Notes on relativistic quantum electrodynamics

Matteo Giordano

May 22, 2018

1 Introduction

The appropriate tool to study quantum phenomena at relativistic energy is Quantum Field Theory (QFT). QFT is a synthesis of quantum mechanics and special relativity, which allows to study the quantum properties of Nature in full compliance with Lorentz invariance. Our ultimate purpose in these notes is to develop a quantum relativistic theory of electrodynamics, usually called simply Quantum ElectroDynamics (QED).

While the Q in QFT stands clearly for quantum aspects, it is the F of fields that stands for the relativistic ones. Even in a non-quantum setting, fields are required to give a proper relativistic description of Nature. In a classical non relativistic setting, two particles of electric charge q_1 and q_2 at distance r interact via the Coulomb potential,

$$V_{\text{Coulomb}}(r) = \frac{q_1 q_2}{4\pi r}\,. \tag{1.1}$$

The choice of numerical factors corresponds to the rationalised Gaussian system of units, which we are going to use in these notes. Interaction through a potential is however problematic in a relativistic setting because, being an istantaneous interaction at a distance, it violates locality and therefore causality (the order in time of spacelike-separated events depends on the reference frame). For this reason the electromagnetic field is introduced, with which particles interact locally: the particle-particle interaction is effectively mediated by the field, it propagates from one particle to the other in some finite time due to the finiteness of the speed of light, and causality is saved.

In certain situations, the electromagnetic radiation shows the behaviour typical of particles, instead of that of waves: prime examples are the photoelectric effect and Compton scattering. The particles corresponding to the electromagnetic fields are called *photons*. As we will see, a similar particle/field duality applies to matter particles as well. The particle/wave, or particle/field duality was one of the most relevant and most striking new aspects of quantum mechanics. The field in question there, though, is the particle wave function, or rather its absolute value square, which represents a probability field. The fields in QFT are of a completely different type: they do not represent the probability amplitude of a particle being somewhere in spacetime, and in fact they are not even real-valued quantities: they are instead linear operators acting on the Hilbert space of the system. What they do in practice is to create and destroy particles at some point in spacetime. An operator with these properties is certainly needed in a relativistic setting, where particles can be created and destroyed as a consequence of the mass-energy equivalence.

The fact that the number of particles is not conserved in a relativistic setting has two important consequences. First of all, it shows how any attempt at formulating relativistic quantum mechanics in terms of single-particle wave functions is bound to fail: creation and destruction of particles are not admitted in this kind of formalism. Relativistic wave equations will prove nonetheless very useful for our tasks. Moreover, from the historical point of view QFT appeared after quantising the wave functions appearing in relativistic quantum mechanics, a procedure called "second quantisation". Since we will follow a logical rather than a historical thread in these lectures, we will not adopt this interpretation; we will nevertheless make use of the same formalism when developing the canonical approach to quantisation.

The other important consequence is the impossibility to define consistently a position operator. Recall that when moving from classical to quantum mechanics we had to change the nature of the position of particles from ordinary numbers to Hermitian operators. When moving from quantum mechanics to QFT, space coordinates become just labels for the degrees of freedom of our system, namely the quantum field (more on this later). The difficulties in defining a position operator in QFT show up when we try to measure the position of a relativistic particle with a precision better than the particle's Compton length. The uncertainty principle tells us that the uncertainty on the momentum will be $\Delta p \sim \hbar/\Delta x$, which for a relativistic particle for which $E \simeq |\vec{p}|c$ means that $\Delta E \sim \hbar c/\Delta x$. For $\Delta x \sim \hbar/(mc)$ we then have $\Delta E \sim mc^2$, i.e, the uncertainty is large enough for a new particle to be created: at this point, whose position would we be talking about?

Using fields that create and destroy particles, we can describe interactions in terms of emission and absorption of intermediate particles. To make a concrete example: electromagnetic scattering of two electrons can be seen, to lowest order, as one of the electrons emitting a photon (which is created by the electromagnetic field operator), which subsequently travels to meet the other electron and gets absorbed (i.e., the photon gets destroyed by the corresponding field). One can then imagine multiple exchanges, and build up the electromagnetic interaction in terms of them. This picture is what emerges in the framework of perturbative quantisation, on which we will mostly focus.

The main advantages in using fields instead of particles in building up interacting theories is that the fields are local objects, and that they can be contructed so that they transform in a simple manner under Lorentz transformations. In this way we can rest assured that the physics we are building will look the same in any inertial reference frame, as it should. Despite all the emphasis on fields, our starting point will be the particles, more precisely how one characterises particles, and how their states can be constructed. The key point here will be developing the representation theory of the Poincaré group, i.e., the group of translations and Lorentz transformations.

An extremely convenient way to formulate our field theories is in terms of the Lagrangian formalism and of the action principle. This approach has the advantage that Lorentz invariance, and in general invariance under any symmetry, can be imposed at the outset. Activating the machinery of canonical quantisation we will then obtain a theory with the desired symmetry and locality properties: whether this has anything to do with Nature is of course a completely different question, one that can be answered only by means of experiments.

The canonical formalism for fields requires an extension of Lagrangian mechanics. In fact, the main difference between the mechanics of particles and that of fields is that the number of degrees of freedom of the system is finite in the first case, and infinite in the second case. This is true already at the classical level: the degrees of freedom of the electromagnetic field, for

example, are the values of the electric and magnetic fields at each point in space, which come in an uncountably infinite amount. Similarly, as noted above, in QFT the spatial coordinates will label the degrees of freedom of our quantum field operators.

1.1 A brief history of Quantum Field Theory

The history of QFT,[1] or rather its pre-history, begins with the relativistic wave equation derived by Schrödinger in 1926 (even before that, de Broglie's particle/wave duality had been formulated in terms of a relativistic wave). Discouraged by the fact that it gave the wrong fine structure for the hydrogen atom, he dismissed it and kept only its non-relativistic version, which worked well in the non-relativistic limit. By the time he finally came to publish it, it had been rediscovered independently by Klein and Gordon, and so it became known as the Klein-Gordon equation:

$$(\Box + m^2)\phi = 0\,. \tag{1.2}$$

As an equation for wave functions, this has the problem that the only conserved current that one can build from ϕ, in terms of which one would like to express the conservation of probability, has a non-positive-definite probability density, that makes no physical sense. In 1928 Dirac "fixed" this problem by deriving the equation that bears his name for a relativistic electron,

$$(i\slashed{\partial} - m)\psi = 0\,. \tag{1.3}$$

In this case it is possible to build a positive-definite probability density. Nevertheless, this equation admits negative-energy solutions (as the Klein-Gordon equation does), and one has to explain why the electrons do not fall into states of more and more negative energy, making everything unstable. This was done by Dirac in 1930, postulating the existence of a "sea" (the "Dirac sea") of occupied negative-energy states, and invoking the exclusion principle, that forbids two electrons to occupy the same state. When one of the negative-energy states of the negatively-charged electron becomes unoccupied due to some interaction, the corresponding "hole" behaves in practice as a positive-energy state for a positively-charged particle. This led[2] to the prediction of a positively-charged electron, or *positron*, which was experimentally detected by Anderson in 1932.

Dirac's theory correctly accounted for the spin of the electron; coupling the electron to the electromagnetic field led to correctly describe the fine structure of hydrogen; correctly predicted the electron magnetic moment; and, as mentioned above, predicted the existence of the positron. Nevertheless, this did not solve all the problems: obtaining positive probabilities for a spin-$\frac{1}{2}$ particle will not make them positive in the spin-0 case of the Klein-Gordon equation. Moreover, taking care of negative energy states via a Dirac sea cannot possibly work for bosons, which do not obey any exclusion principle. Finally, the correct prediction of the electron magnetic moment is due to neglecting a term, the Pauli term, in the wave equation: such a term is perfectly legitimate from the point of view of relativistic quantum mechanics, and comes with an arbitrary coefficient that makes the magnetic moment of the electron whatever one pleases. To solve these problems, relativistic quantum mechanics had to give way to QFT.

The history proper of QFT begins with Born, Heisenberg and Jordan who quantised the electromagnetic field in 1926, imposing canonical commutation relations on the coeffcients of

---

[1]This brief history of QFT is based on the introduction to Weinberg's *The Quantum Theory of Fields*, vol. 1, to which we refer the reader for a more exhaustive description.

[2]More or less: initially Dirac proposed that the positively-charged states corresponded to the proton.

the field's normal modes. In this way the field is seen to behave as a collection of independent harmonic oscillators. In 1927 this was used by Dirac, coupled to the usual non-relativistic description of atomic spectra, to explain spontaneous emission satisfactorily. Work in refining the quantisation of the electromagnetic field was done by Weisskopf and Wigner, by Fermi, and by Jordan and Pauli. The quantisation of fields was applied to other systems as well, regarding it as a way to quantise the wave functions of relativistic quantum mechanics: for this reasons, the procedure acquired the name of "second quantisation". For the Dirac field this was accomplished by Jordan and Wigner, who had to use anticommutation relations instead of the usual commutation ones to obtain a satisfactory result. The general theory was first laid out by Heisenberg and Pauli in 1929, using the canonical formalism on the fields themselves instead of on their normal modes. This is the canonical formalism that we will develop in due course. The problem of negative-energy electron states was solved by Fock and by Furry and Oppenheimer (1933-1934) essentially by calling them the positive-energy states of a positron, treated as a legitimate different particle. A similar point of view led Pauli and Weisskopf to solve the same problem for spin-0 particles by introducing corresponding antiparticles. This led to clarify that quantum fields are not probability amplitudes but rather operators which create and destroy particles and antiparticles.

QFT was then applied[3] in the period 1929-1936 to the calculations of cross sections for various processes (elastic $e^-\gamma$ scattering, Klein-Nishina; $e^-e^+ \to 2\gamma$, Dirac; elastic e^-e^- scattering, Møller; elastic e^+e^- scattering, Bhabha) to lowest order of approximation, obtaining good agreement with experiment. In 1934 it was used by Fermi to correctly describe the energy distribution of electrons in β-decays; in 1935 Yukawa used it to predict the existence of the pion. Despite these successes, the 1930s were also the period in which the problem of infinities showed up: calculations beyond lowest order were plagued by all kinds of nonsensical divergent results. This was noted by, among others, Heisenberg, Pauli, Oppenheimer, and Waller. The problem of infinities was present already classically (e.g., the electron self-energy). The fundamental problem seems to be the inability of field theories to describe properly what happens at very short scales, or equivalently at very high energies. Although the problem of infinities can be cured, this inability of field theory remains somehow an issue. The cure for infinities is *renormalisation*, which "sweeps" the infinities "under the carpet" of a redefinition of the parameters of the theory. This idea was developed systematically by Tomonaga, Schwinger, and Feynman (who shared the 1965 Nobel prize for physics), and by Dyson (who did not). Notably, renormalisation cannot be done for all theories; in particular it prevents including the Pauli term in the interaction between electrons and electromagnetic fields. The idea of renormalisation was disliked by many (most notably Dirac), but it allowed in practice to make precise calculations in spectacular agreement with experiments: the Lamb shift (the energy difference between the $2s_{\frac{1}{2}}$ and $2p_{\frac{1}{2}}$ levels of hydrogen) and the anomalous magnetic moment of the electron (the deviation from the value 2 predicted by Dirac's theory) are accurate to 9 significant figures.

The quantum theory of electrodynamics (QED) was finally complete: an extremely successful theory describing subatomic phenomena with astonishing precision, elegant in its symmetrical nature. Although it was not the final chapter in the history of theoretical physics, it nevertheless provided the model on which were built the other quantum field theories currently believed to describe Nature. The purpose of this course is to build the tools of QFT up to the point where the full formulation of QED is possible, and lowest-order calculations can be performed. The

---

[3]Not exclusively: the sea/hole point of view still lingered for a few years.

issue of renormalisation is not treated here.

2   Preliminaries

In this section we review a few basic concepts and tools that will be needed during the course.

2.1   Natural units

For most of the time we will be using the system of *natural units*, where the fundamental constants c (speed of light) and $\hbar$ (Planck's constant) are set to 1, $c = \hbar = 1$. In this system of units, length and time have the same dimensions since velocities are now dimensionless. Moreover, energies, momenta and masses have all the same dimensions, as it follows from the mass-energy relation $mc^2 = E^2 + (pc)^2$. Finally, since the action is now dimensionless, so is angular momentum, which implies that length has the dimension of inverse momentum. Summarising, everything can be expressed in terms of mass dimensions as follows:

$$[E] = [p] = [m] = M, \qquad [l] = [t] = M^{-1}. \tag{2.1}$$

From the Coulomb potential Eq. (1.1) it then follows that in this system of units the electric charge is dimensionless. This also follows from the fact that the *fine structure constant* $\alpha = \frac{e^2}{4\pi\hbar c} \simeq 1/137$ is a dimensionless number, which reads simply $\alpha = \frac{e^2}{4\pi}$ in natural units. Similarly, in this system of units the Compton length associated to a particle, $\lambda_C = \frac{\hbar}{mc}$, reads simply $\lambda_C = \frac{1}{m}$.

The typical energy unit used in high energy physics is the electronvolt, (eV), 1 eV = $1.6 \cdot 10^{-19}\, J$, i.e., the energy acquired by an electron after travelling through an electric potential difference of one volt, or rather its multiples, the megaelectronvol (MeV = 10^6 eV), the gigaelectronvolt (GeV = 10^9 eV), and the teraelectronvolt (TeV = 10^{12} eV). The typical length scale in ordinary units is the fermi, 1 fm = 10^{-15} m, which is the approximate size of a proton. The conversion between a length expressed in MeV^{-1} in natural units and its value in fermi in ordinary units is easily obtained via the relation $\hbar c \simeq 197$ MeV $\cdot$ fm. In natural units the left-hand side is one, so 1 fm $\approx (1/5)\, \mathrm{GeV}^{-1}$.

2.2   Special relativity in a nutshell

Special relativity states the equivalence of all inertial frames for the description of physical phenomena. Inertial frames are moving at constant speed with respect to each other. The descriptions of physical phenomena in different frames are obtained from each other by means of *Lorentz transformations*. These are most easily characterised in terms of the geometry of Minkowski space.

Minkowski space is the vector space $\mathbb{R}^4$ endowed with the Minkowski (pseudo)metric: given vectors X^μ, Y^μ with $\mu = 0, 1, 2, 3$, their distance, or rather their *interval*, is defined as (here comes some notation) $(X - Y)^2 = (X - Y) \cdot (X - Y) = X^2 + Y^2 - 2X \cdot Y$, where the scalar product is defined as $X \cdot Y = X^\mu Y^\nu \eta_{\mu\nu} = X^\mu Y_\mu$ (summation over repeated indices is understood). Here $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$ is the Minkowski tensor. We will often use the notation $X = (X^0, \vec{X})$ for the spatial part of four-vectors, which we will refer to as three-vectors, and $\vec{v} \cdot \vec{w} = \vec{v}^i \vec{w}^i$ for the usual Euclidean scalar product of three-vectors, where it is understood that latin indices run over $i = 1, 2, 3$ only. Lowering or raising of indices is realised through multiplication by

the Minkowski tensor, i.e., $X_\mu = \eta_{\mu\nu}X^\nu$, and its inverse $\eta^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1) = \eta_{\mu\nu}$, $X^\mu = \eta^{\mu\nu}X_\nu$. Notice that $\eta^{\mu\alpha}\eta_{\alpha\nu} = \delta^\mu{}_\nu$. Vectors are characterised as *timelike*, *spacelike*, or *null*, according to the sign of their square: $X^2 > 0$ for timelike vectors, $X^2 < 0$ for spacelike vectors, and $X^2 = 0$ for null vectors.

Lorentz transformations are linear transformations of Minkowski space that leave intervals invariant. This is equivalent to ask that all scalar products are left invariant, since $4X \cdot Y = (X+Y)^2 - (X-Y)^2$. Given a transformation $X'^\mu = \Lambda^\mu{}_\nu X^\nu$ we then require

$$X'^\alpha Y'^\beta \eta_{\alpha\beta} = X^\mu Y^\nu \eta_{\mu\nu}\,,$$
$$\Lambda^\alpha{}_\mu X^\mu \Lambda^\beta{}_\nu Y^\nu \eta_{\alpha\beta} = X^\mu Y^\nu \eta_{\mu\nu}\,,$$

(2.2)

and since this has to hold for any X^μ, Y^μ,

$$\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \eta_{\alpha\beta} = \eta_{\mu\nu}\,.$$

(2.3)

This identifies the Lorentz transformations as the elements of the group O(3,1). In fact, adopting the matrix notation $(\Lambda)_{\mu\nu} = \Lambda^\mu{}_\nu$ and $(\eta)_{\mu\nu} = \eta_{\mu\nu}$, Eq. (2.3) becomes (no brackets when writing matrix operations)

$$\Lambda^T \eta \Lambda = \eta\,.$$

(2.4)

From this relation one immediately obtains $\det\Lambda = \pm 1$, so that Lorentz transformations are invertible, and $\Lambda^{-1} = \eta^{-1}\Lambda^T\eta$, which reads also

$$\eta^{\alpha\mu}\Lambda^\nu{}_\mu \eta_{\nu\rho}\Lambda^\rho{}_\beta = \delta^\alpha{}_\beta\,,$$
$$\Lambda_\rho{}^\alpha \Lambda^\rho{}_\beta = \delta^\alpha{}_\beta\,,$$
$$(\Lambda^{-1})_{\alpha\beta} = \Lambda^{-1\alpha}{}_\beta = \Lambda_\beta{}^\alpha\,.$$

(2.5)

Lorentz transformations clearly do not change the nature of a vector, since they do not change its (pseudo)length. If we take the 00 component of Eq. (2.3) we obtain

$$1 = \Lambda^0{}_0 \Lambda^0{}_0 - \Lambda^i{}_0 \Lambda^i{}_0\,,$$
$$(\Lambda^0{}_0)^2 = 1 + \Lambda^i{}_0 \Lambda^i{}_0\,.$$

(2.6)

Similarly, the same equation for the inverse transformation Λ^{-1} implies

$$1 = (\Lambda^{-1})^0{}_0 (\Lambda^{-1})^0{}_0 - (\Lambda^{-1})^i{}_0 (\Lambda^{-1})^i{}_0 = \Lambda_0{}^0 \Lambda_0{}^0 - \Lambda_0{}^i \Lambda_0{}^i = \Lambda^0{}_0 \Lambda^0{}_0 - \Lambda^0{}_i \Lambda^0{}_i\,,$$
$$(\Lambda^0{}_0)^2 = 1 + \Lambda^0{}_i \Lambda^0{}_i\,,$$

(2.7)

where we used $\Lambda_0{}^0 = \Lambda^0{}_0$ and $\Lambda_0{}^i = -\Lambda^0{}_i$. Eq. (2.6) implies that either $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq -1$. We can then classify Lorentz transformations in four types according to the sign of their determinant and of their 00 component. Transformations with $\det\Lambda = 1$ are called *proper*, while those with $\det\Lambda = -1$ are *improper*; transformations with positive $\Lambda^0{}_0$ are called *orthocronous*. The reason for this nomenclature is that proper Lorentz transformations keep the orientation of space unchanged, i.e., if we take a set of three right-handed orthogonal spatial vectors they will be transformed in another such set, while orthocronous transformations keep unchanged the direction of time for timelike events, the only ones for which this has an intrinsic meaning independent of the reference frame. To see this, consider a timelike vector X^μ, $X^2 > 0$, and act

	$\det \Lambda = +1$	$\det \Lambda = -1$
$\Lambda^0{}_0 \geq +1$	proper orthocronous	improper orthocronous
$\Lambda^0{}_0 \leq -1$	proper non orthocronous	improper non orthocronous

Table 1: Components of the Lorentz group.

with an orthocronous transformation on it. Denoting with $\vec{\Lambda}$ the three-vector $\vec{\Lambda}^i = \Lambda^{0i}$, from Eq. (2.7) and making use of Schwartz's inequality $|\vec{v} \cdot \vec{w}| \leq |\vec{v}||\vec{w}|$ we find for $X^0 > 0$

$$
\begin{aligned}
X'^0 = \Lambda^0{}_\nu X^\nu = \Lambda^0{}_0 X^0 + \Lambda^0{}_i X^i = \Lambda^0{}_0 X^0 - \vec{\Lambda} \cdot \vec{X} \\
\geq \Lambda^0{}_0 X^0 - |\vec{\Lambda}||\vec{X}| \geq \left(\Lambda^0{}_0 - |\vec{\Lambda}| \right)|\vec{X}| \geq |\vec{X}| \geq 0 \,.
\end{aligned}
\tag{2.8}
$$

One then has $X'^0 > 0$ unless $\vec{X}$, but in this case $X'^0 = \Lambda^0{}_0 X^0 > 0$. Similarly, for $X^0 < 0$

$$
X'^0 = \Lambda^0{}_\nu X^\nu = \Lambda^0{}_0 X^0 + \leq \Lambda^0{}_0 X^0 + |\vec{\Lambda}||\vec{X}| \leq - \left(\Lambda^0{}_0 - |\vec{\Lambda}| \right)|\vec{X}| \leq -|\vec{X}| \leq 0 \,,
\tag{2.9}
$$

so $X'^0 < 0$ unless $\vec{X}$, in which case $X'^0 = \Lambda^0{}_0 X^0 < 0$. Given a proper orthocronous transformation, $\det \Lambda = 1$, $\Lambda^0{}_0 > 0$, we can get an improper orthocronous one by multiplying with a *parity* transformation, i.e., $P = \mathrm{diag}(1, -1 - 1 - 1)$, and an improper non-orthocronous one by multiplication with the *time reversal* transformation $T = \mathrm{diag}(-1, 1, 1, 1)$. The last component of the group is obtained from the proper orthocronous one via multiplication with $PT = \mathrm{diag}(-1, -1, -1, -1)$. The subset of proper transformations is denoted with SO(3,1), while the subset of proper orthocronous transformations is denoted as SO$^\uparrow$(3, 1).

2.3 Quantum mechanics on the tip of a pin

We will not review here quantum mechanics but for a very few points, mostly to fix the notation. To the states of physical systems are associated vectors (or more precisely rays) in a Hilbert space $\mathcal{H}$. The transition amplitude from a state vector ψ to a state vector ϕ is given by their scalar product, which is denoted with $(\phi, \psi) = (\psi, \phi)^*$, or in Dirac notation $\langle \phi | \psi \rangle$. The corresponding transition probability is given by $|(\phi, \psi)|^2$. Observables are associated to self-adjoint operators $A = A^\dagger$ in $\mathcal{H}$. The possible outcomes of measurements of A are its (real) eigenvalues a_n, $A\psi_n = a_n \psi_n$; if the value a_n is obtained, the state vector of the system is projected on the corresponding eigenvector ψ_n, and the probability for this to happen is $|(\psi_n, \psi)|^2$. The expectation value of an observable A on a state ψ is given by $(\psi, A\psi)$, or $\langle \psi | A | \psi \rangle$ in Dirac notation. In particular, transition probabilities are the expectation values of projectors, $|(\phi, \psi)|^2 = (\psi, \phi \otimes \phi^\dagger \psi) = \langle \psi | \phi \rangle \langle \phi | \psi \rangle$. Transformations U that leave invariant the norm $\|\psi\| = (\psi, \psi)^{\frac{1}{2}}$ of vectors are called *isometric*. Such transformations actually leave all the scalar products invariant,[4] and are therefore characterised by the relation $U^\dagger U = \mathbf{1}$. Isometric transformation that have the whole Hilbert space as image are called *unitary*, and they satisfy also $UU^\dagger = \mathbf{1}$. Conversely, $U^\dagger U = UU^\dagger = \mathbf{1}$ implies that U is unitary. Unitary transformations play a very important role in quantum mechanics in connection with symmetries.

---

[4]This is because $\|\phi + \psi\|^2 - \|\phi - \psi\|^2 + i\|\phi - i\psi\|^2 - i\|\phi + i\psi\|^2 = 4(\phi, \psi)$.

2.4 Classical Lagrangian mechanics for particles and fields

In Lagrangian mechanics the equations of motion for particles are obtained by means of a variational principle starting from the so-called *action functional*. Given (generalised) coordinates q_i for the particles and their time derivatives $\dot{q}_i$, the action functional, $S[q]$, is given by the integral of the Lagrangian function or simply *Lagrangian*, $L(q, \dot{q}, t)$, with respect to time:[5]

$$S[q] = \int_{t_0}^{t_1} dt\, L(q, \dot{q}, t)\,. \tag{2.10}$$

In the argument of L we have dropped the discrete index i, and q and $\dot{q}$ denote collectively all such variables. We also dropped the time dependence of q and $\dot{q}$. The trajectories of particles are those that extremise the action functional, i.e., those for which the action does not change under any small variation $q_i \to q_i + \delta q_i$. This is the action principle in brief, but some qualifications must be added. Given the choice of a time interval $[t_0, t_1]$, the variations we consider are those that leave the initial and the final points fixed, i.e., $\delta q_i(t_0) = \delta q_i(t_1) = 0$. The variation of the action is simply $\delta S = S[q + \delta q] - S[q]$, and our purpose is to find q_i such that $\delta S = 0$. For small variations, retaining only first order corrections, we have (notice that $\delta\dot{q} = (\dot{\delta q})$)

$$
\begin{aligned}
\delta S &= \int_{t_0}^{t_1} dt\, \left(L(q + \delta q, \dot{q} + \delta\dot{q}_i, t) - L(q, \dot{q}, t)\right) \\
&= \int_{t_0}^{t_1} dt\, \sum_i \left(\delta q_i \frac{\partial}{\partial q_i} L(q, \dot{q}, t) + (\dot{\delta q}) \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t)\right) \\
&= \int_{t_0}^{t_1} dt\, \left\{\sum_i \delta q_i \left(\frac{\partial}{\partial q_i} L(q, \dot{q}, t) - \frac{d}{dt}\frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t)\right) + \frac{d}{dt}\left(\sum_i \delta q_i \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t)\right)\right\} \\
&= \left(\sum_i \delta q_i \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t)\right)\bigg|_{t_0}^{t_1} + \int_{t_0}^{t_1} dt\, \sum_i \delta q_i \left(\frac{\partial}{\partial q_i} L(q, \dot{q}, t) - \frac{d}{dt}\frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t)\right)\,,
\end{aligned}
\tag{2.11}
$$

where we have used integration by parts. Since δq_i vanishes at $t_{0,1}$, the first term is zero. On the other hand, if we want δS to vanish for arbitrary small variations, then the quantities multiplying δq_i in the integrand of the second term have to vanish separately. We thus obtain the *Euler-Lagrange equations*:

$$\frac{\partial}{\partial q_i} L(q, \dot{q}, t) - \frac{d}{dt}\frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t) = 0\,. \tag{2.12}$$

From Lagrangian mechanics we can go over to Hamiltonian mechanics by means of a Legendre transform. Define the momentum p_i conjugated to the coordinate q_i as

$$p_i = \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t)\,, \tag{2.13}$$

an the *Hamiltonian* as

$$H(q, p, t) = \sum_i p_i \dot{q}_i - L(q, \dot{q}, t)\,, \tag{2.14}$$

---

[5]While a function maps numbers into numbers, a functional maps a function into a number: in the case at hand, the action functional maps any trajectory $q(t) : [t_0, t_1] \to \mathbb{R}$, and so it depends on the whole trajectory.

where the q_i and p_i are treated as independent variables, and it is understood that $\dot{q}_i$ have to be expressed as functions of the q_j and p_j. We find that

$$\frac{\partial}{\partial p_i} H(q,p,t) = \dot{q}_i + \sum_j p_j \frac{\partial}{\partial p_i}\dot{q}_j - \frac{\partial \dot{q}_j}{\partial p_i}\frac{\partial}{\partial \dot{q}_j}L(q,\dot{q},t) = \dot{q}_i + \sum_j p_j \frac{\partial}{\partial p_i}\dot{q}_j - \frac{\partial \dot{q}_j}{\partial p_i}p_j = \dot{q}_i, \quad (2.15)$$

and moreover,

$$\begin{aligned}
\frac{\partial}{\partial q_i} H(q,p,t) &= -\frac{\partial}{\partial q_i}L(q,\dot{q},t) + \sum_j p_j \frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial \dot{q}_j}{\partial q_i}\frac{\partial}{\partial \dot{q}_j}L(q,\dot{q},t) \\
&= -\frac{\partial}{\partial q_i}L(q,\dot{q},t) = -\frac{d}{dt}\frac{\partial}{\partial \dot{q}_i}L(q,\dot{q},t) = -\dot{p}_i,
\end{aligned} \quad (2.16)$$

which holds for solutions of the Lagrangian equations of motion, which we used in the last passage. Finally, again using the Lagrangian equations of motion $\dot{p}_i = \frac{\partial}{\partial q_i}L(q,\dot{q},t)$ we find

$$\begin{aligned}
\frac{d}{dt} H(q,p,t) &= -\frac{\partial}{\partial t}L(q,\dot{q},t) + \sum_i \dot{p}_i\dot{q}_i + p_i\ddot{q}_i - \dot{q}_i\frac{\partial}{\partial q_i}L(q,\dot{q},t) - \ddot{q}_i\frac{\partial}{\partial \dot{q}_i}L(q,\dot{q},t) \\
&= -\frac{\partial}{\partial t}L(q,\dot{q},t) + \sum_i \dot{p}_i\dot{q}_i + p_i\ddot{q}_i - \dot{q}_i\dot{p}_i - \ddot{q}_i p_i = -\frac{\partial}{\partial t}L(q,\dot{q},t) \\
&= \frac{\partial}{\partial t}H(q,\dot{q},t).
\end{aligned} \quad (2.17)$$

Summarising, we have the *Hamilton equations*,

$$\begin{aligned}
\dot{p}_i &= -\frac{\partial H}{\partial q_i}, \\
\dot{q}_i &= \frac{\partial H}{\partial p_i}, \\
\frac{dH}{dt} &= \frac{\partial H}{\partial t}.
\end{aligned} \quad (2.18)$$

These can be recast in yet another way by making use of the *Poisson brackets*,

$$\{f,g\}_{\mathrm{PB}} \equiv \sum_i \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}. \quad (2.19)$$

Notice first of all that

$$\{q_i,p_j\}_{\mathrm{PB}} = \delta_{ij}, \qquad \{q_i,q_j\}_{\mathrm{PB}} = \{p_i,p_j\}_{\mathrm{PB}} = 0. \quad (2.20)$$

Consider next a generic observable $\mathcal{O} = \mathcal{O}(q,p,t)$. Its time evolution is governed by the equation

$$\frac{d\mathcal{O}}{dt} = \frac{\partial \mathcal{O}}{\partial t} + \sum_i \dot{q}_i\frac{\partial \mathcal{O}}{\partial q_i} + \dot{p}_i\frac{\partial \mathcal{O}}{\partial p_i} = \frac{\partial \mathcal{O}}{\partial t} + \sum_i \frac{\partial H}{\partial p_i}\frac{\partial \mathcal{O}}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial \mathcal{O}}{\partial p_i} = \frac{\partial \mathcal{O}}{\partial t} + \{\mathcal{O},H\}_{\mathrm{PB}}. \quad (2.21)$$

This applies in particular to q_i and p_i, for which

$$\begin{aligned}
\dot{p}_i &= \{p_i,H\}_{\mathrm{PB}}, \\
\dot{q}_i &= \{q_i,H\}_{\mathrm{PB}}.
\end{aligned} \quad (2.22)$$

The formalism discussed so far applies to particles, or more generally to systems with a finite number of degrees of freedom, both in a nonrelativistic and a relativistic setting. Fields, on the other hand, have infinitely many degrees of freedom, since they are definited at every point of spacetime. The extension of the formalism is however quite straightforward. The starting point is again an action functional which depends on the configuration of the fields $\phi_i(x)$. Such functional is defined starting from a Lagrangian density $\mathscr{L}(\phi_i(x), \partial_\mu \phi_i(x), x)$ which depends on the fields and their derivatives and possibly on the spacetime point x,

$$S[\phi] = \int_D d^4x \, \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) \,. \tag{2.23}$$

Here D denotes some domain in spacetime, and

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x^0}, \vec{\nabla} \right) \,. \tag{2.24}$$

The equations of motions are obtained again via an action principle: we require that the variation of S vanishes for all small variations $\delta\phi_i$ of the fields, subject to the condition $\delta\phi_i = 0$ on the boundary ∂D of D. Since $\delta(\partial_\mu \phi) = \partial_\mu(\delta\phi)$, we have

$$
\begin{aligned}
0 = \delta S &= \int_D d^4x \, [\mathscr{L}(\phi + \delta\phi, \partial_\mu \phi + \partial_\mu \delta\phi, x) - \mathscr{L}(\phi, \partial_\mu \phi, x)] \\
&= \int_D d^4x \, \left[\sum_i \delta\phi_i \frac{\partial \mathscr{L}}{\partial \phi_i} + (\partial_\mu \delta\phi_i) \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \right] \\
&= \int_D d^4x \, \left\{ \sum_i \delta\phi_i \left[\frac{\partial \mathscr{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \right] + \partial_\mu \left[\delta\phi_i \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \right] \right\} \\
&= \int_{\partial D} d\Sigma_\mu \, \delta\phi_i \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} + \int_D d^4x \, \sum_i \delta\phi_i \left[\frac{\partial \mathscr{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} \right] \,.
\end{aligned} \tag{2.25}
$$

Since $\delta\phi_i = 0$ on ∂D the first terms vanishes, and since $\delta\phi_i$ is arbitrary inside D, each of the terms in the integrand of the second term has to vanish separately. We then obtain the Euler-Lagrange equations for fields:

$$\frac{\partial \mathscr{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} = 0 \,. \tag{2.26}$$

One can develop a Hamiltonian formalism exactly as in the case of finitely many degrees of freedom. Defining the conjugate momenta

$$\pi_i(x) = \frac{\partial \mathscr{L}}{\partial(\partial_0 \phi_i(x))} \,, \tag{2.27}$$

the Hamiltonian reads

$$
\begin{aligned}
H(t) &= \int d^3x \left[\sum_i \pi_i(t, \vec{x}) \partial_0 \phi_i(t, \vec{x}) - \mathscr{L}(\phi(t, \vec{x}), \partial_\mu \phi(t, \vec{x}), t, \vec{x}) \right] \\
&= \int d^3x \, \mathscr{H}(\pi_i(t, \vec{x}), \phi_i(t, \vec{x}), t, \vec{x}) \,,
\end{aligned} \tag{2.28}
$$

10

where we have made explicit the dependence on time $t = x^0$ and spatial coordinates $\vec{x}$, and in the last passage we have defined the Hamiltonian density $\mathscr{H}$. Notice, however, that upon imposing the equations of motion we find that H is time-independent if it is not depending explicitly on time. More precisely, if we compute explicitly the derivative we find

$$
\begin{aligned}
\frac{d}{dt}H(t) - \frac{\partial}{\partial t}H(t) &= \int d^3x \left[\sum_i \dot{\pi}_i \dot{\phi}_i + \pi_i \ddot{\phi}_i - \dot{\phi}_i \frac{\partial \mathscr{L}}{\partial \phi_i} - (\dot{\partial_\mu \phi})_i \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i(t,\vec{x}))} \right] \\
&= \int d^3x \left[\sum_i \partial_0 \left(\dot{\phi} \frac{\partial \mathscr{L}}{\partial(\partial_0 \phi)} \right) - \partial_\mu \left(\dot{\phi} \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)} \right) \right] \\
&= -\int d^3x \, \partial_j \left(\dot{\phi} \frac{\partial \mathscr{L}}{\partial(\partial_j \phi)} \right) = -\int_{\partial V \to \infty} dn_j \left(\dot{\phi} \frac{\partial \mathscr{L}}{\partial(\partial_j \phi)} \right) = 0 \,,
\end{aligned}
\tag{2.29}
$$

i.e., it equals the flux through the surface at infinity of a certain function of fields and momenta, which vanishes if fields are sufficiently well behaved at infinity. In full analogy with the finite-dimensional case we can derive from here the Hamilton equations[6]

$$
\begin{aligned}
\frac{\delta H}{\delta \pi_i(t,\vec{x})} &\equiv \frac{\partial \mathscr{H}}{\partial \pi_i(t,\vec{x})} = \dot{\phi}_i(t,\vec{x}) \,, \\
\frac{\delta H}{\delta \phi_i(t,\vec{x})} &\equiv \frac{\partial \mathscr{H}}{\partial \phi_i(t,\vec{x})} = -\dot{\pi}_i(t,\vec{x}) \,,
\end{aligned}
\tag{2.30}
$$

together with

$$
\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\int d^3x \, \frac{\partial \mathscr{L}}{\partial t} \,.
\tag{2.31}
$$

Finally, we can define Poisson brackets for functionals of fields and their conjugate momenta as

$$
\{F, G\}_{\text{PB}} = \int d^3z \sum_k \left[\frac{\delta F}{\delta \phi_k(t,\vec{z})} \frac{\delta G}{\delta \pi_k(t,\vec{z})} - \frac{\delta G}{\delta \phi_k(t,\vec{z})} \frac{\delta F}{\delta \pi_k(t,\vec{z})} \right] \,.
\tag{2.32}
$$

The fundamental Poisson brackets are

$$
\begin{aligned}
\{\phi_i(t,\vec{x}), \pi_j(t,\vec{y})\}_{\text{PB}} &= \int d^3z \sum_k \delta_{ik} \delta^{(3)}(\vec{x} - \vec{z}) \delta_{jk} \delta^{(3)}(\vec{y} - \vec{z}) = \delta_{ij} \delta^{(3)}(\vec{x} - \vec{y}) \,, \\
\{\phi_i(t,\vec{x}), \phi_j(t,\vec{y})\}_{\text{PB}} &= \{\pi_i(t,\vec{x}), \pi_j(t,\vec{y})\}_{\text{PB}} = 0 \,.
\end{aligned}
\tag{2.33}
$$

---

[6]Functional derivatives are defined by the relation

$$
\lim_{\eta \to 0} \frac{F[\phi + \eta \delta \phi] - F[\phi]}{\eta} = \int d^3x \, \delta \phi(t,\vec{x}) \frac{\delta F}{\delta \phi(t,\vec{x})} \,,
$$

and correspond essentially to the variation of the functional under an infinitesimal change of a field at one point. The functional derivative can be extracted from the relation above by taking $\delta \phi$ to be a delta function, or, more formally, using a representation of the delta function like $\Delta_\epsilon(z) = \frac{1}{\pi} \frac{\epsilon}{z^2 + \epsilon^2}$, as

$$
\frac{\delta F}{\delta \phi(t_0, \vec{x}_0)} = \lim_{\epsilon \to 0} \lim_{\eta \to 0} \frac{F[\phi + \eta \delta_\epsilon^{t_0, \vec{x}_0} \phi] - F[\phi]}{\eta} \,,
$$

where $\delta_\epsilon^{t_0, \vec{x}_0} \phi(t,\vec{x}) = \Delta_\epsilon(t - t_0) \prod_{j=1}^3 \Delta_\epsilon(\vec{x}_j - \vec{x}_{0j})$.

Finally, the time evolution of any such functional is given by

$$
\begin{aligned}
\frac{dF}{dt} &= \frac{\partial F}{\partial t} + \int d^3z \sum_k \left[\dot{\phi}_k(t, \vec{z}) \frac{\delta F}{\delta \phi_k(t, \vec{z})} + \dot{\pi}_k(t, \vec{z}) \frac{\delta F}{\delta \pi_k(t, \vec{z})} \right] \\
&= \frac{\partial F}{\partial t} + \int d^3z \sum_k \left[\frac{\delta F}{\delta \phi_k(t, \vec{z})} \frac{\delta H}{\delta \pi_i(t, \vec{x})} - \frac{\delta H}{\delta \phi_i(t, \vec{x})} \frac{\delta F}{\delta \pi_k(t, \vec{z})} \right] = \frac{\partial F}{\partial t} + \{F, H\}_{\mathrm{PB}} \, .
\end{aligned}
\tag{2.34}
$$

This provides a simpler proof that $\dot{H} = 0$ if there is no explicit time dependence, since obviously $\{H, H\}_{\mathrm{PB}} = 0$.

2.5 Classical symmetries and Noether's theorem

One of the main advantages of the Lagrangian formalism (and of the Hamiltonian formalism as well) over Newtonian mechanics is that it makes more transparent the role and the consequences of symmetry. This subsection is about symmetry and their consequences in the classical setting. The quantum mechanical case is discussed below.

Quoting almost verbatim from Weinberg's book, a *symmetry* is a change in the experimenter's point of view that does not change the results of possible experiments. Let us explain this in more detail. Consider two experimenters $\mathcal{O}$ and $\mathcal{O}'$ making measurements on the same physical system. They subscribe to the same operative rules concerning the measurement of the various observables, but they use in general different reference frames, so that in general they find different values for the various physical quantities, thus producing two different descriptions of the same system. In mathematical terms, this means that they will assign two different points Z and Z', respectively, to the system in its configuration space $\mathcal{C}$. Here Z denotes collectively all the coordinate and velocities characterising the system.[7] Although the two descriptions are in general not the same, for certain pairs of observers they will be *equivalent*, i.e., the physical laws implied by the measurements will be the same for both observers. In other words, it will be impossible for an observer to determine her or his reference frame using only her or his measurements. If two observers give equivalent descriptions, then the set of possible outcomes for measurements on the physical system must be the same, and the same $\mathcal{C}$ will describe the states of the system. Moreover, if the same laws of physics have to apply for the two observers, then the temporal evolution must be governed by the same equations.

Let us formalise the two statements above. Establishing a relation between the two descriptions corresponds to defining a mapping $\mathcal{M}$ from $\mathcal{C}$ to itself. Since the same outcomes are possible for both observers, such a mapping must be surjective (or onto), i.e., its image is the whole of $\mathcal{C}$. On top of that, different outcomes for $\mathcal{O}$ will show up as different outcomes for $\mathcal{O}'$, so $\mathcal{M}$ must also be injective (or one-to-one), and so $\mathcal{M}$ is a bijective, i.e., invertible, mapping. This can be seen more directly by observing that if the two observers must be equivalent, then if there is a mapping from $\mathcal{O}$ to $\mathcal{O}'$ there must be one also from $\mathcal{O}'$ to $\mathcal{O}$.

Next, if the time evolution of the system is governed by the same equations for both observers, then the way that the initial conditions Z_0 at time t_0 for $\mathcal{O}$, and the corresponding initial conditions Z'_0 at the corresponding time t'_0 for $\mathcal{O}'$, are mapped into the state Z of the system at t for $\mathcal{O}$ and Z' at t' for $\mathcal{O}'$ must be the same. Formally, denoting the evolution "operator" with $u_{t_0 \to t}$, we must have both

$$
Z(t) = u_{t_0 \to t}(Z(t_0)), \qquad Z'(t') = u_{t'_0 \to t'}(Z'(t'_0)), \tag{2.35}
$$

---

[7]Analogously, one can use coordinates and momenta in phase space Φ.

where $Z'(t') = \mathcal{M}_t(Z(t))$, allowing for a time-dependence in the mapping, and $t' = T(t)$ is the mapping between times for the two observers. But then

$$Z'(t') = \mathcal{M}_t(Z(t)) = \mathcal{M}_t(u_{t_0 \to t}(Z(t_0))) = u_{t_0' \to t'}(\mathcal{M}_{t_0}(Z(t_0))) = u_{T(t_0) \to T(t)}(\mathcal{M}_{t_0}(Z(t_0))),$$
(2.36)

which expresses in an almost unintelligible way the request that *the evolved of the transformed equals the transformed of the evolved*, which is a quite catchy formulation of our second requirement.

Summarising, a *symmetry transformation* is a transformation on the space of states of the system that relates the description of two equivalent observers, which by definition see the same and all the same possible states of the system, and see its temporal evolution being governed by the same law. The composition $\mathcal{M}_2 \circ \mathcal{M}_1$ of two symmetry transformations $\mathcal{M}_1$ and $\mathcal{M}_2$ is clearly still a symmetry transformation, and such a composition is associative, $\mathcal{M}_3 \circ (\mathcal{M}_2 \circ \mathcal{M}_1) = (\mathcal{M}_3 \circ \mathcal{M}_2) \circ \mathcal{M}_1$. Moreover, the identity transformation is obviously a symmetry, and symmetry transformations are invertible by definition. This sums up to symmetry transformation of a physical system forming a *group*, the (you don't say...) group of symmetry transformations of the system.

