_{fi}|^2 \tag{5.3}
$$

and is sometimes called Fermi's golden rule. It shows that these factors V in general are not canceling out, and the question remains whether this really makes sense in the infinite volume limit.

In a decay the initial state $|i\rangle$ is a one-particle state and the V does cancel. The differential rate (5.3) can be written as

$$
d\Gamma = \frac{1}{2E_i}\left(\prod_f (dQ_f)\right)(2\pi)^4\,\delta^4(\mathcal{P}_i-\mathcal{P}_f)\,|\mathcal{M}_{fi}|^2
$$

and this is almost Lorentz-invariant except for the factor $\frac{1}{2E_i}$. However, the decay rate should not be Lorentz-invariant because it is a rate, and this factor is exactly the time dilation factor needed by Special Relativity.

The factor $d\Phi(\mathcal{P}_i)$ defined as

$$d\Phi(\mathcal{P}_i) = \left(\prod_f (dQ_f)\right) (2\pi)^4 \, \delta^4(\mathcal{P}_i - \mathcal{P}_f) \tag{5.4}$$

is called the Lorentz-invariant phase space with momentum $\mathcal{P}_i$ and depends on the total initial momentum. The total decay rate in the rest frame is

$$\Gamma = \int d\Gamma = \frac{1}{2M} \int d\Phi_f(P) \, |\mathcal{M}_{fi}|^2$$

where the energy is just the mass M.

5.6 Total Decay Rate in the Toy Model

In the toy model $|\mathcal{M}_{fi}|^2 = g^2$ is a constant and Γ becomes

$$\Gamma = \frac{g^2}{2M} \int \frac{d^3\vec{q}_1}{(2\pi)^3} \frac{1}{2E_{\vec{q}_1}} \int \frac{d^3\vec{q}_2}{(2\pi)^3} \frac{1}{2E_{\vec{q}_2}} \, (2\pi)^4 \delta^4\big(P_i - (Q_1 + Q_2)\big)$$

with $P_i^\mu = (M, 0, 0, 0)$ in the rest frame. One can use δ^3 to do the $d^3\vec{q}_2$ integral and that gives $\vec{q}_2 = -\vec{q}_1$ and $E_{\vec{q}_2} = E_{\vec{q}_1}$ because $\vec{p}_i$ is zero in the rest frame. Thus, one gets

$$\begin{aligned}
\Gamma &= \frac{g^2}{2M} \int \frac{d^3\vec{q}_1}{(2\pi)^3} \left(\frac{1}{2E_{\vec{q}_1}}\right) (2\pi) \, \delta\left(M - 2\sqrt{\vec{q}_1^{\,2} + m^2}\right) \\
&= \frac{g^2}{2M} \int_0^\infty \frac{4\pi \, q^2 \, dq}{(2\pi)^3} \frac{1}{4(q^2 + m^2)} \, 2\pi \, \delta\left(2\sqrt{\vec{q}_1^{\,2} + m^2} - M\right) \\
&= \frac{g^2}{8\pi \, M} \int_0^\infty dq \, \frac{q^2}{q^2 + m^2} \, \delta\left(2\sqrt{\vec{q}_1^{\,2} + m^2} - M\right)
\end{aligned}$$

where only the time delta function responsible for energy conservation remains in the first step and where spherical coordinates in momentum space with the radius q can be used in the second step because the equation only depends on $\vec{q}_1^{\,2}$ and not on $\vec{q}_1$. In the last step the argument in the delta function can be brought into the form

$$q = \frac{1}{2}\sqrt{M^2 - 4m^2}$$

and it follows $M > 2m$. This just means that there has to be enough energy in the meson at rest to create a nucleon and an antinucleon. With this formula for q the integral becomes

$$\Gamma = \frac{g^2}{8\pi \, M} \left[\frac{q^2}{q^2 + m^2} \frac{\sqrt{q^2 + m^2}}{4q}\right]_{q = \frac{1}{2}\sqrt{M^2 - 4m^2}} = \frac{g^2}{32\pi \, M} \left(1 - \frac{4m^2}{M^2}\right)^{\frac{1}{2}}$$

where the factor $\sqrt{q^2 + m^2}/4q$ is the Jacobian coming from the fact that the argument of the delta function is a function.

This is the first actual calculation of a physical quantity shown so far. The theory has three parameters where g is the coupling indicating the strength of the interaction, and where M and m are the two masses. The result is the total decay rate Γ for the decay of a meson into a nucleon and an antinucleon.