2.5.1 Particle mechanics

The Lagrangian formalism allows us to identify from the outset mappings which are symmetry transformations of our system. Let S and S' be the following action functionals,

$$S = \int_{t_0}^{t_1} dt\, L(q(t), \dot{q}(t), t),$$
$$S' = \int_{t_0'}^{t_1'} dt'\, L(q'(t'), \dot{q}'(t'), t'),$$
(2.37)

where primed quantites are obtained via the (generally time-dependent) mapping $\mathcal{M}$ from $\mathcal{O}$ to $\mathcal{O}'$, $q_i'(t') = \mathcal{M}_{i\,t}(q_i(t))$. Notice that the *same* Lagrangian appears in both expressions: in general it is possible to find some L' for which the new action functional for $\mathcal{O}'$ will equal that of $\mathcal{O}$, but this will imply different equations of motions (EOM) for the two observers. If, on the other hand, $S = S'$ up to boundary terms (see below), then the *same* equations of motions will govern the evolution of the system for the two observers. This is seen as follows. Equality up to boundary terms means that

$$S' = S + F(q(t_1), t_1) - F(q(t_0), t_0),$$
(2.38)

for some function $F(q, t)$. This means that the *variation* of the two functionals, which is done at fixed $q_i(t_0), q_i(t_0)$, and correspondingly at fixed $q_i'(t_0'), q_i'(t_1')$, will be the same, i.e., $\delta S = \delta S'$. Indeed,

$$\delta q_i'(t_{0,1}') = \mathcal{M}_{i\,t_{0,1}}(q(t_{0,1}) + \delta q(t_{0,1})) - \mathcal{M}_{i\,t_{0,1}}(q(t_{0,1})) = \mathcal{M}_{i\,t_{0,1}}(q(t_{0,1})) - \mathcal{M}_{i\,t_{0,1}}(q(t_{0,1})) = 0,$$
(2.39)

13

so the statement follows. Consider now $q_i(t)$ solving the EOM for $\mathcal{O}$. We have [see Eq. (2.12)]

$$
\begin{aligned}
0 = \delta S = \delta S' = {}& \left(\sum_i \delta q_i' \frac{\partial}{\partial \dot{q}_i'} L(q', \dot{q}', t') \right) \Bigg|_{t_0'}^{t_1'} \\
& + \int_{t_0'}^{t_1'} dt' \sum_i \delta q_i' \left(\frac{\partial}{\partial q_i'} L(q', \dot{q}', t') - \frac{d}{dt'} \frac{\partial}{\partial \dot{q}_i'} L(q', \dot{q}', t') \right),
\end{aligned}
\tag{2.40}
$$

but the first term in Eq. (2.40) vanishes due to Eq. (2.39), and since $\delta S = 0$ for arbitrary δq_i, which in turn imply arbitrary $\delta q_i'$, the second term can vanish only if

$$
\frac{\partial}{\partial q_i'} L(q', \dot{q}', t') - \frac{d}{dt'} \frac{\partial}{\partial \dot{q}_i'} L(q', \dot{q}', t') = 0,
\tag{2.41}
$$

i.e., if $q'(t')$ (which is the trajectory of the system as seen by $\mathcal{O}'$) satisfies the same EOM as $q(t)$.

The equality up to boundary terms means in practice that the change in the action corresponds to adding a total derivative to the Lagrangian:[8]

$$
S' = S + F(q(t_1), t_1) - F(q(t_0), t_0) = S + \int_{t_0}^{t_1} dt \, \frac{d}{dt} F(q(t), t).
\tag{2.42}
$$

A well known example of this situation is that of a free particle under change of frame through a Galilei transformation $q \to q' = q - vt$, $t \to t' = t$:

$$
\begin{aligned}
L(q, \dot{q}) = {}& \frac{1}{2} m \dot{q}^2, \\
L(q', \dot{q}') = {}& \frac{1}{2} m \dot{q}'^2 = L(q, \dot{q}) + mv\dot{q} + \frac{1}{2} mv^2 = L(q, \dot{q}) + \frac{d}{dt} \left(mvq + \frac{1}{2} mv^2 t \right).
\end{aligned}
\tag{2.43}
$$

There is more to invariance of the action up to boundary terms than the fact that the EOM are the same for the two observers. In fact, if we have invariance under a continuous group of transformations, Noether's theorem guarantees the existence of a conserved charge. Consider an infinitesimal such transformation,

$$
\begin{aligned}
q_i'(t') &= q_i(t) + \delta q_i(t) = q_i(t) + \epsilon M_i(q, \dot{q}, t), \\
t' &= t + \delta t = t + \epsilon T(q, \dot{q}, t),
\end{aligned}
\tag{2.44}
$$

and assume invariance up to boundary terms of the action, as above:

$$
\int_{t_0}^{t_1} dt \, [L(q(t), \dot{q}(t), t) + \epsilon F(q(t), t)] = \int_{t_0'}^{t_1'} dt' \, L(q'(t'), \dot{q}'(t'), t').
\tag{2.45}
$$

Since the transformation is infinitesimal, also the boundary term must be (it has to vanish as $\epsilon \to 0$). Changing variables back to t in the right-hand side, and expanding up to first order in the variations, we find

$$
\begin{aligned}
\int_{t_0'}^{t_1'} dt' \, L(q'(t'), \dot{q}'(t'), t') = {}& \int_{t_0}^{t_1} dt \left(1 + \frac{d\delta t}{dt} \right) L(q'(t + \delta t), \dot{q}'(t + \delta t), t + \delta t) = \\
& \int_{t_0}^{t_1} dt \left(1 + \frac{d\delta t}{dt} \right) \left[L(q'(t), \dot{q}'(t), t) + \delta t \frac{d}{dt} L(q(t), \dot{q}(t), t) \right].
\end{aligned}
\tag{2.46}
$$

---

[8]We assume that $dt'/dt > 0$.

From Eq. (2.44) we find (notice that although they differ in general, $\dot{q}'(t') = \dot{q}(t)$ to zeroth order)

$$q'_i(t) = q'_i(t' - \delta t) = q'_i(t') - \delta t \dot{q}_i(t) = q_i(t) + \delta q_i(t) - \delta t \dot{q}_i(t)$$
$$\equiv q_i(t) + \overline{\delta q}_i(t) = q_i(t) + \epsilon[M_i(q, \dot{q}, t) - \dot{q}_i(t)T(q, \dot{q}, t)],$$

(2.47)

and plugging this back into Eq. (2.46) and expanding we get

$$\int_{t'_0}^{t'_1} dt' \, L(q'(t'), \dot{q}'(t'), t')$$

$$= \int_{t_0}^{t_1} dt \left(1 + \frac{d\delta t}{dt} \right) \left[L(q(t), \dot{q}(t), t) + \sum_i \overline{\delta q}_i(t) \frac{\partial}{\partial q_i} L(q(t), \dot{q}(t), t) \right.$$

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

(2.48)

$$= \int_{t_0}^{t_1} dt \left\{ L(q(t), \dot{q}(t), t) + \sum_i \overline{\delta q}_i(t) \left[\frac{\partial}{\partial q_i} L(q(t), \dot{q}(t), t) - \frac{d}{dt} \frac{\partial}{\partial \dot{q}_i} L(q(t), \dot{q}(t), t) \right] \right.$$

$$\left. + \frac{d}{dt} \left[\sum_i \overline{\delta q}_i(t) \frac{\partial}{\partial \dot{q}_i} L(q(t), \dot{q}(t), t) + \delta t L(q(t), \dot{q}(t), t) \right] \right\}.$$

Comparing this with Eq. (2.45), and imposing EOM to make the term proportional to $\overline{\delta q}_i$ vanish, we find

$$0 = \int_{t_0}^{t_1} dt \frac{d}{dt} \left[\sum_i \overline{\delta q}_i(t) \frac{\partial}{\partial \dot{q}_i} L(q(t), \dot{q}(t), t) + \delta t L(q(t), \dot{q}(t), t) - \epsilon F(q(t), t) \right],$$

$$= \epsilon \int_{t_0}^{t_1} dt \frac{d}{dt} \left[\sum_i [M_i(q, \dot{q}, t) - \dot{q}_i(t)T(q, \dot{q}, t)] \frac{\partial}{\partial \dot{q}_i} L(q(t), \dot{q}(t), t) \right.$$

(2.49)

$$\left. + T(q, \dot{q}, t)L(q(t), \dot{q}(t), t) - F(t, q(t)) \right],$$

which entails that the following quantity, named *Noether charge*, is a constant of motion (we drop the t-dependence of q and $\dot{q}$ from the notation for simplicity):

$$Q(q, \dot{q}, t) = \sum_i \left\{ [M_i(q, \dot{q}, t) - \dot{q}_i T(q, \dot{q}, t)] \frac{\partial}{\partial \dot{q}_i} L(q, \dot{q}, t) \right\} + T(q, \dot{q}, t)L(q, \dot{q}, t) - F(q, t)$$

$$= \sum_i \left\{ [M_i(q, \dot{q}, t) - \dot{q}_i T(q, \dot{q}, t)] p_i(q, \dot{q}, t) \right\} + T(q, \dot{q}, t)L(q, \dot{q}, t) - F(q, t).$$

(2.50)

In the case when $M_i = M_i(q)$, we have that Q generates the transformation of the coordinates through the Poisson brackets:

$$\{q_i, Q\}_{\mathrm{PB}} = M_i(q).$$

(2.51)

Notice that in this case

$$\{p_i, Q\}_{\mathrm{PB}} = -\sum_j \frac{\partial}{\partial q_i} M_j(q) p_j,$$

(2.52)

and the combined transformation of q_i and p_i is an infinitesimal canonical transformation, preserving the Poisson brackets.

Let us now consider a few examples.

Time translations In this case $M_i = 0$, $T = 1$. If the Lagrangian itself is invariant, i.e., independent of time, then there is no F, and we find

$$Q = L - \sum_i \dot{q}_i p_i = -H \,. \tag{2.53}$$

The Hamiltonian, i.e., the energy of the system is constant if there is invariance under time translations.

Coordinate translations In this case $M_i = \delta_{ij}$ if we are changing only q_j, and $T = 0$. Assuming again that the Lagrangian itself is invariant then $F = 0$, and we find

$$Q_j = p_j \,. \tag{2.54}$$

This is obvious from the EOM: if the Lagrangian is independent of q_j, then one finds immediately $\dot{p}_j = 0$. If q_j are Cartesian coordinates, invariance under space translations is associated with conservation of spatial momentum.

Rotations Let q_i, $i = 1, 2, 3$ be Cartesian coordinates. For an infinitesimal rotation around axis j we have $M_i^{(j)} = \sum_k \varepsilon_{ijk} q_k$, $T = 0$. When $F = 0$

$$Q^{(j)} = \sum_{ij} \varepsilon_{ijk} q_k p_i = (\vec{q} \wedge \vec{p})_j = \vec{L}_j \,, \tag{2.55}$$

and we see that rotational invariance is associated with conservation of angular momentum.

Galilei boost This is a different example, as it is impossible to write a Lagrangian, function of q and $\dot{q}$, exactly invariant under Galilei boosts.[9] Consider a Lagrangian of the form

$$L(q_i, \dot{q}_i) = \sum_i \frac{1}{2} m_i \dot{\vec{q}}_i^2 - V(\vec{q}_i - \vec{q}_j) \,, \tag{2.56}$$

where $\vec{q}_i$ are the spatial coordinates of particle i, and V some potential that depends only on the relative positions of particles. Under $\vec{q}_i \rightarrow \vec{q}_i' = \vec{q}_i - \vec{v}t$, $t \rightarrow t' = t$ we find

$$L(q_i', \dot{q}_i') = \sum_i \frac{1}{2} m_i \dot{\vec{q}}_i'^2 - V(\vec{q}_i' - \vec{q}_j') = L(q_i, \dot{q}_i) + \vec{v} \cdot \frac{d}{dt} \sum_i \left(m_i \vec{q}_i + \frac{1}{2} m_i \vec{v}t \right) \,. \tag{2.57}$$

For infinitesimal $\vec{v}_j$ we have $M_{ij} = -t\delta_{ij}$, $T = 0$, $F_j = \sum_i m_i (\vec{q}_i)_j$ (the $\mathcal{O}(v)$ term is of higher order), and so

$$\vec{Q} = \sum_i (m_i \vec{q}_i - \vec{p}_i t) \,. \tag{2.58}$$

This conservation law expresses the fact that the center of mass moves at constant speed on a straight line:

$$\frac{\sum_i m_i \vec{q}_i}{\sum_i m_i} = t \frac{\sum_i \vec{p}_i}{\sum_i m_i} + \vec{Q} = \frac{\vec{P}_{\text{tot}}}{M_{\text{tot}}} t + \vec{Q} \,. \tag{2.59}$$

---

[9]Imposing $L(q-vt, \dot{q}-v) = L(q, \dot{q})$ for all v, t implies $0 = \frac{\partial}{\partial t} L(q, \dot{q}) = -v \frac{\partial}{\partial q} L(q-vt, \dot{q}-v)$, and taking one further derivative with respect to v and setting $v = 0$ we find $\frac{\partial}{\partial q} L(q, \dot{q}) = 0$, i.e., $\mathscr{L}$ is independent of its first argument. Taking instead the derivative with respect to v and setting $v = 0$ we find $0 = t \frac{\partial}{\partial q} L(q, \dot{q}) + \frac{\partial}{\partial \dot{q}} L(q, \dot{q}) = \frac{\partial}{\partial \dot{q}} L(q, \dot{q})$.

2.5.2 Field theory

The same game can be played when working with fields. We assume again invariance of the action under a symmetry transformation up to boundary terms, which are now given by the integral of a total divergence. Let

$$
\begin{aligned}
S &= \int_D d^4x \, \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) \,, \\
S' &= \int_{D'} d^4x' \, \mathscr{L}(\phi'(x'), \partial'_\mu \phi'(x'), x') \,,
\end{aligned}
\tag{2.60}
$$

where $\phi'_i(x') = \mathcal{M}_i(\phi(x), \partial_\mu \phi(x), x)$ and $x' = X(x)$ for some transformation $(\mathcal{M}, X)$ mapping from observer $\mathcal{O}$ to observer $\mathcal{O}'$. If

$$
S' = S + \int_D d^4x \, \partial_\mu F^\mu(\phi(x), x) \,,
\tag{2.61}
$$

then the same EOM apply for both observers. The argument is identical to the one given above for finitely many degrees of freedom, once that the appropriate generalisations are made, and will not be given here again. Consider then an infinitesimal such transformation,

$$
\begin{aligned}
\phi'_i(x') &= \phi_i(x) + \delta\phi_i(x) = \phi_i(x) + \epsilon M_i(\phi, \partial_\mu \phi, x) \,, \\
x'^\mu &= x^\mu + \delta x^\mu = x^\mu + \epsilon \mathcal{A}^\mu(x) \,.
\end{aligned}
\tag{2.62}
$$

For future utility, let us note that

$$
\begin{aligned}
\phi'_i(x) &= \phi'_i(x' - \delta x) = \phi'_i(x') - \delta x^\mu \partial_\mu \phi_i(x) = \phi_i(x) + \delta\phi_i(x) - \delta x^\mu \partial_\mu \phi_i(x) \\
&\equiv \phi_i(x) + \overline{\delta\phi_i}(x) = \phi_i(x) + \epsilon[M_i(\phi, \partial_\mu \phi, x) - \mathcal{A}^\mu(x)\partial_\mu \phi_i(x)] \,.
\end{aligned}
\tag{2.63}
$$

We then proceed as above, changing variables in S' back to x,

$$
S' = \int_D d^4x \left| \det_{\mu\nu} \frac{\partial x'^\mu}{\partial x^\nu} \right| \mathscr{L}(\phi'(x + \delta x), \partial_\mu \phi'(x + \delta x), x + \delta x) \,,
\tag{2.64}
$$

and then expanding in ϵ. Notice that to lowest order

$$
\left| \det_{\mu\nu} \frac{\partial x'^\mu}{\partial x^\nu} \right| = \left| \det_{\mu\nu} \left(\delta^\mu{}_\nu + \frac{\partial \delta x^\mu}{\partial x^\nu} \right) \right| = \left| 1 + \mathrm{tr}\left(\frac{\partial \delta x^\mu}{\partial x^\nu} \right) \right| = \left| 1 + \partial_\mu \delta x^\mu \right| = 1 + \partial_\mu \delta x^\mu \,.
\tag{2.65}
$$

We then find

$$
\begin{aligned}
S' = \int_D d^4x \Bigg\{ &\mathscr{L}(\phi(x), \partial_\mu \phi(x), x) + \partial_\mu \left[\delta x^\mu \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) \right] \\
&+ \sum_i \left[\overline{\delta\phi_i}(x) \frac{\partial}{\partial \phi_i} \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) + \partial_\mu(\overline{\delta\phi_i}(x)) \frac{\partial}{\partial(\partial_\mu \phi_i)} \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) \right] \Bigg\} \,.
\end{aligned}
\tag{2.66}
$$

Using Eq. (2.61) and integrating by parts we find

$$
\begin{aligned}
0 = \int_D d^4x \Bigg\{ &\partial_\mu \left[\delta x^\mu \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) + \sum_i \overline{\delta\phi_i}(x) \frac{\partial}{\partial(\partial_\mu \phi_i)} \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) - F^\mu(\phi, x) \right] \\
&+ \sum_i \overline{\delta\phi_i}(x) \left[\frac{\partial}{\partial \phi_i} \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) - \partial_\mu \frac{\partial}{\partial(\partial_\mu \phi_i)} \mathscr{L}(\phi(x), \partial_\mu \phi(x), x) \right] \Bigg\} \,.
\end{aligned}
\tag{2.67}
$$

Imposing now the EOM, and given the arbitrariness of D and ϵ, we conclude that the *Noether current* J^μ,

$$J^\mu(\phi,\partial_\mu\phi,x) =$$
$$\mathcal{A}^\mu\mathscr{L}(\phi,\partial_\mu\phi,x) + \sum_i [M_i(\phi,\partial_\mu\phi,x) - \mathcal{A}^\nu\partial_\nu\phi_i]\frac{\partial}{\partial(\partial_\mu\phi_i)}\mathscr{L}(\phi,\partial_\mu\phi,x) - F^\mu(\phi,x)\,, \quad (2.68)$$

is a conserved current,

$$\partial_\mu J^\mu = 0\,. \tag{2.69}$$

From the Noether current one can easily construct a conserved Noether charge,

$$Q = \int d^3x\, J^0(x)\,. \tag{2.70}$$

A simple calculation shows that

$$\frac{d}{dt}Q = \int d^3x\, \partial_0 J^0(x) = -\int d^3x\, \partial_j J^j(x) = -\lim_{V\to\infty}\int_V d^3x\, \partial_j J^j(x)$$
$$= -\lim_{V\to\infty}\int_{\partial V} dn_j\, J^j(x) = 0\,, \tag{2.71}$$

assuming that the fields are sufficiently well behaved at infinity, so that the flux of J^j at infinity vanishes. Explicitly,

$$Q = \int d^3x\left[\mathcal{A}^0\mathscr{L} + \sum_i [M_i - \mathcal{A}^\nu\partial_\nu\phi_i]\frac{\partial}{\partial(\partial_0\phi_i)}\mathscr{L} - F^0\right]$$
$$= \int d^3x\left[\mathcal{A}^0\left(\mathscr{L} - \sum_i \partial_0\phi_i\pi_i\right) + \sum_i [M_i - \mathcal{A}^j\partial_j\phi_i]\pi_i - F^0\right] \tag{2.72}$$
$$= \int d^3x\left[\sum_i [M_i - \mathcal{A}^j\partial_j\phi_i]\pi_i - \mathcal{A}^0\mathscr{H} - F^0\right].$$

Also in this case, for $M_i = M_i(\phi)$ we find (recall that $Q(t) = Q(0)$ can be taken at any time since it is conserved)

$$\{\phi_i(t,\vec{x}), Q\}_{\mathrm{PB}} = \int d^3x \int d^3z \sum_k \frac{\delta\phi_i(t,\vec{x})}{\delta\phi_k(t,\vec{z})}\frac{\delta J^0(\phi,\partial\phi,x)}{\delta\pi_k(t,\vec{z})} = \int d^3x\frac{\delta J^0(\phi,\partial\phi,x)}{\delta\pi_i(t,\vec{x})}$$
$$= [M_i(\phi(t,\vec{x})) - \mathcal{A}^j(t,\vec{x})\partial_j\phi_i(t,\vec{x})] - \mathcal{A}^0(t,\vec{x})\frac{\partial\mathscr{H}}{\partial\pi(t,\vec{x})} \tag{2.73}$$
$$= [M_i(\phi(t,\vec{x})) - \mathcal{A}^j(t,\vec{x})\partial_j\phi_i(t,\vec{x})] - \mathcal{A}^0(t,\vec{x})\partial_0\phi(t,\vec{x})$$
$$= [M_i(\phi(t,\vec{x})) - \mathcal{A}^\nu(t,\vec{x})\partial_\nu\phi_i(t,\vec{x})]\,,$$

which is precisely the transformation of the fields (up to the parameter ϵ).

Also in this case let us consider a few examples.

Translations The simplest case is that of invariance under translations. We assume again that no boundary term appears, which is the case if there is no explicit dependence of the Lagrangian density on the coordinates (the only case that we will be considering in practice). In this case the action is obviously invariant. There are four kinds of translations, corresponding to four $\mathcal{A}^{\mu}_{(\nu)} = \delta^{\mu}{}_{\nu}$, with $M_i = 0$. The components of the corresponding conserved currents form the *canonical energy-momentum tensor*:

$$\Theta^{\mu}{}_{\nu} = -J^{\mu}_{(\nu)} = \sum_i \partial_\nu \phi_i \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} - \delta^{\mu}{}_{\nu} \mathscr{L} \,, \qquad \Theta^{\mu\nu} = \sum_i \partial^\nu \phi_i \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} - \eta^{\mu\nu} \mathscr{L} \,. \qquad (2.74)$$

The reason for this nomenclature is easy: the $\mu = 0$ components of this tensor represent the energy and momentum density of the system, from which energy and momentum are obtained via integration:

$$\int d^3x \, \Theta^{00} = \int d^3x \sum_i \partial^0 \phi_i \pi_i - \mathscr{L} = \int d^3x \, \mathscr{H} = H = P^0 \,,$$

$$\int d^3x \, \Theta^{0j} = \int d^3x \sum_i \partial^j \phi_i \pi_i = -\int d^3x \sum_i \partial_j \phi_i \pi_i = P^j \,. \qquad (2.75)$$

Lorentz transformations The next, and most interesting case would be that of Lorentz transformations, but we are not ready yet, since we do not know yet enough about them to discuss this point. This will have to wait a bit.

Internal symmetries Certain theories are symmetric under transformations that involve only the fields and not the spacetime coordinates. Such transformations, mixing the various fields, are called internal transformations, and the corresponding symmetry is an *internal symmetry*. The simplest such transformations are linear in the fields, i.e.,

$$\delta \phi_i(x) = \epsilon \sum_j K_{ij} \phi_j(x) \,. \qquad (2.76)$$

The corresponding conserved current and charge are easily determined,

$$J^\mu = \sum_{ij} \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_i)} K_{ij} \phi_j \,,$$

$$Q = \int d^3x \sum_{ij} \frac{\partial \mathscr{L}}{\partial(\partial_0 \phi_i)} K_{ij} \phi_j = \int d^3x \sum_{ij} \pi_i K_{ij} \phi_j \,, \qquad (2.77)$$

and one immediately observes that

$$\{\phi_i, Q\}_{\mathrm{PB}} = \sum_j K_{ij} \phi_j \,. \qquad (2.78)$$

2.6 Classical electrodynamics

Let us apply the formalism developed above to a concrete case, that of classical electrodynamics. The classical theory of electromagnetism is contained in the four Maxwell equations:

$$\vec{\nabla} \cdot \vec{E} = \rho, \qquad \vec{\nabla} \wedge \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0,$$
$$\vec{\nabla} \cdot \vec{B} = 0, \qquad \vec{\nabla} \wedge \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{j}. \tag{2.79}$$

Here $\vec{E}$ and $\vec{B}$ are the electric and magnetic fields, ρ the charge density and $\vec{j}$ the electric current. These equations can be written in compact, and manifestly Lorentz-covariant, form, using the *four-potential* $A^\mu = (\varphi, \vec{A})$, in terms of which the fields read

$$\vec{E} = -\vec{\nabla}\varphi - \frac{\partial \vec{A}}{\partial t}, \qquad \vec{B} = \vec{\nabla} \wedge \vec{A}. \tag{2.80}$$

Defining the *field strength tensor*, $F_{\mu\nu}$, and the *four-current*, J^μ, as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \qquad J^\mu = (\rho, \vec{j}) \tag{2.81}$$

the inhomogeneous Maxwell equations are written as

$$\partial_\mu F^{\mu\nu} = J^\mu, \tag{2.82}$$

while the homogeneous equations, which are automatically solved thanks to the representation Eq. (2.80), are encoded in the *Bianchi identities*,

$$\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} = 0, \tag{2.83}$$

identically satisfied by the field strength tensor. Electric and magnetic field components are related to the field strength tensor components as follows:

$$\vec{E}_i = F_{0i} = \partial_0 A_i - \partial_i A_0 = -\frac{\partial}{\partial t}\vec{A}_i - \vec{\nabla}_i\varphi,$$
$$\vec{B}_i = -\frac{1}{2}\varepsilon_{ijk}F_{jk} = -\varepsilon_{ijk}\partial_j A_k = \varepsilon_{ijk}\vec{\nabla}_j\vec{A}_k = (\vec{\nabla} \wedge \vec{A})_i. \tag{2.84}$$

Notice also the inverse relations

$$F_{ij} = -\epsilon_{ijk}\vec{B}_k. \tag{2.85}$$

Let us now check the inhomogeneous equations:

$$\partial_\mu F^{\mu 0} = J^0 = \partial_j F_{0j} = \vec{\nabla} \cdot \vec{E} = \rho,$$
$$\partial_\mu F^{\mu j} = J^j = -\partial_0 F_{0j} + \partial_i F_{ij} = -\frac{\partial}{\partial t}\vec{E}_j + \epsilon_{jik}\vec{\nabla}_i\vec{B}_k = -\frac{\partial}{\partial t}\vec{E}_j + (\vec{\nabla} \wedge \vec{B})_j = \vec{j}_j. \tag{2.86}$$

The equations of motion Eq. (2.82) can be obtained from the following action by means of the usual variational principle:

$$S = -\frac{1}{4}\int d^4x \, F_{\mu\nu}F^{\mu\nu} - \int d^4x \, J_\mu A^\mu. \tag{2.87}$$

Indeed,

$$\frac{\partial \mathscr{L}}{\partial(\partial_\mu A_\nu)} = -\frac{1}{2}F^{\alpha\beta}\frac{\partial F_{\alpha\beta}}{\partial(\partial_\mu A_\nu)} = -F^{\alpha\beta}\frac{\partial_\alpha A_\beta}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu}, \tag{2.88}$$

and

$$\frac{\partial \mathscr{L}}{\partial A_\mu} = -J^\mu, \tag{2.89}$$

and so we get

$$-J^\nu + \partial_\mu F^{\mu\nu} = 0. \tag{2.90}$$

This equation shows that not any current can be coupled to the electromagnetic field, but only a conserved one: indeed, taking the divergence ∂_ν of Eq. (2.90) we obtain $\partial_\nu J^\nu = 0$, since $\partial_\mu \partial_\nu F^{\mu\nu} = 0$ identically due to the antisymmetry of $F^{\mu\nu}$.

Let us now express the EOM in terms of the four-potential,

$$\Box A^\nu - \partial^\nu \partial_\mu A^\mu = J^\nu, \tag{2.91}$$

where $\Box = \partial_\mu \partial^\mu = \partial_0^2 - \vec{\nabla}^2$ is the D'Alembert operator. Writing separately the $\nu = 0$ and the $\nu = 1, 2, 3$ equations we find

$$\begin{aligned} \rho &= -(\vec{\nabla}^2\varphi + \partial_0\vec{\nabla}\cdot\vec{A}) = -\vec{\nabla}\cdot(\vec{\nabla}\varphi + \partial_0\vec{A}), \\ \vec{j} &= \Box\vec{A} + \vec{\nabla}(\partial_0\varphi + \vec{\nabla}\cdot\vec{A}) = \partial_0(\partial_0\vec{A} + \vec{\nabla}\varphi) - \vec{\nabla}^2\vec{A} + \vec{\nabla}\vec{\nabla}\cdot\vec{A} \end{aligned} \tag{2.92}$$

which in terms of physical fields are just Gauss' law and the Maxwell-Faraday equation, respectively,

$$\begin{aligned} \vec{\nabla}\cdot\vec{E} &= \rho, \\ -\partial_0\vec{E} + \vec{\nabla}\wedge\vec{B} &= \vec{j}. \end{aligned} \tag{2.93}$$

The first equation in Eq. (2.92) is not a dynamical equation for φ, since its time derivative does not appear. The other three equations involve four functions, and so their solutions are undetermined by one arbitrary function. Let us choose to parameterise this indeterminacy with φ, and write $\vec{A} = \vec{A}(\varphi)$ for the vector potential solving the equations, where φ is so far arbitrary. We might hope to fix φ by imposing the constraint provided by the first equation, but here we run into a problem. Let us take the time derivative of the constraint equation, and impose that the equations of motion are solved. Using the continuity equation $\partial_0\rho + \vec{\nabla}\cdot\vec{j} = 0$ (i.e., conservation of the current), we find

$$\begin{aligned} \partial_0(\vec{\nabla}^2\varphi + \partial_0\vec{\nabla}\cdot\vec{A} + \rho) &= \vec{\nabla}\cdot(\vec{\nabla}\partial_0\varphi + \partial_0^2\vec{A} - \vec{j}) \\ &= \vec{\nabla}\cdot[\Box\vec{A} + \vec{\nabla}^2\vec{A} + \vec{\nabla}\partial_0\varphi - \vec{j}) = \vec{\nabla}\cdot[\Box\vec{A} + \vec{\nabla}(\vec{\nabla}\cdot\vec{A} + \partial_0\varphi) - \vec{j}] = 0. \end{aligned} \tag{2.94}$$

This means that if we impose the constraint at a certain time, then by virtue of the equations of motion it will be satisfied automatically at all times. Stated differently, the functional relation between $\vec{A}$ and φ implied by the equations of motion entails that the constraint will be satisfied no matter what choice we make for φ, as soon as we enforce the constraint at one given time. The system of differential equations Eq. (2.92) is thus underdetermined.

The difficulty found above can be rephrased in a number of ways. The fact that the temporal derivative of the constraint vanishes by virtue of the equations of motions implies that the system of equations Eq. (2.92) does not specify the dynamics of φ: if we try to recast one of the equations

of the second set as an equation for $\partial_0\varphi$, we end up with an empty identity. Another way to see that problem is to solve explicitly the constraint for φ and plug it into the dynamical equations. Solving for φ requires inverting the Laplacian $\Delta = \vec{\nabla}^2$ (something we will learn to do in due time), and gives

$$\varphi = -\frac{1}{\Delta}(\partial_0 \vec{\nabla} \cdot \vec{A} + \rho) \, . \tag{2.95}$$

Substituting it in the other equations, we find after a little manipulation [using the fact that $\vec{\nabla}(\frac{1}{\Delta}f) = \frac{1}{\Delta}\vec{\nabla}f$]

$$\Box\left(\vec{A} - \vec{\nabla}\frac{1}{\Delta}\vec{\nabla}\cdot\vec{A}\right) = \vec{j} - \vec{\nabla}\frac{1}{\Delta}\vec{\nabla}\cdot\vec{j} \Rightarrow \Box\vec{A}_\perp = \vec{j}_\perp \, . \tag{2.96}$$

The quantities $\vec{A}_\perp \equiv \vec{A} - \vec{\nabla}\frac{1}{\Delta}\vec{\nabla}\cdot\vec{A}$ and $\vec{j}_\perp \equiv \vec{j} - \vec{\nabla}\frac{1}{\Delta}\vec{\nabla}\cdot\vec{j}$ are automatically *transverse*, i.e., divergenceless, so Eq. (2.96) is an equation for two dynamical degrees of freedom. The remaining degrees of freedom, $\vec{\nabla}\cdot\vec{A}$ and φ, are related by Eq. (2.95), but are otherwise arbitrary.

In fact, the constraint equation is solved by infinitely many choices: given any solution $(\varphi, \vec{A})$ of the constraint, one can construct a new one by sending $\varphi \to \varphi + \partial_0\Lambda$, $\vec{A} \to \vec{A} - \vec{\nabla}\Lambda$ for an arbitrary function Λ. Since the transverse part of the vector potential is unchanged by this transformation, $\vec{A}_\perp \to \vec{A} - \vec{\nabla}\Lambda - \vec{\nabla}\frac{1}{\Delta}\vec{\nabla}\cdot(\vec{A} - \vec{\nabla}\Lambda) = \vec{A}_\perp$, what we obtain is a new solution of the system of equations. In summary, the equations of motion are invariant under the transformation

$$A_\mu(x) \to A_\mu(x) + \partial_\mu\Lambda(x) \, , \tag{2.97}$$

for arbitrary $\Lambda(x)$. This is a *gauge transformation*, and the invariance under them is called *gauge invariance*. This is actually an invariance of the Lagrangian. To see that this is an invariance, notice that $F_{\mu\nu}$ is manifestly invariant, while for the current term the change is given by

$$\int d^4x \, J^\mu\partial_\mu\Lambda = \int d^4x \, [\partial_\mu(J^\mu\Lambda) - \Lambda\partial_\mu J_\mu] = \int d^4x \, \partial_\mu(J^\mu\Lambda) \, , \tag{2.98}$$

due to current conservation. The action changes by a boundary term only, so the equations of motion do not change. Gauge invariance shows once again that the system of equations Eq. (2.92) is underdetermined, and solutions to the Cauchy problem, i.e., solutions of the EOM with prescribed values for fields and their spatial derivatives at $t = 0$, are not unique: we can always add an arbitrary Λ that vanishes at $t = 0$ and with vanishing spatial derivative there, and still obtain a solution. Our description of the system in terms of the four-potential is then redundant, and to specify the solution uniquely we have to impose extra conditions on A_μ, i.e., we have to *fix the gauge*.

Gauge invariance and the presence of constraints among the EOM are related facts, and, as we will see, they require a careful treatment when quantising the theory. As a matter of fact, the special nature of the equation Eq. (2.91) for $\nu = 0$ can be seen directly from the Lagrangian. If we compute the conjugate momenta we find

$$\pi^\mu = \frac{\partial\mathscr{L}}{\partial(\partial_0 A_\mu)} = -F^{0\mu} \, . \tag{2.99}$$

In particular, $\pi^0 = 0$, i.e., there is no conjugate momentum for A^0. This is a consequence of gauge invariance, and the reason why the time derivative of A^0 does not appear in the EOM for A^0, which is therefore a constraint for A^0 and not a dynamical equation. Moreover, since $\partial_0 A_0$

does not appear in the Lagrangian, the second time derivative of A^0 never appears in the EOM. The other momenta, on the other hand, are well defined and are simply

$$\pi^j = -F^{0j} = F_{0j} = -(\partial_0 A^j + \partial_j A^0) = \vec{E}_j \,. \tag{2.100}$$

In terms of them, the EOM for A^0 reads

$$\vec{\nabla} \cdot \vec{\pi} = J^0 \,. \tag{2.101}$$

As a Hamiltonian system, the one corresponding to the Maxwell Lagrangian is a constrained system, with constraints

$$\pi^0 = 0 \,, \qquad \vec{\nabla} \cdot \vec{\pi} - J^0 = 0 \,. \tag{2.102}$$

As we will see, this makes things not quite straightforward when the time comes for quantisation.

Let us now compute the energy-momentum tensor, setting the current $J^\mu = 0$. From Eq. (2.74) we obtain

$$\Theta^{\mu\nu} = -\partial^\nu A_\sigma F^{\mu\sigma} + \frac{1}{4}\eta^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} \,. \tag{2.103}$$

It is possible to define a new energy-momentum tensor which is symmetric and conserved, differing from the canonical one by a divergenceless term:

$$\Theta^{\mu\nu} = -\partial_\sigma A^\nu F^{\mu\sigma} - F^\nu{}_\sigma F^{\mu\sigma} + \frac{1}{4}\eta^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} = -\partial_\sigma A^\nu F^{\mu\sigma} + \Theta_S^{\mu\nu} \,. \tag{2.104}$$

The symmetric energy-momentum tensor is also conserved:

$$0 = \partial_\mu \Theta^{\mu\nu} = -\partial_\mu(\partial_\sigma A^\nu F^{\mu\sigma}) + \partial_\mu \Theta_S^{\mu\nu} = -\partial_\mu \partial_\sigma A^\nu F^{\mu\sigma} - \partial_\sigma A^\nu \partial_\mu F^{\mu\sigma} + \partial_\mu \Theta_S^{\mu\nu} = \partial_\mu \Theta_S^{\mu\nu} \,, \tag{2.105}$$

since $\partial_\mu F^{\mu\sigma}$ upon imposing the EOM, and $\partial_\mu \partial_\sigma A^\nu F^{\mu\sigma} = 0$ since it is the contraction of a symmetric and an antisymmetric tensor. The symmetric energy-momentum tensor is known as the *Belinfante-Rosenfeld* tensor. We will see later how the possibility to define a symmetric energy-momentum tensor is related to the Lorentz invariance of the theory. Let us work out explicitly the components of the symmetric energy-momentum tensor:

$$\begin{aligned} \Theta_S^{00} &= \frac{1}{2}(\vec{E}^2 + \vec{B}^2) \,, \\ \Theta_S^{0i} &= (\vec{E} \wedge \vec{B})_i \,, \\ \Theta_S^{ij} &= \frac{1}{2}\delta_{ij}(\vec{E}^2 + \vec{B}^2) - \vec{E}_i \vec{E}_j - \vec{B}_i \vec{B}_j \,. \end{aligned} \tag{2.106}$$

The first two lines correspond to the energy density and the momentum density of the electromagnetic field, respectively. The vector $\vec{S} = \vec{E} \wedge \vec{B}$ is known as the Poynting vector. As a final remark, notice that a nonzero current would only contribute a term $\eta^{\mu\nu} A_\alpha J^\alpha$ to the tensor, a term which is diagonal (so symmetric). This gives a contribution to the energy density, but not to the momentum density of the system: this can be understood from the fact that in our setting the current is not dynamical.

2.7 Canonical quantisation in quantum mechanics

The Hamiltonian formalism, especially when expressed in terms of Poisson brackets, is the starting point for *canonical quantisation*. In the finite-dimensional case this is just quantum mechanics as you have learnt it: Poisson brackets of classical (c-number) observables get replaced by $-i/\hbar$ times the commutator of their quantum mechanical (operator) counterparts. Let us look at this quickly, introducing in passing the concept of picture, and reviewing Heisenberg's, Schrödinger's, and Dirac's.

Quantisation of a classical system with coordinates q_i and momenta p_i, and a time-independent Hamiltonian $H(q, p)$ is achieved by replacing the classical observables with self-adjoint operators on a Hilbert space, and the fundamental Poisson brackets with commutators:

$$[q_i, p_j] = i\delta_{ij}, \qquad [q_i, q_j] = [p_i, p_j] = 0,$$
$$\frac{d\mathcal{O}}{dt} = i[H, \mathcal{O}], \tag{2.107}$$

where for simplicity we take observables $\mathcal{O}$ that do not depend explicitly on time. The equations of motion, i.e., Heisenberg's equation, can be solved:

$$\mathcal{O}(t) = e^{iHt}\mathcal{O}(0)e^{-iHt} = U(t)^\dagger \mathcal{O}(0) U(t), \tag{2.108}$$

where $U(t)$ is unitary. The states of the system are associated to vectors ψ (or rather rays $e^{i\alpha}\psi$) in a Hilbert space $\mathcal{H}$, and the expectation value of observables reads

$$\langle \mathcal{O} \rangle_\psi(t) = {}_H\langle \psi | \mathcal{O}_H(t) | \psi \rangle_H. \tag{2.109}$$

Here we have added a subscript H to signal that we are working in *Heisenberg's picture*: the state of the system is given and fixed in time, while observables evolve according to the Hamiltonian of the system. One can equivalently write Eq. (2.109) as

$$\langle \mathcal{O} \rangle_\psi(t) = {}_H\langle \psi | e^{iHt}\mathcal{O}_H(0)e^{-iHT} | \psi \rangle_H = {}_S\langle \psi(t) | \mathcal{O}_S | \psi(t) \rangle_S, \tag{2.110}$$

where $\mathcal{O}_S = \mathcal{O}_H(0)$ and $|\psi(0)\rangle_S = |\psi\rangle_H$. This provides the *Schrödinger's picture* of quantum mechanics, in which observables are fixed operators, while states evolve in time according to the Schrödinger equation,

$$i\frac{\partial}{\partial t}|\psi(t)\rangle_S = H|\psi(t)\rangle_S. \tag{2.111}$$

The two pictures are physically equivalent by construction. In Schrödinger's it is easier to express the rules of measurements: the probability of obtaining the value a_n for a measurement of $\mathcal{O}$ at time t is given by the absolute value squared of the transition amplitude to the corresponding eigenvector

$$P = |\langle a_n | \psi(t)\rangle|^2, \tag{2.112}$$

where $\mathcal{O}|a_n\rangle = a_n|a_n\rangle$. On the other hand, Heisenberg picture turns out to be more convenient when dealing with field operators.

A third picture is the *interaction* or *Dirac's picture*, in which obervables and states take different "parts" of the temporal evolution, with observables evolving freely and states (essentially) with the interactions. This picture presumes that the Hamiltonian can be split into a free and an interaction part, $H = H_0 + V$, and then sets

$$\langle \mathcal{O} \rangle_\psi(t) = {}_H\langle \psi | e^{iHt}e^{-iH_0 t}e^{iH_0 t}\mathcal{O}_H(0)e^{-iH_0 t}e^{iH_0 t}e^{-iHT} | \psi \rangle_H = {}_I\langle \psi(t) | \mathcal{O}_I(t) | \psi(t) \rangle_I, \tag{2.113}$$

with

$$\mathcal{O}_I(t) = e^{iH_0 t}\mathcal{O}_H(0)e^{-iH_0 t} = e^{iH_0 t}\mathcal{O}_I(0)e^{-iH_0 t},$$
$$|\psi(t)\rangle_I = e^{iH_0 t}e^{-iHt}|\psi\rangle_H = e^{iH_0 t}e^{-iHt}|\psi(0)\rangle_I. \qquad (2.114)$$

States and observables coincide in all pictures at $t = 0$. The key object here is the temporal evolution

$$\mathcal{U}(t_2, t_1) = e^{iH_0 t_2}e^{-iHt_2}e^{iHt_1}e^{-iH_0 t_1}, \qquad (2.115)$$

in terms of which $|\psi(t)\rangle_I = \mathcal{U}(t,0)|\psi(0)\rangle_I$. To obtain an explicit expression for it, we solve the differential equation

$$\frac{\partial}{\partial t_2}\mathcal{U}(t_2,t_1) = e^{iH_0 t_2}i(H_0 - H)e^{-iH_0 t_2}\mathcal{U}(t_2,t_1) = -ie^{iH_0 t_2}Ve^{-iH_0 t_2}\mathcal{U}(t_2,t_1) = -iV_I(t_2)\mathcal{U}(t_2,t_1), \qquad (2.116)$$

subject to the boundary condition $\mathcal{U}(t_1,t_1) = \mathbf{1}$. For $t_2 > t_1$ the solution is provided by Dyson's formula

$$\mathcal{U}(t_2,t_1) = \mathrm{Texp}\left\{-i\int_{t_1}^{t_2} dt\, V_I(t)\right\}$$
$$= \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int_{t_1}^{t_2} d\tau_1 \ldots \int_{t_1}^{t_2} d\tau_n\, T\{V_I(\tau_1)\ldots V_I(\tau_n)\}, \qquad (2.117)$$

where the *time-ordering symbol* T places the operators in descending order with respect to time starting from the left:

$$T\{A_1(t_1)A_2(t_2)\} = \theta(t_1 - t_2)A_1(t_1)A_2(t_2) + \theta(t_2 - t_1)A_2(t_2)A_1(t_1),$$
$$T\{A_1(t_1)\ldots A_n(t_n)\} = \sum_P \theta(t_{P(1)} - t_{P(2)})\ldots\theta(t_{P(n-1)} - t_{P(n)})A_{P(1)}(t_{P(1)})\ldots A_{P(n)}(t_{P(n)}), \qquad (2.118)$$

where the sum is over all the distinct permutations P of $\{1,\ldots,n\}$. For $t_1 > t_2$ we can use the relation $\mathcal{U}(t_2,t_1) = \mathcal{U}(t_1,t_2)^\dagger$ and obtain the solution from Eq. (2.118).

All the above is known stuff for systems with finitely many degrees of freedom. The practical realisation is obtained by representing the canonical commutation relations on some Hilbert space, and teh Stone-Von Neumann theorem tells us that this can be done in an essentially unique way: for a single variable, $\mathcal{H} = \mathscr{L}(\mathbb{R})$ and

$$(q\psi)(x) = x\psi(x)$$
$$(p\psi)(x) = -i\frac{d}{dx}\psi(x), \qquad (2.119)$$

provides the desired representation of $[q,p] = i$; all other representations are unitarily equivalent to this one.

We would like to follow a similar procedure with infinitely many degrees of freedom, replacing the canonical Poisson brackets with commutators, i.e., performing canonical quantisation of fields. Does this work? Yes, we know it does. Do we understand why? At this stage no, not at all. Also: for infinitely many degrees of freedom the representation of the canonical algebra is not unique, and we have to motivate physically which one we are going to use. For free fields we will see easily that it is the Fock representation to be used. For interacting fields we would first have to understand the dynamics, a task which is almost always impossible to tackle, and

so we will need to develop approximation techniques. Finally, for particles with half-integer spin commutation relations will not work, and we will have to use anticommutation relations instead.

Instead of quantising things that we do not understand properly, it is better from the logical point of view to start from what we do understand properly, namely the particles, and then proceed to develop quantum fields as a very convenient (Weinberg claims necessary) tool to develop what we need, which is a relativistic, local, unitary quantum theory.[10] A key point in this approach is invariance under symmetry transformations. We now extend the discussion on this point from the classical to the quantum case.

2.8   Symmetries in quantum mechanics

Consider again two observers $\mathcal{O}$ and $\mathcal{O}'$ dealing with the same physical system. In quantum mechanics, the results of their measurements, defining the state of the system, are used to assign a representative vector to the system in the corresponding Hilbert space. We work here in the same spirit of Schrödinger's picture for time evolution: the operators corresponding to the observables are defined once and for all for all observers, reflecting the fact that their operative rules for measurements are the same.[11] Since measurements in general differ for the two observers, they will assign different representative vectors to the system, in such a way as to obtain the correct expectation values of observables. For an observable A, for which $\mathcal{O}$ gets the expectation value a, she/he will choose a vector ψ such that $(\psi, A\psi) = a$, while $\mathcal{O}'$ who gets the expectation value a' will choose a vector ψ' such that $(\psi', A\psi') = a'$.[12] More precisely, they will assign $rays$[13] $\mathcal{R}$ and $\mathcal{R}'$ to the state of the system, as to reproduce correctly the expectation values. Establishing a relation between the descriptions of the two observers corresponds to providing a mapping $\mathcal{M}$ from the space of rays $\underline{\mathcal{H}}$ to itself, i.e., if $\mathcal{O}$ assigns the ray $\mathcal{R}$ to the state of the system then $\mathcal{O}'$ assigns the ray $\mathcal{R}' = \mathcal{M}\mathcal{R}$. As we discussed in the classical case, for equivalent observers such a mapping has to be invertible.

Suppose now that an experiment is performed on the system, and a transition from a state to another is observed. The two observers $\mathcal{O}$ and $\mathcal{O}'$ will see respectively the transitions

$$\mathcal{O} : \mathcal{R}_i \longrightarrow \mathcal{R}_f \,, \qquad \mathcal{O}' : \mathcal{R}'_i \longrightarrow \mathcal{R}'_f \,, \tag{2.120}$$

occurring with probabilities P and P',

$$P = (\mathcal{R}_i \cdot \mathcal{R}_f) \,, \qquad P' = (\mathcal{R}'_i \cdot \mathcal{R}'_f) \,, \tag{2.121}$$

where

$$\mathcal{R}_1 \cdot \mathcal{R}_2 = |(\psi_1, \psi_2)|^2 \,, \tag{2.122}$$

with $\psi_{1,2}$ any normalised vector belonging to $\mathcal{R}_{1,2}$. Since $\mathcal{O}$ and $\mathcal{O}'$ are looking at the same physical process, the transition probabilities they observe must be the same, $P = P'$, and so

$$\mathcal{R}_i \cdot \mathcal{R}_f = \mathcal{R}'_i \cdot \mathcal{R}'_f \,. \tag{2.123}$$

---

[10]In Weinberg's argument, a key role is played by the cluster decomposition principle, which basically says that the outcomes of far away experiments should be uncorrelated. Weinberg shows how using interactions built out of local fields one automatically satisfies the cluster decomposition principle.

[11]Time evolution can be seen as a change in the origin of time, and in Schrödinger's picture the observables do not change under it. What we do here is applying the same approach to all sorts of changes of reference frame.

[12]It is understood that vectors are normalised to 1.

[13]A ray is an equivalence class of vectors $\psi \in \mathcal{H}$ with respect to the equivalence relation $\{\psi \sim \phi \ \text{if} \ \psi = e^{i\zeta}\phi\}$.

Since $\mathcal{R}'_{i,f} = \mathcal{M}\mathcal{R}_{i,f}$, we have that

$$\mathcal{R}_i \cdot \mathcal{R}_f = (\mathcal{M}\mathcal{R}_i) \cdot (\mathcal{M}\mathcal{R}_f) . \tag{2.124}$$

A theorem due to Wigner guarantees that an invertible transformation $\mathcal{M}$ on the space of rays $\underline{\mathcal{H}}$ that conserves probabilities as in Eq. (2.124) can be implemented as a transformation M on the space of vectors $\mathcal{H}$ that is either linear and unitary or antilinear and antiunitary. Linear and antilinear operators are defined as follows:

$$\begin{aligned}
\text{linear}: \ & U(\alpha\psi + \beta\phi) = \alpha U\psi + \beta U\phi , \\
\text{antilinear}: \ & T(\alpha\psi + \beta\phi) = \alpha^* T\psi + \beta^* T\phi .
\end{aligned} \tag{2.125}$$

The *adjoint* of an operator is defined as follows in the two cases:

$$\begin{aligned}
\text{linear}: \ & (\psi, U\phi) = (U^\dagger\psi, \phi) , \\
\text{antilinear}: \ & (\psi, T\phi) = (T^\dagger\psi, \phi)^* .
\end{aligned} \tag{2.126}$$

Finally, a linear unitary and an antilinear antiunitary operator are operators that are onto $\mathcal{H}$ (i.e., their image is the whole space), and that satisfy

$$\begin{aligned}
\text{linear and unitary}: \ & (U\psi, U\phi) = (\psi, \phi) , \\
\text{antilinear and antiunitary}: \ & (T\psi, T\phi) = (\psi, \phi)^* .
\end{aligned} \tag{2.127}$$

In both cases, these requirements are equivalent to $U^\dagger U = UU^\dagger = \mathbf{1}$, $T^\dagger T = TT^\dagger = \mathbf{1}$. In practice, then, a symmetry transformation, relating two equivalent observers, is a unitary or an antiunitary mapping M of the Hilbert space onto itself. As we discussed previously, the symmetry transformations of a system form a group.

So far we have discussed only the kinematical aspect of symmetries. There is also the dynamical aspect to be taken into account, i.e., the evolved of the transformed must equal the transformed of the evolved for a transformation to be a symmetry. In quantum mechanics the temporal evolution is determined by the Hamiltonian H of the system. If H is not depending explicitly on time, then the temporal evolution of a system is obtained via the unitary transformation $U(t) = e^{-iHt}$. For equivalent observers this unitary operator must be the same, and for a transformation $M(t)$ (which we momentarily take to depend explicitly on time for generality) to be a symmetry the following relation must hold,

$$M(t)\psi(t) = M(t)U(t)\psi(0) = U(t)M(0)\psi(0) , \tag{2.128}$$

for all initial vectors $\psi(0)$, and therefore

$$M(t)U(t) = U(t)M(0) . \tag{2.129}$$

This equation can be recast as $U(t) = M(t)U(t)M(0)^\dagger$, expressing the invariance of the evolution operator. Eq. (2.129) can be recast also as $M(t) = U(t)M(0)U(t)^\dagger$, which means that the temporal dependence of the transformation $M(t)$ has to be entirely determined by the Hamiltonian of the system, with no room for some extra explicit dependence.

Let us now make our life simpler, and let us focus on time-independent symmetry transformations. From Eq. (2.129) we find $[U(t), M] = 0$ at all times, and so taking the derivative and setting $t = 0$ we find

$$[H, M] = 0 . \tag{2.130}$$

A time-independent transformation is then a symmetry if it commutes with the Hamiltonian. Symmetries are classified in two big groups, namely continuous and discrete: in the first case there is a continuous family of symmetry transformations depending on some real parameter, in the second case the transformation is "isolated". If $M = M(\alpha)$ is an element of a continuous family of symmetry transformations dependent on some parameter α and connected to the identity, $M(0) = \mathbf{1}$, then it has to be unitary. Indeed, since $(M(\alpha)\psi, M(\alpha)\phi)$ has to be equal to (ψ, ϕ) or $(\psi, \phi)^*$, it must be a constant and so equal to $(M(0)\psi, M(0)\phi) = (\psi, \phi)$. We can then write $M(\alpha) = e^{i\alpha Q}$ for some self-adjoint operator $Q = Q^\dagger$, which is therefore a good observable. If $M(\alpha)$ is a symmetry for all α, we can take the derivative with respect to α to find $[H, Q] = 0$. This means that Q is a conserved physical quantity, i.e., its expectation value on a state does not change with time. In Heisenberg's picture this reads $dQ(t)/dt = -i[H, Q(t)] = 0$. Moreover, since Q commutes with the Hamiltonian it can be diagonalised simultaneously with it. The construction above obviously does not apply to discrete symmetries, for which an analogous conserved quantity cannot be constructed. Examples of continuous transformations are translations and rotations, to which correspond respectively the conserved four-momentum and angular momentum. Discrete symmetries include most notably parity (P), charge conjugation (C), and time reversal (T).

In the classical case we have seen how we can find conserved charges using the Lagrangian formalism and exploiting Noether's theorem. The same kind of approach can be used here, although care is needed because of possible operator-ordering issues. Notice that since $\epsilon\{q, Q\}_{PB} = \delta q$, which upon canonical quantisation becomes $i\epsilon[Q, q] = \delta q$ (modulo ordering problems), these charges would indeed generate the transformation. A similar argument goes for quantum fields, as we will see in due time.

3 Particles as irreducible representations of the Poincaré group

Experimental evidence shows the existence of particles, small localised objects characterised by a few constant quantities, mainly mass and total spin, plus others like electric charge, labelling different types; and by their energy, momenta and spin components, or equivalently energy, total angular momentum and angular momentum components. In the absence of interactions, particles travel freely on straight lines determined by their momenta. What we will discuss here is the theoretical characterisation of particles as irreducible representations of the Poincaré group, from which the same features of the experimental characterisation (i.e., the same observable quantities) will emerge.

The basic principle of special relativity is that all observers related by a Lorentz transformation are equivalent. Moreover, one of our fundamental assumptions (supported by countless experimental results) is that physics is invariant with respect to temporal and spatial translations. These two facts are summarised geometrically by taking Minkowski space as the arena for our physical theories: in fact, the isometries[14] of Minkowski space are precisely translations and Lorentz transformations, and what we assume is that physics is invariant under this family of transformations.

The group of isometries of Minkowski space is the *Poincaré group*, and we postulate that

---

[14]The isometries of a metric space are those mapping of the space into itself that preserve distances. For spaces, like Minkowski, were we have instead a pseudo-Euclidean metric, it is the interval $(x - y)^\mu (x - y)^\nu \eta_{\mu\nu}$ that has to be preserved.

physics is invariant under its action. Due to Wigner's theorem, this implies that the Hilbert space of the states of a physical system provides the basis for a representation of the Poincaré group in terms of unitary or antiunitary operators. In particular, the component of the group connected to the identity must be represented unitarily. We then define a particle as a system providing an *irreducible* representation of the Poincaré group, i.e., the space of states has no subspace left invariant by all the transformations. This means that we cannot subdivide the space of states into subspaces which are not transformed into each other by some Poincaré transformation, and so we cannot obtain a simpler description by dealing with these two subsets of states separately.

A remark is in order here. While for systems with finitely many degrees of freedom it is always possible to realise a symmetry unitarily (or antiunitarily), for systems with infinitely many degrees of freedom (such as, e.g., statistical systems in the infinite volume limit, or systems of fields), there is also a distinct possibility. It may happen that a symmetry of the (algebra of) observables is in fact not a symmetry realised in the system, a fact usually (and sloppily) expressed by saying that the ground state is not invariant under the symmetry transformation. In this case one speaks of *spontaneous symmetry breaking*. In the cases we will be dealing with, it is assumed that symmetries are (anti)unitarily realised, or realised *à la Wigner-Weyl*.

The plan of this section is the following. We discuss first representation theory in some generality, and give some details for the familiar case of rotations in three dimensions. We then discuss the Poincaré group and its representation. Finally, we will construct the space describing any number of noninteracting particles (Fock space), introduce the creation and annihilation operators, and easily derive their transformation properties under the Poincaré group. This will be the starting point for the developments of the following section, where we discuss how to manage creation and annihilation of particles by means of local operators, i.e., fields, with simple transformation properties under symmetry transformations. We will then (finally!) show how all of this can be obtained via canonical quantisation starting from a Lagrangian. As discussed in the Introduction, this is quite the opposite of the historical approach.

3.1 Representation theory: a brief introduction

Given a group G and a vector space V over the real or complex numbers, a representation of G is a mapping from G to the space of invertible matrices $GL(V)$ that preserves the group multiplication. If V has dimension n, the representation is said to be n-dimensional; V is said to provide a basis for the representation. The mapping $D : G \to GL(V)$ must then satisfy

$$D(g_1)D(g_2) = D(g_1 g_2) \,, \tag{3.1}$$

for all elements $g_1, g_2 \in G$. Using Eq. (3.1) for $g_1 = e$, the identity element of G, we find

$$D(e)D(g) = D(g) \Longrightarrow D(e) = \mathbf{1}_n \,, \tag{3.2}$$

since $D(g)$ is invertible. Moreover,

$$D(g)D(g^{-1}) = D(e) = \mathbf{1}_n \Longrightarrow D(g^{-1}) = D(g)^{-1} \,. \tag{3.3}$$

An irreducible representation is one such that there is no proper subspace $W \subset V$, $W \neq \varnothing$ with $D(g)W = W$ for all g, i.e., no subspace is left invariant by all matrices $D(g)$.

The group we are interested in is a *Lie group*, i.e., a group which is also a smooth manifold, with multiplication and inversion being smooth maps on the manifold. Elements of a Lie group can then be parameterised locally in terms of parameters α_a, denoted collectively by α, whose number is the dimension of the manifold. In this case, representations are required to be smooth in the group elements. For Lie groups, representations can be (essentially) obtained from representations of the corresponding *Lie algebra*. Parameterising as $g(\alpha)$ the elements of the group in a neighbourhood of the identity, the tangent space of the Lie group at the identity element is generated as a vector space by the elements

$$\tilde{\tau}_a = \frac{\partial g(\alpha)}{\partial \alpha_a}\bigg|_{\alpha=0}, \tag{3.4}$$

which are called the *generators* of the group. The generic element A of the tangent space can be written as a linear combination with real coefficients of the generators, $A = \sum_a c_a \tilde{\tau}_a$. The Lie groups we will be dealing with are matrix groups, i.e., g are matrices, and so will be the $\tilde{\tau}_a$ and the generic A. It can be shown that the tangent space at the identity is closed under the commutator $[A, B] = AB - BA$. The Lie algebra of the Lie group is precisely this tangent space with the commutator operation. Being closed under commutators implies that

$$[\tilde{\tau}_a, \tilde{\tau}_b] = -f_{ab}{}^c \tilde{\tau}_c, \tag{3.5}$$

for some real constants $f_{ab}{}^c$, called the *structure constants* of the group. Here sum over repeated indices is understood. A representation d of the algebra is again a mapping of the algebra to some matrix space $GL(V)$, which this time preserves commutators,

$$[d(A), d(B)] = d([A, B]), \tag{3.6}$$

and it is irreducible if there is no subspace W for which $d(A)W = W$ for all A in the algebra. The correspondence between irreducible representations of the algebra and irreducible representations of the group is the following:[15]

$$D(e^A) = e^{d(A)}. \tag{3.7}$$

Physicists prefer to redefine the generators as $\tilde{\tau}_a = i\tau_a$, so that Eq. (3.5) is recast as

$$[\tau_a, \tau_b] = i f_{ab}{}^c \tau_c, \tag{3.8}$$

and Eq. (3.7) becomes

$$D(e^{iA}) = e^{id(A)}. \tag{3.9}$$

For connected compact groups, all finite-dimensional irreducible representations are equivalent to *unitary representations*, i.e., representations in which the matrices $D(g)$ are unitary, and so the τ_a are Hermitian. This is not true for non-compact groups: for the Lorentz group there are no finite-dimensional unitary representations, except for the *trivial representation* $D(g) = 1\ \forall g$. Finite-dimensional representations do exist, but they are necessarily non-unitary.

---

[15]This is strictly true only for simply connected Lie groups. For connected but not simply connected groups, this formula provides an irreducible representation of the universal cover of the group, which descends to a projective representation of the group, i.e. a representation in which Eq. (3.1) is relaxed to $D(g_1)D(g_2) = e^{i\phi(g_1,g_2)}D(g_1 g_2)$. Since adding an extra phase to all the state vectors of a physical system has no observable consequence, projective representations are perfectly good for us, and we will not care in making the distinction.

Let us discuss these concepts in the familiar case of the three-dimensional rotation group SO(3). The group O(3) is the group of transformations of $\mathbb{R}^3$ that keeps the Euclidean length of a vector unchanged, and is characterised by the relation

$$O^T O = \mathbf{1} \,, \tag{3.10}$$

where T denotes the transpose. This implies $\det O = \pm 1$. The subgroup with $\det O = 1$ forms the group of (proper) rotations SO(3), preserving the orientation of space. Since O is a real matrix, Eq. (3.10) implies that O (seen as a complex matrix) is unitary, and thus can be written as $O = e^A$ with antihermitian A. The condition $\det O = 1$ implies $\operatorname{tr} A = 0$. Reality of O then implies that A is antisymmetric real.[16] The most general such matrix can be written as

$$A_{ij} = \theta^a \epsilon_{aij} = i\theta^a \tau_a \,, \qquad (\tau_a)_{ij} = -i\epsilon_{aij} \,, \tag{3.11}$$

since there are precisely $\frac{3 \cdot 2}{2} = 3$ independent real entries in an antisymmetric real matrix. The τ_a are the generators of the rotation group, and the generic element of the group can be written as $O = e^{i\theta^a \tau_a}$. For an infinitesimal rotation $\vec{x} \to \vec{O}x$,

$$\vec{O}x_i = O_{ij}x_j = x_i + i\theta^a(\tau_a)_{ij}x_j = x_i + \theta^a \epsilon_{aij}x_j = (\vec{x} - \vec{\theta} \wedge \vec{x})_i \,. \tag{3.12}$$

In the physicist's convention, the Lie algebra $\mathfrak{so}(3)$ of the rotation group is the space of purely imaginary antisymmetric matrices, which is closed under ($-i$ times, in the physicist's convention) commutators. Indeed,

$$[\tau_a, \tau_b]^T = -[\tau_a^T, \tau_b^T] = -[\tau_a, \tau_b] \,, \quad [\tau_a, \tau_b]^* = [\tau_a^*, \tau_b^*] = [\tau_a, \tau_b] \,, \tag{3.13}$$

and therefore

$$[\tau_a, \tau_b] = i f_{ab}{}^c \tau_c \,. \tag{3.14}$$

An explicit computation shows that

$$\begin{aligned}
([\tau_a, \tau_b])_{ij} &= -(\epsilon_{aik}\epsilon_{bkj} - \epsilon_{bik}\epsilon_{akj}) = \delta_{ab}\delta_{ij} - \delta_{aj}\delta_{bi} - \delta_{ab}\delta_{ij} + \delta_{bj}\delta_{ai} \\
&= \delta_{ai}\delta_{bj} - \delta_{aj}\delta_{bi} = \epsilon_{abc}\epsilon_{cij} = i\epsilon_{abc}(\tau_c)_{ij} \,,
\end{aligned} \tag{3.15}$$

i.e., the commutation relations read

$$[\tau_a, \tau_b] = i\epsilon_{abc}\tau_c \,. \tag{3.16}$$

These are the same commutation relations of the algebra of $\mathfrak{su}(2)$.[17] The irreducible representations of this algebra are well known: they are obtained by representing the algebra on the linear spaces generated by vectors $|j, j_3\rangle$, with $2j$ being an integer and $j_3 = -j, -j+1, \ldots, j-1, j$, so that the dimension of the representation is $(2j + 1)$. Here $d(\tau_3)|j, j_3\rangle = j_3|j, j_3\rangle$, and $j(j + 1)$ is the eigenvalue of the quadratic Casimir operator $C_2 = \sum_a d(\tau_a)^2$: in other words, they are the eigenvectors of the third component J_3 of the angular momentum and of the total angular momentum $\vec{J}^2$. Casimir operators commute with the representatives of all the generators, thus with all the elements of an irreducible representation of a group, and by Schur's lemma they

---

[16]In general $O^* = O$, i.e., $e^{A^*} = e^A$, implies $A^* = A + i2\pi k$, or equivalently $(A + i\pi k)^* = A + i\pi k$, but the tracelessness condition fixes k to zero.

[17]The group SU(2) is the universal covering group of SO(3).

must be multiples of the identity in each irreducible representation. This means that Casimir operators can be used to label the various irreducible representations. The quadratic Casimir clearly commutes with all the generators:[18]

$$[C_2, d(\tau_b)] = \sum_a [d(\tau_a)^2, d(\tau_b)] = \sum_a \{d(\tau_a), [d(\tau_a), d(\tau_b)]\} = \sum_{ac} i\epsilon_{abc}\{d(\tau_a), d(\tau_c)\} = 0\,, \quad (3.17)$$

since ϵ_{abc} is antisymmetric in bc while the anticommutator is symmetric.

For future utility, we discuss also how representations appear in the transformation laws of fields. Our fields are operators in Hilbert spaces, and their transformed is defined formally as

$$\phi'(x) = U(R)^{-1}\phi(x)U(R)\,, \quad (3.18)$$

where $U(R)$ is the representation acting on the states of the system. The representations we are interested in are unitary, $U^{-1} = U^\dagger$, and fields will be the basic building blocks of observables. The law Eq. (3.18) is then motivated by the request that the new field describes on the old states the same physics that the old field describes on the new states:

$$\langle\psi_2'|\phi(x)|\psi_1'\rangle = \langle\psi_2|U^\dagger\phi(x)U|\psi_1\rangle = \langle\psi_2|\phi'(x)|\psi_1\rangle\,. \quad (3.19)$$

In other words, we can go from a Schrödinger to a Heisenberg picture of symmetry transformations in which we keep the states fixed and transform the fields. Different possibilities arise for the actual right-hand side member of Eq. (3.18):

$$\begin{aligned}
&\text{scalar: } \phi'(x) = \phi(R^{-1}x)\,, \\
&\text{vector: } \phi_i'(x) = R_{ij}\phi_j(R^{-1}x)\,, \\
&\text{spinor: } \phi_a'(x) = D^{(\frac{1}{2})}(R)_{ab}\phi_b(R^{-1}x)\,, \\
&\text{general: } \phi_a'(x) = D^{(j)}(R)_{ab}\phi_b(R^{-1}x)\,,
\end{aligned} \quad (3.20)$$

where $D^{(\frac{1}{2})}(e^{i\theta^a\tau_a}) = e^{i\theta^a\frac{\sigma_a}{2}}$ with σ_a the Pauli matrices, and $D^{(j)}(R)$ stands for a general, spin-j representation.

To see why scalar and vectors transform the way they do, one can proceed this way. A scalar field is a field that is not modified by rotations, besides the fact that the coordinate system has been rotated. This means that the fields ϕ and ϕ' at the physical point in space identified by x in the old coordinates and by $x' = Rx$ in the new ones have to coincide:

$$\phi'(x') = \phi(x) \Rightarrow \phi'(Rx) = \phi(x) \Rightarrow \phi'(x) = \phi(R^{-1}x)\,. \quad (3.21)$$

Similarly, a vector field has an orientation in space, whose description in a given coordinate system has to change under rotations of the coordinates in order to keep the physical orientation unchanged. Therefore, the scalar product of a vector field $\phi(x)$ with some coordinate vector y has to remain invariant under changes of coordinates, so that the physical, intrinsic orientation of the field at some physical point does not change. In formulas,

$$\begin{aligned}
&y_i'\phi_i'(x') = y_i\phi_i(x) \Rightarrow R_{ji}y_i\phi_j'(Rx) = y_i\phi_i(x) \Rightarrow \\
&y_i R_{ij}^{-1}\phi_j'(Rx) = y_i\phi_i(x) \Rightarrow \phi_i'(x) = R_{ij}\phi_j(R^{-1}x)\,.
\end{aligned} \quad (3.22)$$

---

[18]This proof immediately extends to compact semisimple algebras, i.e., compact algebras for which there is no invariant Abelian subalgebra (i.e., no subalgebra of commuting elements which also commutes with any other element of the algebra), since for these algebras the structure constants are totally antisymmetric under permutations of the indices: $f_{ab}{}^c = -f_{ba}{}^c = -f_{ac}{}^b$.

The fact that $D^{(j)}$ has to be a representation of the group in the general case can be seen composing two transformations, as follows. Since U is a representation,

$$U^{-1}(R_2)U^{-1}(R_1)\phi_a(x)U(R_1)U(R_2) = U^{-1}(R_1R_2)\phi_a(x)U(R_1R_2)\,. \tag{3.23}$$

But then

$$
\begin{aligned}
U^{-1}(R_1R_2)\phi_a(x)U(R_1R_2) &= D^{(j)}(R_1R_2)_{ab}\phi_b((R_1R_2)^{-1}x)\\
&= U^{-1}(R_2)U^{-1}(R_1)\phi_a(x)U(R_1)U(R_2) = U^{-1}(R_2)D^{(j)}(R_1)_{ab}\phi_b(R_1^{-1}x)U(R_2)\\
&= D^{(j)}(R_1)_{ab}D^{(j)}(R_2)_{bc}\phi_c(R_2^{-1}R_1^{-1}x)\,,
\end{aligned} \tag{3.24}
$$

i.e., $D^{(j)}(R_1R_2) = D^{(j)}(R_1)D^{(j)}(R_2)$.

The same formalism for the transformations of fields will apply also in the case of the Poincaré group. Notice however that in that case, since the matrices D provide a finite-dimensional representation of the group, they cannot be unitary matrices.

3.2 The Lorentz group

We now discuss in some detail the structure of the Lorentz group. Recall from Section 2.2 that Lorentz transformations are defined as the linear transformations $x'^\mu = \Lambda^\mu{}_\nu x^\nu$ that satisfy

$$\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \eta_{\alpha\beta} = \eta_{\mu\nu}\,, \tag{3.25}$$

or in matrix notation, with $(\Lambda)_{\mu\nu} = \Lambda^\mu{}_\nu$ and $(\eta)_{\mu\nu} = \eta_{\mu\nu}$,

$$\Lambda^T \eta \Lambda = \eta\,. \tag{3.26}$$

We have seen how these relations entail that $\det \Lambda = \pm 1$ and $\Lambda^0{}_0$ is either greater than or equal to 1, or less than or equal to -1, so that the group can be divided in four components. The component containing the identity ($\det \Lambda = 1$, $\Lambda^0{}_0 \geq 1$) forms the proper orthocronous Lorentz group $SO^\uparrow(3,1)$.

Let us now parameterise a general transformation Λ as follows,

$$\Lambda = \left(\begin{array}{c|c} \alpha & b^T \\ \hline c & D \end{array}\right)\,, \tag{3.27}$$

where $\alpha \in \mathbb{R}$ is a real number, $b, c \in \mathbb{R}^3$ are three-dimensional vectors and $D \in {}^3\mathbb{R}^3$ is a 3×3 real matrix. Eq. (3.26) reads then

$$\left(\begin{array}{c|c} \alpha & c^T \\ \hline b & D^T \end{array}\right)\left(\begin{array}{c|c} \alpha & b^T \\ \hline -c & -D \end{array}\right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & -\mathbf{1}_3 \end{array}\right)\,. \tag{3.28}$$

This is equivalent to the following three equations:

$$
\begin{aligned}
\alpha^2 - c^T c &= 1\,,\\
\alpha b - D^T c &= 0\,,\\
bb^T - D^T D &= -\mathbf{1}_3\,.
\end{aligned} \tag{3.29}
$$

33

Let us start from the last equation. Setting $b^T b = B^2$ with positive B, and $\hat{b} = b/B$, it can be recast as

$$D^T D = \mathbf{1}_3 + bb^T = \mathbf{1}_3 + B^2 \hat{b}\hat{b}^T = (1 + B^2)\hat{b}\hat{b}^T + (\mathbf{1}_3 - \hat{b}\hat{b}^T). \tag{3.30}$$

The quantities $\Pi_{\hat{b}} = \hat{b}\hat{b}^T$ and $\Pi_{\hat{b}_\perp} = \mathbf{1}_3 - \hat{b}\hat{b}^T$ are orthogonal projectors on $\hat{b}$ and on its orthogonal complement, respectively, i.e., $\Pi_{\hat{b}} = \Pi_{\hat{b}}^T = \Pi_{\hat{b}}^2$, $\Pi_{\hat{b}_\perp} = \Pi_{\hat{b}_\perp}^T = \Pi_{\hat{b}_\perp}^2$, and $\Pi_{\hat{b}}\Pi_{\hat{b}_\perp} = \Pi_{\hat{b}_\perp}\Pi_{\hat{b}} = 0$. It then follows that

$$D^T D = (\sqrt{1 + B^2}\Pi_{\hat{b}} + \Pi_{\hat{b}_\perp})^2 = S^2, \tag{3.31}$$

with $S = \sqrt{1 + B^2}\Pi_{\hat{b}} + \Pi_{\hat{b}_\perp}$ a symmetric and positive-definite matrix. Since $\det S = \sqrt{1 + B^2}$, we have $\det D \neq 0$, and so that both S and D are invertible matrices. Setting $O = DS^{-1}$, we see that $O^T O = S^{-1T}D^T D S^{-1} = S^{-1}S^2 S^{-1} = \mathbf{1}_3$, i.e., $O \in \mathrm{O}(3)$, and so we can write $D = OS$.[19] This concludes the treatment of the third equation. We now plug this into the second one to obtain

$$\alpha b = D^T c = SO^T c \Longrightarrow c = \alpha OS^{-1}b = \frac{\alpha B}{\sqrt{1 + B^2}} O\hat{b}. \tag{3.32}$$

Finally, we plug this into the first equation to obtain

$$1 = \alpha^2 - c^T c = \alpha^2 \left(1 - \frac{B^2}{1 + B^2}\right) = \alpha^2 \frac{1}{1 + B^2} \Longrightarrow \alpha = \pm\sqrt{1 + B^2}, \tag{3.33}$$

and using this in Eq. (3.32) we get

$$c = \pm BO\hat{b}. \tag{3.34}$$

We have now completely characterised the Lorentz transformations Λ. Let us clean up the calculation a bit to make the result more transparent. Set $B = \sinh\Theta$, from which it follows $\alpha = (-1)^{k_T}\cosh\Theta$, with $k_T = 0, 1$ taking into account the two possible signs of α, and let $\hat{b} = (-1)^{k_T}n$. Finally, set $O = (-1)^{k_P}O_0$ with $k_P = 0, 1$ and $O_0 \in \mathrm{SO}(3)$, so that $\det O = (-1)^{k_P}$. These are all redefinitions which do not change the nature of the quantities involved. We can now write

$$\begin{aligned}
\Lambda &= \left(\begin{array}{c|c} (-1)^{k_T}\cosh\Theta & (-1)^{k_T}\sinh\Theta n^T \\ \hline (-1)^{k_P}O_0\sinh\Theta n & (-1)^{k_P}O_0[(\mathbf{1}_3 - nn^T) + \cosh\Theta nn^T] \end{array}\right) \\
&= \left(\begin{array}{c|c} (-1)^{k_T} & 0 \\ \hline 0 & (-1)^{k_P}\mathbf{1}_3 \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & O_0 \end{array}\right) \left(\begin{array}{c|c} \cosh\Theta & \sinh\Theta n^T \\ \hline \sinh\Theta n & (\mathbf{1}_3 - nn^T) + \cosh\Theta nn^T \end{array}\right).
\end{aligned} \tag{3.35}$$

Setting $T = \mathrm{diag}(-1, 1, 1, 1)$ and $P = \mathrm{diag}(1, -1, -1, -1)$, we have that the most general Lorentz transformation can be written as the product of discrete transformation, times a rotation, times a boost in direction n:

$$\Lambda = T^{k_T}P^{k_P}R(\theta, r)L(\Theta, n) \tag{3.36}$$

where

$$R(\theta, r) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & O_0(\theta, r) \end{array}\right), \tag{3.37}$$

with $O_0(\theta, r)$ a three-dimensional rotation around axis r of angle θ, and

$$L(\Theta, n) = \left(\begin{array}{c|c} \cosh\Theta & \sinh\Theta n^T \\ \hline \sinh\Theta n & (\mathbf{1}_3 - nn^T) + \cosh\Theta nn^T \end{array}\right), \tag{3.38}$$

---

[19]This is an example of the polar decomposition of an invertible matrix into the product of an orthogonal and a symmetric positive-definite matrix.

is a boost of *rapidity* Θ in the spatial direction direction n. The last expression looks probably more familiar if we set $\sinh\Theta = \gamma\beta$, $\cosh\Theta = \gamma$ with $\gamma = 1/\sqrt{1-\beta^2}$, and choose coordinates so that $n = (1,0,0)$. In this case

$$L = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.39}$$

is the transformation from the frame in which we are at rest to a second frame moving at speed β in the negative x direction, or equivalently, the transformation from the frame moving at speed β in the positive x direction to the frame in which we are at rest.

From Eq. (3.36) we see that k_T and k_P determine whether the transforamtion is proper or not and orthocronous or not: $k_T = 0$ implies orthocronous, $k_T + k_P \mod 2 = 0$ implies proper. The remaining two factors give transformations in the proper orthocronous group, and it is clear that they form a continuous group. Its parameters are the two angles determining the direction of n in three-dimensional space and the magnitude of Θ, or alternatively the three components of the vector Θn, and the three parameters of the proper rotation group. The proper orthocronous Lorentz group $\mathrm{SO}^{\uparrow}(3,1)$ is thus a six-dimensional Lie group. This group is furthermore connected, as any of its elements can be reached from the identity moving along a continuous path in parameter space. However, Θ is any real positive number, which means that the group is not compact.[20]

3.3   The algebra of the Lorentz group

Now that we have worked out the global structure of the Lorentz group, let us look at its local structure, or in other words at the Lie algebra corresponding to its component connected to the identity, i.e., the proper orthocronous subgroup. To this end, let us consider an infinitesimal Lorentz transformation,

$$\Lambda = \mathbf{1} + \Omega\,, \tag{3.40}$$

where the matrix Ω is small, and again $(\Omega)_{\mu\nu} = \Omega^{\mu}{}_{\nu}$. Plugging this into the defining relation for Lorentz transformations we find to first order in Ω

$$\eta = \Lambda^T \eta \Lambda = \eta + \Omega^T \eta + \eta\Omega \implies \Omega^T \eta + \eta\Omega = 0\,, \tag{3.41}$$

or putting back indices

$$\begin{aligned} (\Omega)_{\rho\mu}(\eta)_{\rho\nu} + (\eta)_{\mu\rho}(\Omega)_{\rho\nu} &= 0\,, \\ \Omega^{\rho}{}_{\mu}\eta_{\rho\nu} + \eta_{\mu\rho}\Omega^{\rho}{}_{\nu} &= 0\,, \\ \Omega_{\nu\mu} + \Omega_{\mu\nu} &= 0\,, \end{aligned} \tag{3.42}$$

i.e., the tensor $\Omega_{\mu\nu}$ is antisymmetric. Such a tensor has six independent entries (i.e., $\frac{4\cdot3}{2} = 6$), which precisely correspond to the six parameters of the group. The most general infinitesimal transformation can then be written as

$$\Omega_{\mu\nu} = \frac{1}{2}\omega_{\rho\sigma}\left(\delta^{\rho}{}_{\mu}\delta^{\sigma}{}_{\nu} - \delta^{\rho}{}_{\nu}\delta^{\sigma}{}_{\mu}\right) = \frac{1}{2}\omega_{\rho\sigma}M^{(\rho\sigma)}_{\mu\nu}\,, \tag{3.43}$$

---

[20]Moreover, although connected, this group is not simply connected. From a topological point of view, it is the direct product $\mathbb{R}^3 \times \mathrm{SO}(3) = \mathbb{R}^3 \times (S^2/\mathbb{Z}_2)$, from which both non-compactness and non-simple-connectedness follow.

where $M^{(\rho\sigma)}_{\mu\nu}$ are antisymmetric both in μ, ν and ρ, σ. To make contact with the matrix notation used above, we need

$$(M^{(\rho\sigma)})_{\mu\nu} = M^{(\rho\sigma)\,\mu}{}_{\nu} = \eta^{\rho\mu}\delta^{\sigma}{}_{\nu} - \eta^{\sigma\mu}\delta^{\rho}{}_{\nu}\,. \tag{3.44}$$

The matrices $M^{(\rho\sigma)}$ are the generators of the Lorentz group for mathematicians. Physicists prefer to use a different convention by introducing a factor i,

$$(J^{(\rho\sigma)})_{\mu\nu} = J^{(\rho\sigma)\,\mu}{}_{\nu} = -i(M^{(\rho\sigma)})_{\mu\nu}\,. \tag{3.45}$$

These matrices are not all Hermitian, as we will see below. To complete the study of the group algebra we need the commutation relations of the generators, which can be worked explicity quite easily by direct computation, and read

$$[J^{(\mu\nu)}, J^{(\rho\sigma)}] = i\left(\eta^{\mu\rho}J^{(\nu\sigma)} + \eta^{\nu\sigma}J^{(\mu\rho)} - \eta^{\mu\sigma}J^{(\nu\rho)} - \eta^{\nu\rho}J^{(\mu\sigma)}\right)\,. \tag{3.46}$$

To make them more transparent, set

$$\begin{aligned} K^i &= J^{(0i)}\,, \\ J^i &= -\frac{1}{2}\epsilon^{0ijk}J^{(jk)}\,, \end{aligned} \tag{3.47}$$

where Latin indices run on $i, j, k = 1, 2, 3$, and $\epsilon^{0123} = 1$. Notice the inverse relation

$$J^{(ij)} = -\epsilon_{ijk}J^k = -\epsilon^{0ijk}J^k\,, \tag{3.48}$$

where ϵ_{ijk} is the three-dimensional Levi-Civita tensor. One can easily be convinced that K^i and J^i correspond respectively to boosts and three-dimensional rotations. Indeed,

$$(K^i)_{\mu\nu} = -i\left(\begin{array}{c|c} 0 & \delta_{\nu i} \\ \hline \delta_{\mu i} & 0 \end{array}\right) = -i(\mathbf{1}_{\mu i}\mathbf{1}_{\nu 0} + \mathbf{1}_{\mu 0}\mathbf{1}_{\nu i})\,, \tag{3.49}$$

and since $(iK^i)^2 = \mathbf{1}$, a simple calculation shows that (no sum over repeated i)

$$(e^{i\alpha K^i})_{\mu\nu} = \mathbf{1}_{\mu\nu} + (\cosh\alpha - 1)\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \delta_{\mu i}\delta_{\nu i} \end{array}\right) + \sinh\alpha\left(\begin{array}{c|c} 0 & \delta_{\nu i} \\ \hline \delta_{\mu i} & 0 \end{array}\right)\,. \tag{3.50}$$

These generators are anti-Hermitian. On the other hand,

$$(J^i)_{jk} = -i\epsilon^{0ijk} = -i\epsilon_{ijk}\,, \tag{3.51}$$

all other entries being zeros, which are the Hermitian generators of rotations discussed in Section 3.1. In terms of K^i and J^i, the commutation relations read

$$\begin{aligned} [J^i, J^j] &= i\epsilon_{ijk}J^k\,, \\ [J^i, K^j] &= i\epsilon_{ijk}K^k\,, \\ [K^i, K^j] &= -i\epsilon_{ijk}J^k\,, \end{aligned} \tag{3.52}$$

as one can obtain by direct calculation. These tell us that $\vec{J}$ generate rotations, that the boosts $\vec{K}$ transform like vectors under rotations, and that the commutator of two infinitesimal boosts is

a rotation. These commutation relations can be further simplified if we consider the *complexified* Lie algebra of the group. As we discussed above in Section 3.1, the algebra of a Lie group is a real vector space, meaning that only linear combinations of the generators with real coefficients are allowed. If we allow also complex coefficients we deal instead with the complexification of the algebra. This is actually not new: think of representations of SU(2) and the use of raising and lowering operators. Complexified algebras are of practical utility since from a representation of the complexified algebra we can get a representation of the algebra itself However, one should keep in mind the conceptual distinction between the two. With this caveat in mind, define

$$\vec{J}_\pm = \frac{1}{2}(\vec{J} \pm i\vec{K})\,. \tag{3.53}$$

A straighforward calculation shows that

$$[J_\pm^i, J_\pm^j] = i\epsilon_{ijk}J_\pm^k\,, \qquad [J_\pm^i, J_\mp^j] = 0\,. \tag{3.54}$$

The complexified algebra $\mathfrak{so}_\mathbb{C}(3,1)$ of $\mathrm{SO}^\uparrow(3,1)$ is then isomorphic to the complexified algebra $\mathfrak{su}_\mathbb{C}(2) \oplus \mathfrak{su}_\mathbb{C}(2)$ of $\mathrm{SU}(2) \otimes \mathrm{SU}(2)$. This allows to classify the finite-dimensional irreducible representations of $\mathfrak{so}(3,1)$ and thus those of $\mathrm{SO}^\uparrow(3,1)$ using our knowledge of SU(2): finite-dimensional representations are labelled by two half-integers (j_1, j_2), and are $(2j_1 + 1)(2j_2 + 1)$ dimensional. The representations of lowest dimension are the following:

$$
\begin{array}{lll}
(j_1, j_2) & \text{dimensionality} & \\
(0,0) & 1 & \text{scalar} \\
(\frac{1}{2}, 0) & 2 & \text{left-handed Weyl spinor} \\
(0, \frac{1}{2}) & 2 & \text{right-handed Weyl spinor} \\
(\frac{1}{2}, 0) \oplus (0, \frac{1}{2}) & 4 & \text{Dirac spinor} \\
(\frac{1}{2}, \frac{1}{2}) & 4 & \text{vector}
\end{array}
\tag{3.55}
$$

Notice that Dirac spinors do not provide an irreducible representation of the proper orthocronous Lorentz group. However, since $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ are transformed into each other by parity, it provides an irreducible representation of the full group. Even though we have used the representation theory of SU(2) to derive this, we should not confuse the Lorentz group with $\mathrm{SU}(2) \otimes \mathrm{SU}(2)$. The algebras $\mathfrak{so}(3,1)$ and $\mathfrak{su}(2) \otimes \mathfrak{su}(2)$ are *not* isomorphic [while, parenthetically $\mathfrak{su}(2) \otimes \mathfrak{su}(2) \approx \mathfrak{so}(4)$], and in fact while $\mathfrak{su}(2) \otimes \mathfrak{su}(2)$ is compact, $\mathfrak{so}(3,1)$ is not.

3.4 The Poincaré group and its algebra

As we have said above, the Poincaré group consists of the isometries of Minkowski space, and includes Lorentz transformations and translations. As such, it is a Lie group with ten parameters. The most general transformation reads

$$M(\Lambda, a)x = \Lambda x + a\,, \tag{3.56}$$

where Λ is a Lorentz transformation and a a four-vector. This can be represented in matrix form using five-dimensional vectors, i.e.,

$$\mathcal{M}(\Lambda, a) \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} \Lambda x + a \\ 1 \end{pmatrix} = \left(\begin{array}{c|c} \Lambda & a \\ \hline 0 & 1 \end{array} \right) \begin{pmatrix} x \\ 1 \end{pmatrix}\,, \tag{3.57}$$

where $\mathcal{M} = \mathcal{M}_{AB}$ has indices running on $A, B = 0, 1, 2, 3, 4$. It is easy to work out the composition law of the group:

$$M(\Lambda_2, a_2)M(\Lambda_1, a_1) = M(\Lambda_2\Lambda_1, a_2 + \Lambda_2 a_1). \tag{3.58}$$

The identity element is clearly $M(\mathbf{1}, 0)$, while the inverse of an element is given by

$$M(\Lambda, a)^{-1} = M(\Lambda^{-1}, -\Lambda^{-1}a). \tag{3.59}$$

Infinitesimal Lorentz transformations are generated by the matrices $J^{(\rho\sigma)}$ discussed in the previous subsection,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + i\frac{1}{2}\omega_{\rho\sigma}J^{(\rho\sigma)\mu}{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu. \tag{3.60}$$

In the matrix representation of Eq. (3.57) these matrices have to be embedded in the top-left corner of $\mathcal{M}$, so the generators read

$$\mathcal{J}^{(\rho\sigma)}_{\mu\nu} = J^{(\rho\sigma)\mu}{}_\nu, \quad \mathcal{J}^{(\rho\sigma)}_{A4} = \mathcal{J}^{(\rho\sigma)}_{4B} = 0. \tag{3.61}$$

In the same representation, infinitesimal translations are generated instead by the four matrices $\mathcal{P}^\mu$,

$$(\mathcal{P}^\mu)_{AB} = \delta_{A\mu}\delta_{B4}. \tag{3.62}$$

Alternatively, we can think of translations as being generated by the derivative operator $P_\mu = i\partial_\mu$,

$$e^{-ia_\mu P^\mu}x^\nu = x^\nu + a^\nu. \tag{3.63}$$

This is however more appropriate when representing translations on a space of functions.

The representation discussed above is the defining representation of the group. We are going now to work out the commutation relations of the generators. This of course works in any representation – by the very definition of representation. With some abuse of notation, we will denote with $J^{(\rho\sigma)}$ and P^μ the generators of the Poincaré group in any representation. We begin by working out the effect of a finite transformation on the generators. For a generic representation $U(\Lambda, a)$ of the group element $M(\Lambda, a)$ we write

$$U(\Lambda, a) = e^{-ia_\mu P^\mu}e^{\frac{i}{2}\omega_{\rho\sigma}J^{(\rho\sigma)}} \tag{3.64}$$

for a Lorentz transformation $\Lambda = e^{\frac{i}{2}\omega_{\rho\sigma}J^{(\rho\sigma)}} = e^{i\omega}$, see Eq. (3.60), followed by a translation. Using only the group composition law, we find[21]

$$U(\Lambda, a)^{-1}U(\Lambda_0, a_0)U(\Lambda, a) = U(\Lambda^{-1}\Lambda_0\Lambda, \Lambda^{-1}(\Lambda_0 - \mathbf{1})a + \Lambda^{-1}a_0). \tag{3.65}$$

Let us now $\Lambda_0 = 1 + \omega$ be an infinitesimal transformation, and let $a_0 = \varepsilon$ be infinitesimal as well. Expanding Eq. (3.65) we find

$$U(\Lambda, a)^{-1}(\mathbf{1} + \frac{i}{2}\omega_{\rho\sigma}J^{(\rho\sigma)} - i\varepsilon_\mu P^\mu)U(\Lambda, a) = U(\mathbf{1} + \Lambda^{-1}\omega\Lambda, \Lambda^{-1}\omega a + \Lambda^{-1}\varepsilon)$$

$$\mathbf{1} + U(\Lambda, a)^{-1}(\frac{i}{2}\omega_{\rho\sigma}J^{(\rho\sigma)} - i\varepsilon_\mu P^\mu)U(\Lambda, a) = \mathbf{1} + \frac{i}{2}[\Lambda^{-1}\omega\Lambda]_{\rho\sigma}J^{(\rho\sigma)} - i[\Lambda^{-1}\omega a + \Lambda^{-1}\varepsilon]_\mu P^\mu. \tag{3.66}$$

---

[21]Notice the use of the inverse in this equation: this means that we are not (yet) asking for U to be unitary.