To make sure that the result makes sense one can check a few things. A first check may see whether the units are correct. The decay rate Γ is one over time and has therefore dimension $[\Gamma] = +1$ in units of mass. This works out because g also has dimension of mass. Further the dimensionless factor

$$\left(1 - \frac{4m^2}{M^2}\right)^{\frac{1}{2}} \to 0$$

for $M \to 2m$. This illustrates the fact that one has to sum over all of the final states to find the total decay rate, and when $M \approx 2m$ then there are far fewer states satisfying energy-momentum conservation to sum over than for $M \gg m$. In this case the decay rate is said to be phase-space suppressed. The decay rate of the neutron, for example, is extremely suppressed because of a factor like this, and this is responsible for the long lifetime of the neutron.

6 Particles with Spin

6.1 Relation between the Poincaré Group and Spin

It turns out that particles with spin can be associated with non-minimal representations of the Poincaré group. The Poincaré group is the union of the Lorentz transformations and the spacetime translations. For particles with spin the translations in spacetime play an important role because they become intertwined with Lorentz transformations. Thus, the goal is to find all the possible realizations of Poincaré invariance on states. That will tell all the different kinds of particles possible leading to Wigner's classification.

Elements of the Poincaré group act as $X^\mu \to \Lambda^\mu_{\ \nu} X^\nu + A^\mu$ on a point in spacetime or as $X \to \Lambda X + A$ in index-free notation. The elements of the Poincaré group are labeled by (Λ, A) and satisfy the group composition law

$$\begin{aligned} &X \overset{(\Lambda_1,A_1)}{\longrightarrow} \Lambda_1 X + A_1 = X' \overset{(\Lambda_2,A_2)}{\longrightarrow} \Lambda_2 X' + A_2 = \Lambda_2(\Lambda_1 X + A_1) + A_2 \\ &(\Lambda_2, A_2) \cdot (\Lambda_1, A_1) = (\Lambda_2\Lambda_1, \Lambda_2 A_1 + A_2) \end{aligned} \tag{6.1}$$

where the second Lorentz transformation acts on the first translation.

The Lorentz transformations are $\Lambda^\mu_{\ \nu} = \delta^\mu_{\ \nu} + \omega^\mu_{\ \nu} + O(\omega^2)$ where $\delta^\mu_{\ \nu}$ is the identity, $\omega^\mu_{\ \nu}$ is a deviation and $O(\omega^2)$ are the higher terms. For infinitesimal Lorentz transformations only the linear term ω is kept and the higher terms are ignored. The infinitesimal variation is $\delta X^\mu = X'^\mu - X^\mu$ with only the linear terms and is therefore $\delta X^\mu = \omega^\mu_{\ \nu} X^\nu$.

Because Lorentz transformations keep the metric the linear transformation ω with lowered indices is antisymmetric and therefore satisfies $\omega_{\mu\nu} = -\omega_{\nu\mu}$. Thus, $\omega_{\mu\nu}$ as an antisymmetric 4×4 matrix has six parameters corresponding to the three parameters for the rotations and the three parameters for the boosts.

A representation of the Poincaré group is needed that acts on states as $|\psi\rangle \to \boldsymbol{U}(\Lambda, A)\,|\psi\rangle$ in Schrödinger picture. In order to be a symmetry $\boldsymbol{U}(\Lambda, A)$, a unitary representation of the Poincaré group is needed, and in order to be a representation it must obey the group multiplication law

$$\boldsymbol{U}(\Lambda_2, A_2) \cdot \boldsymbol{U}(\Lambda_1, A_1) = \boldsymbol{U}(\Lambda_2\Lambda_1, \Lambda_2 A_1 + A_2)$$

according to (6.1). The goal is to classify all unitary matrices that satisfy this composition law. Due to Wigner one can pretty much give a complete answer to this mathematical question.

Instead of working with the finite group elements only the infinitesimal transformations are used. Without higher terms the operator $\boldsymbol{U}(\Lambda, A)$ is

$$\boldsymbol{U}(\Lambda, A) = \mathbf{1} + \frac{i}{2}\omega^{\mu\nu}\boldsymbol{M}_{\mu\nu} + iA^\mu \boldsymbol{P}_\mu + ... \tag{6.2}$$

and the two generators $\boldsymbol{M}_{\mu\nu}$ and $\boldsymbol{P}_\mu$ must be Hermitian because $\boldsymbol{U}(\Lambda, A)$ is unitary and ω as well as A are real. In addition $\boldsymbol{M}$ is antisymmetric since ω is antisymmetric. Thus, to summarize $\boldsymbol{M}^\dagger_{\mu\nu} = \boldsymbol{M}_{\mu\nu}$, $\boldsymbol{M}_{\mu\nu} = -\boldsymbol{M}_{\nu\mu}$ and $\boldsymbol{P}^\dagger_\mu = \boldsymbol{P}_\mu$ must be satisfied.