Here square brackets $[\ldots]_{\mu\nu}$ are used to indicate that the object in question is a fully covariant tensor, so the relation with the matrix notation used previously $(A)_{\mu\nu} = A^\mu{}_\nu$ is $[A]_{\mu\nu} = (\eta)_{\mu\alpha}(A)_{\alpha\nu} = A_{\mu\nu}$. A similar notation is used for covariant vectors. Matching terms of the same order in ω and ϵ we then find

$$
\begin{aligned}
U(\Lambda, a)^{-1}\omega_{\rho\sigma}J^{(\rho\sigma)}U(\Lambda, a) &= [\Lambda^{-1}\omega\Lambda]_{\rho\sigma}J^{(\rho\sigma)} - 2[\Lambda^{-1}\omega a]_\mu P^\mu \\
&= (\eta)_{\rho\gamma}(\Lambda^{-1})_{\gamma\alpha}(\omega)_{\alpha\beta}(\Lambda)_{\beta\sigma}J^{(\rho\sigma)} - 2(\eta)_{\mu\gamma}(\Lambda^{-1})_{\gamma\alpha}(\omega)_{\alpha\beta}a^\beta P^\mu \\
&= \Lambda_{\alpha\rho}\omega^\alpha{}_\beta\Lambda^\beta{}_\sigma J^{(\rho\sigma)} - 2\Lambda_{\alpha\mu}\omega^\alpha{}_\beta a^\beta P^\mu \\
&= \omega_{\rho\sigma}(\Lambda^\rho{}_\alpha\Lambda^\sigma{}_\beta J^{(\alpha\beta)} - \Lambda^\rho{}_\mu a^\sigma P^\mu + \Lambda^\sigma{}_\mu a^\rho P^\mu) \\
U(\Lambda, a)^{-1}\varepsilon_\mu P^\mu U(\Lambda, a) &= [\Lambda^{-1}\varepsilon]_\mu P^\mu = (\eta)_{\mu\gamma}(\Lambda^{-1})_{\gamma\alpha}\varepsilon^\alpha P^\mu = \epsilon^\alpha\Lambda_{\alpha\mu}P^\mu = \epsilon_\mu\Lambda^\mu{}_\alpha P^\alpha \,.
\end{aligned}
$$
(3.67)

From this we readily obtain the transformation laws of the generators:

$$
\begin{aligned}
U(\Lambda, a)^{-1}J^{(\rho\sigma)}U(\Lambda, a) &= \Lambda^\rho{}_\alpha\Lambda^\sigma{}_\beta J^{(\alpha\beta)} - \Lambda^\rho{}_\mu a^\sigma P^\mu + \Lambda^\sigma{}_\mu a^\rho P^\mu \,, \\
U(\Lambda, a)^{-1}P^\mu U(\Lambda, a) &= \Lambda^\mu{}_\alpha P^\alpha \,.
\end{aligned}
$$
(3.68)

This shows that under Lorentz transformation $J^{(\rho\sigma)}$ transforms as a doubly contravariant tensor, while P^μ as a contravariant vector. Under translations P^μ is unchanged, which means that the P^μ commute, while $J^{(\rho\sigma)}$ acquires a term linear in a, analogous to the change in the three-dimensional angular momentum under change of the reference point.

If we now make also Λ and a infinitesimal we can read off the commutation relations of the generators. Let again $\Lambda = 1 + \omega$ and $a = \varepsilon$. We find

$$
\begin{aligned}
U(1 + \omega, \varepsilon)^{-1}J^{(\rho\sigma)}U(1 + \omega, \varepsilon) &= J^{(\rho\sigma)} - \frac{i}{2}\omega_{\mu\nu}[J^{(\mu\nu)}, J^{(\rho\sigma)}] + i\varepsilon_\mu[P^\mu, J^{(\rho\sigma)}] \\
&= J^{(\rho\sigma)} + (\omega^\rho{}_\alpha J^{(\alpha\sigma)} + \omega^\sigma{}_\beta J^{(\rho\beta)}) - \varepsilon^\sigma P^\rho + \varepsilon^\rho P^\sigma \\
&= J^{(\rho\sigma)} + \frac{1}{2}\omega_{\mu\nu}(\eta^{\mu\rho}J^{(\nu\sigma)} + \eta^{(\nu\sigma)}J^{\mu\rho} - \eta^{\nu\rho}J^{(\mu\sigma)} - \eta^{(\mu\sigma)}J^{\nu\rho}) - \varepsilon_\mu(\eta^{\mu\sigma}P^\rho - \eta^{\mu\rho}P^\sigma) \,,
\end{aligned}
$$
(3.69)

from which we conclude that the commutation relations are

$$
\begin{aligned}
[J^{(\mu\nu)}, J^{(\rho\sigma)}] &= i(\eta^{\mu\rho}J^{(\nu\sigma)} + \eta^{\nu\sigma}J^{\mu\rho} - \eta^{\nu\rho}J^{(\mu\sigma)} - \eta^{\mu\sigma}J^{(\nu\rho)}) \,, \\
[J^{(\rho\sigma)}, P^\mu] &= i(\eta^{\rho\mu}P^\sigma - \eta^{\sigma\mu}P^\rho) \,, \\
[P^\mu, P^\nu] &= 0 \,.
\end{aligned}
$$
(3.70)

Incidentally, these commutation relations show us how a rank-2 tensor and a vector should transform under infinitesimal Lorentz transformation: it suffices to replace $J^{(\rho\sigma)}$ in the first line with $T^{\rho\sigma}$, and P^μ in the second line with V^μ.

3.5   Unitary representations of the Poincaré group

We now build unitary representations of the Poincaré group, or more precisely of its component connected to the identity. Since this is a non-compact group, it has no finite-dimensional unitary representation that are faithful, i.e., that map different elements of the group into different representatives. This is true in particular for the proper orthocronous Lorentz subgroup, which we want to represent faithfully and unitarily. We have therefore to look for infinite-dimensional representations.

Irreducible representation of a group are labelled by the values of its Casimir invariants, i.e., operators that commute with all the generators and therefore are multiple of the identity in an irreducible representation due to Schur's lemma. One such operator is easily built for the Poincaré group, and it is the mass operator $P^2 = P_\mu P^\mu$: it obviously commutes with translations, and it commmutes with the Lorentz generators since it is a scalar. Its eigenvalue will be denoted with m^2. In general this can be any real number, but we will be interested only in representations where $m^2 \geq 0$.[22] Since the translation generators all commute with each other, they can be diagonalised together. We can then take the basis of a representation to be of the form

$$|\vec{p}, \sigma\rangle, \qquad P^2|\vec{p}, \sigma\rangle = m^2|\vec{p}, \sigma\rangle, \quad P^\mu|\vec{p}, \sigma\rangle = p^\mu|\vec{p}, \sigma\rangle, \tag{3.71}$$

where $p^0 = \pm\sqrt{\vec{p}^2 + m^2}$, and σ labels whatever other degrees of freedom we might need, which we assume to correspond to some finite-dimensional space. The sign of p^0 is left invariant by proper orthochronous Lorentz transformations, so it is another invariant of the representation. We choose the plus sign for physical reason; a completely analogous analysis can be carried out for the choice of negative sign. Imposing that the $\vec{p}$ are real we have that P^μ is Hermitian, and we obtain a unitary representation of translations:[23]

$$e^{-ia_\mu P^\mu}|\vec{p}, \sigma\rangle = e^{-ia_\mu p^\mu}|\vec{p}, \sigma\rangle. \tag{3.72}$$

The states $|\vec{p}, \sigma\rangle$ are identified with the *momentum eigenstates of a particle of mass m*. Basis vectors corresponding to different $\vec{p}$ must be orthogonal. We choose their normalisation such that

$$\langle\vec{p}', \sigma'|\vec{p}, \sigma\rangle = (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}' - \vec{p})\delta_{\sigma'\sigma}. \tag{3.73}$$

This normalisation is Lorentz-invariant, in the sense that[24]

$$(2\pi)^3 2(\Lambda p)^0 \delta^{(3)}(\Lambda\vec{p}' - \Lambda\vec{p})\delta_{\sigma'\sigma} = (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}' - \vec{p})\delta_{\sigma'\sigma}. \tag{3.74}$$

We now have to represent Lorentz transformations in our vector space. If we now apply a Lorentz transformation on a basis vector, we find that

$$\begin{aligned} P^\mu U(\Lambda)|\vec{p}, \sigma\rangle &= U(\Lambda)U(\Lambda)^{-1}P^\mu U(\Lambda)|\vec{p}, \sigma\rangle = U(\Lambda)\Lambda^\mu{}_\nu P^\nu|\vec{p}, \sigma\rangle \\ &= \Lambda^\mu{}_\nu p^\nu U(\Lambda)|\vec{p}, \sigma\rangle, \end{aligned} \tag{3.75}$$

and so

$$U(\Lambda)|\vec{p}, \sigma\rangle = \sum_{\bar{\sigma}} C_{\bar{\sigma}\sigma}(\Lambda, \vec{p})|\Lambda\vec{p}, \bar{\sigma}\rangle, \tag{3.76}$$

_____

[22]Representations with $m^2 < 0$ correspond to tachions, which are not welcome since they travel faster than light.

[23]More precisely, each basis vector provides a different representation. These are actually all the possible unitary representations of translations: being an Abelian group its irreducible representations must be one-dimensional, and the request of unitarity combined with the request of smoothness of the representation imposes that $\partial U(a)/\partial a^\mu = U(a)\partial U(a)/\partial a^\mu|_0$, so $U(a) = \exp\{-ia_\mu(i\partial U(a)/\partial a^\mu|_0)\} = \exp\{-ia_\mu p^\mu\}$ with real p^μ.

[24]A quick way to see this is that

$$1 = \int \frac{d^3p}{p^0} p^0 \delta^{(3)}(\vec{p}' - \vec{p}) = \int \frac{d^3(\Lambda p)}{(\Lambda p)^0}(\Lambda p)^0 \delta^{(3)}(\Lambda\vec{p}' - \Lambda\vec{p}) = \int \frac{d^3p}{p^0}(\Lambda p)^0 \delta^{(3)}(\Lambda\vec{p}' - \Lambda\vec{p})$$

since the integration measure is invariant. This in turn follows from the fact that $d^4p\,\theta(p^0)\delta(p^2 - m^2) = \frac{d^3p}{2p^0}$. Notice also that we denote with $\Lambda\vec{p}$ the spatial components of the transformed four-vector.

for some matrix $C_{\bar{\sigma}\sigma}(\Lambda, \vec{p})$. If we require now that the representation be unitary, then

$$
\begin{aligned}
(2\pi)^3 2p^0 \delta^{(3)}(\vec{p}' - \vec{p})\delta_{\sigma'\sigma} &= \langle \vec{p}', \sigma' | U(\Lambda)^\dagger U(\Lambda) | \vec{p}, \sigma \rangle = \sum_{\bar{\sigma}', \bar{\sigma}} C_{\bar{\sigma}'\sigma'}(\Lambda, \vec{p}')^* C_{\bar{\sigma}\sigma}(\Lambda, \vec{p}) \langle \Lambda \vec{p}', \bar{\sigma}' | \Lambda \vec{p}, \bar{\sigma} \rangle \\
&= (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}' - \vec{p}) \sum_{\bar{\sigma}', \bar{\sigma}} C_{\bar{\sigma}'\sigma'}(\Lambda, \vec{p})^* C_{\bar{\sigma}\sigma}(\Lambda, \vec{p}) \delta_{\bar{\sigma}'\bar{\sigma}} \\
&= (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}' - \vec{p}) \sum_{\bar{\sigma}} C_{\sigma'\bar{\sigma}}(\Lambda, \vec{p})^\dagger C_{\bar{\sigma}\sigma}(\Lambda, \vec{p}) \,,
\end{aligned}
$$

$$(3.77)$$

and it follows that the matrices $C(\Lambda, \vec{p})$ must be unitary. Up to a unitary transformation in the "internal" space, the representative of a Lorentz transformation sends a state of momentum $\vec{p}$ into a state of momentum $\vec{\Lambda p}$. For a given $\vec{p}$ and a given Λ, the internal unitary transformation can always be undone by a judicious choice of basis, i.e., we can always rename $\sum_{\bar{\sigma}} C_{\bar{\sigma}\sigma}(\Lambda, \vec{p}) | \Lambda \vec{p}, \bar{\sigma} \rangle \to | \Lambda \vec{p}, \sigma \rangle$. In general, any vector of a given type (timelike, spacelike, or null) can be reached via a Lorentz transformation from any other vector of the same type. This is because we can always transform any vector into a vector of the form $(p^0, 0, 0, p^3)$ by means of a rotation, and by means of a boost we can transform this into $(m, 0, 0, 0)$ if the vector is timelike, or into $(0, 0, 0, m)$ if the vector is spacelike. If it is null, after rotation it is of the form $(\bar{k}, 0, 0, \bar{k})$, which cannot be further simplified by means of boosts. We can then always write a general four-vector p as

$$p^\mu = \Lambda_p{}^\mu{}_\nu k^\nu \,, \tag{3.78}$$

for a given reference vector k, and a transformation Λ_p (which implicitly depends on k as well). We are then tempted to choose the basis so that $C_{\bar{\sigma}\sigma}(\Lambda_p, \vec{k})$ are all the identity matrix. This however is not possible: the transformation Λ_p is not unique, as the same final p is obtained if we multiply Λ_p on the right by a transformation Λ_k^0 that leaves k invariant, and on the left by a transformation Λ_p^0 that leaves p invariant. A transformation Λ_p^0 that leaves p invariant is said to belong to the *little group* of p. In general, $C(\Lambda_p, \vec{k})$ and $C(\Lambda_p^0 \Lambda_p \Lambda_k^0, \vec{k})$ are not the same, and cannot be both replaced by the identity. We can however do this for a single, definite choice for Λ_p:

- for timelike vectors we will choose Λ_p to be a pure boost that transforms $k = (m, 0, 0, 0)$ into p;

- for null vectors we will choose Λ_p to be a a pure boost that transforms the reference vector $k = (\bar{k}, 0, 0, \bar{k})$ into $(|\vec{p}|, 0, 0, |\vec{p}|)$, followed by a rotation that transforms this into $p = (|\vec{p}|, \vec{p})$.

Now we can finally choose our basis such that

$$U(\Lambda_p) | \vec{k}, \sigma \rangle = | \vec{p}, \sigma \rangle \,. \tag{3.79}$$

Once this is done, every other matrix $C(\Lambda, \vec{p})$ has to be worked out and cannot be simply put to the identity. To this end, consider a general transforamtion on a general state, and write

$$
\begin{aligned}
U(\Lambda) | \vec{p}, \sigma \rangle &= U(\Lambda) U(\Lambda_p) | \vec{k}, \sigma \rangle = U(\Lambda_{\Lambda p}) U(\Lambda_{\Lambda p})^\dagger U(\Lambda) U(\Lambda_p) | \vec{k}, \sigma \rangle \\
&= U(\Lambda_{\Lambda p}) U(\Lambda_{\Lambda p}^{-1} \Lambda \Lambda_p) | \vec{k}, \sigma \rangle \,.
\end{aligned}
\tag{3.80}
$$

The transformation $W(\Lambda, \vec{p}) \equiv \Lambda_{\Lambda p}^{-1} \Lambda \Lambda_p$ sends first $k \to p$, then $p \to \Lambda p$, and finally $\Lambda p \to k$, which means that it belongs to the little group of k. We have then

$$
\begin{aligned}
U(\Lambda)|\vec{p}, \sigma\rangle &= U(\Lambda_{\Lambda p}) \sum_{\bar{\sigma}} C_{\bar{\sigma}\sigma}(\Lambda_{\Lambda p}^{-1} \Lambda \Lambda_p, \vec{k})|\vec{k}, \bar{\sigma}\rangle \equiv U(\Lambda_{\Lambda p}) \sum_{\bar{\sigma}} \mathscr{D}_{\bar{\sigma}\sigma}(W(\Lambda, \vec{p}))|\vec{k}, \bar{\sigma}\rangle \\
&= \sum_{\bar{\sigma}} \mathscr{D}_{\bar{\sigma}\sigma}(W(\Lambda, \vec{p}))|\Lambda \vec{p}, \bar{\sigma}\rangle \,.
\end{aligned}
\tag{3.81}
$$

The matrix $\mathscr{D}_{\bar{\sigma}\sigma}(W(\Lambda, \vec{p}))$ is well defined once that we have made our choice for the reference vector, for Λ_p and for the basis vectors, since $W(\Lambda, \vec{p}) = \Lambda_{\Lambda p}^{-1} \Lambda \Lambda_p$ is a well-defined element of the little group of k. Furthermore, such matrices must provide a representation of the little group. To make further progress we need to find out what the little group is in the physically relevant cases, and what are its irreducible, finite-dimensional representations. At that point our construction will be complete. We first discuss the little group in some generality.

3.5.1 The little group

The little group of k is defined as the subgroup that leaves k invariant. In our case the group in question is the (proper orthocronous) Lorentz group, and we look for Lorentz transformations W such that

$$
W^\mu{}_\nu k^\nu = k^\mu \,.
\tag{3.82}
$$

For infinitesimal transformations $W^\mu{}_\nu = \delta^\mu{}_\nu + \Omega^\mu{}_\nu$, Eq. (3.82) becomes

$$
\Omega^\mu{}_\nu k^\nu = 0 \,.
\tag{3.83}
$$

Out of the six independent antisymmetric tensors $\Omega_{\mu\nu}$, this condition singles out three,[25] and we can write the most general $\Omega_{\mu\nu}$ obeying our equation as

$$
\Omega_{\mu\nu}(\theta) = \theta^\tau \epsilon_{\tau\mu\nu\rho} k^\rho = \frac{i}{2} \theta^\tau \epsilon_{\tau\alpha\beta\rho} J_{\mu\nu}^{(\alpha\beta)} k^\rho \,,
\tag{3.84}
$$

with $\theta^\tau k_\tau = 0$, or in matrix notation

$$
\Omega(\theta) = \frac{i}{2} \theta^\tau \epsilon_{\tau\alpha\beta\rho} J^{(\alpha\beta)} k^\rho \,.
\tag{3.85}
$$

Here and above $J^{(\alpha\beta)}$ are the 4×4 matrices of the defining representation of the Lorentz group. Finite transformations of the little group are obtained by exponentiating Ω, i.e., as e^Ω. If we now define for a generic representation the quantity

$$
W_\tau = \frac{1}{2} \epsilon_{\tau\alpha\beta\rho} J^{(\alpha\beta)} P^\rho \,,
\tag{3.86}
$$

we see that

$$
e^{i\theta^\tau W_\tau}|\vec{k}, \sigma\rangle = e^{\frac{i}{2}\theta^\tau \epsilon_{\tau\alpha\beta\rho} J^{(\alpha\beta)} k^\rho}|\vec{k}, \sigma\rangle = \sum_{\bar{\sigma}} \mathscr{D}_{\bar{\sigma}\sigma}(\theta)|\vec{k}, \bar{\sigma}\rangle \,,
\tag{3.87}
$$

---

[25] For a generic tensor $t_{\mu\nu}$, the condition Eq. (3.83) leaves three out of four columns still independent; antisymmetry implies that only three out of four rows are independent. The independent entries are then those of a 3×3 antisymmetric matrix, which are precisely 3.

since k is left invariant. Here $\mathscr{D}(\theta) = e^{i\theta^\tau \Gamma_\tau}$ is a unitary matrix which forms part of a representation of the little group, and Γ_τ are Hermitian matrices representing the corresponding algebra. The quantity in Eq. (3.86) is the *Pauli-Lubański (pseudo)vector*, and it generates the transformation of the little group. Notice that

$$
\begin{aligned}
W_\tau &= \frac{1}{2}\epsilon_{\tau\alpha\beta\rho}J^{(\alpha\beta)}P^\rho = \frac{1}{2}\epsilon_{\tau\alpha\beta\rho}\left\{ P^\rho J^{(\alpha\beta)} + [J^{(\alpha\beta)}, P^\rho] \right\} \\
&= \frac{1}{2}\epsilon_{\tau\alpha\beta\rho}\left\{ P^\rho J^{(\alpha\beta)} + i\left(\eta^{\alpha\rho}P^\beta - \eta^{\beta\rho}P^\alpha \right) \right\} \\
&= \frac{1}{2}\epsilon_{\tau\alpha\beta\rho}P^\rho J^{(\alpha\beta)} \,,
\end{aligned}
$$

$$
P^\tau W_\tau = \frac{1}{2}\epsilon_{\tau\alpha\beta\rho}P^\tau P^\rho J^{(\alpha\beta)} = 0 \,. \tag{3.88}
$$

The transformation properties of W_τ under (proper) Lorentz transformations are clearly those of a vector. Moreover,

$$
[W_\tau, P_\mu] = \frac{1}{2}\epsilon_{\tau\alpha\beta\rho}P^\rho[J^{(\alpha\beta)}, P_\mu] = \frac{i}{2}\epsilon_{\tau\alpha\beta\rho}P^\rho\left(\delta^\alpha{}_\mu P^\beta - \delta^\beta{}_\mu P^\alpha \right) = 0 \,. \tag{3.89}
$$

It is then clear that the quantity $W^2 = W_\mu W^\mu$ is a Casimir operator. Its explicit expression can be obtain by direct calculation, making use of the identity

$$
\epsilon_{\mu\alpha\beta\rho}\epsilon^{\mu\gamma\delta\sigma} = -\begin{vmatrix} \delta^\gamma{}_\alpha & \delta^\gamma{}_\beta & \delta^\gamma{}_\rho \\ \delta^\delta{}_\alpha & \delta^\delta{}_\beta & \delta^\delta{}_\rho \\ \delta^\sigma{}_\alpha & \delta^\sigma{}_\beta & \delta^\sigma{}_\rho \end{vmatrix} \,, \tag{3.90}
$$

and it reads

$$
W_\mu W^\mu = -\frac{1}{2}P_\mu P^\mu J^{(\alpha\beta)}J_{(\alpha\beta)} + J^{(\alpha\beta)}P_\beta J_{(\alpha\gamma)}P^\gamma \,. \tag{3.91}
$$

Finally, note that on a general vector

$$
\begin{aligned}
W^\mu|\vec{p}, \sigma\rangle &= W^\mu U(\Lambda_p)|\vec{k}, \sigma\rangle = U(\Lambda_p)U(\Lambda_p)^\dagger W^\mu U(\Lambda_p)|\vec{k}, \sigma\rangle = U(\Lambda_p)\Lambda_p{}^\mu{}_\nu W^\nu|\vec{k}, \sigma\rangle \\
&= U(\Lambda_p)\sum_{\bar\sigma}\Lambda_p{}^\mu{}_\nu\Gamma^\nu_{\bar\sigma\sigma}|\vec{k}, \bar\sigma\rangle = \sum_{\bar\sigma}\Lambda_p{}^\mu{}_\nu\Gamma^\nu_{\bar\sigma\sigma}|\vec{p}, \bar\sigma\rangle \,.
\end{aligned} \tag{3.92}
$$

3.5.2 Massive representations

Consider now the case $m > 0$. In this case $k = (m, 0, 0, 0)$, corresponding to the rest frame of the particle, and the generic state is given by $|\vec{p}, \sigma\rangle = U(\Lambda_p)|\vec{0}, \sigma\rangle$ with Λ_p a pure boost in the direction of $\vec{p}$ of velocity $\tanh\Theta = |\vec{p}|/p^0$. This can be written as

$$
\Lambda_p{}^\mu{}_\nu = \delta^\mu{}_\nu + (\cosh\Theta - 1)\left(\delta^\mu{}_0\delta^0{}_\nu - n^\mu n_\nu \right) + \sinh\theta(n^\mu\delta^0{}_\nu - \delta^\mu{}_0 n_\nu) \,, \tag{3.93}
$$

where $n^\mu = (0, \vec{n})$ and $\vec{n} = \vec{p}/|\vec{p}|$. The little group is clearly the group of proper rotations, SO(3), of which we know all the irreducible representations. This can be seen from the Pauli-Lubański vector as well: denoting $w_\tau = \frac{1}{2}\epsilon_{\tau\alpha\beta\rho}J^{(\alpha\beta)}k^\rho$ we have

$$
w_\tau = \frac{1}{2}\epsilon_{\tau\alpha\beta 0}J^{(\alpha\beta)}m = -\frac{m}{2}\epsilon_{0\tau jk}J^{(jk)} \,, \tag{3.94}
$$

43

so $w_0 = 0$ and $w^i = -\frac{m}{2}\epsilon^{0ijk}J^{(jk)} = mJ^i$, i.e.,

$$w^\tau = \left(0, m\vec{J}\right). \qquad (3.95)$$

The angular momentum in the rest frame of the particle is by definition its *spin*. The Casimir invariant is just $W^2 = -m^2\vec{J}^2$ evaluated in the rest frame. Equivalently, we can take the value of $\vec{J}^2 = s(s+1)$ to label the irreducible representations of the little group. These are just the spin-s representations of the rotation group. The representatives of $\vec{J}$ are denoted by $\vec{s}$, and the index σ can then be taken as the eigenvalue of, say, the third component of $\vec{s}$, s_3; the vectors $|\vec{0}, s_3\rangle$ provide an irreducible representation of the rotation group of dimension $(2s+1)$, and the Pauli-Lubański vector is represented by $w^\tau = (0, m\vec{s})$. From an irreducible representation of the little group we obtain an irreducible representation of the full group and viceversa: in fact, since every vector $|\vec{p}, \sigma\rangle$ in the representation is obtained from some vector of the form $\sum_\sigma c_\sigma |\vec{p}, \sigma\rangle$, the representation of the full group could reduce only if the representation of the little group on that subspace did, and viceversa if the representation of the full group has to be irreducible, then that of the little group cannot be reducible. Finally, notice that in a general frame, from Eqs. (3.92) and (3.95) we have that the representative of w^τ is replaced by $\Lambda_p{}^\tau{}_\rho w^\rho$,

$$\Lambda_p{}^\tau{}_\rho w^\rho = \left(\vec{p}\cdot\vec{s}, m\vec{s} + \frac{\vec{p}\cdot\vec{s}}{p^0 + m}\vec{p}\right). \qquad (3.96)$$

This completes the construction of the irreducible unitary representations of the Poincaré group for massive particles: as announced, particles are characterised by their mass and spin. The characterisation of particles as irreducible representations of the symmetry group actually explains the origin itself of spin: massive particles do not have to transform trivially under rotations in their rest frame, and the relevant representations are labelled precisely by the value of spin.

3.5.3   Massless representations

If $m = 0$ it is not possible to make a Lorentz transformation to make the spatial momentum vanish. We then choose our reference four-momentum to be of the form $k^\mu = (\bar{k}, 0, 0, \bar{k})$. The corresponding little group is generated by

$$w_\tau = \frac{\bar{k}}{2}\left(\epsilon_{\tau\alpha\beta 0}J^{(\alpha\beta)} + \epsilon_{\tau\alpha\beta 3}J^{(\alpha\beta)}\right). \qquad (3.97)$$

Explicitly,
$$
\begin{aligned}
w_0 &= -\bar{k}J^{(12)} = \bar{k}J^3, \\
w_1 &= \bar{k}\left(J^{(23)} + J^{(02)}\right) = \bar{k}\left(-J^1 + K^2\right), \\
w_2 &= -\bar{k}\left(J^{(13)} + J^{(01)}\right) = -\bar{k}\left(J^2 + K^1\right), \\
w_3 &= \bar{k}J^{(12)} = -\bar{k}J^3,
\end{aligned}
\qquad (3.98)
$$

and so
$$w^\mu = \left(J^3, J^1 - K^2, J^2 + K^1, J^3\right). \qquad (3.99)$$

The Casimir operator is therefore $w_\mu w^\mu = -(w^1)^2 - (w^2)^2$. The commutation relations of the generators of the little group are comuted easily:

$$[w^3, w^1] = [J^3, J^1 - K^2] = i(J^2 + K^1) = iw^2\,,$$
$$[w^3, w^2] = [J^3, J^2 + K^1] = i(-J^1 + K^2) = -iw^1\,,$$
$$[w^1, w^2] = [J^1 - K^2, J^2 + K^1] = i(J^3 - J^3) = 0\,.$$

(3.100)

This is the algebra of the group ISO(2) of ismoetries of the two-dimensional Euclidean plane. Indeed, denoting with $P^{1,2}$ the generators of translations and with J that of rotations [i.e., the generator of SO(2)], a simple calculation shows that they obey the commutation relations

$$[J, P^i] = i\epsilon_{ij} P^j\,, \qquad [P^1, P^2] = 0\,,$$

(3.101)

where $\epsilon_{12} = -\epsilon_{21} = 1$, $\epsilon_1 ii = 0$ is the two-dimensional Levi-Civita tensor.

In order to build representations we can proceed in two ways. We can diagonalise both w^1 and w^2, taking basis wectors of the form $|\vec{w}_\perp\rangle$ with $w^i|\vec{w}_\perp\rangle = \vec{w}_\perp^i|\vec{w}_\perp\rangle$, $i = 1, 2$; the value of the Casimir operator is then just $-\vec{w}_\perp^2$. The space of all the $|\vec{w}_\perp\rangle$ with the same $\vec{w}_\perp^2$ is clearly left invariant by rotations, and it has no proper subspace which is: we have then identified an irreducible representation. Alternatively, we can notice that $[w^3, w_\pm] = \pm w_\pm$ where $w_\pm = w^1 \pm iw^2$, and take a basis of eigenvectors of w^3, $w^3|\sigma\rangle = \sigma|\sigma\rangle$. From the commutation relations we find that $w^3 w_\pm|\sigma\rangle = (\sigma \pm 1)w_\pm|\sigma\rangle$, and moreover $\langle\sigma|w_\mp w_\pm|\sigma\rangle = \langle\sigma|(w^1)^2 + (w^2)^2|\sigma\rangle = \vec{w}_\perp^2\langle\sigma|\sigma\rangle$, so that all the vectors generated with raising and lowering operators belong to the same irreducible representation. No matter how we proceed, we find that in general there is an infinity of basis vectors, except when $\vec{w}_\perp^2 = 0$: in this case there is a single basis vector, which we take to be an eigenvector of w^3. We redefine it to be $|\lambda\rangle$ with $w^3|\lambda\rangle = \bar{k}\lambda|\lambda\rangle$, and moreover $w_\pm|\lambda\rangle = 0$. This means that w^μ is diagonal with $w^\mu = \lambda(\bar{k}, 0, 0\bar{k}) = \lambda k^\mu$. Since w^μ is a pseudovector and k^μ is a vector, λ must be a pseudoscalar. Since J^3 is the component of angular momentum along the direction of motion for our reference vector, we can write it as $\frac{\vec{J}\cdot\vec{k}}{|\vec{k}|}$ in a frame-independent way (this can be seen also from the very definition of w^μ). This quantity is called *helicity*, it is a relativistic invariant for massless particles, and it plays a role analogous to spin in labelling different irreducible representations. While $+\lambda$ and $-\lambda$ correspond to distinct representations of the proper orthochronous group, if we include parity we have to put them together, as helicity changes sign under a spatial reflection.

Up to this point there is no reason not to consider arbitrary values of λ, but there is topological argument to restrict it to integer or half-integer values only. The Lorentz group is a doubly connected group, as it contains the doubly connected subgroup of rotations. Lie groups are also manifolds, and the one corresponding to SO(3) is the three-dimensional ball with antipodal points identified (also called the three-dimensional projective space). Consider a path that winds one around the manifold, starting from, say, the identity, reaching the surface of the ball, a reaching out again to the origin continuing from the antipodal point. Such a path cannot be shrunk to a point. This means that in a generic representation the composition of two transformations $U(\Lambda_1)U(\Lambda_2)$ will be equal to $U(\Lambda_1\Lambda_2)$ only up to a phase, as we are not guaranteed that $U(\Lambda_1\Lambda_2)^{-1}U(\Lambda_1)U(\Lambda_2)$ must be equal to the identity since we cannot deform the corresponding path into the trivial path. On the other hand, if we wind around the manifold twice, then we can deform the path into the trivial one, and so $[U(\Lambda_1\Lambda_2)^{-1}U(\Lambda_1)U(\Lambda_2)]^2 = \mathbf{1}$, or so $U(\Lambda_1\Lambda_2)^{-1}U(\Lambda_1)U(\Lambda_2) = \pm\mathbf{1}$. Consider now a transformation $e^{i\alpha w_3} = e^{-i\alpha\bar{k}J^3} = e^{-i\alpha\bar{k}\lambda}$

and choose $\alpha \bar{k} = 2\pi$, so that we wind once around the group manifold. We must have that $\mathbf{1} = (e^{i2\pi\lambda})^2 = e^{i4\pi\lambda}$, from which it follows that $2\lambda \in \mathbb{Z}$.

To complete the construction for massless particles, we have to connect the reference vector to a generic one. This is done by acting first with a boost in direction 3, and then by a rotation that sends the direction 3 in the direction of $\vec{p} = |\vec{p}|(\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$. This can be done in many different ways; following Weinberg, we choose it to be a rotation around axis 2 of angle $\theta \in [0, \pi]$, followed by a rotation around the axis 3 of angle $\phi \in [0, 2\pi)$. Everything is now well defined.

3.6 Discrete transformations

So far we have built representations of that component of the Poincaré group connected to the identity (translations and proper orthocronous Lorentz transformations). It was believed for a long time, and it still is approximately true in certain physical situations, that parity and time reversal were symmetries of Nature as well. In this case, it would be possible to represent them as unitary or antiunitary operators in the Hilbert space of physical systems.

Let us first discuss what is the effect of parity and time reversal in the defining representation of the group. This is perhaps most clearly seen using the five-dimensional matrix representation of the group. Here parity and time reversal read

$$\mathcal{P} = \left(\begin{array}{c|c} P & 0 \\ \hline 0 & 1 \end{array} \right), \qquad \mathcal{T} = \left(\begin{array}{c|c} T & 0 \\ \hline 0 & 1 \end{array} \right), \tag{3.102}$$

where P and T are the four-dimensional matrices

$$P = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & -\mathbf{1}_3 \end{array} \right), \qquad T = \left(\begin{array}{c|c} -1 & 0 \\ \hline 0 & \mathbf{1}_3 \end{array} \right). \tag{3.103}$$

The generators of the connected component are then easily seen to transform as follows:

$$\begin{aligned} \mathcal{P}^{-1}\mathcal{K}^i\mathcal{P} = -\mathcal{K}^i\,, \qquad \mathcal{P}^{-1}\mathcal{J}^i\mathcal{P} = \mathcal{J}^i\,, \\ \mathcal{T}^{-1}\mathcal{K}^i\mathcal{T} = -\mathcal{K}^i\,, \qquad \mathcal{T}^{-1}\mathcal{J}^i\mathcal{T} = \mathcal{J}^i\,, \end{aligned} \tag{3.104}$$

for the Lorentz generators, and

$$\begin{aligned} \mathcal{P}^{-1}\mathcal{P}^0\mathcal{P} = \mathcal{P}^0\,, \qquad \mathcal{P}^{-1}\mathcal{P}^i\mathcal{P} = -\mathcal{P}^i\,, \\ \mathcal{T}^{-1}\mathcal{P}^0\mathcal{T} = -\mathcal{P}^0\,, \qquad \mathcal{T}^{-1}\mathcal{P}^i\mathcal{T} = \mathcal{P}^i\,, \end{aligned} \tag{3.105}$$

for the translation generators. We have seen how when representing the Poincaré group on a Hilbert space we have to use unitary operators for the component connected to the identity; for the discrete transformations that we are discussing here there is however no *a priori* reason to discard the antiunitary possibility. This means that when trying to represent the transformation properties Eqs. (3.104) and (3.105) of the generators in one of the unitary representations discussed in the previous subsection, we have to take into account that finite transformations are of the form $U(\Lambda) = e^{\frac{i}{2}\omega_{\rho\sigma}J^{(\rho\sigma)}}$ and $U(a) = e^{-ia^\mu P_\mu}$, i.e., the generators are effectively $iJ^{(\rho\sigma)}$ and iP^μ. We have then that the operators $U(P)$ and $U(T)$ realising parity and time-reversal on our Hilbert space must satisfy

$$\begin{aligned} U(P)^\dagger iK^i U(P) = -iK^i\,, \qquad U(P)^\dagger iJ^i U(P) = iJ^i\,, \\ U(T)^\dagger iK^i U(T) = -iK^i\,, \qquad U(T)^\dagger iJ^i U(T) = iJ^i\,, \end{aligned} \tag{3.106}$$

for the Lorentz generators, and

$$U(P)^\dagger i P^0 U(P) = i P^0 \,, \qquad U(P)^\dagger i P^i U(P) = -i P^i \,,$$
$$U(T)^\dagger i P^0 U(T) = -i P^0 \,, \qquad U(T)^\dagger i P^i U(T) = i P^i \,,$$
(3.107)

for translations. The choice whether $U(P)$ is linear or antilinear is done based on physical reasons. If $U(P)$ were antilinear, then we would have $U(P)^\dagger P^0 U(P) = -P^0$, or equivalently $\{P^0, U(P)\} = 0$. This would mean that for any positive-energy state $|\vec{p}, \sigma\rangle$ we would have a negative-energy state $U(P)|\vec{p}, \sigma\rangle$, which contradicts experience. We are then forced to choose the linear unitary possibility, and that $[P^0, U(P)] = 0$. For the same reason, we are forced to choose $U(T)$ to be antilinear antiunitary, and that $[P^0, U(T)] = 0$. In fact, if $U(T)$ were linear we would find again $U(T)^\dagger P^0 U(T) = -P^0$, or equivalently $\{P^0, U(T)\} = 0$, which would imply the existence of negative-energy states. With these choices we have the following transformation properties:

$$U(P)^\dagger K^i U(P) = -K^i \,, \qquad U(P)^\dagger J^i U(P) = J^i \,,$$
$$U(T)^\dagger K^i U(T) = K^i \,, \qquad U(T)^\dagger J^i U(T) = -J^i \,,$$
(3.108)

for the Lorentz generators, and

$$U(P)^\dagger P^0 U(P) = P^0 \,, \qquad U(P)^\dagger P^i U(P) = -P^i \,,$$
$$U(T)^\dagger P^0 U(T) = P^0 \,, \qquad U(T)^\dagger P^i U(T) = -P^i \,,$$
(3.109)

for translations. These are in agreement wuith our physical intuition: angular momentum is unchanged under parity but changes sign under time reversal (objects spin the same way in the mirror, and in the opposite way if we move backwards in time), and spatial momenta change sign under both parity and time reversal.

We can now work out the effect of parity on the one-particle states, confining our discussion to the massive case. The discussion of time reversal and the extension to the massless case present some minor complications; a detailed discussion is found in Weinberg, *op. cit.* Consider the state $U(P)|\vec{0}, \sigma\rangle$ and act on it with P^μ. We find

$$P^\mu U(P)|\vec{0}, \sigma\rangle = U(P) U(P)^\dagger P^\mu U(P)|\vec{0}, \sigma\rangle = U(P) \eta^{\mu\mu} P^\mu |\vec{0}, \sigma\rangle = (m, \vec{0}) U(P)|\vec{0}, \sigma\rangle \,. \quad (3.110)$$

Furthermore,

$$J^3 U(P)|\vec{0}, \sigma\rangle = U(P) J^3 |\vec{0}, \sigma\rangle = \sigma U(P)|\vec{0}, \sigma\rangle \,. \tag{3.111}$$

This means that $U(P)|\vec{0}, \sigma\rangle = \eta_\sigma |\vec{0}, \sigma\rangle$ for some phase η_σ. Since $U(P)$ commutes with the raising and lowering operators $J_\pm$, one sees that η_σ is actually independent of σ, $\eta_\sigma = \eta$. The phase η is the *intrinsic parity* of the particle. Finally,

$$U(P)|\vec{p}, \sigma\rangle = U(P) U(\Lambda_p)|\vec{0}, \sigma\rangle = U(P) U(\Lambda_p) U(P)^\dagger U(P)|\vec{0}, \sigma\rangle$$
$$= U(P\Lambda_p P^{-1}) \eta |\vec{0}, \sigma\rangle = \eta U(\Lambda_{Pp})|\vec{0}, \sigma\rangle = \eta| -\vec{p}, \sigma\rangle \,,$$
(3.112)

since $P\Lambda_p P^{-1} = \Lambda_{Pp}$ is precisely the boost bringing $k = (m, \vec{0})$ to Pp, where $p = (p^0, \vec{p})$ and $Pp = (p^0, -\vec{p})$.

3.7 Noether's theorem and the Lorentz group

We can now fill in a gap that remains since our discussion of Noether's theorem, i.e., its application to the invariance under Lorentz transformations. The transformation law of coordinates under an infinitesimal transformation is

$$x'^{\mu} = x^{\mu} + \frac{1}{2}\omega_{\rho\sigma}M^{(\rho\sigma)\mu}{}_{\nu}x^{\nu}\,, \tag{3.113}$$

where the matrices $M^{(\rho\sigma)}$ are given in Eq. (3.44) and read

$$M^{(\rho\sigma)\,\mu}{}_{\nu} = \eta^{\rho\mu}\delta^{\sigma}{}_{\nu} - \eta^{\sigma\mu}\delta^{\rho}{}_{\nu}\,. \tag{3.114}$$

A general multicomponent field $\phi_a(x)$ will transform as

$$\phi'_a(x') = \phi_a(x) + \frac{1}{2}\omega_{\rho\sigma}S^{(\rho\sigma)}_{ab}\phi_b(x)\,, \tag{3.115}$$

where $S^{(\rho\sigma)}_{ab} = -S^{(\sigma\rho)}_{ab}$ are the generators of the Lorentz group in some finite-dimensional representation. The reason for this form of the transformation law is the same discussed in the case of rotations in Section 3.1: the new and the old field at the same physical point should be the same, up to a mixing of the various components which has to provide a representation of the symmetry group. Plugging Eqs. (3.113) and (3.115) in the general expression for the Noether's current, Eq. (2.68), we find the conserved current $J^{(\rho\sigma)\mu}$ associated to the transformation with only $\omega_{\rho\sigma} \neq 0$,

$$\begin{aligned}
\mathcal{M}^{\mu,\rho\sigma} \equiv J^{(\rho\sigma)\mu} &= M^{(\rho\sigma)\,\mu}{}_{\nu}x^{\nu}\mathcal{L} + \left(S^{(\rho\sigma)}_{ab}\phi_b - M^{(\rho\sigma)\,\nu}{}_{\alpha}x^{\alpha}\partial_{\nu}\phi_a\right)\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)} \\
&= x^{\rho}\left(\partial^{\sigma}\phi_a\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)} - \eta^{\mu\sigma}\mathcal{L}\right) - x^{\sigma}\left(\partial^{\rho}\phi_a\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)} - \eta^{\mu\rho}\mathcal{L}\right) + S^{(\rho\sigma)}_{ab}\phi_b\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)} \\
&= x^{\rho}\Theta^{\mu\sigma} - x^{\sigma}\Theta^{\mu\rho} + \mathcal{S}^{\mu,\rho\sigma}\,,
\end{aligned} \tag{3.116}$$

where $\Theta^{\mu\nu}$ is the energy-momentum tensor and we have defined

$$\mathcal{S}^{\mu,\rho\sigma} \equiv \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\phi_a)}S^{(\rho\sigma)}_{ab}\phi_b\,. \tag{3.117}$$

The conserved charges are the generators of Lorentz transformations,

$$\int d^3x\,\mathcal{M}^{0,\rho\sigma} = \int d^3x\,\left\{x^{\rho}\Theta^{0\sigma} - x^{\sigma}\Theta^{0\rho} + \mathcal{S}^{0,\rho\sigma}\right\} \tag{3.118}$$

where the first two terms are seen to correspond to the integral of the orbital angular momentum density ($\sim x^{\rho}p^{\sigma} - x^{\sigma}p^{\rho}$), while the last term is the contribution of spin.

Taking the divergence of Eq. (3.116) and recalling that also the energy-momentum tensor is conserved one finds the relation

$$0 = \partial_{\mu}\mathcal{M}^{\mu,\rho\sigma} = \Theta^{\rho\sigma} - \Theta^{\sigma\rho} + \partial_{\mu}\mathcal{S}^{\mu,\rho\sigma}\,, \tag{3.119}$$

i.e., $\Theta^{\rho\sigma}$ is not symmetric in the general case. However, thanks to the antisymmetry of $\mathcal{S}^{\mu,\rho\sigma}$ in ρ, σ, Eq. (3.119) can be recast as

$$\Theta^{\rho\sigma} + \frac{1}{2}\partial_\mu \mathcal{S}^{\mu,\rho\sigma} = \Theta^{\sigma\rho} + \frac{1}{2}\partial_\mu \mathcal{S}^{\mu,\sigma\rho} \,, \tag{3.120}$$

which means that the combination $\Theta^{\rho\sigma} + \frac{1}{2}\partial_\mu \mathcal{S}^{\mu,\rho\sigma}$ is symmetric in ρ, σ. It is possible to exploit this to define a new conserved tensor which is also symmetric. Define the Belinfante-Rosenfeld tensor as

$$\Theta_S^{\mu\nu} \equiv \Theta^{\mu\nu} + \frac{1}{2}\partial_\sigma \left(S^{\mu,\nu\sigma} + S^{\nu,\mu\sigma} - S^{\sigma,\nu\mu}\right) = \Theta^{\mu\nu} + \frac{1}{2}\partial_\sigma \left(S^{\mu,\nu\sigma} + S^{\nu,\mu\sigma} + S^{\sigma,\mu\nu}\right) \,, \tag{3.121}$$

which is symmetric precisely because of Eq. (3.120). To show that it is conserved, notice that

$$\partial_\mu \Theta_S^{\mu\nu} = \frac{1}{2}\partial_\mu\partial_\sigma \left[(S^{\mu,\nu\sigma} - S^{\sigma,\nu\mu}) + S^{\nu,\mu\sigma}\right] = 0 \,, \tag{3.122}$$

since both the term in brackets and $S^{\nu,\mu\sigma}$ are antisymmetric in μ, σ, and they are being contracted with the symmetric tensor $\partial_\mu\partial_\sigma$. Plugging the definition of $\Theta_S^{\mu\nu}$ in Eq. (3.118) we find after an integration by part (in which we assume that wer can neglect boundary terms at infinity)

$$\begin{aligned}
\int d^3x \, \mathcal{M}^{0,\rho\sigma} &= \int d^3x \left\{ x^\rho \Theta_S^{0\sigma} - x^\sigma \Theta_S^{0\rho} - \frac{1}{2}x^\rho \partial_\alpha \left(S^{0,\sigma\alpha} + S^{\sigma,0\alpha} - S^{\alpha,\sigma 0}\right) \right. \\
&\qquad\qquad + \frac{1}{2}x^\sigma \partial_\alpha \left(S^{0,\rho\alpha} + S^{\rho,0\alpha} - S^{\alpha,\rho 0}\right) + \mathcal{S}^{0,\rho\sigma} \Big\} \\
&= \int d^3x \left\{ x^\rho \Theta_S^{0\sigma} - x^\sigma \Theta_S^{0\rho} + \frac{1}{2}\left(S^{0,\sigma\rho} + S^{\sigma,0\rho} - S^{\rho,\sigma 0}\right) \right. \tag{3.123} \\
&\qquad\qquad - \frac{1}{2}\left(S^{0,\rho\sigma} + S^{\rho,0\sigma} - S^{\sigma,\rho 0}\right) + \mathcal{S}^{0,\rho\sigma} \Big\} \\
&= \int d^3x \left\{ x^\rho \Theta_S^{0\sigma} - x^\sigma \Theta_S^{0\rho} \right\}.
\end{aligned}$$

It is now easy to understand the conservation law associated to the $\rho = 0$, $\sigma = j$ generators, i.e., the boost generators. Indeed,

$$\frac{d}{dx^0}\int d^3x \, \mathcal{M}^{0,0j} = \frac{d}{dx^0}\int d^3x \left\{ x^0 \Theta_S^{0j} - x^j \Theta_S^{00} \right\} = \int d^3x \, \Theta_S^{0j} - \frac{d}{dx^0}\int d^3x \, x^j \Theta_S^{00} \,. \tag{3.124}$$

Since $\Theta_S^{\mu\nu}$ is divergenceless, $\int d^3x \, \Theta_S^{0j}$ is a conserved quantity, and the relation Eq. (3.124) expresses the fact that the center of mass $\int d^3x \, x^j \Theta_S^{00} / \int d^3x \, \Theta_S^{00}$ moves on a straight-line trajectory with constant momentum $\int d^3x \, \Theta_S^{0j}$. This is nothing but the center of mass theorem.

3.8 Active and passive points of view

We discuss here a few details about symmetry transformations.

Symmetry transformations can be looked at in two ways: the *active* point of view, in which the system is actively transformed and the reference frame is not, and the *passive* point of view, in which the system is not transformed but the reference frame is. For example, rotating

clockwise a system can be equivalently seen as rotating counterclockwise the reference frame. Suppose we have a local observable $\mathcal{O}(x)$ defined in a certain reference frame, and state vectors ψ and φ. If we transform the system by some transformation M, which transforms its coordinates as $x \to Mx$, the new state vectors will be $\psi_M = U(M)\psi$ and $\varphi_M = U(M)\varphi$ for some operator $U(M)$. The local observable $\mathcal{O}(Mx)$ at the point Mx for the new, transformed system must be related to the local observable $\mathcal{O}(x)$ at point x of the original system. If the local observable is a multicomponent observable, its components will also mix under the transformation. For example, if there is an orientation associated to the observable and we rotate the system, looking in the rotated of a direction at the rotated point for the rotated system should yield the same as looking in the given direction at the original point in the original system. More generally, suppose that we have a set of observables $\{\mathcal{O}_i(x)\}$ where the label i indicates a certain direction in some local internal space of the system S. Suppose furthermore that when making the transformation of coordinates M this internal space is also transformed, and the local coordinates transform as $\sigma_i \to \sigma_i' = \bar{D}_{ij}(M)\sigma_j$ for some invertible $\bar{D}(M)$. For example, in the case of rotations discussed above, $\bar{D}_{ij}(R) = R_{ij}$. What we should have is that for any choice of σ_i, $\sigma_i'\mathcal{O}_i(Mx) = \sigma_j\mathcal{O}_j(x)$ (sum understood), meaning that observing in matching directions we should find the same result. Then for classical systems $\bar{D}_{ij}(M)\sigma_j\mathcal{O}_i(Mx) = \sigma_j\mathcal{O}_j(x)$, from which we conclude $\mathcal{O}_i(Mx) = \bar{D}_{ji}^{-1}(M)\mathcal{O}_j(x) = \bar{D}_{ij}^{T\,-1}(M)\mathcal{O}_j(x)$. Notice that for a rotation R we have $\bar{D}_{ij}^{T\,-1}(R) = R$. Setting $\bar{D}^{T\,-1}(M) = D(M)$, this can be summarised formally as follows in the quantum mechanical case:

$$\langle \varphi_M | \mathcal{O}(Mx) | \psi_M \rangle = \langle \varphi | D(M)\mathcal{O}(x) | \psi \rangle \,, \tag{3.125}$$

where $D(M)$ is some matrix which appropriately mixes the components of the observables. This can be written also as

$$\langle \varphi | U(M)^\dagger \mathcal{O}(Mx) U(M) | \psi \rangle = \langle \varphi | D(M)\mathcal{O}(x) | \psi \rangle \,, \tag{3.126}$$

which should hold for all states, and so

$$U(M)^\dagger \mathcal{O}(Mx) U(M) = D(M)\mathcal{O}(x) \,, \tag{3.127}$$

or also

$$U(M)^\dagger \mathcal{O}(x) U(M) = D(M)\mathcal{O}(M^{-1}x) \,. \tag{3.128}$$

If M is a symmetry of the system, it can be implemented as a unitary (or antiunitary) operator on the Hilbert space of the system, and so $U(M)^\dagger = U(M)^{-1}$, and Eq. (3.128) provides the transformation law for the observable.

Consider translations first. The operator $e^{-ia\cdot P}$ actively translates the system *to the left* (i.e., in the negative a direction), or equivalently it translates the reference frame *to the right*. To see this consider ordinary quantum mechanics and the momentum operator $P_\mu = i\partial_\mu$, or $P^\mu = (i\partial_0, -i\vec{\nabla})$ represented on Lebesgue-integrable functions. We have that the action of the translation operator on a wave function is

$$\psi_a(t,\vec{x}) \equiv e^{-ia\cdot P}\psi(t,\vec{x}) = e^{a^0\partial_0 + \vec{a}\cdot\vec{\nabla}}\psi(t,\vec{x}) = \psi(t + a_0, \vec{x} + \vec{a}) \,, \tag{3.129}$$

i.e., the transformed wave function ψ_a at $(t, \vec{x})$ is the original wave function at $(t + a_0, \vec{x} + \vec{a})$, so that the system has been effectively translated to the left. Then, for $x' = Tx = x - a$ (translation of the system to the left/reference frame to the right)

$$e^{ia\cdot P}\mathcal{O}(x)e^{-ia\cdot P} = \mathcal{O}(x + a) \,. \tag{3.130}$$

Next consider rotations in three dimensions. A counterclockwise rotation of the system around the x_3 direction places the system at the new coordinates

$$x_1' = \cos\theta x_1 - \sin\theta x_2\,, \qquad x_2' = \sin\theta x_1 + \cos\theta x_2\,, \qquad x_3' = x_3\,. \qquad (3.131)$$

In matrix notation

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad (3.132)$$

or more compactly $x' = R_3^A(\theta)x$, where A stands for "active". Now, one can show with a simple calculation that

$$R_3^A(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = e^{-i\theta L^3}\,, \qquad L^3 = -i\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,. \qquad (3.133)$$

A passive counterclockwise rotation is simply obtained by sending $\theta \to -\theta$. Rotations around other axes are studied in the same way. We then find that *a counterclockwise rotation of the reference frame* of angle θ around the axis $\hat{\theta}$ is given by

$$R(\vec{\theta}) = e^{i\vec{\theta}\cdot\vec{L}}\,, \qquad (3.134)$$

where $\vec{\theta} = \theta\hat{\theta}$ and

$$(L^a)_{ij} = -i\epsilon_{aij}\,. \qquad (3.135)$$

This can be easily extended to Minkowski spacetime: a spatial rotation will read $R(\theta) = e^{i\vec{\theta}\cdot\vec{J}}$ with

$$J^a = \left(\begin{array}{c|c} 0 & \vec{0}^{\mathrm{T}} \\ \hline \vec{0} & L^a \end{array}\right)\,. \qquad (3.136)$$

An active boost in direction 1 reads

$$\begin{pmatrix} x_0' \\ x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} \cosh\Theta & \sinh\Theta & 0 & 0 \\ \sinh\Theta & \cosh\Theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad (3.137)$$

or $x' = L_1^A(\Theta)x$, with

$$L_1^A(\Theta) = \begin{pmatrix} \cosh\Theta & \sinh\Theta & 0 & 0 \\ \sinh\Theta & \cosh\Theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = e^{i\Theta K^1}\,, \qquad K^1 = -i\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\,. \qquad (3.138)$$

Similarly to what has been done with rotations, a change to a reference frame moving with velocity $\vec{v} = \hat{\Theta}\tanh\Theta$ is obtained via

$$L(\vec{\Theta}) = e^{-i\vec{\Theta}\cdot\vec{K}}\,, \qquad (3.139)$$

where $\vec{\Theta} = \Theta\hat{\Theta}$ and

$$K^a = \left(\begin{array}{c|c} 0 & \vec{n}_a^{\mathrm{T}} \\ \hline \vec{n}_a & \mathbf{0} \end{array} \right) , \tag{3.140}$$

with $(\vec{n}_a)_i = \delta_{ai}$. In covariant notation, choosing the matrices $(K^a)_{\mu\nu}$ and $(J^a)_{\mu\nu}$ to match the tensors $K^{a\mu}{}_\nu$ and $J^{a\mu}{}_\nu$, we have $K^a = J^{(0a)}$ and $J^a = -\frac{1}{2}\epsilon^{0aij}J^{(ij)}$, with $\epsilon^{0123} = 1$ and

$$J^{(\rho\sigma)\mu}{}_\nu = -i(\eta^{\rho\mu}\delta^\sigma{}_\nu - \eta^{\sigma\mu}\delta^\rho{}_\nu) . \tag{3.141}$$

A general Lorentz transformation therefore reads $e^{i(\vec{\theta}\cdot\vec{J} - \vec{\Theta}\cdot\vec{K})}$, which for $\vec{\Theta} = 0$ is *a counter-clockwise rotation of the reference frame* of angle $|\vec{\theta}|$ around $\hat{\theta}$, and for $\theta = 0$ is *a change of frame to one moving* with relative rapidity $|\vec{\Theta}|$ in direction $\hat{\Theta}$. This can also be written as $e^{\frac{i}{2}\omega_{\rho\sigma}J^{(\rho\sigma)}}$, with $\omega_{0a} = -\Theta^a$ and $\omega_{ab} = -\epsilon_{abc}\theta^c = -\epsilon^{0abc}\theta^c$. A general Poincaré transformation reads $e^{\frac{i}{2}\omega_{\rho\sigma}J^{(\rho\sigma)} - ia^\mu P_\mu}$.

4 Scalar particles and scalar fields

We now have all the basic tools to discussion quantum fields. The first step is showing how free fields appear as a convenient way to describe free particles. Our aim is to develop tools which will allow us to describe relativistic, quantum-mechanical processes. Such tools have to comply with two important requests:

- they should allow creation and destruction of particles, since these are processes allowed by special relativity which actually happen in Nature;

- they should allow a local description of interactions, in order to avoid problem with causality.

We begin our discussion with the simplest case of a scalar massive particle.

4.1 Fock space

We have seen in the previous section how the states of a single particle provide the basis for an irreducible representation of the Poincaré group. The simplest such representation is the scalar representation for massive particles, whose states are labelled by the three spatial components of the momentum, $\vec{p}$, and whose energy satisfies the relation $p^0 = \sqrt{\vec{p}^2 + m^2}$. Denoting such states with $|\vec{p}\rangle$ we then have

$$P^\mu|\vec{p}\rangle = p^\mu|\vec{p}\rangle . \tag{4.1}$$

States containing N non-interacting particles are easily constructed as tensor products of single particle states. To take into account the experimental fact that identical particles are indistinguishable, such tensor products have to be symmetrised under exchange of any two momenta, in order to satisfy the Bose-Einstein statistics appropriate for scalar particles. We denote a generic multiparticle state with N particles as $|\vec{p}_1, \ldots, \vec{p}_N\rangle$, satisfying the symmetry condition $|\vec{p}_1, \ldots, \vec{p}_N\rangle = |\vec{p}_{\mathrm{P}(1)}, \ldots, \vec{p}_{\mathrm{P}(N)}\rangle$ for any permutation P of $\{1, \ldots, N\}$. From the mathematical point of view, we construct such states by taking the N-fold tensor product of single-particle states and symmetrising with respect to the momenta,

$$|\vec{p}_1, \ldots, \vec{p}_N\rangle = \sum_{\mathrm{P}} |\vec{p}_{\mathrm{P}(1)}\rangle \otimes \ldots \otimes |\vec{p}_{\mathrm{P}(N)}\rangle . \tag{4.2}$$

Their transformation properties under symmetry transformations are inherited from the single particle states. Such states provide therefore a reducible representation of the Poincaré group. Single-particle states for a scalar particle transform as follows,

$$U(a)|\vec{p}\rangle = e^{-ia \cdot P}|\vec{p}\rangle = e^{-ia \cdot p}|\vec{p}\rangle \,,$$
$$U(\Lambda)|\vec{p}\rangle = e^{\frac{i}{2}\omega_{\mu\nu}J^{(\mu\nu)}}|\vec{p}\rangle = |\Lambda\vec{p}\rangle \,,$$

(4.3)

and so multiparticle states transform as follows,

$$U(a)|\vec{p}_1, \ldots, \vec{p}_N\rangle = e^{-ia \cdot P}|\vec{p}_1, \ldots, \vec{p}_N\rangle = e^{-ia \cdot \sum_j p_j}|\vec{p}_1, \ldots, \vec{p}_N\rangle \,,$$
$$U(\Lambda)|\vec{p}_1, \ldots, \vec{p}_N\rangle = e^{\frac{i}{2}\omega_{\mu\nu}J^{(\mu\nu)}}|\vec{p}_1, \ldots, \vec{p}_N\rangle = |\Lambda\vec{p}_1, \ldots, \Lambda\vec{p}_N\rangle \,.$$

(4.4)

The normalisation of multiparticle states is easily obtained from the single-particle normalisation,

$$\langle\vec{p}'|\vec{p}\rangle = (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}' - \vec{p}) \,,$$

(4.5)

and from the definition Eq. (4.2):

$$\langle\vec{p}'_1, \ldots, \vec{p}'_M|\vec{p}_1, \ldots, \vec{p}_N\rangle = \delta_{MN} \sum_{\mathrm{P}} \prod_{j=1}^{N} (2\pi)^3 2p_j^0 \delta^{(3)}(\vec{p}'_{\mathrm{P}(j)} - \vec{p}_j) \,,$$

(4.6)

having naturally set to zero the scalar product of states with different number of particles. Since particles can be created and destroyed, any number of particles is allowed; in particular, no particles is a possibility. We call the state without particles the *vacuum state*, and denote by $|0\rangle$ the corresponding vector, which we take to be normalised to $\langle 0|0\rangle = 1$ and orthogonal to all other states, $\langle 0|\vec{p}_1, \ldots, \vec{p}_N\rangle = 0$. Furthermore, we assume the vacuum to be invariant under translations and Lorentz transformations, $U(a)|0\rangle = U(\Lambda)|0\rangle = |0\rangle$. The appropriate state space for scalar particles is then the complex vector space (Hilbert space) spanned by the basis $\{|\vec{p}_1, \ldots, \vec{p}_N\rangle\}_{N=0,1,\ldots,\infty}$.[26] This is the so-called *Fock space* for scalar particles.

In order to describe creation and annihilation of particles, we introduce the *creation operator* $a(\vec{p})^\dagger$, defined by its action on the basis vectors:

$$a(\vec{p})^\dagger|\vec{p}_1, \ldots, \vec{p}_N\rangle = |\vec{p}, \vec{p}_1, \ldots, \vec{p}_N\rangle \,.$$

(4.7)

This operator creates a particle with momentum $\vec{p}$, which is then added to the state vector. We can then obtain all the basis vectors by repeated application of creation operators on the vacuum:

$$|\vec{p}_1, \ldots, \vec{p}_N\rangle = a(\vec{p}_1)^\dagger \ldots a(\vec{p}_N)^\dagger|0\rangle \,.$$

(4.8)

The adjoint of the creation operator, $a(\vec{p})$, is the *annihilation operator*, for reasons that will

---

[26]Technically, the appropriate Hilbert space is the completion of the space generated by arbitrary linear combinations of these basis vectors, which includes the limits of any sequence of such linear combinations in the norm derived from the scalar product.

appear clearly as soon as we determine its effect on the basis vectors. By definition of adjoint,[27]

$$(a(\vec{p})\Psi_{\vec{p}'_1,\dots,\vec{p}'_M}, \Psi_{\vec{p}_1,\dots,\vec{p}_N}) = (\Psi_{\vec{p}'_1,\dots,\vec{p}'_M}, a(\vec{p})^\dagger \Psi_{\vec{p}_1,\dots,\vec{p}_N}) = (\Psi_{\vec{p}'_1,\dots,\vec{p}'_M}, \Psi_{\vec{p},\vec{p}_1,\dots,\vec{p}_N})$$

$$= \delta_{M,N+1} \sum_{\mathrm{P}} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}'_{\mathrm{P}(1)} - \vec{p}) \prod_{k=2}^{M} (2\pi)^3 2p_k^0 \delta^{(3)}(\vec{p}'_{\mathrm{P}(k)} - \vec{p}_{k-1}) \qquad (4.9)$$

$$= \sum_{j=1}^{M} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}'_j - \vec{p}) (\Psi_{\vec{p}'_1,\dots,\vec{p}'_M \setminus \vec{p}'_j}, \Psi_{\vec{p}_1,\dots,\vec{p}_N}) \,,$$

where $\Psi_{\vec{p}'_1,\dots,\vec{p}'_M \setminus \vec{p}'_j}$ is obtained from the state $\Psi_{\vec{p}'_1,\dots,\vec{p}'_M}$ by removing a particle with momentum $\vec{p}'_j$. We then conclude

$$a(\vec{p})|\vec{p}_1,\dots,\vec{p}_N\rangle = \sum_{j=1}^{N} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}_j - \vec{p})|\vec{p}_1,\dots,\vec{p}_{j-1},\vec{p}_{j+1},\dots,\vec{p}_N\rangle \,. \qquad (4.10)$$

The annihilation operator $a(\vec{p})$ destroys a particle with momentum $\vec{p}$, which is then removed from the state vector in case it is present; if it is not, the annihilation operator simply yields zero. In particular, any annihilation operator applied on the vacuum state gives zero:

$$a(\vec{p})|0\rangle = 0 \,, \qquad \forall \vec{p}. \qquad (4.11)$$

The proof of this result is straightforward: schematically, $(\Psi_N, a(\vec{p})\Psi_0) = (a(\vec{p})^\dagger \Psi_N, \Psi_0) = (\Psi_{N+1}, \Psi_0) = 0$. The vacuum state is the only state for twhich this happens.

One can determine quite easily the commutation relations of creation and annihilation operators. Indeed,

$$a(\vec{q}_2)^\dagger a(\vec{q}_1)^\dagger |\vec{p}_1,\dots,\vec{p}_N\rangle = a(\vec{q}_2)^\dagger |\vec{q}_1,\vec{p}_1,\dots,\vec{p}_N\rangle = |\vec{q}_2,\vec{q}_1,\vec{p}_1,\dots,\vec{p}_N\rangle = |\vec{q}_1,\vec{q}_2,\vec{p}_1,\dots,\vec{p}_N\rangle$$

$$= a(\vec{q}_1)^\dagger |\vec{q}_2,\vec{p}_1,\dots,\vec{p}_N\rangle = a(\vec{q}_1)^\dagger a(\vec{q}_2)^\dagger |\vec{p}_1,\dots,\vec{p}_N\rangle \,,$$

$$(4.12)$$

for any basis vector, and so

$$[a(\vec{p})^\dagger, a(\vec{q})^\dagger] = 0 \,. \qquad (4.13)$$

Taking the hermitian conjugate we find

$$[a(\vec{p}), a(\vec{q})] = 0 \,. \qquad (4.14)$$

On the other hand,

$$a(\vec{q}_2)a(\vec{q}_1)^\dagger |\vec{p}_1,\dots,\vec{p}_N\rangle = a(\vec{q}_2)|\vec{q}_1,\vec{p}_1,\dots,\vec{p}_N\rangle$$

$$= (2\pi)^3 2p^0 \delta^{(3)}(\vec{q}_1 - \vec{q}_2)|\vec{p}_1,\dots,\vec{p}_N\rangle + \sum_{j=1}^{N}(2\pi)^3 2p^0 \delta^{(3)}(\vec{p}_j - \vec{q}_2)|\vec{q}_1,\vec{p}_1,\dots,\vec{p}_{j-1},\vec{p}_{j+1},\vec{p}_N\rangle$$

$$= (2\pi)^3 2p^0 \delta^{(3)}(\vec{q}_1 - \vec{q}_2)|\vec{p}_1,\dots,\vec{p}_N\rangle + \sum_{j=1}^{N}(2\pi)^3 2p^0 \delta^{(3)}(\vec{p}_j - \vec{q}_2)a(\vec{q}_1)^\dagger|\vec{p}_1,\dots,\vec{p}_{j-1},\vec{p}_{j+1},\vec{p}_N\rangle$$

$$= (2\pi)^3 2p^0 \delta^{(3)}(\vec{q}_1 - \vec{q}_2)|\vec{p}_1,\dots,\vec{p}_N\rangle + a(\vec{q}_1)^\dagger a(\vec{q}_2)|\vec{p}_1,\dots,\vec{p}_N\rangle \,,$$

$$(4.15)$$

---

[27]Here, for clarity reasons, we use the notation $\Psi_{\vec{p},\vec{p}_1,\dots,\vec{p}_N}$ for the vector $|\vec{p}_1,\dots,\vec{p}_N\rangle$, and (Ψ', Ψ) for scalar products.

again for any basis vector, and therefore

$$[a(\vec{p}), a(\vec{q})^\dagger] = (2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{q}) \,. \tag{4.16}$$

This is consistent with our choice of normalisation. Indeed, using Eqs. (4.16) and (4.8),

$$
\begin{aligned}
\langle \vec{p}_1', \ldots, \vec{p}_M' | \vec{p}_1, \ldots, \vec{p}_N \rangle &= \langle 0 | a(\vec{p}_M') \ldots a(\vec{p}_1') a(\vec{p}_1)^\dagger \ldots a(\vec{p}_N)^\dagger | 0 \rangle \\
&= \sum_{j=1}^{N} (2\pi)^3 2p_j^0 \delta^{(3)}(\vec{p}_1' - \vec{p}_j) \langle 0 | a(\vec{p}_M') \ldots a(\vec{p}_2') a(\vec{p}_1)^\dagger \ldots a(\vec{p}_{j-1})^\dagger a(\vec{p}_{j+1})^\dagger \ldots a(\vec{p}_N)^\dagger | 0 \rangle \\
&= \delta_{MN} \sum_{P} \prod_{j=1}^{N} (2\pi)^3 2p_j^0 \delta^{(3)}(\vec{p}_j' - \vec{p}_{P(j)}) \langle 0 | 0 \rangle = \delta_{MN} \sum_{P} \prod_{j=1}^{N} (2\pi)^3 2p_j^0 \delta^{(3)}(\vec{p}_{P(j)}' - \vec{p}_j) \,.
\end{aligned}
\tag{4.17}
$$

The creation and annihilation operators can be used to build the number density operator $N(\vec{p}) = a(\vec{p})^\dagger a(\vec{p})$, which acts as follows on a basis vector:

$$
\begin{aligned}
N(\vec{p}) | \vec{p}_1, \ldots, \vec{p}_N \rangle &= \sum_{j=1}^{N} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}_j - \vec{p}) a(\vec{p})^\dagger | \vec{p}_1, \ldots, \vec{p}_{j-1}, \vec{p}_{j+1}, \vec{p}_N \rangle \\
&= \sum_{j=1}^{N} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}_j - \vec{p}) | \vec{p}_1, \ldots, \vec{p}_N \rangle \,.
\end{aligned}
\tag{4.18}
$$

Basis vectors are then eigenstates of the number operator $\mathcal{N}$,

$$\mathcal{N} = \int \frac{d^3p}{(2\pi)^3 2p^0} N(\vec{p}) = \int d\Omega_p N(\vec{p}) \,, \tag{4.19}$$

where we have introduced the notation $d\Omega_p$ for the Lorentz-invariant integration measure. Indeed,

$$\mathcal{N} | \vec{p}_1, \ldots, \vec{p}_N \rangle = \int \frac{d^3p}{(2\pi)^3 2p^0} \sum_{j=1}^{N} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}_j - \vec{p}) | \vec{p}_1, \ldots, \vec{p}_N \rangle = N | \vec{p}_1, \ldots, \vec{p}_N \rangle \,. \tag{4.20}$$

The number operator is then diagonal in Fock space, together with the momentum components.

It is worth observing at this point that we could work backwards, assuming an abstract point of view in which our purpose is that of representing the algebra of commutators

$$
\begin{aligned}
[a(\vec{p}), a(\vec{q})^\dagger] &= (2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{q}) \,, \\
[a(\vec{p}), a(\vec{q})] &= 0 \,, \\
[a(\vec{p})^\dagger, a(\vec{q})^\dagger] &= 0 \,.
\end{aligned}
\tag{4.21}
$$

This can be done taking a state $|0\rangle$ annihilated by all $a(\vec{p})$, $a(\vec{p})|0\rangle = 0$, assumed to be unique and normalised to 1, $\langle 0|0\rangle = 1$. Repeated application of the creation operators $a(\vec{p})^\dagger$ yields precisely the basis vectors of Fock space. This procedure yields therefore the Hilbert space appropriate for the description of non-interacting scalar particles.

Using now the number density operator we can define the total energy and total spatial momenta,

$$P^\mu = \int d\Omega_p p^\mu N(\vec{p}) \,. \tag{4.22}$$

55

One finds straightforwardly

$$P^\mu|\vec{p}_1,\ldots,\vec{p}_N\rangle = \int \frac{d^3p}{(2\pi)^3 2p^0} p^\mu \sum_{j=1}^{N}(2\pi)^3 2p^0 \delta^{(3)}(\vec{p}_j - \vec{p})|\vec{p}_1,\ldots,\vec{p}_N\rangle$$

$$= \left(\sum_{j=1}^{N}p_j^\mu\right)|\vec{p}_1,\ldots,\vec{p}_N\rangle\,. \tag{4.23}$$

Moreover, since

$$[N(\vec{p}), N(\vec{q})] = [a(\vec{p})^\dagger a(\vec{p}), a(\vec{q})^\dagger a(\vec{q})] = a(\vec{p})^\dagger[a(\vec{p}), a(\vec{q})^\dagger]a(\vec{q}) + a(\vec{q})^\dagger[a(\vec{p})^\dagger, a(\vec{q})]a(\vec{p})$$

$$= (2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{q})(a(\vec{p})^\dagger a(\vec{q}) - a(\vec{q})^\dagger a(\vec{p})) = 0\,, \tag{4.24}$$

we have that the momentum operators commute with each other (as it is obvious from the fact that they are simultaneously diagonal):

$$[P^\mu, P^\nu] = \int d\Omega_p \int d\Omega_q\, p^\mu q^\nu [N(\vec{p}), N(\vec{q})] = 0\,. \tag{4.25}$$

This formalism can be extended easily to several types of bosons, generalising the commutation relations Eq. (4.21) to

$$[a_i(\vec{p}), a_j(\vec{q})^\dagger] = \delta_{ij}(2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{q})\,,$$
$$[a_i(\vec{p}), a_j(\vec{q})] = 0\,, \tag{4.26}$$
$$[a_i(\vec{p})^\dagger, a_j(\vec{q})^\dagger] = 0\,.$$

It is clear the analogy between this formalism and a system of independent harmonic oscillators. It is a matter of simple algebra to show that defining the operators $\tilde{\phi}(\vec{p})$ and $\tilde{\pi}(\vec{p})$ via

$$a(\vec{p}) = (2\pi)^{\frac{3}{2}}[p^0\tilde{\phi}(\vec{p}) + i\tilde{\pi}(\vec{p})]\,, \qquad a(\vec{p})^\dagger = (2\pi)^{\frac{3}{2}}[p^0\tilde{\phi}(\vec{p}) - i\tilde{\pi}(\vec{p})]\,, \tag{4.27}$$

one has the following:

$$[\tilde{\phi}(\vec{p}), \tilde{\pi}(\vec{p})] = i\delta^{(3)}(\vec{p} - \vec{q})\,, \tag{4.28}$$

i.e., $\tilde{\phi}(\vec{p})$ and $\tilde{\pi}(\vec{p})$ have the same commutation relations of coordinate and momenta in quantum mechanics, and moreover

$$H = \int \frac{d^3p}{(2\pi)^3 2p^0} p^0 a(\vec{p})^\dagger a(\vec{p}) = \int d^3p \frac{1}{2}\left\{\tilde{\pi}(\vec{p})^2 + (p^0)^2\tilde{\phi}(\vec{p})^2 + p^0\delta^{(3)}(0)\right\}\,, \tag{4.29}$$

which coincides with the Hamiltonian of a system of infinitely many independent harmonic oscillators, up to an irrelevant (although infinite) constant.

Finally, let us discuss the transformation properties of creation and annihilation operators under translations and Lorentz transformations. These are inherited from the transformation properties of the states. For translations we have

$$U(a)|\vec{p}, \vec{p}_1,\ldots,\vec{p}_N\rangle = e^{-ia\cdot p}e^{-ia\cdot\sum_j p_j}|\vec{p}, \vec{p}_1,\ldots,\vec{p}_N\rangle = e^{-ia\cdot p}e^{-ia\cdot\sum_j p_j}a(\vec{p})^\dagger|\vec{p}_1,\ldots,\vec{p}_N\rangle$$

$$= U(a)a(\vec{p})^\dagger U(a)^\dagger U(a)|\vec{p}_1,\ldots,\vec{p}_N\rangle \tag{4.30}$$

$$= U(a)a(\vec{p})^\dagger U(a)^\dagger e^{-ia\cdot\sum_j p_j}|\vec{p}_1,\ldots,\vec{p}_N\rangle\,,$$

which then implies, since it holds for all states,[28] that

$$U(a)a(\vec{p})^{\dagger}U(a)^{\dagger} = e^{-ia\cdot p}a(\vec{p})^{\dagger}, \qquad U(a)a(\vec{p})U(a)^{\dagger} = e^{ia\cdot p}a(\vec{p}), \qquad (4.31)$$

or equivalently

$$U(a)^{\dagger}a(\vec{p})^{\dagger}U(a) = e^{ia\cdot p}a(\vec{p})^{\dagger}, \qquad U(a)^{\dagger}a(\vec{p})U(a) = e^{-ia\cdot p}a(\vec{p}). \qquad (4.32)$$

For Lorentz transformations,

$$\begin{aligned}
U(\Lambda)|\vec{p}, \vec{p}_1, \ldots, \vec{p}_N\rangle &= |\Lambda\vec{p}, \Lambda\vec{p}_1, \ldots, \Lambda\vec{p}_N\rangle = a(\Lambda\vec{p})^{\dagger}|\Lambda\vec{p}_1, \ldots, \Lambda\vec{p}_N\rangle \\
&= U(\Lambda)a(\vec{p})^{\dagger}U(\Lambda)^{\dagger}U(\Lambda)|\vec{p}_1, \ldots, \vec{p}_N\rangle \\
&= U(\Lambda)a(\vec{p})^{\dagger}U(\Lambda)^{\dagger}|\Lambda\vec{p}_1, \ldots, \Lambda\vec{p}_N\rangle,
\end{aligned} \qquad (4.33)$$

which then implies

$$U(\Lambda)a(\vec{p})^{\dagger}U(\Lambda)^{\dagger} = a(\Lambda\vec{p})^{\dagger}, \qquad U(\Lambda)a(\vec{p})U(\Lambda)^{\dagger} = a(\Lambda\vec{p}), \qquad (4.34)$$

or equivalently

$$U(\Lambda)^{\dagger}a(\vec{p})^{\dagger}U(\Lambda) = a(\Lambda^{-1}\vec{p})^{\dagger}, \qquad U(\Lambda)^{\dagger}a(\vec{p})U(\Lambda) = a(\Lambda^{-1}\vec{p}). \qquad (4.35)$$

4.2   Causal scalar fields

The discussion of the previous subsection completes the first part of the program outlined at the beginning of this section. To describe interactions, we could add to the Hamiltonian linear combinations of products of creation and annihilation operators. In this way one can build in principle any kind of Hamiltonian, which can then be adjusted to fit the experiments. More generally, one can build any observable out of the creation and annihilation operators. On the other hand, the most general Hamiltonian built in this way will violate locality and Poincaré symmetry, and it is not easy to monitor these properties (especially locality) in momentum space. Lorentz invariance of the theory and locality of interactions are most easily kept under control working in coordinate space, so our aim is now to develop a coordinate-space formalism equivalent to the formalism of creation and annihilation operators.

Coordinate space is also the appropriate setting to discuss the issue of causality. Consider two experimenters making measurements at spacetime points x and y, and suppose that they are measuring two local observables $\mathcal{O}(x)$ and $\mathcal{O}'(y)$, i.e., observables that are measured at a given point in time and space. For example, they could be using detectors to reveal the presence of a particle at a given point in space. One of the tenets of special relativity is that no information can travel faster than light. We then expect that measurements made at spacelike separated points (which are outside of each other's lightcone) do not affect each other. In other words, such measurements have to be compatible, in the sense that the corresponding Hilbert space operators commute. We can summarise this discussion in the *microcausality* postulate: for any local observables and any pair of spacelike separated points, we must have

$$[\mathcal{O}(x), \mathcal{O}'(y)] = 0, \quad (x-y)^2 < 0. \qquad (4.36)$$

---

[28]In deriving these results we make use of the fact that the vacuum is invariant under translations and Lorentz transformations.

The development of our coordinate-space formalism must take this requirement into account.

The construction of operators that are local in coordinate space and that have simple transformation properties under symmetry transformations is actually quite straightforward, and simply amounts to taking a Fourier transform. More precisely, we define the following operators,

$$\varphi_+(x) = \int d\Omega_p \, e^{-ip\cdot x} a(\vec{p}) \,, \qquad \varphi_-(x) = \int d\Omega_p \, e^{ip\cdot x} a(\vec{p})^\dagger \,. \tag{4.37}$$

Clearly $\varphi_\pm(x)^\dagger = \varphi_\mp(x)$. These relations can be inverted to give

$$a(\vec{p}) = \int d^3x \, e^{ip\cdot x} i\overleftrightarrow{\partial_0}\varphi_+(x) \,, \qquad a(\vec{p})^\dagger = \int d^3x \, e^{-ip\cdot x}(-i\overleftrightarrow{\partial_0})\varphi_-(x) \,, \tag{4.38}$$

independently of the choice of x^0. Indeed,

$$\int d^3x \, e^{ip\cdot x} i\overleftrightarrow{\partial_0}\varphi_+(x) = \int d^3x \int d\Omega_q \, e^{ip\cdot x} i \left\{ -iq^0 a(\vec{q})e^{-iq\cdot x} - ip^0 a(\vec{q})e^{-iq\cdot x} \right\}$$

$$= \int d^3x \int d\Omega_q (p^0 + q^0)a(\vec{q})e^{i(p-q)\cdot x} = \int d\Omega_q \, (2\pi)^3\delta^{(3)}(\vec{p} - \vec{q})(p^0 + q^0)a(\vec{q})e^{i(p^0-q^0)x^0} \tag{4.39}$$

$$= \int d\Omega_q \, (2\pi)^3 2q^0 a(\vec{q})\delta^{(3)}(\vec{p} - \vec{q}) = a(\vec{p}) \,.$$

The other formula is proved similarly.[29] One can also show that

$$\int d^3x \, e^{ip\cdot x} i\overleftrightarrow{\partial_0}\varphi_-(x) = \int d^3x \, e^{-ip\cdot x}(-i\overleftrightarrow{\partial_0})\varphi_+(x) = 0 \,. \tag{4.40}$$

The independence of time of these formulas is a consequence of a general result, which we now discuss in detail. First of all, notice that $\varphi_\pm(x)$ are solutions of the Klein-Gordon (KG) equation,

$$(\Box + m^2)\varphi_\pm(x) = 0 \,. \tag{4.41}$$

This is because the KG equation is linear, and $\varphi_\pm(x)$ are linear combinations (with operator-valued coefficients) of $e^{\pm ip\cdot x}$, which are solutions of the equation. For any two functions f, g one has that $\partial_0(f\overleftrightarrow{\partial_0}g) = f(\partial_0^2 g) - (\partial_0^2 f)g$. If f, g are solutions of the KG equation, then

$$\partial_0 \int d^3x \, f(x)\overleftrightarrow{\partial_0}g(x) = \int d^3x \, [f(x)(\partial_0^2 g(x)) - (\partial_0^2 f(x))g(x)]$$

$$= \int d^3x \, \{f(x)[(\vec{\nabla}^2 - m^2)g(x)] - [(\vec{\nabla}^2 - m^2)f(x)]g(x)\}$$

$$= \int d^3x \, [f(x)(\vec{\nabla}^2 g(x)) - (\vec{\nabla}^2 f(x))g(x)] = \int d^3x \, \vec{\nabla}[f(x)(\vec{\nabla}g(x)) - (\vec{\nabla}f(x))g(x)] \tag{4.42}$$

$$= \lim_{V\to\infty} \int_{\partial V} d\vec{n} \, [f(x)(\vec{\nabla}g(x)) - (\vec{\nabla}f(x))g(x)] = 0 \,.$$

In the last step we have assumed that f is normalisable and g is bounded. Take now for g the matrix elements of $\varphi_\pm(x)$, and for f some solution of the KG equation $f_{p,\varepsilon}(x) = \int d\Omega_q \, \tilde{f}_{p,\varepsilon}(\vec{q})e^{\mp iq\cdot x}$

---

[29]Here it would suffice to take the Hermitian conjugate of this formula.

with $\tilde{f}_{p,\varepsilon}(\vec{q})$ strongly peaked around $\vec{p}$ and becoming a delta as ε goes to zero. Since the relation Eq. (4.42) holds for all ε, taking $\varepsilon \to 0$ we show that all the matrix elements of the time derivative of the operators on the right-hand side of the relations in Eq. (4.38) are zero, which finally proves that they are time-independent.

To understand what the operators $\varphi_{\pm}(x)$ do, consider a wave packet,

$$|\psi\rangle = \int d\Omega_p \, \tilde{f}(\vec{p})|\vec{p}\rangle \,, \qquad 1 = \langle\psi|\psi\rangle = \int d\Omega_p \, |\tilde{f}(\vec{p})|^2 \,. \tag{4.43}$$

A little manipulation shows that

$$
\begin{aligned}
|\psi\rangle = \int d\Omega_p \, \tilde{f}(\vec{p})|\vec{p}\rangle = \int d\Omega_p \, \tilde{f}(\vec{p})a(\vec{p})^\dagger|0\rangle &= \int d\Omega_p \, \tilde{f}(\vec{p}) \int d^3x \, e^{-ip\cdot x}(-i\overleftrightarrow{\partial_0})\varphi_-(x)|0\rangle \\
&= \int d^3x \left(\int d\Omega_p \, \tilde{f}(\vec{p})e^{-ip\cdot x} \right) (-i\overleftrightarrow{\partial_0})\varphi_-(x)|0\rangle \\
&= \int d^3x \, f(x)(-i\overleftrightarrow{\partial_0})\varphi_-(x)|0\rangle \,,
\end{aligned}
\tag{4.44}
$$

where $f(x) = \int d\Omega_p \, \tilde{f}(\vec{p})e^{-ip\cdot x}$ is a solution of the Klein-Gordon equation, $(\Box + m^2)f(x) = 0$. The main contribution to $f(x)$ comes from momenta that make the phase of the integrand stationary (thus reducing the effect of cancellations due to the oscillatory nature of the integrand). Setting $\tilde{f}(\vec{p}) = |\tilde{f}(\vec{p})|e^{-i\phi(\vec{p})}$, the phase is stationary when

$$0 = \vec{\nabla}_p[\phi(\vec{p}) + p\cdot x] = \vec{\nabla}_p\phi(\vec{p}) + \frac{\vec{p}}{p^0}x^0 - \vec{x} \,. \tag{4.45}$$

On the other hand, if we take a wave packet strongly peaked around some momentum $\vec{p}_*$, the integral will receive contributions only from momenta near $\vec{p}_*$. This means that the function $f(x)$ will be peaked at values of x where $\vec{p}_*$ solves the stationary phase equation, i.e., on the straight-line trajectory

$$\vec{x} = \frac{\vec{p}_*}{p_*^0}x^0 + \vec{\nabla}_p\phi(\vec{p}_*) = \vec{\beta}_*x^0 + \vec{x}_* \,. \tag{4.46}$$

The function $f(x)$ therefore describes the evolution of the wave packet in coordinate space. We then conclude that the operator

$$\varphi_-^f(x) = \int d^3x \, f(x)(-i\overleftrightarrow{\partial_0})\varphi_-(x) \tag{4.47}$$

creates a wave packet with coordinate-space wave function $f(x)$, and thus $\varphi_-(x)$ essentially creates a particle at x. Acting on $|\psi\rangle$ with the operator

$$\varphi_+^f(x) = \int d^3x \, f(x)^* i\overleftrightarrow{\partial_0}\varphi_+(x) = \int d^3x \int d\Omega_p \, \tilde{f}(\vec{p})^* e^{ip\cdot x} i\overleftrightarrow{\partial_0}\varphi_+(x) = \int d\Omega_p \, \tilde{f}(\vec{p})^* a(\vec{p}) \,, \tag{4.48}$$

we find that

$$\varphi_+^f(x)|\psi\rangle = \int d\Omega_p \, \tilde{f}(\vec{p})^* a(\vec{p}) \int d\Omega_q \, \tilde{f}(\vec{q})a(\vec{q})^\dagger|0\rangle = \int d\Omega_p \, |\tilde{f}(\vec{p})|^2|0\rangle = |0\rangle \,, \tag{4.49}$$

so it destroys the wave packet with coordinate-space wave function $f(x)$, and thus $\varphi_+(x)$ essentially destroys a particle at x.

The transformation properties of $\varphi_\pm(x)$ are easily obtained from Eq. (4.35), and read

$$
\begin{aligned}
U(\Lambda)^\dagger \varphi_\pm(x) U(\Lambda) &= \int d\Omega_p\, U(\Lambda)^\dagger a(\vec{p})^{(\dagger)} U(\Lambda) e^{\mp ip\cdot x} = \int d\Omega_p\, a(\Lambda^{-1}\vec{p})^{(\dagger)} e^{\mp ip\cdot x} \\
&= \int d\Omega_p\, a(\vec{p})^{(\dagger)} e^{\mp i\Lambda p\cdot x} = \int d\Omega_p\, a(\vec{p})^{(\dagger)} e^{\mp ip\cdot \Lambda^{-1}x} = \varphi_\pm(\Lambda^{-1}x)\,, \\
U(a)^\dagger \varphi_\pm(x) U(a) &= \int d\Omega_p\, U(a)^\dagger a(\vec{p})^{(\dagger)} U(a) e^{\mp ip\cdot x} = \int d\Omega_p\, a(\vec{p})^{(\dagger)} e^{\mp ip\cdot a} e^{\mp ip\cdot x} \\
&= \varphi_\pm(x+a)\,.
\end{aligned}
\tag{4.50}
$$

4.2.1 Microcausality

Since all the observables can be built from the creation and annihilation operators, multiplying them by some function of the momenta and integrating over momenta, they can equivalently be built out of $\varphi_\pm(x)$ and its derivatives. The most general local observable will thus be a sum of products of $\varphi_\pm(x)$ and their derivatives. Using independently both $\varphi_+(x)$ and $\varphi_-(x)$ however leads to troubles in satisfying the microcausality condition, since their commutator does not vanish for spacelike separations. To see this, compute first

$$
\begin{aligned}
[\varphi_+(x), \varphi_-(y)] &= \int d\Omega_p \int d\Omega_q\, e^{-ip\cdot x} e^{iq\cdot y} [a(\vec{p}), a(\vec{q})^\dagger] \\
&= \int d\Omega_p \int d\Omega_q\, e^{-ip\cdot x} e^{iq\cdot y} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}-\vec{q}) \\
&= \int d\Omega_p\, e^{-ip\cdot(x-y)} \equiv \Delta(x-y)\,.
\end{aligned}
\tag{4.51}
$$

Since $d\Omega_p$ is a Lorentz-invariant integration measure, $\Delta(x)$ is a Lorentz-invariant function of its argument, and can be evaluated in any reference frame. For spacelike separations, $(x-y)^2 < 0$, there is always a reference frame in which $x'^0 = y'^0$, and in this frame one sees immediately that interchanging $\vec{x}'$ and $\vec{y}'$ can be compensated by a change of variables $\vec{p} \to -\vec{p}$. Since $\Delta(x)$ is Lorentz-invariant, one concludes that for spacelike separations it is an even function of its argument, $\Delta(x-y) = \Delta(y-x)$ if $(x-y)^2 < 0$. However, this function does not vanish in general. For future utility, we introduce the following notation for the equal-time commutator, $[A(x), B(y)]_{\mathrm{ET}} = [A(x^0, \vec{x}), B(y^0 = x^0, \vec{y})]$. On the other hand, the commutators $[\varphi_\pm(x), \varphi_\pm(y)] = 0$ identically.