6.2 Lie Groups and Lie Algebras

Using a general notation for groups with continuous parameters θ_a and generators $\boldsymbol{T}_a$ where a is an index, the parameters are $\theta_a = \{\omega^{\mu\nu}, A^\mu\}$ and the generators are $\boldsymbol{T}_a = \{\frac{i}{2}\boldsymbol{M}_{\mu\nu}, i\boldsymbol{P}_\mu\}$ for the Poincaré group.

(Generators are sometimes defined without the factor i because there is no general agreement about the notation.) A group element g is, if infinitesimal, $g = \mathbf{1} + \boldsymbol{X} + ...$ where $\boldsymbol{X} = \theta_a \boldsymbol{T}_a$ is a linear combination (using summation over a). If g_1 and g_2 are two infinitesimal transformations then $g_1 g_2 = \mathbf{1} + \boldsymbol{X}_1 + \boldsymbol{X}_2 + ...$ is a generator. In fact linear combinations of generators are generators, and one can think of them as vectors in a vector space.

For a fixed group element g_0 any group element g can be transformed by g_0 as $g \to g' = g_0\, g\, g_0^{-1}$. This map preserves group composition $g_1 g_2 \to g_0\, g_1 g_2\, g_0^{-1}$ because $g_0\, g_1\, g_0^{-1}\, g_0 g_2\, g_0^{-1} = g_0\, g_1 g_2\, g_0^{-1}$. Mathematically this is called a homomorphism. It gives $g = \mathbf{1} + \boldsymbol{X} \to \mathbf{1} + g_0\, \boldsymbol{X}\, g_0^{-1}$ when applied to an infinitesimal group element g, and this is also a generator. The result is already a representation of the group on $\boldsymbol{X}$. The next case to consider is that $g_0 = \mathbf{1} + \boldsymbol{X}_0$ is infinitesimal where $g_0\, \boldsymbol{X}\, g_0^{-1} = \boldsymbol{X} + [\boldsymbol{X}_0, \boldsymbol{X}] + O(\boldsymbol{X}_0^2)$. Because $\boldsymbol{X}$ and $g_0\, \boldsymbol{X}\, g_0^{-1}$ are generators also $[\boldsymbol{X}_0, \boldsymbol{X}]$ must be a generator. Since $\boldsymbol{X}$ and $\boldsymbol{X}_0$ are completely arbitrary generators any commutator of generators must also be a generator. In other words, the generators close under commutation.

This is the basic fact that allows to reconstruct the finite transformations from the infinitesimal ones. To show this, a group element $\gamma(\varepsilon) = g \cdot (\mathbf{1} + \varepsilon \boldsymbol{X} + O(\varepsilon^2))$ is chosen. This is actually a curve in the group because there is a group element for every ε. It has the property $\frac{d}{d\varepsilon}\gamma(\varepsilon) = g\boldsymbol{X} + O(\varepsilon)$. This means that through the point g there is a curve $\gamma(\varepsilon)$ whose slope is $g\boldsymbol{X}$. It follows that there is a curve $\gamma_L(\lambda)$ where λ is a real parameter such that $\frac{d}{d\lambda}\gamma_L(\lambda) = \gamma_L(\lambda) \cdot \boldsymbol{X}$. This is a differential equation that defines a curve in the group. Similarly there is a curve $\gamma_R(\lambda)$ defined by $\frac{d}{d\lambda}\gamma_R(\lambda) = \boldsymbol{X} \cdot \gamma_R(\lambda)$. The initial condition is assumed to be $\gamma_L(0) = \gamma_R(0) = 1$. The solution is $\gamma_L(\lambda) = \gamma_R(\lambda) = e^{\lambda \boldsymbol{X}}$ and is therefore the same for both differential equations. The exponential $e^{\boldsymbol{X}}$ for matrices $\boldsymbol{X}$ is defined by

$$e^{\boldsymbol{X}} = \sum_{n=0}^{\infty} \frac{1}{n!} \boldsymbol{X}^n$$

as a power series. This shows that an exponential of a generator is a group element. In other words, the exponential of a generator gives a finite transformation one can think of iterating the infinitesimal transformations. The property without proof

$$e^{\boldsymbol{X}} e^{\boldsymbol{Y}} = \exp\left\{ \boldsymbol{X} + \boldsymbol{Y} + \frac{1}{2}[\boldsymbol{X}, \boldsymbol{Y}] + \frac{1}{12}\big([\boldsymbol{X}[\boldsymbol{X}, \boldsymbol{Y}] + [\boldsymbol{Y}, [\boldsymbol{Y}, \boldsymbol{X}]]\big) + ... \right\}$$

for two generators $\boldsymbol{X}$ and $\boldsymbol{Y}$ is called the Baker-Campbell-Hausdorff formula. The higher terms are all nested commutators and are therefore also generators. Thus, $\boldsymbol{Z}$ in $e^{\boldsymbol{Z}} = e^{\boldsymbol{X}} e^{\boldsymbol{Y}}$ is also a generator, and these exponentials form a group.

This kind of results usually holds at the level of infinite series and is only valid locally. An infinite series is an expansion around some point, and if it is true to all orders in a series expansion around that point, it usually means at best that it is true for some finite neighborhood around that point. Here the expansion is around the identity. A continuous group like this is called a Lie group, and the generators form a Lie algebra.

6.3 Generator Algebra for the Poincaré Group

The algebra of the generators is the same for all representations, and one can take the simplest representation which is the representation of the Poincaré group on scalar fields $\varphi(X)$. With

$$\varphi(X) \xrightarrow{(\Lambda, A)} \varphi'(X) \qquad \varphi'(X') = \varphi'(\Lambda X + A) = \varphi(X) \qquad \varphi'(X) = \varphi(\Lambda^{-1}(X - A))$$

and with $\Lambda^{-1} = 1 - \omega + ...$ in first order

$$\begin{aligned} \delta\varphi(X) &= \varphi'(X) - \varphi(X) = \varphi(X^\mu - \omega^\mu{}_\nu X^\nu - A^\mu) - \varphi(X) \\ &= -\omega^\mu{}_\nu X^\nu \partial_\mu \varphi(X) - A^\mu \partial_\mu \varphi(X) = \frac{i}{2} \omega^{\mu\nu} \boldsymbol{M}_{\mu\nu}\, \varphi(X) + i A^\mu \boldsymbol{P}_\mu\, \varphi(X) \end{aligned}$$

is the infinitesimal element of the Poincaré group using (6.2) for the generators of the Poincaré group. Matching the expressions in the last line gives

$$\boldsymbol{M}_{\mu\nu} = -\boldsymbol{M}_{\nu\mu} = i(X_\mu \partial_\nu - X_\nu \partial_\mu) \qquad \boldsymbol{P}_\mu = i\partial_\mu \tag{6.3}$$

for the generators that act naturally on scalar fields. These equations look familiar because if this scalar field happens to be a wave function in Quantum Mechanics then they are exactly the things that generate the transformations on the wave function. Taking $\boldsymbol{P}_\mu$ gives $-i\vec{\nabla}$ for the momentum operator where the minus sign comes from changing $\boldsymbol{P}_\mu$ to $\boldsymbol{P}^\mu$, and $\boldsymbol{P}_0 = \boldsymbol{P}^0 = i\frac{\partial}{\partial t}$ is the energy operator and corresponds to the Schrödinger equation. There are usually no boosts in Quantum Mechanics but there are rotations with $\boldsymbol{J}^i = \frac{1}{2}\varepsilon^{ijk}\boldsymbol{M}_{jk}$ where i, j, k are spatial indices, ε^{ijk} is fully antisymmetric and $\vec{\boldsymbol{J}}$ is angular momentum. Equations (6.3) are written in a completely covariant notation, and $\boldsymbol{P}_0$, $\boldsymbol{P}_i$ generate translations, $\boldsymbol{M}_{12}$, $\boldsymbol{M}_{23}$, $\boldsymbol{M}_{13}$ generate rotations, and $\boldsymbol{M}_{01}$, $\boldsymbol{M}_{02}$, $\boldsymbol{M}_{03}$ generate boosts.

With the explicit expressions (6.3) one can work out the commutator algebra of the generators. The logic is that commutator algebra is completely independent of the representation, but one can use this representation to work it out. The commutator $[\boldsymbol{P}_\mu, \boldsymbol{P}_\nu] = 0$ is easy because derivatives commute. The commutator $[\boldsymbol{M}_{\mu\nu}, \boldsymbol{P}_\rho] = -[X_\mu\partial_\nu - X_\nu\partial_\mu, \partial_\rho] = \eta_{\mu\rho}\partial_\nu - \eta_{\nu\rho}\partial_\mu$ is not zero because ∂_ρ can act on X_μ and X_ν. Using the corresponding $\boldsymbol{P}$ instead of ∂ gives the complete commutator algebra

$$\begin{aligned}
&[\boldsymbol{P}_\mu, \boldsymbol{P}_\nu] = 0 \\
&[\boldsymbol{M}_{\mu\nu}, \boldsymbol{P}_\rho] = i(\eta_{\mu\rho}\boldsymbol{P}_\nu - \eta_{\nu\rho}\boldsymbol{P}_\mu) \\
&[\boldsymbol{M}_{\mu\nu}, \boldsymbol{M}_{\rho\sigma}] = i(\eta_{\mu\rho}\boldsymbol{M}_{\nu\sigma} - \eta_{\mu\sigma}\boldsymbol{M}_{\nu\rho} - \eta_{\nu\rho}\boldsymbol{M}_{\mu\sigma} + \eta_{\nu\sigma}\boldsymbol{M}_{\mu\rho})
\end{aligned} \tag{6.4}$$

but without the derivation for the last commutator.