We cannot use $\varphi_\pm(x)$ independently, but of course we do not want to use only one of them, since that would certainly make it impossible to built all the observables. We then seek for a linear combination of $\varphi_\pm(x)$ with nonvanishing coefficients, such that the microcausality conditions are satisfied. Let us set

$$
\varphi(x) = \alpha_+ \varphi_+(x) + \alpha_- \varphi_-(x)\,,
\tag{4.52}
$$

and find what relations $\alpha_\pm$ have to satisfy in order to not violate microcausality. Let us consider first the commutator of $\varphi(x)$ with itself and its Hermitian conjugate. in general

$$
\begin{aligned}
[\varphi(x), \varphi(y)] &= \alpha_+ \alpha_- \left([\varphi_+(x), \varphi_-(y)] + [\varphi_-(x), \varphi_+(y)]\right) \\
&= \alpha_+ \alpha_- \left([\varphi_+(x), \varphi_-(y)] - [\varphi_+(y), \varphi_-(x)]\right)\,.
\end{aligned}
\tag{4.53}
$$

We then find for spacelike separations

$$
[\varphi(x), \varphi(y)] = \alpha_+ \alpha_- \left[\Delta(x-y) - \Delta(y-x)\right] = 0\,,
\tag{4.54}
$$

for any $\alpha_\pm$. Similarly, $[\varphi(x)^\dagger, \varphi(y)^\dagger] = 0$. Instead,

$$[\varphi(x), \varphi(y)^\dagger] = |\alpha_+|^2 [\varphi_+(x), \varphi_-(y)] - |\alpha_-|^2 [\varphi_-(x), \varphi_+(y)], \qquad (4.55)$$

and for spacelike separations

$$[\varphi(x), \varphi(y)^\dagger] = |\alpha_+|^2 \Delta(x - y) - |\alpha_-|^2 \Delta(y - x) = (|\alpha_+|^2 - |\alpha_-|^2)\Delta(x - y). \qquad (4.56)$$

To make this vanish for all spacelike x, y we need to set $|\alpha_+| = |\alpha_-|$. Parameterising $\alpha_\pm = |\alpha| e^{i\zeta} e^{\pm i\kappa}$, we see that up to an irrelevant overall factor we have to take φ of the form

$$\varphi_\kappa(x) = e^{i\kappa} \varphi_+(x) + e^{-i\kappa} \varphi_-(x). \qquad (4.57)$$

However, in building our observables we cannot use at the same time $\varphi_\kappa(x)$ and some other $\varphi_{\kappa'}(x)$, because that would be equivalent to use both $\varphi_\pm(x)$, which is troublesome for causality. Finally, since we can always redefine our particle states by a phase without changing the physics, we can reabsorb $e^{i\kappa} a(\vec{p}) \to a(\vec{p})$. In conclusion, the appropriate local object to use here, in order to satisfy automatically the microcausality condition, is the *Hermitian scalar field* $\varphi(x) = \varphi(x)^\dagger$,

$$\varphi(x) = \varphi_+(x) + \varphi_-(x) = \int d\Omega_p \left(a(\vec{p}) e^{-ip\cdot x} + a(\vec{p})^\dagger e^{ip\cdot x} \right). \qquad (4.58)$$

The fields $\varphi_+(x)$ and $\varphi_-(x)$ are called the *positive-frequency component* and the *negative-frequency component* of the field.

So far we have checked the microcausality condition only for the field itself, but in general we want to use also its spatial and temporal derivatives. Spatial derivatives are treated most easily. In fact, since by definition

$$\partial_i \varphi(y) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [\varphi(y + \varepsilon \hat{\imath}) - \varphi(y)], \qquad (4.59)$$

we have for equal-time commutators

$$[\varphi(x), \partial_i \varphi(y)]_{\mathrm{ET}} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [\varphi(x), \varphi(y + \varepsilon \hat{\imath}) - \varphi(y)]_{\mathrm{ET}} = 0, \qquad (4.60)$$

which then extends to general spacelike x, y by covariance. This is easily generalised to any order, and we can write

$$[\partial_i^n \varphi(x), \partial_j^m \varphi(y)] = 0, \quad (x - y)^2 < 0. \qquad (4.61)$$

This trick does not work for temporal derivatives, since in that case the commutator of the field at different times is involved.[30] We can nevertheless compute the commutator explicitly, obtaining

$$[\varphi(x), \partial_0 \varphi(y)] = [\varphi_+(x), \partial_0 \varphi_-(y)] - [\partial_0 \varphi_+(y), \varphi_-(x)]. \qquad (4.62)$$

The basic object is $[\varphi_+(x), \partial_0 \varphi_-(y)]$ which reads

$$[\varphi_+(x), \partial_0 \varphi_-(y)] = \int d\Omega_p \int d\Omega_q \, e^{-ip\cdot x} e^{iq\cdot y} iq^0 [a(\vec{p}), a(\vec{q})^\dagger] = i \int d\Omega_p \, p^0 e^{-ip\cdot(x-y)}. \qquad (4.63)$$

---

[30] Actually, if x and y are spacelike-separated, then one can always find ϵ_0 small enough such that x and $y + \varepsilon \hat{\mu}$ are still spacelike-separated for all $\epsilon < \epsilon_0$, even for $\mu = 0$. This cannot be done when x and y are the same point, in which case for any ϵ the points x and $x + \epsilon \hat{0}$ are timelike-separated.

More generally,

$$[\varphi_+(x), \partial_\mu \varphi_-(y)] = i \int d\Omega_p \, p_\mu e^{-ip\cdot(x-y)} = \partial_{y\mu} \int d\Omega_p \, e^{-ip\cdot(x-y)} = \partial_{y\mu}\Delta(x-y)\,, \qquad (4.64)$$

and so

$$[\varphi(x), \partial_\mu \varphi(y)] = [\varphi_+(x), \partial_\mu \varphi_-(y)] - [\partial_\mu \varphi_+(y), \varphi_-(x)] = \partial_{y\mu}(\Delta(x-y) - \Delta(y-x))\,, \qquad (4.65)$$

since the second term is just the Hermitian conjugate of the first one and so of Eq. (4.64). This object is a Lorentz vector, and so if it vanishes in one reference frame it will vanish in any other. For spacelike separations we can then choose the frame where $x^0 = y^0$, and find

$$[\varphi_+(x), \partial_0 \varphi_-(y)]_{\mathrm{ET}} = i \int d\Omega_p \, p^0 e^{i\vec{p}\cdot(\vec{x}-\vec{y})} = \frac{i}{2}\delta^{(3)}(\vec{x}-\vec{y})\,. \qquad (4.66)$$

The commutator $[\partial_0 \varphi_-(y), \varphi_+(x)]_{\mathrm{ET}}$ is just the Hermitian conjugate of Eq. (4.66), and so we find

$$[\varphi(x), \partial_0 \varphi(y)]_{\mathrm{ET}} = i\delta^{(3)}(\vec{x}-\vec{y})\,. \qquad (4.67)$$

At equal times, spacelike separation means $\vec{x} \neq \vec{y}$, and in this case the commutator vanishes. We consider next

$$[\partial_0 \varphi(x), \partial_0 \varphi(y)] = [\partial_0 \varphi_+(x), \partial_0 \varphi_-(y)] - [\partial_0 \varphi_+(y), \partial_0 \varphi_-(x)]\,. \qquad (4.68)$$

The basic object is $[\partial_0 \varphi_+(x), \partial_0 \varphi_-(y)]$ which reads

$$\begin{aligned}
[\partial_0 \varphi_+(x), \partial_0 \varphi_-(y)] &= \int d\Omega_p \int d\Omega_q \, e^{-ip\cdot x} e^{iq\cdot y}(-ip^0)iq^0[a(\vec{p}), a(\vec{q})^\dagger] \\
&= \int d\Omega_p \,(p^0)^2 e^{-ip\cdot(x-y)} = \tilde{\Delta}(x-y)\,.
\end{aligned} \qquad (4.69)$$

More generally, we have

$$\begin{aligned}
[\partial_\mu \varphi_+(x), \partial_\nu \varphi_-(y)] &= \int d\Omega_p \int d\Omega_q \, e^{-ip\cdot x} e^{iq\cdot y}(-ip_\mu)iq_\nu[a(\vec{p}), a(\vec{q})^\dagger] \\
&= \int d\Omega_p \, p_\mu p_\nu e^{-ip\cdot(x-y)} = \partial_{x\mu}\partial_{y\nu}\Delta(x-y)\,,
\end{aligned} \qquad (4.70)$$

and so

$$[\partial_\mu \varphi(x), \partial_\nu \varphi(y)] = [\partial_\mu \varphi_+(x), \partial_\nu \varphi_-(y)] - [\partial_\nu \varphi_+(y), \partial_\mu \varphi_-(x)] = \partial_{x\mu}\partial_{y\nu}[\Delta(x-y) - \Delta(y-x)]\,. \qquad (4.71)$$

We can then rest assured that if Eq. (4.71) vanishes in a frame, it will do so in any other. For spacelike separations we then choose the frame with $x^0 = y^0$ and find that $\tilde{\Delta}(x-y)|_{x^0=y^0} = \partial_{x0}\partial_{y0}\Delta(x-y)|_{x^0=y^0}$ is symmetric under exchange of $\vec{x}$ and $\vec{y}$. From this we conclude that

$$[\partial_0 \varphi(x), \partial_0 \varphi(y)]_{\mathrm{ET}} = \tilde{\Delta}(x-y)|_{x^0=y^0} - \tilde{\Delta}(y-x)|_{x^0=y^0} = 0\,. \qquad (4.72)$$

Finally, since $\varphi(x)$ satisfies the KG equation, commutators involving time derivatives of fields of order 2 or higher can be reduced to combinations of commutators involving no more than single

time derivatives of fields. This is because we can always replace $\partial_0^2\varphi = (\vec{\nabla}^2 - m^2)\varphi$. We can then conclude that also these commutators will satisfy microcausality, as they are in general a combination of spatial delta functions and their derivatives. We can summarise our findings in the following equations:

$$[\varphi(x), \varphi(y)]_{\mathrm{ET}} = 0\,, \qquad\qquad [\partial_0\varphi(x), \partial_0\varphi(y)]_{\mathrm{ET}} = 0\,,$$
$$[\varphi(x), \partial_0\varphi(y)]_{\mathrm{ET}} = i\delta^{(3)}(\vec{x} - \vec{y})\,, \tag{4.73}$$

and

$$(\Box + m^2)\varphi(x) = 0\,. \tag{4.74}$$

Together, they guarantee that any commutator invoving fields and its derivatives will vanish for spacelike separations. As a final remark, notice that the commutation relations found above remains unchanged for any of the fields φ_κ, since the phase factors always cancel out. Therefore, any of the φ_κ is a good causal field, but we cannot use more than one at a time.

We conclude this subsection presenting the transformation properties of the field, which are easily obtained from those of its positive and negative frequency components, and read

$$U(\Lambda)^\dagger \varphi(x) U(\Lambda) = \varphi(\Lambda^{-1}x)\,,$$
$$U(a)^\dagger \varphi(x) U(a) = \varphi(x + a)\,. \tag{4.75}$$

4.2.2 Energy and momentum operators in terms of fields: normal ordering

Let us now express the basic observables, namely energy and momentum, in terms of the field and its derivatives. We start by reformulating Eq. (4.38) in terms of the full field,

$$a(\vec{p}) = \int d^3x\, e^{ip\cdot x} i\overleftrightarrow{\partial_0}\varphi(x)\,, \qquad a(\vec{p})^\dagger = \int d^3x\, e^{-ip\cdot x}(-i\overleftrightarrow{\partial_0})\varphi(x)\,, \tag{4.76}$$

where we made use of Eq. (4.40). Next, we use this to write

$$a(\vec{p})^\dagger a(\vec{p}) = \int d^3x \int d^3y\, \left[e^{-ip\cdot x}(-i\overleftrightarrow{\partial_0})\varphi(x)\right]\left[e^{ip\cdot y}i\overleftrightarrow{\partial_0}\varphi(y)\right]$$
$$= \int d^3x \int d^3y\, e^{i\vec{p}\cdot(\vec{x}-\vec{y})}\left[(p^0)^2\varphi(x)\varphi(y) + \dot\varphi(x)\dot\varphi(y) + ip^0\left(\varphi(x)\dot\varphi(y) - \dot\varphi(x)\varphi(y)\right)\right]\,,$$
$$a(\vec{p})a(\vec{p})^\dagger$$
$$= \int d^3x \int d^3y\, e^{i\vec{p}\cdot(\vec{x}-\vec{y})}\left[(p^0)^2\varphi(y)\varphi(x) + \dot\varphi(y)\dot\varphi(x) - ip^0\left(\varphi(y)\dot\varphi(x) - \dot\varphi(y)\varphi(x)\right)\right]\,, \tag{4.77}$$

where $\dot\varphi(x) = \partial_0\varphi(x)$ and it is understood that $x^0 = y^0$. The first expression above could be integrated over momenta after multiplying with p^0 to yield the Hamiltonian H. The last term in square brackets is however not very appealing, and it would be nice to find a way to remove it. Since φ and $\dot\varphi$ commute with themselves for spacelike separations, the first two terms in square brackets are the same for both expressions. The last term differs only by a sign and by the order in which x and y appear, which can be reversed since they are being integrated over, with the consequence that the sign of $\vec{p}$ will change in the phase factor, i.e.

$$a(\vec{p})a(\vec{p})^\dagger$$
$$= \int d^3x \int d^3y\, \left\{e^{i\vec{p}\cdot(\vec{x}-\vec{y})}\left[(p^0)^2\varphi(x)\varphi(y) + \dot\varphi(x)\dot\varphi(y)\right] - ip^0 e^{-i\vec{p}\cdot(\vec{x}-\vec{y})}\left(\varphi(x)\dot\varphi(y) - \dot\varphi(x)\varphi(y)\right)\right\}\,. \tag{4.78}$$

If we now take the sum of the two expressions, multiply by p^0 and integrate over momenta with the invariant measure we get

$$\int d\Omega_p \, \tfrac{1}{2}\{a(\vec{p}), a(\vec{p})^\dagger\} = \int d\Omega_p \int d^3x \int d^3y \, e^{i\vec{p}\cdot(\vec{x}-\vec{y})} p^0 \left[(p^0)^2 \varphi(x)\varphi(y) + \dot\varphi(x)\dot\varphi(y)\right] , \quad (4.79)$$

having dropped the term proportional to p^0 since it is odd under $\vec{p} \to -\vec{p}$. A simple manipulation gives us

$$\int d\Omega_p \, p^0 \tfrac{1}{2}\{a(\vec{p}), a(\vec{p})^\dagger\} = \int d\Omega_p \int d^3x \int d^3y \, e^{i\vec{p}\cdot(\vec{x}-\vec{y})} p^0 \left[(\vec{p}^2 + m^2)\varphi(x)\varphi(y) + \dot\varphi(x)\dot\varphi(y)\right]$$

$$= \int d\Omega_p \int d^3x \int d^3y \, p^0 \left[\varphi(x)\varphi(y)(\vec{\nabla}_x \cdot \vec{\nabla}_y + m^2)e^{i\vec{p}\cdot(\vec{x}-\vec{y})} + e^{i\vec{p}\cdot(\vec{x}-\vec{y})}\dot\varphi(x)\dot\varphi(y)\right]$$

$$= \int d\Omega_p \int d^3x \int d^3y \, e^{i\vec{p}\cdot(\vec{x}-\vec{y})} p^0 \left[\dot\varphi(x)\dot\varphi(y) + \vec{\nabla}\varphi(x) \cdot \vec{\nabla}\varphi(y) + m^2\varphi(x)\varphi(y)\right]$$

$$= \int d^3x \int d^3y \, \delta^{(3)}(\vec{x} - \vec{y})\tfrac{1}{2} \left[\dot\varphi(x)\dot\varphi(y) + \vec{\nabla}\varphi(x) \cdot \vec{\nabla}\varphi(y) + m^2\varphi(x)\varphi(y)\right]$$

$$= \int d^3x \, \tfrac{1}{2} \left[\dot\varphi(x)^2 + \vec{\nabla}\varphi(x)^2 + m^2\varphi(x)^2\right] .$$

$$(4.80)$$

The left-hand side is almost H, where unfortunately the "almost" involves an infinite constant:

$$\int d\Omega_p \, p^0 \tfrac{1}{2}\{a(\vec{p}), a(\vec{p})^\dagger\} = \int d\Omega_p \, p^0 \left\{a(\vec{p})^\dagger a(\vec{p}) + \tfrac{1}{2}(2\pi)^3 2p^0 \delta^{(3)}(0)\right\} . \quad (4.81)$$

This is not a big deal from the physical point of view: in any case, what we can measure are energy differences, so we might as well remove the (infinite) constant contribution. Mathematically, however, our expressions must make good sense. To this end, we define the *normal ordered product* of fields as follows. Expand the fields in creation and annihilation operators, and write the product of fields as an integral over momenta of a sum of terms, each of which is a product of creation and annihilation operators. The normal ordered product is obtained by replacing each of these products with the product obtained using the same creation and annihilation operators, but with all creation operators being placed to the *left* of the annihilation operators. This defines the normal ordered product of creation and annihilation operators. Equivalently, the normal ordere product of fields is obtained by placing all the negative frequency components on the left of the positive frequency ones. A few examples will be helpful. We denote the normal ordered product by placing colons at the sides, i.e., $: \ldots :$. Then

$$: a(\vec{p})^\dagger a(\vec{q}) : \,= a(\vec{p})^\dagger a(\vec{q}) , \qquad\qquad : a(\vec{p})a(\vec{q})^\dagger : \,= a(\vec{q})^\dagger a(\vec{p}) ,$$
$$: a(\vec{p})a(\vec{q}) : \,= a(\vec{p})a(\vec{q}) = a(\vec{q})a(\vec{p}) , \qquad : a(\vec{p})^\dagger a(\vec{q})^\dagger : \,= a(\vec{p})^\dagger a(\vec{q})^\dagger = a(\vec{q})^\dagger a(\vec{p})^\dagger ,$$

$$(4.82)$$

and similarly for products of more operators, while for fields

$$: \varphi(x)\varphi(y) : \,=: [\varphi_+(x) + \varphi_-(x)][\varphi_+(y) + \varphi_-(y)] :$$
$$= \varphi_+(x)\varphi_+(y) + \varphi_-(x)\varphi_+(y) + \varphi_-(y)\varphi_+(x) + \varphi_-(x)\varphi_-(y) . \quad (4.83)$$

It is worth noting that

$$: \varphi(x)\varphi(y) : \,= \varphi(x)\varphi(y) - [\varphi_+(x), \varphi_-(y)] = \varphi(x)\varphi(y) - \langle 0|[\varphi_+(x), \varphi_-(y)]|0\rangle$$
$$= \varphi(x)\varphi(y) - \langle 0|\varphi_+(x)\varphi_-(y)|0\rangle = \varphi(x)\varphi(y) - \langle 0|\varphi(x)\varphi(y)|0\rangle , \quad (4.84)$$

64

where we have used the fact that $[\varphi_+(x), \varphi_-(y)]$ is a c-number to replace it with its vacuum expectation value, and then used the fact that φ_+ annihilates the vacuum. It is also worth noting that the order of terms inside a normal product is irrelevant, as the normal ordering always reorders the various positive and negative frequency components in the same way, i.e., $: \varphi(x)\varphi(y) :=: \varphi(y)\varphi(x) :$. If we now use normal ordering on the integrand on both sides of Eq. (4.80), we find

$$
\begin{aligned}
H &= \int d\Omega_p\, p^0 a(\vec{p})^\dagger a(\vec{p})^\dagger = \int d\Omega_p\, p^0 \tfrac{1}{2} : \{a(\vec{p}), a(\vec{p})^\dagger\} : \\
&= \int d^3x\, \frac{1}{2} : \left[\dot{\varphi}(x)^2 + \vec{\nabla}\varphi(x)^2 + m^2\varphi(x)^2\right] : .
\end{aligned}
\tag{4.85}
$$

Given Eq. (4.84), valid for products of two fields, we also have

$$
H = \int d^3x\, \frac{1}{2} \left\{ \left[\dot{\varphi}(x)^2 + \vec{\nabla}\varphi(x)^2 + m^2\varphi(x)^2\right] - \langle 0|\dot{\varphi}(x)^2 + \vec{\nabla}\varphi(x)^2 + m^2\varphi(x)^2|0\rangle \right\} . \tag{4.86}
$$

If we consider now the spatial momentum operator, we have that

$$
\begin{aligned}
\vec{P} &= \int d\Omega_p\, \vec{p}\, a(\vec{p})^\dagger a(\vec{p})^\dagger = \int d\Omega_p\, \vec{p}\, \tfrac{1}{2} : \{a(\vec{p}), a(\vec{p})^\dagger\} : \\
&= \int d\Omega_p \int d^3x \int d^3y\, \vec{p}\, e^{i\vec{p}\cdot(\vec{x}-\vec{y})} i p^0 \tfrac{1}{2} : \varphi(x)\dot{\varphi}(y) - \dot{\varphi}(x)\varphi(y) - \varphi(y)\dot{\varphi}(x) + \dot{\varphi}(y)\varphi(x) :,
\end{aligned}
\tag{4.87}
$$

while the other term drops since it is even under exchange of x and y, and so after x and y integration gives a term which is even in $\vec{p}$. Using the symmetry properties of the integrand, and the fact that the order is irrelevant inside normal ordered products, we find

$$
\begin{aligned}
\vec{P} &= \int d\Omega_p \int d^3x \int d^3y\, \vec{p}\, e^{i\vec{p}\cdot(\vec{x}-\vec{y})} i p^0 : \varphi(x)\dot{\varphi}(y) - \dot{\varphi}(x)\varphi(y) : \\
&= \int d\Omega_p \int d^3x \int d^3y\, \vec{p}\, e^{i\vec{p}\cdot(\vec{x}-\vec{y})} i p^0 : \varphi(x)\dot{\varphi}(y) - \varphi(y)\dot{\varphi}(x) : \\
&= \int d\Omega_p \int d^3x \int d^3y\, \vec{\nabla}_x e^{i\vec{p}\cdot(\vec{x}-\vec{y})} 2p^0 : \varphi(x)\dot{\varphi}(y) : \\
&= -\int d^3x \int d^3y\, \delta^{(3)}(\vec{x}-\vec{y}) : \dot{\varphi}(y)\vec{\nabla}_x\varphi(x) := -\int d^3x : \dot{\varphi}(x)\vec{\nabla}\varphi(x) : .
\end{aligned}
\tag{4.88}
$$

The need to use normal ordering comes from a general problem of quantum fields, namely that their products at coinciding spacetime points are ill-defined. To see this, consider the vacuum expectation value of the product $\varphi(x)\varphi(y)$. One finds easily that

$$
\langle 0|\varphi(x)\varphi(y)|0\rangle = \int d\Omega_p e^{-ip\cdot(x-y)} , \tag{4.89}
$$

which is badly divergent as $x \to y$. On the other hand, the vacuum expectation value of the normal ordered product is zero.

4.2.3 Hamiltonian and Lagrangian formalism

The equal-time commutation relations Eq. (4.73) are precisely the commutation relations of a Hamiltonian system with infinitely many degrees of freedom, with $\varphi(t, \vec{x})$ for all $\vec{x}$ playing the

role of canonical coordinates, and $\dot{\varphi}(t,\vec{x}) \equiv \pi(t,\vec{x})$ for all $\vec{x}$ that of their conjugated momenta. The Klein-Gordon equation can be recast as an equation of motion for π,

$$0 = (\Box + m^2)\varphi = (\partial_0^2 - \vec{\nabla}^2 + m^2)\varphi \Rightarrow \dot{\pi} = (\vec{\nabla}^2 - m^2)\varphi \,. \tag{4.90}$$

In terms of φ and π the Hamiltonian reads

$$H = \int d^3x\, \mathcal{H} = \int d^3x\, \frac{1}{2} : \left[\pi^2 + \vec{\nabla}\varphi^2 + m^2\varphi^2\right] := \int d^3x\, \frac{1}{2} : \left[\pi^2 - \varphi\vec{\nabla}^2\varphi + m^2\varphi^2\right] : \,. \tag{4.91}$$

The associated Hamilton equations are

$$\dot{\varphi} = \frac{\delta H}{\delta \pi} = \pi \,,$$
$$\dot{\pi} = -\frac{\delta H}{\delta \varphi} = (\vec{\nabla}^2 - m^2)\varphi \,, \tag{4.92}$$

which are precisely the equations satisfied by φ and π. This Hamiltonian system can be converted into a Lagrangian system by means of a Legendre transform, as usual, with the added detail that fields are taken to be normal-ordered. We find for the Lagrangian

$$\mathcal{L} =: \pi\dot{\varphi} : -\mathcal{H} = \frac{1}{2} : \left[\dot{\varphi}^2 - \vec{\nabla}\varphi \cdot \vec{\nabla}\varphi - m^2\varphi^2\right] := \frac{1}{2} : \partial_\mu\varphi\partial^\mu\varphi - m^2\varphi^2 : \,. \tag{4.93}$$

Let us now reverse the direction. Take the classical Klein-Gordon Lagrangian,

$$\mathcal{L} = \frac{1}{2}\left[\partial_\mu\varphi\partial^\mu\varphi - m^2\varphi^2\right] \,, \tag{4.94}$$

where classically φ is a real field, and define the action functional as usual. We set up the procedure of *canonical quantisation* through the following steps:

- find the EOM via an action principle (these are just the KG equation, in this case);

- define the canonical momentum as $\pi = \frac{\partial \mathcal{L}}{\partial(\partial_0\varphi)}$ (this is just $\dot{\varphi}$ in this case);

- find a *Hermitian operator*-valued field $\varphi(x)$ that solves the EOM;

- impose the *canonical commutation relations* (CCR),

$$[\varphi(x), \pi(y)]_{\mathrm{ET}} = i\delta^{(3)}(\vec{x} - \vec{y}) \,, \qquad [\varphi(x), \varphi(y)]_{\mathrm{ET}} = [\pi(x), \pi(y)]_{\mathrm{ET}} = 0 \,. \tag{4.95}$$

We already now that the scalar field built in the previous subsections is a solution of this procedure, but it is instructive to see how this comes about following the procedure step by step. We first solve the KG equation: this is most easily done in momentum space, setting

$$\varphi(x) = \int \frac{d^4p}{(2\pi)^4}\, e^{-ip\cdot x}\tilde{\varphi}(p) \,. \tag{4.96}$$

The KG equation for $\tilde{\varphi}$ is simply $(p^2 - m^2)\tilde{\varphi} = 0$, which is solved by $\tilde{\varphi}(p) = 2\pi\delta(p^2 - m^2)A(p^0, \vec{p})$, for an arbitrary operator-valued A. Plugging this into Eq. (4.96) we find

$$\varphi(x) = \int \frac{d^4p}{(2\pi)^3}\, e^{-ip\cdot x}\frac{1}{2p^0}\left[\delta(p^0 - \sqrt{\vec{p}^2 + m^2}) + \delta(p^0 + \sqrt{\vec{p}^2 + m^2})\right]A(p^0, \vec{p})$$

$$= \int d\Omega_p\left[e^{-i(p^0x^0 - \vec{p}\cdot\vec{x})}A(p^0, \vec{p}) + e^{-i(-p^0x^0 - \vec{p}\cdot\vec{x})}A(-p^0, \vec{p})\right] \tag{4.97}$$

$$= \int d\Omega_p\left[e^{-ip\cdot x}A(p^0, \vec{p}) + e^{ip\cdot x}A(-p^0, -\vec{p})\right] = \int d\Omega_p\left[e^{-ip\cdot x}a(\vec{p}) + e^{ip\cdot x}a(\vec{p})^\dagger\right] \,,$$

where in the last passage we have set $a(\vec{p}) \equiv A(p^0, \vec{p})$ and imposed Hermiticity. It remains to impose the CCR. Let us see what this implies for $a(\vec{p})$ and $a(\vec{p})^\dagger$. We can still use Eq. (4.76) to extract them from the field, since those relations depend only on the fact that φ solves the KG equation. We have then

$$
\begin{aligned}
a(\vec{p}) &= \int d^3x \, e^{ip \cdot x} \left[p^0 \varphi(x) + i \dot{\varphi}(x) \right] = \int d^3x \, e^{ip \cdot x} \left[p^0 \varphi(x) + i\pi(x) \right] , \\
a(\vec{p})^\dagger &= \int d^3x \, e^{-ip \cdot x} \left[p^0 \varphi(x) - i \dot{\varphi}(x) \right] = \int d^3x \, e^{-ip \cdot x} \left[p^0 \varphi(x) - i\pi(x) \right] .
\end{aligned}
\tag{4.98}
$$

We then find, setting $x^0 = y^0$,

$$
\begin{aligned}
[a(\vec{p}), a(\vec{q})] &= \int d^3x \int d^3y \, e^{i(p \cdot x + q \cdot y)} \left[p^0 \varphi(x) + i\pi(x), q^0 \varphi(y) + i\pi(y) \right]_{\mathrm{ET}} \\
&= \int d^3x \int d^3y \, e^{i(p \cdot x + q \cdot y)} (q^0 - p^0) \delta^{(3)}(\vec{x} - \vec{y}) \\
&= (2\pi)^3 \delta^{(3)}(\vec{p} + \vec{q}) e^{i(p^0 + q^0)x^0} (q^0 - p^0) = 0 ,
\end{aligned}
\tag{4.99}
$$

and similarly $[a(\vec{p})^\dagger, a(\vec{q})^\dagger] = 0$. Moreover,

$$
\begin{aligned}
[a(\vec{p}), a(\vec{q})^\dagger] &= \int d^3x \int d^3y \, e^{i(p \cdot x - q \cdot y)} \left[p^0 \varphi(x) + i\pi(x), q^0 \varphi(y) - i\pi(y) \right]_{\mathrm{ET}} \\
&= \int d^3x \int d^3y \, e^{i(p \cdot x - q \cdot y)} (q^0 + p^0) \delta^{(3)}(\vec{x} - \vec{y}) \\
&= (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q})(q^0 + p^0) = (2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{q}) .
\end{aligned}
\tag{4.100}
$$

We obtain the commutation relation of creation and annihilation operators, from which we now how to build up particle states once that we have taken a vacuum state $|0\rangle$ annihilated by all annihilation operators. From here on the construction is the same of subsection 4.1. It is worth noting that Eq. (4.99) automatically implies that our particles will obey Bose-Einstein statistics.

We now build the canonical Hamiltonian via Legendre transform,

$$
H = \int d^3x \left[\pi \dot{\varphi} - \mathscr{L} \right] .
\tag{4.101}
$$

Since we are dealing with operators, in general there might be ordering ambiguities. In this case however $\pi = \dot{\varphi}$, and no ambiguity arises. There is however a problem with taking products of fields at the same spacetime points, which yields divergent terms. This can be cured by imposing normal ordering of fields, and so we set

$$
H = \int d^3x \, : \left[\pi \dot{\varphi} - \mathscr{L} \right] := \int d^3x \, \frac{1}{2} : \left[\pi^2 + \vec{\nabla}\varphi^2 + m^2 \varphi^2 \right] : ,
\tag{4.102}
$$

which coincides with Eq. (4.91). The Lagrangian formalism allows to build conserved quantities associated to the symmetries of the Lagrangian. The Klein-Gordon Lagrangian is invariant under translations and Lorentz transformations. The Noether current associated to translation symmetry is

$$
\Theta^{\mu\nu} = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \varphi)} \partial^\nu \varphi - \eta^{\mu\nu} \mathscr{L} ,
\tag{4.103}
$$

and the corresponding charges are obtained integrating over spacetime the 0 components,

$$\Theta^{0\mu} = \frac{\partial \mathscr{L}}{\partial(\partial_0 \varphi)} \partial^\mu \varphi - \eta^{\mu 0} \mathscr{L} = \partial^\mu \varphi \pi - \eta^{\mu 0} \mathscr{L} \,. \tag{4.104}$$

Ordering ambiguities and singular contributions are both dealt with by normal ordering. We then define the quantum Noether current and charges as

$$\Theta^{\mu\nu} =: \frac{\partial \mathscr{L}}{\partial(\partial_\mu \varphi)} \partial^\nu \varphi - \eta^{\mu\nu} \mathscr{L} : \,,$$
$$P^\mu = \int d^3x : \pi \partial^\mu \varphi - \eta^{\mu 0} \mathscr{L} : \,. \tag{4.105}$$

These are easily seen to be Hermitian operators. We then find

$$P^0 = \int d^3x : \pi \partial^0 \varphi - \mathscr{L} := \int d^3x \frac{1}{2} : \left[\pi^2 + \vec{\nabla}\varphi^2 + m^2 \varphi^2\right] := H \,,$$
$$P^j = \int d^3x : \pi \partial^j \varphi := -\int d^3x : \pi \vec{\nabla}_j \varphi : \,, \tag{4.106}$$

i.e., P^0 coincides with the canonical Hamiltonian, and all the P^μ coincide with the operators constructed in the creation/annihilation operator formalism. In order to study their commutator with the fields, notice that

$$[\Theta^{0\mu}(y), \varphi(x)] = [: \pi(y)\partial^\mu \varphi(y) - \eta^{\mu 0}\mathscr{L}(y) :, \varphi(x)] = [\pi(y)\partial^\mu \varphi(y) - \eta^{\mu 0}\mathscr{L}(y), \varphi(x)] \,, \tag{4.107}$$

since the normal ordered product differs from the product by a constant [see Eq. (4.84)]. Moreover

$$[\Theta^{0\mu}(y), \varphi(x)]_{\mathrm{ET}} = [\pi(y)\partial^\mu \varphi(y) - \eta^{\mu 0}\frac{1}{2}\pi(y)^2, \varphi(x)]_{\mathrm{ET}} = -i\delta^{(3)}(\vec{x} - \vec{y})\partial^\mu \varphi(x) \,. \tag{4.108}$$

Since P^μ are time-independent we can compute $[P^\mu, \varphi(x)]$ using Eq. (4.106) with fields at time x^0, and so

$$[P^\mu, \varphi(x)] = \int d^3y \, [\Theta^{0\mu}(y), \varphi(x)]_{\mathrm{ET}} = -i\int d^3y \, \delta^{(3)}(\vec{x} - \vec{y})\partial^\mu \varphi(x) = -i\partial^\mu \varphi(x) \,. \tag{4.109}$$

Consider next Lorentz transformations. The classical charges are

$$J^{(\rho\sigma)} = -\int d^3x \left[x^\rho \Theta^{0\sigma}(x) - x^\sigma \Theta^{0\rho}(x)\right] \,, \tag{4.110}$$

which we adapt to the quantum case by imposing normal ordering on the products of fields. These makes them Hermitian operators. We find (here again $y^0 = x^0$)

$$[J^{(\rho\sigma)}, \varphi(x)] = \int d^3y \left[y^\rho \Theta^{0\sigma}(y) - y^\sigma \Theta^{0\rho}(y), \varphi(x)\right] = i\left(x^\rho \partial^\sigma - x^\sigma \partial^\rho\right)\varphi(x) \,. \tag{4.111}$$

These are precisely the commutation relations that lead to the transformation properties Eq. (4.75). Indeed, for infinitesimal transformations generated by P_μ we find

$$e^{iP_\mu a^\mu}\varphi(x)e^{-iP_\mu a^\mu} = \varphi(x) + ia^\mu[P_\mu, \varphi(x)] = \varphi(x) + a^\mu \partial_\mu \varphi(x) \,, \tag{4.112}$$

and

$$\varphi(x + a) = \varphi(x) + a^\mu \partial_\mu \varphi(x) \,. \tag{4.113}$$

For a Lorentz transformation $\Lambda = e^{\frac{i}{2}\omega_{\mu\nu}\tilde{J}^{(\mu\nu)}}$, where $\tilde{J}^{(\mu\nu)\alpha}{}_\beta = -i(\eta^{\mu\alpha}\delta^\nu{}_\beta - \eta^{\nu\alpha}\delta^\mu{}_\beta)$, we have

$$e^{-\frac{i}{2}\omega_{\mu\nu}J^{(\mu\nu)}}\varphi(x)e^{-\frac{i}{2}\omega_{\mu\nu}J^{(\mu\nu)}} = \varphi(x) - \frac{i}{2}\omega_{\mu\nu}[J^{(\mu\nu)}, \varphi(x)] = \varphi(x) + \frac{1}{2}\omega_{\mu\nu}\left(x^\mu \partial^\nu - x^\nu \partial^\mu\right)\varphi(x) \,, \tag{4.114}$$

and also

$$\varphi(\Lambda^{-1}x) = \varphi(x) - \frac{i}{2}\omega_{\mu\nu}\tilde{J}^{(\mu\nu)\alpha}{}_\beta x^\beta \partial_\alpha \varphi(x) = \varphi(x) - \frac{1}{2}\omega_{\mu\nu}\left(x^\nu \partial^\mu - x^\mu \partial^\nu\right)\varphi(x) \,. \tag{4.115}$$

The operators P_μ and $J^{(\mu\nu)}$ are therefore the Hermitian generators of translations and Lorentz transformations, and we have explicitly built the represention of the Poincaré group in terms of field operators.

4.2.4 Additive charges and complex scalar fields

Even within the setting of scalar particles, the formalism developed above is rather limited. In fact, Hermitian scalar fields do not allow to describe particles which possess additive quantum numbers. Suppose that we want to describe particles which are eigenstates of some conserved Hermitian charge operator Q, with

$$Q|\vec{p}_1, \ldots, \vec{p}_N\rangle = Nq|\vec{p}_1, \ldots, \vec{p}_N\rangle \,, \tag{4.116}$$

for some $q \in \mathbb{R}$, and $Q|0\rangle = 0$. Then for any basis vector we find

$$\begin{aligned}
Qa(\vec{p})^\dagger|\vec{p}_1, \ldots, \vec{p}_N\rangle &= Q|\vec{p}, \vec{p}_1, \ldots, \vec{p}_N\rangle = (N+1)q|\vec{p}, \vec{p}_1, \ldots, \vec{p}_N\rangle \\
&= (N+1)qa(\vec{p})^\dagger|\vec{p}_1, \ldots, \vec{p}_N\rangle = \left\{[Q, a(\vec{p})^\dagger] + Nqa(\vec{p})^\dagger\right\}|\vec{p}_1, \ldots, \vec{p}_N\rangle \,,
\end{aligned} \tag{4.117}$$

which implies

$$[Q, a(\vec{p})^\dagger] = qa(\vec{p})^\dagger \,, \qquad [Q, a(\vec{p})] = -qa(\vec{p}) \,. \tag{4.118}$$

This implies for the commutator of a Hermitian scalar field with Q

$$[Q, \varphi(x)] = [Q, \varphi_+(x) + \varphi_-(x)] = -q[\varphi_+(x) - \varphi_-(x)] = -q\bar{\varphi}(x) \,. \tag{4.119}$$

If our observables are built out of $\varphi(x)$, and contain also Q, we are forced to introduce also $\bar{\varphi}(x)$ in our algebra of observables: for example, if the Hamiltonian were built out of φ and its derivatives only, $H = H[\varphi, \partial\varphi]$, then $[Q, H]$ would contain $\bar{\varphi}$ as well, and it could hardly vanish, as it should for a conserved charge. On the other hand, while $\bar{\varphi}(x)$ is by itself a perfectly good causal field, it cannot be used together with $\varphi(x)$ without leading to violations of microcausality. The only consistent way to have a Hermitian scalar field and a charge operator is then that $q = 0$. On the other hand we know that particles with charges exist in Nature, so we have to find some nontrivial way out.

This is actually not that complicated. For suppose that a second set of creation and annihilation operators exist, which satisfy the same commutator algebra as the a and $a^\dagger$ and commute

with these. More precisely, let $b(\vec{p})^\dagger$ and $b(\vec{p})$ be respectively creation and annihilation operators satisfying

$$
\begin{aligned}
[b(\vec{p}), b(\vec{q})^\dagger] &= (2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{q})\,, \\
[b(\vec{p}), b(\vec{q})] &= [b(\vec{p})^\dagger, b(\vec{q})^\dagger] = 0\,, \\
[a(\vec{p}), b(\vec{q})] &= [a(\vec{p}), b(\vec{q})^\dagger] = [a(\vec{p})^\dagger, b(\vec{q})] = [a(\vec{p})^\dagger, b(\vec{q})^\dagger] = 0\,,
\end{aligned}
\tag{4.120}
$$

where $p^0 = \sqrt{\vec{p}^2 + m^2}$ with the *same* mass m appearing in the commutation relations of a and $a^\dagger$, and furthermore such that

$$
[Q, b(\vec{p})^\dagger] = -q b(\vec{p})^\dagger\,, \qquad [Q, b(\vec{p})] = q b(\vec{p})\,.
\tag{4.121}
$$

We can now build a *complex scalar field*,

$$
\varphi(x) = \int d\Omega_p \left\{ a(\vec{p}) e^{-ip\cdot x} + b(\vec{p})^\dagger e^{ip\cdot x} \right\}\,,
\tag{4.122}
$$

which satisfies the simple commutation relation

$$
[Q, \varphi(x)] = -q\varphi(x)\,.
\tag{4.123}
$$

Notice that now $\varphi(x) \neq \varphi(x)^\dagger$, but it still obeys the KG equation, $(\Box + m^2)\varphi = 0$. Therefore, if $\varphi(x)$ is causal there is no problem with having it together with the charge operator Q.

To show that microcausality is respected we have to compute various commutation relations. We easily find that

$$
[\varphi(x), \varphi(y)] = [\varphi(x)^\dagger, \varphi(y)^\dagger] = 0\,,
\tag{4.124}
$$

for all x and y, since no nontrivial commutators of creation and annihilation operators are involved. Similarly, one shows that

$$
[\partial_0\varphi(x), \partial_0\varphi(y)] = [\partial_0\varphi(x)^\dagger, \partial_0\varphi(y)^\dagger] = [\varphi(x), \partial_0\varphi(y)] = [\varphi(x)^\dagger, \partial_0\varphi(y)^\dagger] = 0\,.
\tag{4.125}
$$

Next

$$
\begin{aligned}
[\varphi(x), \varphi(y)^\dagger] &= \int d\Omega_p \int d\Omega_q \left\{ e^{-i(p\cdot x - q\cdot y)} [a(\vec{p}), a(\vec{q})^\dagger] - e^{i(p\cdot x - q\cdot y)} [b(\vec{q}), b(\vec{p})^\dagger] \right\} \\
&= \int d\Omega_p \left\{ e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)} \right\} = \Delta(x-y) - \Delta(y-x)\,,
\end{aligned}
\tag{4.126}
$$

and thus we already know that for spacelike separations $(x-y)^2 < 0$

$$
[\varphi(x), \varphi(y)^\dagger] = 0\,.
\tag{4.127}
$$

The next one to check is

$$
\begin{aligned}
[\partial_0\varphi(x), \partial_0\varphi(y)^\dagger] &= \int d\Omega_p \int d\Omega_q\, p_0 q_0 \left\{ e^{-i(p\cdot x - q\cdot y)} [a(\vec{p}), a(\vec{q})^\dagger] - e^{i(p\cdot x - q\cdot y)} [b(\vec{q}), b(\vec{p})^\dagger] \right\} \\
&= \int d\Omega_p\, p_0^2 \left\{ e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)} \right\} = \partial_{x0}\partial_{y0}[\Delta(x-y) - \Delta(y-x)]\,,
\end{aligned}
\tag{4.128}
$$

70

which is the 00 component of a rank-2 tensor. Its value is therefore determined by covariance after we evaluate it in one reference frame. For spacelike separations we can choose to have $x^0 = y^0$, in which case

$$[\partial_0\varphi(x), \partial_0\varphi(y)^\dagger]_{\mathrm{ET}} = \int d\Omega_p \, p_0^2 \left\{ e^{i\vec{p}\cdot(\vec{x}-\vec{y})} - e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \right\} = 0 \, . \qquad (4.129)$$

There is only one commutator left to check, namely

$$[\varphi(x), \partial_0\varphi(y)^\dagger] = \int d\Omega_p \int d\Omega_q \, iq_0 \left\{ e^{-i(p\cdot x - q\cdot y)}[a(\vec{p}), a(\vec{q})^\dagger] + e^{i(p\cdot x - q\cdot y)}[b(\vec{q}), b(\vec{p})^\dagger] \right\}$$
$$= i \int d\Omega_p \, p_0 \left\{ e^{-ip\cdot(x-y)} + e^{ip\cdot(x-y)} \right\} = \partial_{y^0}[\Delta(x-y) + \Delta(y-x)] \, , \qquad (4.130)$$

which is the 0 component of a Lorentz four-vector. For spacelike separations we evaluate it in the frame where $x^0 = y^0$ and find

$$[\varphi(x), \partial_0\varphi(y)^\dagger]_{\mathrm{ET}} = i \int d\Omega_p \, p_0 \left\{ e^{i\vec{p}\cdot(\vec{x}-\vec{y})} + e^{-i\vec{p}\cdot(\vec{x}-\vec{y})} \right\} = i\delta^{(3)}(\vec{x}-\vec{y}) \, , \qquad (4.131)$$

which vanishes for $\vec{x} \neq \vec{y}$, and thus by covariance it vanishes for any $(x-y)^2 < 0$. By taking the Hermitian conjugate of Eq. (4.131) we also find

$$[\varphi(x)^\dagger, \partial_0\varphi(y)]_{\mathrm{ET}} = i\delta^{(3)}(\vec{x}-\vec{y}) \, . \qquad (4.132)$$

Commutators involving spatial derivatives and higher-order time derivatives can be dealt with in the same way we did in the case of a hermitian scalar field. This completes the proof that $\varphi(x)$ is a causal field.

The field $\varphi(x)$ describes not one, but two particles with the same mass and opposite charges: these are a particle and its corresponding *antiparticle*. The concept of antiparticles emerges quite naturally in the framework of QFT, as they are needed to avoid troubles with causality whenever some (nontrivial) additive quantum number can be assigned to particles.

At this point one can develop the Hamiltonian and the Lagrangian formalism by constructing first the particle Hamiltonian in terms of creation and annihilation operators,

$$H = \int d\Omega_p \, p^0 \left\{ a(\vec{p})^\dagger a(\vec{p}) + b(\vec{p})^\dagger (\vec{p}) \right\} \, , \qquad (4.133)$$

expressing this in terms of fields, and finally performing a Legendre transform to obtain the Lagrangian. To make it quicker we proceed here backwards, guessing the Lagrangian and then going over to the Hamiltonian, and finally showing that we obtain Eq. (4.133). The appropriate classical Lagrangian is easily seen to be

$$\mathscr{L} = \partial_\mu\varphi\partial^\mu\varphi^\dagger - m^2\varphi\varphi^\dagger \, , \qquad (4.134)$$

which by the usual action principle applied treating φ and $\varphi^\dagger$ as independent variables (corresponding to independent real and imaginary parts of the complex field φ) yields the expected equations of motion,

$$(\Box + m^2)\varphi = (\Box + m^2)\varphi^\dagger = 0 \, , \qquad (4.135)$$

71

i.e., the KG equation. The conjugated momenta read

$$\pi_\varphi = \partial_0 \varphi^\dagger \equiv \pi \,, \qquad \pi_{\varphi^\dagger} = \partial_0 \varphi = \pi_\varphi^\dagger = \pi^\dagger \,. \tag{4.136}$$

Canonical quantisation requires that we impose the equal time commutation relations

$$
\begin{aligned}
&[\varphi(x), \pi(y)]_{\mathrm{ET}} = i\delta^{(3)}(\vec{x} - \vec{y}) \,, \qquad [\varphi(x)^\dagger, \pi(y)^\dagger]_{\mathrm{ET}} = i\delta^{(3)}(\vec{x} - \vec{y}) \,, \\
&[\varphi(x), \varphi(y)]_{\mathrm{ET}} = [\varphi(x), \varphi(y)^\dagger]_{\mathrm{ET}} = [\varphi(x)^\dagger, \varphi(y)^\dagger]_{\mathrm{ET}} = 0 \,, \\
&[\pi(x), \pi(y)]_{\mathrm{ET}} = [\pi(x), \pi(y)^\dagger]_{\mathrm{ET}} = [\pi(x)^\dagger, \pi(y)^\dagger]_{\mathrm{ET}} = 0 \,.
\end{aligned}
\tag{4.137}
$$

Solving the equations of motion by means of a Fourier transform we find

$$\varphi(x) = \int d\Omega_p \left\{ a(\vec{p}) e^{-ip \cdot x} + b(\vec{p})^\dagger e^{ip \cdot x} \right\} \tag{4.138}$$

for yet unspecified operator-valued coefficients $a(\vec{p})$ and $b(\vec{p})^\dagger$. Expressing the canonical commutation relations in terms of these via the relations

$$
\begin{aligned}
a(\vec{p}) &= \int d^3x \, e^{ip \cdot x} i\overleftrightarrow{\partial_0} \varphi(x) \,, \qquad & a(\vec{p})^\dagger &= \int d^3x \, e^{-ip \cdot x} (-i\overleftrightarrow{\partial_0}) \varphi(x)^\dagger \,, \\
b(\vec{p}) &= \int d^3x \, e^{-ip \cdot x} i\overleftrightarrow{\partial_0} \varphi(x)^\dagger \,, \qquad & b(\vec{p})^\dagger &= \int d^3x \, e^{-ip \cdot x} (-i\overleftrightarrow{\partial_0}) \varphi(x) \,,
\end{aligned}
\tag{4.139}
$$

we see that they must satisfy the commutation relations Eq. (4.120), i.e., we get back the complex scalar field we wanted. The Hamiltonian is obtained via Legendre transform as

$$H = \int d^3x : \pi(x)\pi(x)^\dagger + \vec{\nabla}\varphi\vec{\nabla}\varphi^\dagger + m^2\varphi\varphi^\dagger : \,, \tag{4.140}$$

where we have already imposed normal ordering to take care of a divergent contribution (and also to settle the issue of ordering ambiguities). It is then straightforward to show that Eq. (4.140) coincides with Eq. (4.133) when expressed in terms of creation and annihilation operators. The Noether current associated with the translation invariance of the Lagrangian Eq. (4.134) reads

$$\Theta^{\mu\nu} =: \frac{\partial\mathscr{L}}{\partial(\partial_\mu\varphi)}\partial^\nu\varphi + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\varphi^\dagger)}\partial^\nu\varphi^\dagger - \eta^{\mu\nu}\mathscr{L} :=: \partial^\mu\varphi(x)\partial^\nu\varphi(x)^\dagger + \partial^\nu\varphi(x)\partial^\mu\varphi(x)^\dagger - \eta^{\mu\nu}\mathscr{L} :\,, \tag{4.141}$$

where by normal ordering we have modified the classical expression by an irrelevant (although infinite) constant. The associated Noether charges $P^\mu = \int d^3x \, \Theta^{0\mu}$ are then

$$
\begin{aligned}
P^0 &= \int d^3x : \partial_0\varphi(x)\partial_0\varphi(x)^\dagger + \vec{\nabla}\varphi\vec{\nabla}\varphi^\dagger + m^2\varphi\varphi^\dagger := H \,, \\
P^j &= \int d^3x : \partial^0\varphi(x)\partial^j\varphi(x)^\dagger + \partial^0\varphi(x)^\dagger\partial^j\varphi(x) : \\
&= -\int d^3x : \pi(x)\partial_j\varphi(x)^\dagger + \pi(x)^\dagger\partial_j\varphi(x) : \,.
\end{aligned}
\tag{4.142}
$$

It is again straightforward to show that

$$P^j = \int d\Omega_p \, p^j \left\{ a(\vec{p})^\dagger a(\vec{p}) + b(\vec{p})^\dagger(\vec{p}) \right\} \,. \tag{4.143}$$

One immediately proves that P^μ are precisely the generators of translations (in doing so, the normal ordering of fields can be ignored in commutators since it just amounts to a constant shift). The construction of the Lorentz generators via Noether's theorem proceeds in a similar way, and the same result Eq. (4.110) is obtained [with $\Theta^{\mu\nu}$ given now by Eq. (4.141)].

The Lagrangian Eq. (4.134) possesses a new symmetry under the U(1) transformation

$$\varphi \to e^{i\alpha}\varphi\,, \qquad \varphi^\dagger \to e^{-i\alpha}\varphi^\dagger\,, \tag{4.144}$$

with real, x-independent α. The infinitesimal form of the transformation is

$$\delta x = 0\,, \qquad \delta\varphi = i\varphi\,, \qquad \delta\varphi^\dagger = -i\varphi^\dagger\,, \tag{4.145}$$

from which the associated Noether current is obtained:

$$J^\mu = i : \varphi^\dagger \overset{\leftrightarrow}{\partial^\mu} \varphi : \,. \tag{4.146}$$

Expressing the associated Noether charge in terms of creation and annihilation operators we find

$$Q = \int d^3x\, J^0 = i : \varphi^\dagger \pi^\dagger - \pi\varphi := \int d\Omega_p \left\{ a(\vec{p})^\dagger a(\vec{p}) - b(\vec{p})^\dagger (\vec{p}) \right\}\,, \tag{4.147}$$

which is a Hermitian operator. The following commutation relations then follow:

$$
\begin{aligned}
[Q, \varphi(x)] &= -\varphi(x)\,, & [Q, \varphi(x)^\dagger] &= \varphi(x)^\dagger\,, \\
[Q, a(\vec{p})^\dagger] &= a(\vec{p})^\dagger\,, & [Q, b(\vec{p})^\dagger] &= -b(\vec{p})^\dagger\,.
\end{aligned}
\tag{4.148}
$$

The charge Q is then precisely an additive conserved charge of the type that motivated the introduction of the complex scalar field in the first place. Although the complex scalar field can be seen as a combination of two real fields, $\varphi = \varphi_1 + i\varphi_2$, the creation and annihilation operators corresponding to $\varphi_{1,2}$ do not transform simply under U(1), which means that the associated particles are not charge eigenstates.

4.2.5 Discrete symmetries

Let us now discuss the effect of discrete symmetries on scalar fields, focussing on parity and charge conjugation.

If parity is a symmetry of our theory, it can be implemented on the Hilbert space of states as a unitary operator $U(P)$ (unitarity rather than antiunitarity is required to avoid negative-energy states). Assuming that it commutes with Q, it can be implemented via [see Eq. (3.112)]

$$
\begin{aligned}
U(P)a(\vec{p})^\dagger U(P)^\dagger &= \eta a(-\vec{p})^\dagger\,, \\
U(P)b(\vec{p})^\dagger U(P)^\dagger &= \eta' b(-\vec{p})^\dagger\,,
\end{aligned}
\tag{4.149}
$$

having imposed $U(P)|0\rangle = |0\rangle$.[31] Here η, η' are phase factors, i.e., the intrinsic parities of the particles. Making simple manipulations of Eq. (4.149) we can then obtain the transformation

---

[31] The assumed unicity of the vacuum state, combined with its invariance under Poincaré transformations, implies that $U(P)|0\rangle = U(P)U(\Lambda)|0\rangle = U(P)U(\Lambda)U(P)^\dagger U(P)|0\rangle = U(P\Lambda P^{-1})U(P)|0\rangle$, and similarly $U(P)|0\rangle = U(P)U(a)|0\rangle = U(P)U(a)U(P)^\dagger U(P)|0\rangle = U(Pa)U(P)|0\rangle$, i.e., $U(P)|0\rangle$ is invariant and thus must be equal to $|0\rangle$ up to a phase. Redefining $U(P)$ we can set this phase to 0.

law for the scalar field,

$$U(P)^\dagger\varphi(x)U(P) = \int d\Omega_p \left\{ \eta a(-\vec{p})e^{-ip\cdot x} + b(-\vec{p})^\dagger \eta'^* e^{ip\cdot x} \right\}$$
$$= \eta \int d\Omega_p \left\{ a(\vec{p})e^{-ip\cdot(Px)} + b(\vec{p})^\dagger e^{ip\cdot(Px)}(\eta\eta')^* \right\} \tag{4.150}$$

To avoid problems with causality we cannot have in our theory a second field with a phase relation between its positive and negative frequency parts differing from the one we have for φ, and this requires that we have

$$\eta\eta' = 1 \,, \tag{4.151}$$

if a quantum field theory is to describe Nature. We then obtain the simple transformation law

$$U(P)^\dagger\varphi(x)U(P) = \eta\varphi(Px) \,. \tag{4.152}$$

For a Hermitian scalar field we have a single type of particle, and we conclude that $\eta^2 = 1$, i.e., $\eta = \pm 1$.

Having now two types of particles, we can consider a transformation that exchanges their roles, called *charge conjugation*. In case this is a symmetry, it can be implemented via a unitary operator $U(C)$ on the Hilbert space of states (the unitary choice being made again to avoid negative-energy states). We then have

$$U(C)|\vec{p}\rangle = \xi|\bar{\vec{p}}\rangle \,,$$
$$U(C)|\bar{\vec{p}}\rangle = \xi'|\vec{p}\rangle \,, \tag{4.153}$$

where $|\bar{\vec{p}}\rangle$ denote the momentum eigenstates of the antiparticle. Again, we set $U(C)|0\rangle = |0\rangle$ (see footnote 31). We then obtain the transformation laws for creation operators

$$U(C)a(\vec{p})^\dagger U(C)^\dagger = \xi b(\vec{p})^\dagger \,,$$
$$U(C)b(\vec{p})^\dagger U(C)^\dagger = \xi' a(\vec{p})^\dagger \,, \tag{4.154}$$

and from this we can work out the transformation law for the field. Simple manipulations yield

$$U(C)^\dagger\varphi(x)U(C) = \int d\Omega_p \left\{ \xi' b(\vec{p})e^{-ip\cdot x} + \xi^* a(\vec{p})^\dagger e^{ip\cdot x} \right\}$$
$$= \xi^* \left\{ \int d\Omega_p \left\{ a(\vec{p})e^{-ip\cdot x} + (\xi\xi')^* b(\vec{p})^\dagger e^{ip\cdot x} \right\} \right\}^\dagger \,. \tag{4.155}$$

In order to avoid problems with causality we have to impose

$$\xi\xi' = 1 \,, \tag{4.156}$$

if a quantum field theory is to describe Nature. We then obtain the simple transformation law

$$U(C)^\dagger\varphi(x)U(C) = \xi^*\varphi(x)^\dagger \,. \tag{4.157}$$

For a Hermitian scalar field we have a single type of particle, which is then said to be self-conjugate. As such, a self-conjugate particle cannot have any nonzero value for a conserved charge; moreover, it must be $\xi^2 = 1$, i.e., $\xi = \pm 1$.

This concludes our discussion of scalar fields. We now move on to the more complicated case of spin $\frac{1}{2}$.

5 Spin-$\frac{1}{2}$ particles and Dirac fields

It is an experimental fact that spin-$\frac{1}{2}$ particles obey Fermi-Dirac statistics, i.e., their state vectors change sign under the exchange of the quantum numbers of any two such particles. Starting from single-particle states $|\vec{p}s\rangle$, where $s = \pm\frac{1}{2}$ is the value of the spin in a given particular direction, which we take to be direction 3, we can build the Fock space of such particles in analogy to the scalar case, taking the N-fold tensor product of single-particle states and antisymmetrising with respect to the momenta,

$$|\vec{p}_1 s_1, \ldots, \vec{p}_N s_N\rangle = \sum_{\mathrm{P}} (-1)^{\sigma_{\mathrm{P}}} |\vec{p}_{\mathrm{P}(1)} s_{\mathrm{P}(1)}\rangle \otimes \ldots \otimes |\vec{p}_{\mathrm{P}(N)} s_{\mathrm{P}(N)}\rangle. \tag{5.1}$$

Here σ_{P} is the signature of permutation P, i.e., the number of transpositions required to obtain P modulo 2. The vacuum state $|0\rangle$ is also included. In this way

$$|\vec{p}_1 s_1, \ldots, \vec{p}_N s_N\rangle = (-1)^{\sigma_{\mathrm{P}}} |\vec{p}_{\mathrm{P}(1)} s_{\mathrm{P}(1)}, \ldots, \vec{p}_{\mathrm{P}(N)} s_{\mathrm{P}(N)}\rangle. \tag{5.2}$$

The normalisation of multiparticle states is again inherited form the single particle normalisation,

$$\langle \vec{p}' s' | \vec{p}s\rangle = \delta_{s's}(2\pi)^3 2p^0 \delta^{(3)}(\vec{p}' - \vec{p}), \tag{5.3}$$

and reads

$$\langle \vec{p}'_1 s'_1, \ldots, \vec{p}'_M s'_M | \vec{p}_1 s_1, \ldots, \vec{p}_N s_N\rangle = \delta_{MN} \sum_{\mathrm{P}} (-1)^{\sigma_{\mathrm{P}}} \prod_{j=1}^{N} \delta_{s'_{\mathrm{P}(j)} s_j} (2\pi)^3 2p_j^0 \delta^{(3)}(\vec{p}'_{\mathrm{P}(j)} - \vec{p}_j). \tag{5.4}$$

Transformation properties under symmetry transformations also follow from the single-particle properties. The vacuum state $|0\rangle$ is taken to be invariant under any symmetry transformation, and the only such state.

In full analogy with the scalar case we set up the formalism of creation and annihilation operators. We define the creation operator $b_s(\vec{p})^\dagger$, creating a particle of momentum $\vec{p}$ and third component of the spin s as

$$b_s(\vec{p})^\dagger |\vec{p}_1 s_1, \ldots, \vec{p}_N s_N\rangle \equiv |\vec{p}s, \vec{p}_1 s_1, \ldots, \vec{p}_N s_N\rangle. \tag{5.5}$$

Notice that the position where we place the quantum numbers of the new particle do matter now, as minus signs pop up under permutations. The adjoint operator $b_s(\vec{p})$, i.e., the annihilation operator, removes a particle from a state vector. By definition of the adjoint

$$
\begin{aligned}
&(b_s(\vec{p})\Psi_{\vec{p}'_1 s'_1, \ldots, \vec{p}'_M s'_M}, \Psi_{\vec{p}_1 s_1, \ldots, \vec{p}_N s_N}) = (\Psi_{\vec{p}'_1 s'_1, \ldots, \vec{p}'_M s'_M}, b_s(\vec{p})^\dagger \Psi_{\vec{p}_1 s_1, \ldots, \vec{p}_N s_N})\\
&= (\Psi_{\vec{p}'_1 s'_1, \ldots, \vec{p}'_M s'_M}, \Psi_{\vec{p}s, \vec{p}_1 s_1, \ldots, \vec{p}_N s_N})\\
&= \delta_{M,N+1} \sum_{\mathrm{P}} (-1)^{\sigma_P} \delta_{s'_{\mathrm{P}(1)} s} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}'_{\mathrm{P}(1)} - \vec{p}) \prod_{k=2}^{M} \delta_{s'_{\mathrm{P}(k)} s_{k-1}} (2\pi)^3 2p_k^0 \delta^{(3)}(\vec{p}'_{\mathrm{P}(k)} - \vec{p}_{k-1})\\
&= \delta_{M,N+1} \sum_{j=1}^{M} (-1)^{j-1} \delta_{s'_j s} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}'_j - \vec{p}) \sum_{\mathrm{P}_j} (-1)^{\sigma_{P_j}} \prod_{k=2}^{M} \delta_{s'_{\mathrm{P}_j(k)} s_{k-1}} (2\pi)^3 2p_k^0 \delta^{(3)}(\vec{p}'_{\mathrm{P}_j(k)} - \vec{p}_{k-1})\\
&= \sum_{j=1}^{M} (-1)^{j-1} (2\pi)^3 2p^0 \delta^{(3)}(\vec{p}'_j - \vec{p}) (\Psi_{\vec{p}'_1 s'_1, \ldots, \vec{p}'_M s'_M \setminus \vec{p}'_j s'_j}, \Psi_{\vec{p}_1 s_1, \ldots, \vec{p}_N s_N}),
\end{aligned}
$$

$$\tag{5.6}$$

where P_j is a permutation of $2, \ldots, M$ such that $P_j(k) \neq j$, and where $\Psi_{\vec{p}_1' s_1', \ldots, \vec{p}_M' s_M' \backslash \vec{p}_j' s_j'}$ is obtained from the state $\Psi_{\vec{p}_1' s_1', \ldots, \vec{p}_M' s_M'}$ by removing the particle with momentum $\vec{p}_j'$ and spin s_j'. Here we have written a generic permutation P uniquely as the product of $j - 1$ transpositions that shift $j = P(1)$ to first position, times a permutation P_j. The sum over permutations is then the sum over the assignment $j = P(1)$ and over the permutations P_j for a given j. The signature of P is the signature of P_j plus that of the product of transpositions, and so $(-1)^{\sigma_P} = (-1)^{j-1}(-1)^{\sigma_{P_j}}$. We then conclude

$$
b_s(\vec{p})|\vec{p}_1 s_1, \ldots, \vec{p}_N s_N\rangle =
$$
$$
\sum_{j=1}^{N} (-1)^{j-1} \delta_{s_j s}(2\pi)^3 2p^0 \delta^{(3)}(\vec{p}_j - \vec{p})|\vec{p}_1 s_1, \ldots, \vec{p}_{j-1} s_{j-1}, \vec{p}_{j+1} s_{j+1}, \ldots, \vec{p}_N s_N\rangle \,.
\tag{5.7}
$$

It is now straightforward to compute the *anticommutators* of creation and annihilation operators: the calculation is the same as in the scalar case, except that now there are minus signs to be taken into account. We find

$$
\{b_s(\vec{p}), b_{s'}(\vec{q})^\dagger\} = \delta_{ss'}(2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{q}) \,,
$$
$$
\{b_s(\vec{p}), b_{s'}(\vec{q})\} = \{b_s(\vec{p})^\dagger, b_{s'}(\vec{q})^\dagger\} = 0 \,.
\tag{5.8}
$$

The Hamiltonian and the total momentum operators are also defined straightforwardly as

$$
P^\mu = \int d\Omega_p p^\mu b_s(\vec{p})^\dagger b_s(\vec{p}) \,,
\tag{5.9}
$$

and making use of Eqs. (5.5) and (5.7) one finds

$$
P^\mu|\vec{p}_1 s_1, \ldots, \vec{p}_N s_N\rangle = \left(\sum_{j=1}^{N} p_j^\mu \right) |\vec{p}_1 s_1, \ldots, \vec{p}_N s_N\rangle \,.
\tag{5.10}
$$

From the point of view of the Fock space construction, everything is the same as in the scalar case, except that the creation and annihilation operators satisfy now anticommutation rather than commutation relations. While the transformation properties of the creation and annihilation operators, as obtained from those of the states (using also the invariance of the vacuum), are the same as for scalars for translations, they differ from the scalar case for Lorentz transformations, and read

$$
U(a)b_s(\vec{p})^\dagger U(a)^\dagger = e^{-ip \cdot a} b_s(\vec{p})^\dagger \,,
$$
$$
U(\Lambda)b_s(\vec{p})^\dagger U(\Lambda)^\dagger = \sum_{\bar{s}} \mathscr{D}_{\bar{s}s}^{(\frac{1}{2})}(W(\Lambda, \vec{p})) b_{\bar{s}}(\vec{p})^\dagger \,.
\tag{5.11}
$$

Here we focus on massive particles. Here $W(\Lambda, \vec{p})$ is the Wigner rotation associated to Λ and $\vec{p}$, and $\mathscr{D}^{(\frac{1}{2})}$ is the spin-$\frac{1}{2}$ representation of SO(3), which is the relevant little group for massive representations.

5.1 The Dirac field

We now want to trade the creation and annihilation operators for a local field, subject to a few requirements. First of all, we should be able to get back the creation and annihilation operators

from the field, so that the two formalisms are equivalent. Next, we want the field to have simple transformation properties under transformations of the Poincaré group, so that it will be easy to construct interactions that satisfy Poincaré invariance. Finally, we want the fields to obey some form of microcausality, so that we can build local observables out of them that will commute for spacelike separations. We consider the more general case of charged particles, which we have seen in the scalar case to require the existence of particle/antiparticle pairs. We then introduce a second set of creation and annihilation operators satisfying the anticommutation relations

$$\{d_s(\vec{p}), d_{s'}(\vec{q})^\dagger\} = \delta_{ss'}(2\pi)^3 2p^0 \delta^{(3)}(\vec{p} - \vec{q}),$$
$$\{d_s(\vec{p}), d_{s'}(\vec{q})\} = \{d_s(\vec{p})^\dagger, d_{s'}(\vec{q})^\dagger\} = 0,$$
$$\{b_s(\vec{p}), d_{s'}(\vec{q})\} = \{b_s(\vec{p}), d_{s'}(\vec{q})^\dagger\} = \{b_s(\vec{p})^\dagger, d_{s'}(\vec{q})\} = \{b_s(\vec{p})^\dagger, d_{s'}(\vec{q})^\dagger\} = 0,$$

(5.12)

and transforming as follows under Poincaré transformations,

$$U(a)d_s(\vec{p})^\dagger U(a)^\dagger = e^{-ip\cdot a}d_s(\vec{p})^\dagger,$$
$$U(\Lambda)d_s(\vec{p})^\dagger U(\Lambda)^\dagger = \sum_{\bar{s}} \mathscr{D}_{\bar{s}s}^{(\frac{1}{2})}(W(\Lambda, \vec{p}))d_{\bar{s}}(\vec{p})^\dagger.$$

(5.13)

We then look for fields of the form

$$\psi(x) = \int d\Omega_p \sum_s \left\{ U_s(\vec{p}, x)b_s(\vec{p}) + V_s(\vec{p}, x)d_s(\vec{p})^\dagger \right\},$$

(5.14)

where U_s and V_s are (possibly multicomponent) coefficients yet to be determined. Notice that they can be extracted by taking matrix elements of the field between the vacuum and one particle states,

$$\langle 0|\psi(x)|\vec{p}s\rangle = U_s(\vec{p}, x), \qquad \langle \overline{\vec{p}s}|\psi(x)|0\rangle = V_s(\vec{p}, x),$$

(5.15)

where $|\vec{p}s\rangle$ are particle states and $|\overline{\vec{p}s}\rangle$ are antiparticle states. The request of simple transformation properties is formulated as follows, as appropriate for fields,

$$U(a)^\dagger \psi(x)U(a) = \psi(x + a),$$
$$U(\Lambda)^\dagger \psi(x)U(\Lambda) = S(\Lambda)\psi(\Lambda^{-1}x),$$

(5.16)

where $S(\Lambda)$ is some matrix providing a representation of the proper orthocronous Lorentz group.[32] Such a representation is not necessarily irreducible, and certainly it will not be unitary if it is to be finite-dimensional. Using (manipulations of) Eqs. (5.11) and (5.13) we find

$$U(a)^\dagger \psi(x)U(a) = \int d\Omega_p \sum_s \left\{ U_s(\vec{p}, x)e^{-i\vec{p}\cdot a}b_s(\vec{p}) + V_s(\vec{p}, x)e^{i\vec{p}\cdot a}d_s(\vec{p})^\dagger \right\}$$
$$= \psi(x + a) = \int d\Omega_p \sum_s \left\{ U_s(\vec{p}, x + a)b_s(\vec{p}) + V_s(\vec{p}, x + a)d_s(\vec{p})^\dagger \right\},$$

(5.17)

where the last equality stands actually for the condition we have to impose. Using Eq. (5.15) we then find

$$U_s(\vec{p}, x) = U_s(\vec{p}, 0)e^{-i\vec{p}\cdot x} \equiv u_s(\vec{p})e^{-i\vec{p}\cdot x},$$
$$V_s(\vec{p}, x) = V_s(\vec{p}, 0)e^{i\vec{p}\cdot x} \equiv v_s(\vec{p})e^{i\vec{p}\cdot x}.$$

(5.18)

---

[32]We could ask for different transformation properties, but if we can get a formalism equivalent to the creation/annihilation operator formalism using object that can be easily combined to form Poncaré invariant interactions, why bother?

To study the effect of Lorentz transformation, we first derive a property of the matrices $\mathscr{D}$. By definition

$$U(\Lambda_{\Lambda p})^\dagger U(\Lambda) U(\Lambda_p)|\vec{0},\sigma\rangle = U(W(\Lambda,\vec{p}))|\vec{0},\sigma\rangle = \mathscr{D}_{\bar{\sigma}\sigma}(W(\Lambda,\vec{p}))|\vec{0},\bar{\sigma}\rangle\,,$$

$$[U(\Lambda_{\Lambda p})^\dagger U(\Lambda) U(\Lambda_p)]^\dagger|\vec{0},\sigma\rangle = U(\Lambda_p)^\dagger U(\Lambda)^\dagger U(\Lambda_{\Lambda p})|\vec{0},\sigma\rangle = U(W(\Lambda^{-1},\Lambda\vec{p}))|\vec{0},\sigma\rangle \quad (5.19)$$

$$= \mathscr{D}_{\bar{\sigma}\sigma}(W(\Lambda^{-1},\Lambda\vec{p}))|\vec{0},\bar{\sigma}\rangle\,,$$

from which it follows

$$|\vec{0},\sigma\rangle = U(\Lambda_{\Lambda p})^\dagger U(\Lambda) U(\Lambda_p)[U(\Lambda_{\Lambda p})^\dagger U(\Lambda) U(\Lambda_p)]^\dagger|\vec{0},\sigma\rangle$$

$$= \mathscr{D}_{\bar{\sigma}\sigma}(W(\Lambda^{-1},\Lambda\vec{p})) U(\Lambda_{\Lambda p})^\dagger U(\Lambda) U(\Lambda_p)|\vec{0},\bar{\sigma}\rangle \quad (5.20)$$

$$= \mathscr{D}_{\bar{\sigma}\sigma}(W(\Lambda^{-1},\Lambda\vec{p})) \mathscr{D}_{\bar{\bar{\sigma}}\bar{\sigma}}(W(\Lambda,\vec{p}))|\vec{0},\bar{\bar{\sigma}}\rangle\,,$$

and so

$$\mathscr{D}_{\bar{\bar{\sigma}}\bar{\sigma}}(W(\Lambda,\vec{p})) \mathscr{D}_{\bar{\sigma}\sigma}(W(\Lambda^{-1},\Lambda\vec{p})) = \delta_{\bar{\bar{\sigma}}\sigma} \implies \mathscr{D}(W(\Lambda,\vec{p})) \mathscr{D}(W(\Lambda^{-1},\Lambda\vec{p})) = \mathbf{1}\,, \quad (5.21)$$

or equivalently, since $\mathscr{D}(W(\Lambda,\vec{p}))^{-1} = \mathscr{D}(W(\Lambda,\vec{p}))^\dagger$ due to unitarity of the representation,

$$\mathscr{D}(W(\Lambda^{-1},\Lambda\vec{p})) = \mathscr{D}(W(\Lambda,\vec{p}))^\dagger \implies \mathscr{D}(W(\Lambda,\Lambda^{-1}\vec{p})) = \mathscr{D}(W(\Lambda^{-1},\vec{p}))^\dagger\,. \quad (5.22)$$

This identity allows us to recast the transformation laws under Lorentz transformation in Eqs. (5.11) and (5.13) as follows,

$$U(\Lambda)^\dagger b_s(\vec{p}) U(\Lambda) = \sum_{\bar{s}} \mathscr{D}_{s\bar{s}}^{(\frac{1}{2})}(W(\Lambda,\Lambda^{-1}\vec{p})) b_{\bar{s}}(\Lambda^{-1}\vec{p})\,,$$

$$U(\Lambda)^\dagger d_s^\dagger(\vec{p}) U(\Lambda) = \sum_{\bar{s}} \mathscr{D}_{s\bar{s}}^{(\frac{1}{2})\,*}(W(\Lambda,\Lambda^{-1}\vec{p})) d_{\bar{s}}(\Lambda^{-1}\vec{p})^\dagger\,, \quad (5.23)$$

from which we can determine the transformation law of the field,

$$U(\Lambda)^\dagger \psi(x) U(\Lambda) = \int d\Omega_p \sum_{s,\bar{s}} \left\{ \mathscr{D}_{s\bar{s}}^{(\frac{1}{2})}(W(\Lambda,\Lambda^{-1}\vec{p})) u_s(\vec{p}) e^{-ip\cdot x} b_s(\Lambda^{-1}\vec{p}) \right.$$

$$\left. + \mathscr{D}_{s\bar{s}}^{(\frac{1}{2})\,*}(W(\Lambda,\Lambda^{-1}\vec{p})) v_s(\vec{p}) e^{ip\cdot x} d_s(\Lambda^{-1}\vec{p})^\dagger \right\}$$

$$= \int d\Omega_p \sum_{s,\bar{s}} \left\{ \mathscr{D}_{s\bar{s}}^{(\frac{1}{2})}(W(\Lambda,\vec{p})) u_s(\Lambda\vec{p}) e^{-ip\cdot\Lambda^{-1}x} b_s(\vec{p}) \right. \quad (5.24)$$

$$\left. + \mathscr{D}_{s\bar{s}}^{(\frac{1}{2})\,*}(W(\Lambda,\vec{p})) v_s(\Lambda\vec{p}) e^{ip\cdot\Lambda^{-1}x} d_s(\vec{p})^\dagger \right\}$$

$$= S(\Lambda)\psi(\Lambda^{-1}x)\,,$$

which then leads us to require

$$S(\Lambda) u_s(\vec{p}) = \sum_{\bar{s}} \mathscr{D}_{\bar{s}s}^{(\frac{1}{2})}(W(\Lambda,\vec{p})) u_{\bar{s}}(\Lambda\vec{p})\,,$$

$$S(\Lambda) v_s(\vec{p}) = \sum_{\bar{s}} \mathscr{D}_{\bar{s}s}^{(\frac{1}{2})\,*}(W(\Lambda,\vec{p})) v_{\bar{s}}(\Lambda\vec{p})\,. \quad (5.25)$$

This leads us to look for representations $S(\Lambda)$ that contain spin-$\frac{1}{2}$ representations among their irreducible components. There are two cases in which Eq. (5.25) looks simpler. If we take $\vec{p} = \vec{0}$ and $\Lambda = \Lambda_p$ a pure boost sending $\vec{0}$ to $\vec{p}$, we have by construction that $W(\Lambda, \vec{0}) = \mathbf{1}$ is just the identity, and so $\mathscr{D}^{(\frac{1}{2})}_{\bar{s}s}(W(\Lambda_p, \vec{0})) = \delta_{\bar{s}s}$. Then

$$
\begin{aligned}
S(\Lambda_p)u_s(\vec{0}) &= u_s(\vec{p})\,, \\
S(\Lambda_p)v_s(\vec{0}) &= v_s(\vec{p})\,.
\end{aligned}
\tag{5.26}
$$

Moreover, if we take again $\vec{p} = \vec{0}$ and $\Lambda = R$ a rotation, we have that $W(R, \vec{0}) = R,$[33] and so

$$
\begin{aligned}
S(R)u_s(\vec{0}) &= \sum_{\bar{s}} \mathscr{D}^{(\frac{1}{2})}_{\bar{s}s}(R)u_{\bar{s}}(\vec{0})\,, \\
S(R)v_s(\vec{0}) &= \sum_{\bar{s}} \mathscr{D}^{(\frac{1}{2})\,*}_{\bar{s}s}(R)v_{\bar{s}}(\vec{0})\,,
\end{aligned}
\tag{5.27}
$$

i.e., $u_s(\vec{0})$ and $v_s(\vec{0})$ have to provide a basis for the spin-$\frac{1}{2}$ representation of the rotation group $\mathscr{D}^{(\frac{1}{2})}_{\bar{s}s}(R)$ and the corresponding complex conjugate representation.[34] Once we have found some $S(\Lambda)$ and $u_s(\vec{0})$ and $v_s(\vec{0})$ satisfying Eq. (5.27), defining $u_s(\vec{p})$ and $v_s(\vec{p})$ via Eq. (5.26) implies that Eq. (5.25) will be satisfied, thanks to the fact that $S(\Lambda)$ provides a representation of the Lorentz group: indeed,

$$
\begin{aligned}
S(\Lambda)u_s(\vec{p}) &= S(\Lambda)S(\Lambda_p)u_s(\vec{0}) = S(\Lambda_{\Lambda p})S(\Lambda_{\Lambda p})^{-1}S(\Lambda)S(\Lambda_p)u_s(\vec{0}) \\
&= S(\Lambda_{\Lambda p})S(\Lambda_{\Lambda p}^{-1}\Lambda\Lambda_p)u_s(\vec{0}) = S(\Lambda_{\Lambda p})S(W(\Lambda, \vec{p}))u_s(\vec{0}) = \sum_{\bar{s}}\mathscr{D}^{(\frac{1}{2})}_{\bar{s}s}(W(\Lambda, \vec{p}))u_{\bar{s}}(\Lambda\vec{p})\,, \\
S(\Lambda)v_s(\vec{p}) &= S(\Lambda)S(\Lambda_p)v_s(\vec{0}) = S(\Lambda_{\Lambda p})S(\Lambda_{\Lambda p})^{-1}S(\Lambda)S(\Lambda_p)v_s(\vec{0}) \\
&= S(\Lambda_{\Lambda p})S(\Lambda_{\Lambda p}^{-1}\Lambda\Lambda_p)v_s(\vec{0}) = S(\Lambda_{\Lambda p})S(W(\Lambda, \vec{p}))v_s(\vec{0}) = \sum_{\bar{s}}\mathscr{D}^{(\frac{1}{2})\,*}_{\bar{s}s}(W(\Lambda, \vec{p}))v_{\bar{s}}(\Lambda\vec{p})\,.
\end{aligned}
\tag{5.28}
$$

Let us now look for an appropriate $S(\Lambda)$. As we said above, this must contain the spin-$\frac{1}{2}$ representation of the rotation group among its irreducible components. In general, the irreducible representations of the proper orthocronous Lorentz group are classified by two half-integers (j_+, j_-), labelling the irreducible representations of the two sets of SU(2) generators $\vec{J}_{\pm}$. Since $\vec{J} = \vec{J}_+ + \vec{J}_-$, the irreducible representation of the Lorentz group can be decomposed into irreducible representations of the rotation group (which will *not* be irreducible representations of the Lorentz group!), with spin ranging in $j = j_+ + j_-, j_+ + j_- - 1, \ldots, |j_+ - j_-|$. The simplest irreducible representations containing $j = \frac{1}{2}$ are then the two inequivalent two-dimensional representations $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$. However, we are also interested in representing parity, and since under parity $\vec{J} \to \vec{J}$ and $\vec{K} \to -\vec{K}$, we have that $\vec{J}_{\pm} \to \vec{J}_{\mp}$, and so the two representations are interchanged, $(\frac{1}{2}, 0) \leftrightarrow (\frac{1}{2}, 0)$. If we want to represent parity we have therefore to use both, and we will take $S(\Lambda)$ to belong to the direct sum $(\frac{1}{2}, 0) \oplus (\frac{1}{2}, 0)$, which is a four-dimensional representation not irreducible under the proper orthocronous Lorentz group, but irreducible under

---

[33]We have in general that $W(R, \vec{p}) = \Lambda_{Rp}^{-1}R\Lambda_p = \Lambda_{Rp}^{-1}R\Lambda_p R^{-1}R = \Lambda_{Rp}^{-1}\Lambda_{Rp}R = R$.

[34]These representations are different but not inequivalent, as for any matrix $U \in \text{SU}(2)$ one has $U^* = (-i\sigma_2)U(i\sigma_2)$.

this group and parity. The generators in the $(\frac{1}{2}, 0)$ and $(\frac{1}{2}, 0)$ representations read

$$
\left(\frac{1}{2}, 0\right): \quad \vec{J}_+ = \frac{\vec{\sigma}}{2}, \quad \vec{J}_- = 0, \quad \vec{J} = \frac{\vec{\sigma}}{2}, \quad \vec{K} = -i\frac{\vec{\sigma}}{2}
$$
$$
\left(0, \frac{1}{2}\right): \quad \vec{J}_+ = 0, \quad \vec{J}_- = \frac{\vec{\sigma}}{2}, \quad \vec{J} = \frac{\vec{\sigma}}{2}, \quad \vec{K} = i\frac{\vec{\sigma}}{2}.
$$

(5.29)

We then have for the representatives of rotations R and pure boosts B

$$
S(R) = \begin{pmatrix} e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} & 0 \\ 0 & e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}} \end{pmatrix}, \qquad S(B) = \begin{pmatrix} e^{\vec{\Theta}\cdot\frac{\vec{\sigma}}{2}} & 0 \\ 0 & e^{-\vec{\Theta}\cdot\frac{\vec{\sigma}}{2}} \end{pmatrix}.
$$

(5.30)

Setting

$$
u_s(\vec{0}) = \begin{pmatrix} \chi_{1s} \\ \chi_{2s} \end{pmatrix}, \qquad v_s(\vec{0}) = \begin{pmatrix} \tilde{\chi}_{1s} \\ \tilde{\chi}_{2s} \end{pmatrix},
$$

(5.31)

where χ_{js} and $\tilde{\chi}_{js}$ are two-component spinors, Eq. (5.27) reads

$$
\begin{pmatrix} e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}}\chi_{1s} \\ e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}}\chi_{2s} \end{pmatrix} = \left(e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}}\right)_{\bar{s}s} \begin{pmatrix} \chi_{1\bar{s}} \\ \chi_{2\bar{s}} \end{pmatrix}, \qquad \begin{pmatrix} e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}}\tilde{\chi}_1 \\ e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}}\tilde{\chi}_{2s} \end{pmatrix} = \left(e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}}\right)_{\bar{s}s}^* \begin{pmatrix} \tilde{\chi}_{1\bar{s}} \\ \tilde{\chi}_{2\bar{s}} \end{pmatrix}.
$$

(5.32)

Considering infinitesimal rotations we find

$$
\begin{pmatrix} \sigma_k\chi_{1s} \\ \sigma_k\chi_{2s} \end{pmatrix} = (\sigma_k)_{\bar{s}s} \begin{pmatrix} \chi_{1\bar{s}} \\ \chi_{2\bar{s}} \end{pmatrix},
$$
$$
\begin{pmatrix} \sigma_k\tilde{\chi}_{1s} \\ \sigma_k\tilde{\chi}_{2s} \end{pmatrix} = -(\sigma_k^*)_{\bar{s}s} \begin{pmatrix} \tilde{\chi}_{1\bar{s}} \\ \tilde{\chi}_{2\bar{s}} \end{pmatrix} = (\sigma_2\sigma_k\sigma_2)_{\bar{s}s} \begin{pmatrix} \tilde{\chi}_{1\bar{s}} \\ \tilde{\chi}_{2\bar{s}} \end{pmatrix}.
$$

(5.33)

Writing explicitly all the indices, the two-component spinors read χ_{jms} and $\tilde{\chi}_{jms}$ with $m = 1, 2$. We can then treat χ_j and $\tilde{\chi}_j$ as 2×2 matrices. In this notation Eq. (5.33) reads

$$
\sigma_k\chi_j = \chi_j\sigma_k \qquad \Rightarrow \quad [\sigma_k, \chi_j] = 0,
$$
$$
\sigma_k\tilde{\chi}_j = \tilde{\chi}_j\sigma_2\sigma_k\sigma_2 \quad \Rightarrow \quad [\sigma_k, \tilde{\chi}_j i\sigma_2] = 0,
$$

(5.34)

From this it follows by Schur's lemma (and also by direct calculation) that χ_j and $\tilde{\chi}_j i\sigma_2$ must all be multiples of $\mathbf{1}_2$,

$$
\left(\chi_{j\frac{1}{2}} \mid \chi_{j-\frac{1}{2}} \right) = \alpha_j\mathbf{1}_2, \qquad \left(\tilde{\chi}_{j\frac{1}{2}} \mid \tilde{\chi}_{j-\frac{1}{2}} \right) = \tilde{\alpha}_j(-i\sigma_2).
$$

(5.35)

We conclude that

$$
u_s(\vec{0}) = \begin{pmatrix} \alpha_1 w_s \\ \alpha_2 w_s \end{pmatrix}, \qquad v_s(\vec{0}) = \begin{pmatrix} \tilde{\alpha}_1 \tilde{w}_s \\ \tilde{\alpha}_2 \tilde{w}_s \end{pmatrix},
$$

(5.36)

where the orthonormal sets w_s and $\tilde{w}_s$ are given by

$$
w_{\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad w_{-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},
$$
$$
\tilde{w}_{\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \tilde{w}_{-\frac{1}{2}} = -\begin{pmatrix} 1 \\ 0 \end{pmatrix}.
$$

(5.37)

Eqs. (5.36) and (5.37) correspond to the fact that the coefficients $u_s(\vec{0})$ and $v_s(\vec{0})$ are associated to the annihilation (resp. creation) operator for a particle (resp. antiparticle) at rest with of third component of the angular momentum (i.e., spin since the particle is at rest) equal to s. The properties of w_s and $\tilde{w}_s$ are summarised here:

$$\sigma_3 w_s = (-1)^{\frac{1}{2}-s} w_s\,, \qquad w_{s'}^\dagger w_s = \delta_{s's}\,, \qquad \sum_s w_s w_{s'}^\dagger = \mathbf{1}_2\,,$$

$$\sigma_3 \tilde{w}_s = -(-1)^{\frac{1}{2}-s} \tilde{w}_s\,, \qquad \tilde{w}_{s'}^\dagger \tilde{w}_s = \delta_{s's}\,, \qquad \sum_s \tilde{w}_s \tilde{w}_{s'}^\dagger = \mathbf{1}_2\,,$$

$$(5.38)$$

and the two sets are related as follows,

$$\tilde{w}_s = -i\sigma_2 w_s\,. \tag{5.39}$$

5.1.1 Parity

So far the coefficients α_{js} and $\tilde{\alpha}_{js}$ are undetermined, and the two sets $\tilde{w}_s$ and w_s are unrelated. To make progress, we have now to consider representing parity. If parity is a symmetry, we can represent it unitarily on the space of states. For the creation operators we then have

$$U(P)b_s(\vec{p})^\dagger U(P)^\dagger = \eta b_s(-\vec{p})^\dagger\,,$$

$$U(P)d_s(\vec{p})^\dagger U(P)^\dagger = \eta' d_s(-\vec{p})^\dagger\,,$$

$$(5.40)$$

for appropriate phases η and η'. A little manipulation leads to

$$U(P)^\dagger \psi(x) U(P) = \int d\Omega_p \sum_s \left\{ u_s(\vec{p})e^{-ip\cdot x}\eta b_s(-\vec{p}) + v_s(\vec{p})e^{ip\cdot x}\eta' d_s(-\vec{p})^\dagger \right\}$$

$$= \eta \int d\Omega_p \sum_s \left\{ u_s(-\vec{p})e^{-ip\cdot Px}b_s(\vec{p}) + (\eta\eta')^* v_s(-\vec{p})e^{ip\cdot Px}d_s(\vec{p})^\dagger \right\}\,.$$

$$(5.41)$$

In order to have simple transformation properties under parity, we ask that this field be a linear combination of the components of $\psi(x)$. In this way we will be sure that we will avoid problems with causality, once that we have imposed the microcausality condition on $\psi(x)$. We then set

$$u_s(-\vec{p}) = Pu_s(\vec{p})\,, \qquad (\eta\eta')^* v_s(-\vec{p}) = Pv_s(\vec{p})\,, \tag{5.42}$$

for some matrix P, so that

$$U(P)^\dagger \psi(x) U(P) = \eta P\psi(Px)\,. \tag{5.43}$$

Applying parity twice we obtain the identity

$$[U(P)^\dagger]^2 \psi(x)[U(P)]^2 = \eta^2 P^2 \psi(x) =$$

$$= \eta^2 \int d\Omega_p \sum_s \left\{ u_s(\vec{p})e^{-ip\cdot x}b_s(\vec{p}) + [(\eta\eta')^*]^2 v_s(\vec{p})e^{ip\cdot x}d_s(\vec{p})^\dagger \right\}\,. \tag{5.44}$$

The right-hand side differs from $\psi(x)$ by a proportionality constant and by the possibly different phase of the positive and negative frequency parts. This would however lead to problems with

causality as in the case of the scalar field. Invoking microcausality, we then need to have $(\eta\eta')^2 = 1$, so that also $P^2 = \mathbf{1}$, and moreover $(\eta\eta') = \pm 1$. Eq. (5.42) then becomes

$$u_s(-\vec{p}) = P u_s(\vec{p}), \qquad c v_s(-\vec{p}) = P v_s(\vec{p}), \tag{5.45}$$

where $c = \eta\eta'$ is a yet undetermined sign. The transformation properties of the generators of rotations and boosts under parity imply that $[P, \vec{J}] = \{P, \vec{K}\} = 0$, which together with $P^2 = \mathbf{1}$ lead to the general form

$$P = \begin{pmatrix} \mathbf{0}_2 & \kappa \mathbf{1}_2 \\ \kappa^{-1} \mathbf{1}_2 & \mathbf{0}_2 \end{pmatrix}, \tag{5.46}$$

and moreover to

$$\begin{aligned}
P u_s(\vec{p}) &= P S(\Lambda_p) P^{-1} P u_s(\vec{0}) = S(\Lambda_{-p}) P u_s(\vec{0}) = S(\Lambda_{-p}) u_s(\vec{0}), \\
P v_s(\vec{p}) &= P S(\Lambda_p) P^{-1} P v_s(\vec{0}) = S(\Lambda_{-p}) P v_s(\vec{0}) = c S(\Lambda_{-p}) v_s(\vec{0})
\end{aligned} \tag{5.47}$$

which is equivalent to

$$P u_s(\vec{0}) = u_s(\vec{0}), \qquad P v_s(\vec{0}) = c v_s(\vec{0}). \tag{5.48}$$

Since it is possible to rescale the basis vectors via $M = \mathrm{diag}(\kappa^{\frac{1}{2}}, \kappa^{-\frac{1}{2}})$ while leaving $\vec{J}$ and $\vec{K}$ unchanged, we can set without loss of generality

$$P = \begin{pmatrix} \mathbf{0}_2 & \mathbf{1}_2 \\ \mathbf{1}_2 & \mathbf{0}_2 \end{pmatrix} \equiv \gamma^0, \tag{5.49}$$

where we have introduced a new notation that will become useful in the following. From now on we will drop the subscript 2. Combining Eqs. (5.47) and (5.49) with Eq. (5.36) we find that

$$u_s(\vec{0}) = n_u \begin{pmatrix} w_s \\ w_s \end{pmatrix}, \qquad v_s(\vec{0}) = n_v \begin{pmatrix} \tilde{w}_s \\ c\tilde{w}_s \end{pmatrix}, \tag{5.50}$$

where $n_{u,v}$ are normalisation factors.