For a single particle state $|P\rangle$ of a free scalar field a Lorentz transformation acts as $\boldsymbol{U}(\Lambda, 0)\,|P\rangle = |\Lambda\, P\rangle$ and a translation acts as $\boldsymbol{U}(1, A)\,|P\rangle = e^{iA^\mu \boldsymbol{P}_\mu}\,|P\rangle = e^{iA\cdot P}\,|P\rangle$ with the translation parameters A^μ and the momentum operator $\boldsymbol{P}_\mu$ or the eigenvalue P, respectively. Thus, a general Poincaré transformation can be composed as a Lorentz transformation $(\Lambda, 0)$ first and a translation $(1, A)$ afterwards such that

$$\begin{aligned}
\boldsymbol{U}(\Lambda, A)\,|P\rangle &= \boldsymbol{U}(1, A)\boldsymbol{U}(\Lambda, 0)\,|P\rangle \\
&= \boldsymbol{U}(1, A)\,|\Lambda\, P\rangle \\
&= e^{iA\cdot(\Lambda P)}\,|\Lambda\, P\rangle
\end{aligned}$$

is the unitary representation of the Poincaré group acting on the state of a spinless particle. This representation has an infinite number of dimensions, and the space of states of a single particle is ∞-dimensional because it has to describe a particle at any position or a particle with any momentum. From the perspective of group theory one can say that this is happening because the translation operator is acting non-trivially on the states. If the representation contains state $|P\rangle$ it also must contain all the states $|\Lambda\, P\rangle$.

Before looking at particles with non-zero spins it has to be checked that the particles here actually have spin 0. It seems obvious because there is nothing that could be the spin but it is explicitly shown here that the particles transforms under rotation as a particle with spin 0 assuming that their mass is non-zero such that one can choose the rest frame with $P^\mu = N^\mu = (m, 0, 0, 0)$. Looking for those Lorentz transformations that leave $|N\rangle$ invariant. They satisfy $|N\rangle = \boldsymbol{U}(\Lambda)\,|N\rangle = |\Lambda N\rangle$ and build the so-called little group associated with $|N\rangle$. The Lorentz transformations with $|N\rangle = |\Lambda N\rangle$ are the ones in SO(3) which are the rotations in the three-dimensional space. In other words, the states $|N\rangle$ are left invariant by rotations in their rest frame and this is exactly the statement that they have spin 0.

6.4 Representations of the Poincaré Group

States of particles with spin are $|P, \alpha\rangle$ where α is a spin index. These states are supposed to transform as a non-trivial representation of the little group. With $m \neq 0$ and $N^\mu = (m, 0, 0, 0)$ the state $|N, \alpha\rangle$ should transform as a spin s representation of SO(3) because N^μ is invariant under rotations $R \in \mathrm{SO}(3)$. The unitary transformation $\boldsymbol{U}(R)$ corresponding to R is $\boldsymbol{U}(R)\,|N, \alpha\rangle = D_{\alpha\beta}(R)\,|N, \beta\rangle$ for some representation $D_{\alpha\beta}(R)$. The matrices $D_{\alpha\beta}(R)$ represent the action of rotation R on the wave functions and have been worked out in Quantum Mechanics.

Spin s can be $0, \frac{1}{2}, 1, \frac{3}{2}, ...$ where spin $\frac{1}{2}$ is of main interest here and in the following. Thus, these matrices for spin $s = \frac{1}{2}$ are

$$D_{\alpha\beta}(\vec{\theta}) = \left(e^{2i\,\vec{\theta}\cdot\vec{\sigma}}\right)_{\alpha\beta}$$

where $\vec{\sigma}$ are the Pauli matrices and $\vec{\theta}$ labels the rotation R by three angles. This 2×2 matrix acts in Quantum Mechanics on a two-component wave function. In Quantum Field Theorie there is a two-component particle state instead of this two-component wave function. This means that one knows all about rotations if one always stays in the rest frame but one would like to understand how this relates to general Lorentz transformations. Wigner showed that knowing the representations of the little group which is SO(3) here is enough to completely build up the entire Poincaré representation. Knowing s is enough to completely determine the rest of the representation.

To formalize the relationship between a general momentum P and the N in the rest frame for a particle with mass $m \neq 0$. One can always find a Lorentz transformation that acts on N^μ and gives P^μ as

$$P^\mu = L^\mu{}_\nu(P)\, N^\nu \qquad\qquad P = L(P)N$$

(written with and without indices) where this Lorentz transformation $L^\mu{}_\nu$ obviously depends on P. However, $L^\mu{}_\nu(P)$ is not unique but one can assume that one has picked one that works. It is known how the little group acts on $|N, \alpha\rangle$ and the goal is to build up the properties of $|P, \alpha\rangle$. The states $|P, \alpha\rangle$ have not been defined yet but it is assumed that $|P, \alpha\rangle = \boldsymbol{U}(L(P))\,|N, \alpha\rangle$. This may seem circular because one uses the object $\boldsymbol{U}(L(P))$ one is trying to determine, but the logic is that one assumes that this unitary representation exists and tries to find out what it is. In mathematical language one tries to prove uniqueness.

With $W(\Lambda, P) = L^{-1}(\Lambda P)\,\Lambda\,L(P)$ acting on N one gets

$$W(\Lambda, P)N = L^{-1}(\Lambda P)\,\Lambda L(P)N = L^{-1}(\Lambda P)\,\Lambda P = N \tag{6.5}$$

because $L^{-1}(\Lambda P)$ turns ΛP into N. Thus, $W(\Lambda, P)$ is an element of SO(3) and is called a Wigner rotation. The Wigner rotation is very useful because one can write

$$\begin{aligned}
\boldsymbol{U}(\Lambda)\,|P, \alpha\rangle &= \boldsymbol{U}(\Lambda)\boldsymbol{U}(L(P))\,|N, \alpha\rangle \\
&= \boldsymbol{U}(L(\Lambda P))\,\boldsymbol{U}(L^{-1}(\Lambda P))\,\boldsymbol{U}(\Lambda)\boldsymbol{U}(L(P))\,|N, \alpha\rangle \\
&= \boldsymbol{U}(L(\Lambda P))\,\boldsymbol{U}(W(\Lambda, P))\,|N, \alpha\rangle \\
&= \boldsymbol{U}(L(\Lambda P))\,D_{\alpha\beta}(W(\Lambda, P))\,|N, \beta\rangle \\
&= D_{\alpha\beta}(W(\Lambda, P))\,|\Lambda P, \beta\rangle
\end{aligned}$$

where both $\boldsymbol{U}(\Lambda)$ and $|P, \alpha\rangle$ are general. This shows that any unitary representation of the Poincaré group is determined by the mass m and the spin s. The mass as m^2 fixed the standard momentum and spin s determines which representation $D^{(s)}_{\alpha\beta}$ there is.

The little group representation $D^{(s)}_{\alpha\beta}$ is unitary but it has not been shown that the whole representation is unitary. A complete set of states for a free single particle is $1 = \int (dP) \sum_\alpha |P, \alpha\rangle \langle P, \alpha|$. One calculates $\boldsymbol{U}(\Lambda)\boldsymbol{U}^\dagger(\Lambda)$ and inserts this complete set of states between $\boldsymbol{U}(\Lambda)$ and $\boldsymbol{U}^\dagger(\Lambda)$ and gets

$$\begin{aligned}
\boldsymbol{U}(\Lambda) \int (dP) \sum_\alpha |P, \alpha\rangle \langle P, \alpha|\, \boldsymbol{U}^\dagger(\Lambda) &= \int (dP) \sum_\alpha \boldsymbol{U}(\Lambda)\,|P, \alpha\rangle \langle P, \alpha|\, \boldsymbol{U}^\dagger(\Lambda) \\
&= \int (dP) \sum_{\alpha,\beta,\gamma} D^{(s)}_{\alpha\beta}(W(\Lambda, P))\, D^{(s)*}{}_{\alpha\beta}(W(\Lambda, P))\,|\Lambda P, \beta\rangle \langle \Lambda P, \gamma| \\
&= \int (dP) \sum_{\beta,\gamma} [D^\dagger D]_{\gamma\beta}\,|\Lambda P, \beta\rangle \langle \Lambda P, \gamma| \\
&= \int (dP) \sum_{\beta,\gamma} \delta_{\gamma\beta}\,|\Lambda P, \beta\rangle \langle \Lambda P, \gamma| \\
&= \int (dP) \sum_\alpha |\Lambda P, \alpha\rangle \langle \Lambda P, \alpha| = 1
\end{aligned}$$

because $D^{(s)}_{\alpha\beta}$ is known to be unitary within the little group. Thus, the whole representation is unitary.

As mentioned above $L^\mu{}_\nu(P)N^\nu = P^\mu$ is not unique because one can always, using an element $R(P)$ of SO(3), define a new matrix $\tilde{L}(P) = L(P) \cdot R(P)$ where $R(P)$ as a rotation does not change N. The new

Wigner rotation becomes $W(\Lambda,P)\to R^{-1}(\Lambda P)\,W(\Lambda,P)\,R(P)$, and $\boldsymbol{U}(\Lambda)$ becomes $\boldsymbol{V}^{-1}\boldsymbol{U}(\Lambda)\,\boldsymbol{V}$ where $\boldsymbol{V}\,|P,\alpha\rangle = D_{\alpha\beta}(R(P))\,|P,\beta\rangle$ is unitary.

To summarize, the irreducible representations of the Poincaré group have been constructed out of the representations of the little group which preserves the rest frame of the massive particle. They are just the rotations with their properties known from Quantum Mechanics and are labeled by the spin. Thus, the most general irreducible representations of the Poincaré group are just labeled by the mass and the spin. This is called the method of induced representations and has been introduced by Wigner. The fact that these representations are irreducible means that one can get any state by acting with a Lorentz transformation on a given state. Thus, one can start, for example, from a state $|N,\alpha\rangle$ and get all possible states by applying Lorentz transformations.

In Quantum Mechanics states can be defined to be eigenstates of a maximal set of commuting operators. One can do the same here because all the generators of the Poincaré group are operators, and one can try to diagonalize a maximally commuting set. Since $[\boldsymbol{P}_\mu,\boldsymbol{P}_\nu]=0$, one can start with $\boldsymbol{P}_\mu$ as momentum eigenstates but this does not completely characterize the representation. The Pauli-Lubanski vector defined as $\boldsymbol{W}^\mu=\frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\boldsymbol{P}_\nu\boldsymbol{M}_{\rho\sigma}$ satisfies

$$\begin{aligned}[\boldsymbol{W}^\mu,\boldsymbol{P}_\tau] &= \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}[\boldsymbol{P}_\nu\boldsymbol{M}_{\rho\sigma},\boldsymbol{P}_\tau] = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\boldsymbol{P}_\nu[\boldsymbol{M}_{\rho\sigma},\boldsymbol{P}_\tau]\\ &= \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\boldsymbol{P}_\nu\; i(\eta_{\rho\tau}\boldsymbol{P}_\sigma-\eta_{\sigma\tau}\boldsymbol{P}_\rho)=0\end{aligned}$$

using (6.4). Note that $\varepsilon^{\mu\nu\rho\sigma}\boldsymbol{P}_\mu\boldsymbol{P}_\nu=0$ because $\boldsymbol{P}_\mu\boldsymbol{P}_\nu$ is symmetric and $\varepsilon^{\mu\nu\rho\sigma}$ is antisymmetric with respect to exchanging $\mu\leftrightarrow\nu$. Thus, one can diagonalize $\boldsymbol{P}_\mu$ and $\boldsymbol{W}^\mu$ simultaneously.

The Pauli-Lubanski vector in the rest frame $\boldsymbol{P}^\mu=(m,0,0,0)$ is

$$\boldsymbol{W}^\mu=\frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\boldsymbol{P}_\nu\boldsymbol{M}_{\rho\sigma}=\frac{m}{2}\varepsilon^{\mu0\rho\sigma}\boldsymbol{M}_{\rho\sigma}\qquad\Rightarrow\qquad \boldsymbol{W}^0=0\qquad \boldsymbol{W}^i=-m\boldsymbol{J}^i$$

where $\boldsymbol{J}^i$ is the angular momentum generator. This shows that one can simultaneously diagonalize the angular momentum in the rest frame and the 4-momentum generators.

Alternatively one can also define Casimir operators. They are operators that take the same value on all states in the representation. One Casimir operator is $\boldsymbol{P}^2=\boldsymbol{P}^\mu\boldsymbol{P}_\mu$ and that is equal to m^2 on the representation and therefore m^2 labels the representation just as the Casimir operator $\boldsymbol{J}^2$ in Quantum Mechanics labels the rotation representation. Looking at $\boldsymbol{W}^2\,|P,\alpha\rangle = m^2\,s(s+1)\,|P,\alpha\rangle$ gives another way of understanding that the representations of the Poincaré group are labeled by the two quantities mass m and spin s.

6.5 Massless Particles

Massless particles do not have a rest frame, and therefore one cannot go to their rest frame. The best one can do is to boost to a frame where a massless particle is moving in the z-direction such that $\boldsymbol{P}^\mu=E(1,0,0,1)$. The value E is arbitrary because one can always make a boost in the z-direction to change E. Thus, one can set $E=1$ in some units.

The little group certainly contains SO(2) because rotations in the xy-plane leave this momentum invariant. Its generator is $\boldsymbol{M}_{12}$ and it will be called $\boldsymbol{J}$. Actually there are more elements in the little group. They are $\boldsymbol{T}_1=\boldsymbol{M}_{10}+\boldsymbol{M}_{31}$ and $\boldsymbol{T}_2=\boldsymbol{M}_{20}+\boldsymbol{M}_{23}$. Because $\boldsymbol{M}_{10}$ is a boost in x-direction and $\boldsymbol{M}_{31}$ a rotation in the xz-plane

$$\boldsymbol{M}_{10}=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}\qquad \boldsymbol{M}_{31}=\begin{pmatrix}0&0&0&0\\0&0&0&-1\\0&0&0&0\\0&1&0&0\end{pmatrix}\qquad \boldsymbol{T}_1=\begin{pmatrix}0&1&0&0\\1&0&0&-1\\0&0&0&0\\0&1&0&0\end{pmatrix}$$

$$\begin{pmatrix}1\\0\\0\\1\end{pmatrix}\to \qquad\qquad \begin{pmatrix}1\\0\\0\\1\end{pmatrix}+\begin{pmatrix}0&\varepsilon&0&0\\\varepsilon&0&0&-\varepsilon\\0&0&0&0\\0&\varepsilon&0&0\end{pmatrix}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}+O(\varepsilon^2)$$

shows that $\boldsymbol{T}_1$ leaves $\boldsymbol{P}^\mu$ invariant. There are now the three generators $\boldsymbol{J}$, $\boldsymbol{T}_1$, $\boldsymbol{T}_2$ with the commutation relations

$$[\boldsymbol{J},\boldsymbol{J}]=0 \qquad [\boldsymbol{T}_1,\boldsymbol{T}_2]=0 \qquad [\boldsymbol{T}_1,\boldsymbol{J}]=i\,\boldsymbol{T}_2 \qquad [\boldsymbol{T}_2,\boldsymbol{J}]=-i\,\boldsymbol{T}_1$$

and this is exactly the 2-dimensional Euclidean Poincaré group where $\boldsymbol{T}_{1,2}$ are the translations and $\boldsymbol{J}$ is the rotation.

This gives some intuition of what the representations of this little group are. Wigner's method starts by finding all the generators of the little group and then build up the rest of the group. Because $\boldsymbol{T}_{1,2}$ are translations, if they act non-trivially on the states such that $\boldsymbol{T}_{1,2}\,|N,\alpha\rangle \neq 0$ then there is an infinite representation. Since the $\boldsymbol{T}_{1,2}$ indeed act non-trivially this means that for every state there is also the translated state and this means further that instead of a discrete spin index there is a continuous spin index leading to an infinite number of degrees of freedom in contrast to what one sees in nature. It has usually been assumed that these continuous spin representations here are not physical but recent work by Schuster and Toro suggests that they may in fact be physical and play a role in nature somewhere.

However, particles such as electrons, photons and neutrinos seem to have only a finite number of spin degrees of freedom, and these infinite spin representations will be ignored here. The only way to ignore them is to assume that these $\boldsymbol{T}_{1,2}$ act trivially on the states such that $\boldsymbol{T}_{1,2}\,|N,\alpha\rangle = 0$. That also means that these operators can be ignored too, and the little group collapses back to the well-known SO(2). If θ is the rotation angle then $|N,\lambda\rangle \to e^{i\theta\lambda}\,|N,\lambda\rangle$ where λ is a number telling which representation it is. The allowed values are $\lambda \in \{0, \frac{1}{2}, 1, \frac{3}{2}, ...\}$, and they are the eigenvalues of $\boldsymbol{J}$ which is the z-component $\boldsymbol{J}_z$ of the angular momentum. With $N^\mu = (1,0,0,1)$ this is the eigenvalue of the angular momentum along the direction of motion. It is called helicity and is a Lorentz-invariant property.

Note that λ labels the representation and therefore all the states in a representation have the same λ. In particular, $\pm\lambda$ are different irreducible representations, and there is therefore no rotation or Lorentz transformation to go between $+\lambda$ and $-\lambda$. That is different from the case of massive particles with $\boldsymbol{J}_z = s, s-1, ..., -s$ where it is possible to go from s to $-s$ by rotations. Thus, $+\lambda$ and $-\lambda$ are different particles and one cannot go to the rest frame to turn the spin around. However, it is true that CPT relates $+\lambda$ and $-\lambda$, and therefore any theory that has a helicity $+\lambda$ must also have a helicity $-\lambda$. Since CPT follows from locality of Quantum Field Theory, any local Quantum Field Theory must have both states in it even so one cannot go from one to the other by any Lorentz transformation.

To summarize, particles are associated with irreducible unitary representations of the Poincaré group. These are free particles which are relevant as the initial and final states of scattering amplitudes. These representations are labeled by their mass and their spin. For massless particles there are some interesting subtleties, but as long as one requires that the number of spin states of a massless particle is finite they are labeled by the helicity as the spin along the direction of motion.