5.1.2 Causality

We have now exploited all the symmetry requirements of Poincaré invariance. To fix the values of c and of the normalisation factors $n_{u,v}$ we have now to impose microcausality. In doing this, we will pretend not to know that the creation and annihilation operators satisfy anticommutation relations, instead of commutation relations as in the scalar case: as a matter of fact we have made no use of this property so far. We will see that the requirement of causality imposes the use of anticommutation relations.

In preparation for this task, we study here a few properties of the matrices $S(\Lambda)$. It is clear from Eq. (5.30) and (5.49) that the following relations hold:

$$\begin{aligned}
S(R)^\dagger &= S(R)^{-1}, & S(B)^\dagger &= S(B), \\
\gamma^0 S(R) \gamma^0 &= S(R), & \gamma^0 S(B) \gamma^0 &= S(B)^{-1},
\end{aligned} \tag{5.51}$$

which can be combined in the compact form

$$\gamma^0 S(\Lambda)^\dagger \gamma^0 = S(\Lambda)^{-1}, \tag{5.52}$$

and which also show that

$$S(R)\gamma^0 S(R)^{-1} = S(R)S(R)^{-1}\gamma^0 = \gamma^0 \,,$$
$$S(B)\gamma^0 S(B)^{-1} = S(B)S(B)^\dagger\gamma^0 = S(B)^2\gamma^0 = S(B^2)\gamma^0 \,. \tag{5.53}$$

A simple calculation shows that $(\vec{\Theta} = |\vec{\Theta}|\hat{\Theta} = \Theta\hat{\Theta})$

$$S(B) = \cosh\frac{\Theta}{2} + \sinh\frac{\Theta}{2}\hat{\Theta} \cdot \begin{pmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & -\vec{\sigma} \end{pmatrix}, \tag{5.54}$$

from which it follows

$$S(B)\gamma^0 S(B)^{-1} = \cosh\Theta - \sinh\Theta\hat{\Theta} \cdot \begin{pmatrix} \mathbf{0} & -\vec{\sigma} \\ \vec{\sigma} & \mathbf{0} \end{pmatrix} = \cosh\Theta - \sinh\Theta\hat{\Theta} \cdot \vec{\gamma}, \tag{5.55}$$

where we have introduced the matrices

$$\vec{\gamma} \equiv \begin{pmatrix} \mathbf{0} & -\vec{\sigma} \\ \vec{\sigma} & \mathbf{0} \end{pmatrix}. \tag{5.56}$$

Taking $B = \Lambda_p$, i.e., a boost in the direction of $\vec{p}$ sending $\vec{0}$ to $\vec{p}$, we have to set $\hat{\Theta} = \hat{p}$, $m\sinh\Theta = |\vec{p}|$, $m\cosh\Theta = p^0 = \sqrt{\vec{p}^2 + m^2}$, and we obtain

$$S(\Lambda_p)m\gamma^0 S(\Lambda_p)^{-1} = S(\Lambda_p)k_\mu\gamma^\mu S(\Lambda_p)^{-1} = p_\mu\gamma^\mu \,, \tag{5.57}$$

where $k^\mu = (m,\vec{0})$. The first equation in Eq. (5.53) reads in the same notation

$$S(R)k_\mu\gamma^\mu S(R)^{-1} = k_\mu\gamma^\mu \,. \tag{5.58}$$

We now combine these two results to show that for a general (proper orthocronous) Lorentz transformation Λ we have

$$\begin{aligned} S(\Lambda)p_\mu\gamma^\mu S(\Lambda)^{-1} &= S(\Lambda)S(\Lambda_p)k_\mu\gamma^\mu S(\Lambda_p)^{-1}S(\Lambda)^{-1} \\ &= S(\Lambda_{\Lambda p})S(\Lambda_{\Lambda p}^{-1}\Lambda\Lambda_p)k_\mu\gamma^\mu S(\Lambda_{\Lambda p}^{-1}\Lambda\Lambda_p)^{-1}S(\Lambda_{\Lambda p})^{-1} \\ &= S(\Lambda_{\Lambda p})S(W(\Lambda,\vec{p}))k_\mu\gamma^\mu S(W(\Lambda,\vec{p}))^{-1}S(\Lambda_{\Lambda p})^{-1} \\ &= S(\Lambda_{\Lambda p})k_\mu\gamma^\mu S(\Lambda_{\Lambda p})^{-1} = (\Lambda p)_\mu\gamma^\mu \,, \end{aligned} \tag{5.59}$$

where we used the fact that $W(\Lambda,\vec{p})$ is a rotation. Since $(\Lambda p)_\mu\gamma^\mu = p_\nu\Lambda_\mu{}^\nu\gamma^\mu = p_\nu\Lambda^{-1\nu}{}_\mu\gamma^\mu$, and since Eq. (5.59) holds for any timelike p, we find[35]

$$S(\Lambda)\gamma^\mu S(\Lambda)^{-1} = \Lambda^{-1\mu}{}_\nu\gamma^\nu \,. \tag{5.60}$$

This shows that the matrices γ^μ transform like a vector, $S(\Lambda)^{-1}\gamma^\mu S(\Lambda) = \Lambda^\mu{}_\nu\gamma^\nu$, as anticipated by the notation. These are the *Dirac gamma matrices*, which are seen to satisfy the following anticommutation relations,

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \,, \tag{5.61}$$

---

[35]Use first $p^\mu = (m,\vec{0})$ to show that $S(\Lambda)\gamma^0 S(\Lambda)^{-1} = \Lambda^{-1 0}{}_\nu\gamma^\nu$. Then subtracting p^0 times this relation from this from Eq. (5.59) we find $S(\Lambda)\vec{p}\cdot\vec{\gamma}S(\Lambda)^{-1} = p^j\Lambda^{-1 j}{}_\nu\gamma^\nu$ for any spatial $\vec{p}$.

which define a *Clifford algebra*. Before proceeding we mention another couple of properties of these matrices. The matrix γ^0 is Hermitian, while the matrices $\vec{\gamma}$ are anti-Hermitian. Moreover, $\gamma^{0,1,3}$ are real and γ^2 is purely imaginary. Making use of Eq. (5.61), this can be summarised as

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0\,, \qquad \gamma^{\mu*} = \gamma^2\gamma^\mu\gamma^2\,. \tag{5.62}$$

These identities will be useful later.

We are now ready to investigate microcausality. We will denote with $[\,,\,]_+ = \{\,,\,\}$ the anticommutator and with $[\,,\,]_- = [\,,\,]$ the commutator, and we will assume that the creation and annihilation operators satisfy *only one* of the two sets of relations

$$[b_s(\vec{p}), b_{s'}(\vec{q})^\dagger]_\pm = \delta_{s's}(2\pi)2p^0\delta^{(3)}(\vec{p}-\vec{q})\,, \qquad [d_s(\vec{p}), d_{s'}(\vec{q})^\dagger]_\pm = \delta_{s's}(2\pi)2p^0\delta^{(3)}(\vec{p}-\vec{q})\,,$$
$$[b_s(\vec{p}), b_{s'}(\vec{q})]_\pm = [d_s(\vec{p}), d_{s'}(\vec{q})]_\pm = [b_s(\vec{p}), d_{s'}(\vec{q})]_\pm = [b_s(\vec{p}), d_{s'}(\vec{q})^\dagger]_\pm = 0\,. \tag{5.63}$$

No matter which of the two sets we choose, it is clear that we will have $[\psi(x), \psi(y)]_\pm = 0$ for all x, y. Consider instead the non-trivial quantity

$$[\psi(x), \psi(y)^\dagger]_\pm = \int d\Omega_p \int d\Omega_q \sum_{s,s'} \Big[u_s(\vec{p})e^{-ip\cdot x}b_s(\vec{p}) + v_s(\vec{p})e^{ip\cdot x}d_s(\vec{p})^\dagger,$$

$$u_{s'}(\vec{q})^\dagger e^{iq\cdot y}b_{s'}(\vec{q})^\dagger + v_{s'}(\vec{q})^\dagger e^{-iq\cdot y}d_{s'}(\vec{q})\Big]_\pm$$

$$= \int d\Omega_p \sum_s \Big\{ u_s(\vec{p})u_s(\vec{p})^\dagger e^{-ip\cdot(x-y)} \pm v_s(\vec{p})v_s(\vec{p})^\dagger e^{ip\cdot(x-y)} \Big\}$$

$$= \int d\Omega_p\, S(\Lambda_p) \sum_s \Big\{ u_s(\vec{0})u_s(\vec{0})^\dagger e^{-ip\cdot(x-y)} \pm v_s(\vec{0})v_s(\vec{0})^\dagger e^{ip\cdot(x-y)} \Big\} S(\Lambda_p)^\dagger\,. \tag{5.64}$$

From Eq. (5.50) we find

$$\sum_s u_s(\vec{0})u_s(\vec{0})^\dagger = |n_u|^2 \sum_s \begin{pmatrix} w_s w_s^\dagger & w_s w_s^\dagger \\ w_s w_s^\dagger & w_s w_s^\dagger \end{pmatrix} = |n_u|^2(\mathbf{1} + \gamma^0)\,,$$

$$\sum_s v_s(\vec{0})v_s(\vec{0})^\dagger = |n_v|^2 \sum_s \begin{pmatrix} \tilde{w}_s \tilde{w}_s^\dagger & c\tilde{w}_s \tilde{w}_s^\dagger \\ c\tilde{w}_s \tilde{w}_s^\dagger & \tilde{w}_s \tilde{w}_s^\dagger \end{pmatrix} = |n_v|^2(\mathbf{1} + c\gamma^0)\,, \tag{5.65}$$

where we have used the completeness of the sets w_s and $\tilde{w}_s$. We then obtain

$$[\psi(x), \psi(y)^\dagger]_\pm = \int d\Omega_p\, S(\Lambda_p) \Big\{ |n_u|^2(\mathbf{1} + \gamma^0)e^{-ip\cdot(x-y)} \pm |n_v|^2(\mathbf{1} + c\gamma^0)e^{ip\cdot(x-y)} \Big\} S(\Lambda_p)^\dagger\,. \tag{5.66}$$

Using Eqs. (5.51)-(5.55) we can recast this as ($\mathcal{A} \equiv A_\mu \gamma^\mu$)

$$
\begin{aligned}
[\psi(x), \psi(y)^\dagger]_\pm &= \int d\Omega_p \left\{ |n_u|^2 \left(\frac{\slashed{p}}{m} + 1 \right) e^{-ip\cdot(x-y)} \pm |n_v|^2 \left(\frac{\slashed{p}}{m} + c \right) e^{ip\cdot(x-y)} \right\} \gamma^0 \\
&= \frac{i}{m} \slashed{\partial} \gamma^0 \int d\Omega_p \left\{ |n_u|^2 e^{-ip\cdot(x-y)} \mp |n_v|^2 e^{ip\cdot(x-y)} \right\} \\
&\quad + \gamma^0 \int d\Omega_p \left\{ |n_u|^2 e^{-ip\cdot(x-y)} \pm c|n_v|^2 e^{ip\cdot(x-y)} \right\} \\
&= \frac{i}{m} \slashed{\partial} \gamma^0 \left\{ |n_u|^2 \Delta(x-y) \mp |n_v|^2 \Delta(y-x) \right\} \\
&\quad + \gamma^0 \left\{ |n_u|^2 \Delta(x-y) \pm c|n_v|^2 \Delta(y-x) \right\}.
\end{aligned}
\tag{5.67}
$$

The function $\Delta(x-y)$ is invariant under Lorentz transformations. For spacelike separations we can then evaluate it in the frame where the two events x and y are at equal times, which shows that for $(x-y)^2 < 0$ we have $\Delta(x-y) = \Delta(y-x)$. As a consequence, the first of the two terms in the last two lines of Eq. (5.67) is odd under $x \leftrightarrow y$, while the second one is even, so they have to vanish separately. We then obtain the following constraints:

$$
\begin{aligned}
|n_u|^2 \mp |n_v|^2 &= 0, \\
|n_u|^2 \pm c|n_v|^2 &= 0.
\end{aligned}
\tag{5.68}
$$

These can be satisfied only if we choose the upper sign, i.e., anticommutation relations, and furthermore

$$
|n_u| = |n_v|, \qquad c = -1. \tag{5.69}
$$

The phases of n_u and n_v can both be chosen equal to 0 without loss of generality (any non-trivial phase can be absorbed in a redefinition of one-particle state by an irrelevant phase), and the overall scale can be set at our convenience; we choose $n_u = n_v = \sqrt{m}$. The condition $c = -1$ implies that the product of intrinsic parities $\eta \eta' = -1$, differently from the scalar case.

In conclusion, we find

$$
\{\psi(x), \psi(y)^\dagger\} = (i\slashed{\partial} + m)\gamma^0 [\Delta(x-y) - \Delta(y-x)]. \tag{5.70}
$$

This can be evaluated explicitly at equal times $x^0 = y^0$ to give

$$
\begin{aligned}
\{\psi(x), \psi(y)^\dagger\}_{\mathrm{ET}} &= (i\slashed{\partial} + m)\gamma^0 \int d\Omega_p \left(e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)} \right) \\
&= \int d\Omega_p \left[(\slashed{p} + m)e^{-ip\cdot(x-y)} + (\slashed{p} - m)e^{ip\cdot(x-y)} \right] \gamma^0 \\
&= \int d\Omega_p \left[p^0 (e^{-ip\cdot(x-y)} + e^{ip\cdot(x-y)}) + m\gamma^0 (e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}) \right] \\
&\quad + \vec{p} \cdot \vec{\gamma} \gamma^0 (e^{-ip\cdot(x-y)} + e^{ip\cdot(x-y)}) \right] = \int d\Omega_p \, 2p^0 e^{-ip\cdot(x-y)} \\
&= \delta^{(3)}(\vec{x} - \vec{y}).
\end{aligned}
\tag{5.71}
$$

Obviously, equal-time anticommutators between any amount of spatial derivatives of fields will still vanish (for $\vec{x} \neq \vec{y}$). The analysis of anticommutators involving temporal derivatives is made

easier by the following observations. The relations Eq. (5.48), which we write now as

$$\gamma^0 u_s(\vec{0}) = u_s(\vec{0})\,, \qquad \gamma^0 v_s(\vec{0}) = -v_s(\vec{0})\,, \tag{5.72}$$

imply that [multiply on the left by $mS(\Lambda_p)$]

$$mu_s(\vec{p}) = S(\Lambda_p)m\gamma^0 S(\Lambda_p)^{-1}u_s(\vec{p}) = \not{p}u_s(\vec{p})\,,$$
$$-mv_s(\vec{p}) = S(\Lambda_p)m\gamma^0 S(\Lambda_p)^{-1}v_s(\vec{p}) = \not{p}v_s(\vec{p})\,, \tag{5.73}$$

i.e.,

$$(\not{p} - m)u_s(\vec{p}) = 0\,, \qquad (\not{p} + m)v_s(\vec{p}) = 0\,. \tag{5.74}$$

The functions $U_s(\vec{p}, x) = u_s(\vec{p})e^{-ip \cdot x}$ ad $V_s(\vec{p}, x) = v_s(\vec{p})e^{ip \cdot x}$ are then solutions of the equation

$$(i\not{\partial} - m)U_s(\vec{p}, x) = (i\not{\partial} - m)V_s(\vec{p}, x) = 0\,, \tag{5.75}$$

with positive and negative energy, respectively. The field $\psi(x)$ obeys therefore the *Dirac equation*,

$$(i\not{\partial} - m)\psi(x) = 0\,. \tag{5.76}$$

This automatically guarantees that anticommutators invoving temporal derivatives of the field will vanish for spacelike separations. Applying the operator $(-i\not{\partial} - m)$ on the left, and exploiting the identity $\not{A} = A_\mu A_\nu \gamma^\mu \gamma^\nu = \frac{1}{2} A_\mu A_\nu \{\gamma^\mu, \gamma^\nu\} = A_\mu A^\mu$ we find that $\psi(x)$ satisfies also the KG equation, $(\Box + m^2)\psi(x) = 0$, as appropriate for a field describing particles of mass m.

Let us now collect a few results concerning the *Dirac bispinors* $u_s(\vec{p})$ and $v_s(\vec{p})$. With our choice of normalisation we have $u_{s'}(\vec{0})^\dagger u_s(\vec{0}) = v_{s'}(\vec{0})^\dagger v_s(\vec{0}) = 2m\delta_{s's}$. Moreover, $u_{s'}(\vec{0})^\dagger v_s(\vec{0}) = v_{s'}(\vec{0})^\dagger u_s(\vec{0}) = 0$ since they are eigenvectors of γ^0 with different eigenvalues. For nonzero $\vec{p}$, using Eq. (5.54) and the relations $\frac{p^0}{m} = \cosh\Theta = 2\cosh^2\frac{\Theta}{2} - 1$ and $\sinh\frac{\Theta}{2} = \sqrt{\cosh^2\frac{\Theta}{2} - 1}$ we find

$$u_s(\vec{p}) = \sqrt{\frac{p^0 + m}{2}}\begin{pmatrix} \left(1 + \frac{\vec{p}\cdot\vec{\sigma}}{p^0+m}\right)w_s \\ \left(1 - \frac{\vec{p}\cdot\vec{\sigma}}{p^0+m}\right)w_s \end{pmatrix}\,, \qquad v_s(\vec{p}) = \sqrt{\frac{p^0 + m}{2}}\begin{pmatrix} \left(1 + \frac{\vec{p}\cdot\vec{\sigma}}{p^0+m}\right)\tilde{w}_s \\ -\left(1 - \frac{\vec{p}\cdot\vec{\sigma}}{p^0+m}\right)\tilde{w}_s \end{pmatrix}\,. \tag{5.77}$$

The equations $\gamma^0 u_s(\vec{0}) = u_s(\vec{0})$ and $\gamma^0 v_s(\vec{p}) = -v_s(\vec{0})$ become at finite $\vec{p}$

$$\gamma^0 u_s(\vec{p}) = \gamma^0 S(\Lambda_p)u_s(\vec{0})\gamma^0 S(\Lambda_p)\gamma^0\gamma^0 u_s(\vec{0}) = S(\Lambda_p)^{-1}u_s(\vec{0}) = S(\Lambda_{-p})u_s(\vec{0}) = u_s(-\vec{p})\,,$$
$$\gamma^0 v_s(\vec{p}) = \gamma^0 S(\Lambda_p)v_s(\vec{0})\gamma^0 S(\Lambda_p)\gamma^0\gamma^0 v_s(\vec{0}) = -S(\Lambda_p)^{-1}v_s(\vec{0}) = S - (\Lambda_{-p})v_s(\vec{0}) = -v_s(-\vec{p})\,, \tag{5.78}$$

as they should. Let us introduce the notation $\bar{u}_s(\vec{p}) \equiv u_s(\vec{p})^\dagger\gamma^0$. These quantities satisfy the equations

$$\bar{u}_s(\vec{p})(\not{p} - m) = u_s(\vec{p})^\dagger\gamma^0(\not{p} - m)\gamma^0\gamma^0 = [\gamma^0(\not{p}^\dagger - m)\gamma^0 u_s(\vec{p})]^\dagger\gamma^0 = [(\not{p} - m)u_s(\vec{p})]^\dagger\gamma^0 = 0\,,$$
$$\bar{v}_s(\vec{p})(\not{p} + m) = v_s(\vec{p})^\dagger\gamma^0(\not{p} + m)\gamma^0\gamma^0 = [\gamma^0(\not{p}^\dagger + m)\gamma^0 v_s(\vec{p})]^\dagger\gamma^0 = [(\not{p} + m)v_s(\vec{p})]^\dagger\gamma^0 = 0\,. \tag{5.79}$$

We have the normalisation conditions

$$\bar{u}_{s'}(\vec{p})u_s(\vec{p}) = u_{s'}(\vec{p})^\dagger\gamma^0 u_s(\vec{p}) = u_{s'}(\vec{0})^\dagger S(\Lambda_p)^\dagger\gamma^0 S(\Lambda_p)u_s(\vec{0})$$
$$= u_{s'}(\vec{0})^\dagger\gamma^0 S(\Lambda_p)^{-1}S(\Lambda_p)u_s(\vec{0}) = u_{s'}(\vec{0})^\dagger\gamma^0 u_s(\vec{0}) = u_{s'}(\vec{0})^\dagger u_s(\vec{0}) = 2m\delta_{s's}\,,$$
$$\bar{v}_{s'}(\vec{p})v_s(\vec{p}) = v_{s'}(\vec{p})^\dagger\gamma^0 v_s(\vec{p}) = v_{s'}(\vec{0})^\dagger S(\Lambda_p)^\dagger\gamma^0 S(\Lambda_p)v_s(\vec{0})$$
$$= v_{s'}(\vec{0})^\dagger\gamma^0 S(\Lambda_p)^{-1}S(\Lambda_p)v_s(\vec{0}) = v_{s'}(\vec{0})^\dagger\gamma^0 v_s(\vec{0}) = -v_{s'}(\vec{0})^\dagger v_s(\vec{0}) = -2m\delta_{s's}\,. \tag{5.80}$$

The same kind of calculation leads to

$$\bar{u}_{s'}(\vec{p})v_s(\vec{p}) = u_{s'}(\vec{0})^\dagger \gamma^0 v_s(\vec{0}) = \pm u_{s'}(\vec{0})^\dagger v_s(\vec{0}) = 0\,,$$
$$\bar{v}_{s'}(\vec{p})u_s(\vec{p}) = v_{s'}(\vec{0})^\dagger \gamma^0 u_s(\vec{0}) = \pm v_{s'}(\vec{0})^\dagger u_s(\vec{0}) = 0\,.$$
$$(5.81)$$

The same result of Eq. (5.81) can be obtained by noting that

$$0 = \bar{u}_{s'}(\vec{p})(\not{p} - \not{p})v_s(\vec{p}) = 2m\bar{u}_{s'}(\vec{p})v_s(\vec{p})\,. \tag{5.82}$$

We have the following relations:

$$\begin{aligned}
\bar{u}_{s'}(\vec{p})\gamma^\mu u_s(\vec{p}) &= u_{s'}(\vec{0})^\dagger S(\Lambda_p)^\dagger \gamma^0 \gamma^\mu S(\Lambda_p) u_s(\vec{0}) = u_{s'}(\vec{0})^\dagger \gamma^0 S(\Lambda_p)^{-1} \gamma^\mu S(\Lambda_p) u_s(\vec{0}) \\
&= \Lambda_p{}^\mu{}_\nu \bar{u}_{s'}(\vec{0})\gamma^\nu u_s(\vec{0}) = \Lambda_p{}^\mu{}_\nu \bar{u}_{s'}(\vec{0})\gamma^0\gamma^\nu\gamma^0 u_s(\vec{0}) = \Lambda_p{}^\mu{}_\nu \delta^\nu{}_0 u_{s'}(\vec{0})^\dagger u_s(\vec{0}) \\
&= \Lambda_p{}^\mu{}_\nu 2k^\nu \delta_{s's} = 2p^\mu \delta_{s's}\,, \\
\bar{v}_{s'}(\vec{p})\gamma^\mu v_s(\vec{p}) &= v_{s'}(\vec{0})^\dagger S(\Lambda_p)^\dagger \gamma^0 \gamma^\mu S(\Lambda_p) v_s(\vec{0}) = v_{s'}(\vec{0})^\dagger \gamma^0 S(\Lambda_p)^{-1} \gamma^\mu S(\Lambda_p) v_s(\vec{0}) \\
&= \Lambda_p{}^\mu{}_\nu \bar{v}_{s'}(\vec{0})\gamma^\nu v_s(\vec{0}) = \Lambda_p{}^\mu{}_\nu \bar{v}_{s'}(\vec{0})\gamma^0\gamma^\nu\gamma^0 v_s(\vec{0}) = \Lambda_p{}^\mu{}_\nu \delta^\nu{}_0 v_{s'}(\vec{0})^\dagger v_s(\vec{0}) \\
&= \Lambda_p{}^\mu{}_\nu 2k^\nu \delta_{s's} = 2p^\mu \delta_{s's}\,.
\end{aligned} \tag{5.83}$$

In particular

$$\bar{u}_{s'}(\vec{p})\gamma^0 u_s(\vec{p}) = u_{s'}(\vec{p})^\dagger u_s(\vec{p}) = 2p^0 \delta_{s's}\,,$$
$$\bar{v}_{s'}(\vec{p})\gamma^0 v_s(\vec{p}) = v_{s'}(\vec{p})^\dagger v_s(\vec{p}) = 2p^0 \delta_{s's}\,. \tag{5.84}$$

Notice that

$$\bar{u}_{s'}(\vec{p})v_s(\vec{p}) = u_{s'}(\vec{p})^\dagger \gamma^0 v_s(\vec{p}) = -u_{s'}(\vec{p})^\dagger v_s(-\vec{p}) = 0\,,$$
$$\bar{v}_{s'}(\vec{p})u_s(\vec{p}) = v_{s'}(\vec{p})^\dagger \gamma^0 u_s(\vec{p}) = v_{s'}(\vec{p})^\dagger u_s(-\vec{p}) = 0\,. \tag{5.85}$$

Finally, we have for the positive and negative energy projectors of Eq. (5.65)

$$\sum_s u_s(\vec{p})\bar{u}_s(\vec{p}) = S(\Lambda_p)m(\gamma^0 + \mathbf{1})S(\Lambda_p)^{-1} = \not{p} + m\,,$$
$$\sum_s v_s(\vec{p})\bar{v}_s(\vec{p}) = S(\Lambda_p)m(\gamma^0 - \mathbf{1})S(\Lambda_p)^{-1} = \not{p} - m\,. \tag{5.86}$$

5.1.3 Charge conjugation

The sets w_s and $\tilde{w}_s$ are related by Eq. (5.39) due to the transformation properties of u_s and v_s under rotations. We will see now how the request of charge conjugation invariance leads to the same relation. In full analogy with the scalar case we define the charge conjugation operator $U(C)$ via

$$U(C)b_s(\vec{p})^\dagger U(C)^\dagger = \xi d_s(\vec{p})^\dagger\,,$$
$$U(C)d_s(\vec{p})^\dagger U(C)^\dagger = \xi' b_s(\vec{p})^\dagger\,, \tag{5.87}$$

from which we can derive the transformation law for the field,

$$U(C)^\dagger \psi_\alpha(x) U(C) = \int d\Omega_p \sum_s \left\{ \xi' u_{s\alpha}(\vec{p})d_s(\vec{p})e^{-ip\cdot x} + \xi^* v_{s\alpha}(\vec{p})b_s(\vec{p})^\dagger e^{ip\cdot x} \right\}$$
$$= \xi^* \left\{ \int d\Omega_p \left\{ v_{s\alpha}(\vec{p})^* b_s(\vec{p})e^{-ip\cdot x} + +(\xi\xi')^* u_{s\alpha}(\vec{p})^* d_s(\vec{p})^\dagger e^{ip\cdot x} \right\} \right\}^\dagger\,, \tag{5.88}$$

where we written explicitly the Lorentz index α of the bispinors. If we ask for the right-hand side to be proportional to $\psi(x)^\dagger$, we will obtain simple transformation properties and we will be sure that no problems with causality can arise. This leads to require that

$$
\begin{aligned}
C^* u_s(\vec{p}) &= v_s(\vec{p})^* \\
C^* v_s(\vec{p}) &= (\xi\xi')^* u_s(\vec{p})^* ,
\end{aligned}
\tag{5.89}
$$

for some matrix C. Notice that C and $(\xi\xi')$ must satisfy $CC^* u_s(\vec{p}) = [C^* v_s(\vec{p})]^* = (\xi\xi') u_s(\vec{p})$. We then find

$$
U(C)^\dagger \psi_\alpha(x) U(C) = \xi^* C_{\alpha\beta} \left\{ \int d\Omega_p \left\{ u_{s\beta}(\vec{p}) b_s(\vec{p}) e^{-ip\cdot x} + v_{s\beta}(\vec{p}) d_s(\vec{p})^\dagger e^{ip\cdot x} \right\} \right\}^\dagger .
\tag{5.90}
$$

Applying charge conjugation twice we find

$$
\begin{aligned}
[U(C)^\dagger]^2 \psi_\alpha(x) [U(C)]^2 &= (\xi\xi') \psi(x) \\
&= \xi^* C_{\alpha\beta} \left\{ \int d\Omega_p \sum_s \left\{ \xi' u_{s\beta}(\vec{p}) d_s(\vec{p}) e^{-ip\cdot x} + \xi^* v_{s\beta}(\vec{p}) b_s(\vec{p})^\dagger e^{ip\cdot x} \right\} \right\}^\dagger \\
&= C_{\alpha\beta} \int d\Omega_p \sum_s \left\{ v_{s\beta}(\vec{p})^* b_s(\vec{p}) e^{-ip\cdot x} + (\xi'\xi)^* u_{s\beta}(\vec{p})^* d_s(\vec{p})^\dagger e^{ip\cdot x} \right\} \\
&= C_{\alpha\beta} C_{\beta\gamma}^* \int d\Omega_p \sum_s \left\{ u_{s\gamma}(\vec{p}) b_s(\vec{p}) e^{-ip\cdot x} + (\xi'\xi)^{*2} v_{s\gamma}(\vec{p}) d_s(\vec{p})^\dagger e^{ip\cdot x} \right\} ,
\end{aligned}
\tag{5.91}
$$

which entails $(\xi'\xi)^2 = 1$ and $CC^* = \mathbf{1}$; this further implies $(\xi'\xi) = 1$. We then have to choose u_s and v_s such that a matrix C exists for which

$$
\begin{aligned}
C^* u_s(\vec{p}) &= v_s(\vec{p})^* , \\
C^* v_s(\vec{p}) &= u_s(\vec{p})^* .
\end{aligned}
\tag{5.92}
$$

In order to do this we have to clarify the relation between the matrices $S(\Lambda) = e^{\frac{i}{2}\omega_{\rho\sigma}J^{(\rho\sigma)}}$ and the gamma matrices. It is a matter of simple algebra to show that

$$
\begin{aligned}
J^{(0j)} &= \vec{K}^j = \frac{1}{2} \begin{pmatrix} -i\sigma_j & \mathbf{0} \\ \mathbf{0} & -i\sigma_j \end{pmatrix} = \frac{1}{2} \frac{[\gamma^0, \gamma^j]}{2i} , \\
J^{(ij)} &= -\epsilon_{ijk} \vec{J}^k = -\epsilon_{ijk} \frac{1}{2} \begin{pmatrix} \sigma_k & \mathbf{0} \\ \mathbf{0} & \sigma_k \end{pmatrix} = \frac{1}{2} \frac{[\gamma^i, \gamma^j]}{2i} , \\
J^{(\mu\nu)} &= \frac{1}{2}\sigma^{\mu\nu} , \qquad \sigma^{\mu\nu} \equiv \frac{1}{2i}[\gamma^\mu, \gamma^\nu] .
\end{aligned}
\tag{5.93}
$$

From Eq. (5.62) we find

$$
(i\sigma^{\mu\nu})^* = \frac{1}{2}[\gamma^2\gamma^i\gamma^2, \gamma^2\gamma^j\gamma^2] = -i\gamma^2\sigma^{\mu\nu}\gamma^2 = i(i\gamma^2)\sigma^{\mu\nu}(i\gamma^2) ,
\tag{5.94}
$$

with $(i\gamma^2)^2 = \mathbf{1}$ and $(i\gamma^2)^* = (i\gamma^2)$. We have then

$$
S(\Lambda)^* = [e^{\frac{i}{2}\omega_{\rho\sigma}\sigma^{(\rho\sigma)}}]^* (i\gamma^2) e^{\frac{i}{2}\omega_{\rho\sigma}\sigma^{(\rho\sigma)}}(i\gamma^2) = (i\gamma^2)S(\Lambda)(i\gamma^2) .
\tag{5.95}
$$

Using this in Eq. (5.92) we find

$$
\begin{aligned}
C^{*} S(\Lambda_p) u_s(\vec{0}) &= (i\gamma^2) S(\Lambda_p)(i\gamma^2) v_s(\vec{0})^{*}\,,\\
C^{*} S(\Lambda_p) v_s(\vec{0}) &= (i\gamma^2) S(\Lambda_p)(i\gamma^2) u_s(\vec{0})^{*}\,.
\end{aligned}
\tag{5.96}
$$

The second equation will be satisfied if we take $C = C^{*} = i\gamma^2$ and $v_s(\vec{0}) = (i\gamma^2) u_s(\vec{0})^{*}$, i.e.,

$$
\sqrt{m}\begin{pmatrix} \tilde{w}_s \\ -\tilde{w}_s \end{pmatrix} = \begin{pmatrix} \mathbf{0} & -i\sigma_2 \\ i\sigma_2 & \mathbf{0} \end{pmatrix} \sqrt{m}\begin{pmatrix} w_s^{*} \\ w_s^{*} \end{pmatrix} = \sqrt{m}\begin{pmatrix} -i\sigma_2 w_s^{*} \\ i\sigma_2 w_s^{*} \end{pmatrix}\,.
\tag{5.97}
$$

We would then set

$$
\tilde{w}_s = -i\sigma_2 w_s^{*}\,.
\tag{5.98}
$$

The other equation becomes then $u_s(\vec{0}) = (i\gamma^2) v_s(\vec{0})^{*}$, which is automatically satisfied since taking the complex conjugate of the equation $v_s(\vec{0}) = (i\gamma^2) u_s(\vec{0})^{*}$ and multiplying by $i\gamma^2$ we find $i\gamma^2 (v_s(\vec{0}))^{*} = (i\gamma^2)(i\gamma^2) u_s(\vec{0}) = u_s(\vec{0})$. With this choice we would conclude that

$$
U(C)^{\dagger} \psi_\alpha(x) U(C) = \xi^{*} (i\gamma^2)_{\alpha\beta} \psi_\beta(x)^{\dagger}\,.
\tag{5.99}
$$

Since w_s are real [see Eq. (5.37)] and $\tilde{w}_s = -i\sigma_2 w_s$, we conclude that the field $\psi(x)$ built in the previous subsections satisfies the transformation law Eq. (5.99) under charge conjugation.

5.1.4 The spin-statistics theorem

We conclude the construction of the Dirac field for spin-$\frac{1}{2}$ particles with an important remark. The need to impose anticommutation rather than commutation relations between creation and annihilation operators originates from the request of locality in the form of the microcausality condition. On the other hand, in the case of a scalar field one has to impose commutation relations in order to obtain locality [try to use anticommutation relations in Eqs. (4.53)–(4.56)]. It is a general result in field theory that the requests of locality, Lorentz invariance and positivity of the Hamiltonian, impose that fields of integer spin (i.e., transforming according to representations of the Lorentz group that contain integer spin representations of SU(2)) have to be quantised by commutators, while fields of half-integer spin have to be quantised by anticommutators. This is known as the *spin-statistics theorem*. Since observables have to commute for spacelike separations, this forces us to use an even number of anticommuting fields of a certain type in their construction.

5.2 Hamiltonian and Lagrangian formalism

At this point we can express the Hamiltonian,

$$
H = \int d\Omega_p\, p^0 [b_s(\vec{p})^{\dagger} b_s(\vec{p}) + d_s(\vec{p})^{\dagger} d_s(\vec{p})]\,,
\tag{5.100}
$$

in terms of fields and their derivatives. This can be done directly using the following relations between creation and annihilation operators and fields,

$$
\begin{aligned}
b_s(\vec{p}) &= \int d^3x\, e^{ip\cdot x} u_s(\vec{p})^{\dagger} \psi(x)\,,\\
d_s(\vec{p})^{\dagger} &= \int d^3x\, e^{-ip\cdot x} v_s(\vec{p})^{\dagger} \psi(x)\,.
\end{aligned}
\tag{5.101}
$$

We will follow a faster route by finding a Lagrangian that has the Dirac field as solution, performing canonical quantisation and producing the Hamiltonian by Legendre transform. The Dirac equation can be obtained form the following (Hermitian) Lagrangian,

$$\mathscr{L} = \bar{\psi} \left(\frac{i}{2} \overset{\leftrightarrow}{\partial\!\!\!/} - m \right) \psi = \bar{\psi} \left(i \partial\!\!\!/ - m \right) \psi + \text{total divergence} \,. \tag{5.102}$$

Since a total divergence does not affect the equations of motion, we can use the second, non-Hermitian form of the Lagrangian to derive the EOM performing a variation with respect to $\bar{\psi}$, which yields

$$\frac{\partial \mathscr{L}}{\partial \bar{\psi}} = \left(i \partial\!\!\!/ - m \right) \psi = 0 \,, \tag{5.103}$$

i.e., the Dirac equation, which is solved by

$$\psi(x) = \int d\Omega_p \sum_s \left\{ u_s(\vec{p}) b_s(\vec{p}) e^{-ip \cdot x} + v_s(\vec{p}) d_s(\vec{p})^\dagger e^{ip \cdot x} \right\} \,, \tag{5.104}$$

with $(p\!\!\!/ - m) u_s(\vec{p}) = (p\!\!\!/ + m) v_s(\vec{p}) = 0$. The operator-valued coefficients $b_s(\vec{p})$ and $d_s(\vec{p})^\dagger$ are determined via canonical quantisation. The momentum associated to ψ is

$$\pi = \frac{\partial \mathscr{L}}{\partial(\partial_0 \psi)} = i \psi^\dagger \,, \tag{5.105}$$

and there is no momentum associated to $\psi^\dagger$.[36] Imposing now equal-time *anticommutation* relations,

$$\{\psi(x), \pi(y)\}_{\text{ET}} = \{\psi(x), i\psi(y)^\dagger\}_{\text{ET}} = i\delta^{(3)}(\vec{x} - \vec{y}) \,,$$
$$\{\psi(x), \psi(y)\}_{\text{ET}} = 0 \,, \qquad \{\pi(x), \pi(y)\}_{\text{ET}} = \{i\psi(x)^\dagger, i\psi(y)^\dagger\}_{\text{ET}} = 0 \,, \tag{5.106}$$

we end up with b_s and d_s satisfying the anticommutation relations Eq. (5.8). The Hamiltonian reads

$$H = \int d^3x \left[\pi(x) \partial_0 \psi(x) - \bar{\psi}(x) \left(i \partial\!\!\!/ - m \right) \psi(x) \right]$$
$$= \int d^3x \left[i\bar{\psi}(x) \gamma^0 \partial_0 \psi(x) - \bar{\psi}(x) \left(i \partial\!\!\!/ - m \right) \psi(x) \right] \tag{5.107}$$
$$= \int d^3x \, \bar{\psi}(x) \left(-i\vec{\nabla} \cdot \vec{\gamma} + m \right) \psi(x) \,,$$

---

[36]The Lagrangian Eq. (5.102) is a *singular* Lagrangian, i.e., the Hessian matrix $\frac{\partial^2 \mathscr{L}}{\partial \dot{q}_i \partial \dot{q}_j}$ is non-invertible. Such Lagrangians need a special formal treatment, and correpsond to Hamiltonian systems on which some constraint has been imposed.

which is Hermitian, as it should be. Substituting Eq. (5.104) in this expression, we find

$$H = \int d^3x \int d\Omega_p \int d\Omega_q \sum_{s,s'} \left\{ \bar{u}_s(\vec{p})b_s(\vec{p})^\dagger e^{ip\cdot x} + \bar{v}_s(\vec{p})d_s(\vec{p})e^{-ip\cdot x} \right\}$$

$$\times \left\{ (\vec{q}\cdot\vec{\gamma} + m)\, u_s(\vec{q})b_s(\vec{q})e^{-iq\cdot x} - (\vec{q}\cdot\vec{\gamma} - m)\, v_s(\vec{q})d_s(\vec{q})^\dagger e^{iq\cdot x} \right\}$$

$$= \int d^3x \int d\Omega_p \int d\Omega_q \sum_{s,s'} \left\{ \bar{u}_s(\vec{p})b_s(\vec{p})^\dagger e^{ip\cdot x} + \bar{v}_s(\vec{p})d_s(\vec{p})e^{-ip\cdot x} \right\}$$

$$\times\, q^0\gamma^0 \left\{ u_s(\vec{q})b_s(\vec{q})e^{-iq\cdot x} - v_s(\vec{q})d_s(\vec{q})^\dagger e^{iq\cdot x} \right\} .$$

$$(5.108)$$

Integration over d^3x brings about delta functions $\delta(\vec{p}\mp\vec{q})$, where the minus sign applies to the $b_s^\dagger b_s$ and $d_s d_s^\dagger$ terms, and the plus sign to the $b_s^\dagger d_s^\dagger$ and $d_s b_s$ terms. The coefficients of these terms are then

$$\bar{u}_s(\vec{p})\gamma^0 u_{s'}(\vec{p}) = 2p^0\delta_{ss'}\,,$$
$$\bar{v}_s(\vec{p})\gamma^0 v_{s'}(\vec{p}) = 2p^0\delta_{ss'}\,,$$
$$\bar{u}_s(\vec{p})\gamma^0 v_{s'}(-\vec{p}) = -\bar{u}_s(\vec{p})v_{s'}(\vec{p}) = 0\,,$$
$$\bar{v}_s(\vec{p})\gamma^0 u_{s'}(-\vec{p}) = \bar{v}_s(\vec{p})u_{s'}(\vec{p}) = 0\,.$$

$$(5.109)$$

Eq. (5.108) becomes

$$H = \int d\Omega_p\, p^0 \sum_s [b_s(\vec{p})^\dagger b_s(\vec{p}) - d_s(\vec{p})d_s(\vec{p})^\dagger]\,. \qquad (5.110)$$

This differs from Eq. (5.100) by an infinite constant, which can be removed by imposing normal ordering on the fields. For anticommuting creation and annihilation operators, we define

$$: d_s(\vec{p})d_s(\vec{q})^\dagger := -d_s(\vec{q})^\dagger d_s(\vec{q})\,, \qquad : d_s(\vec{p})^\dagger d_s(\vec{q}) := d_s(\vec{p})^\dagger d_s(\vec{q})\,, \qquad (5.111)$$

i.e., normal ordering brings creation operators to the left of annihilation operators, but a minus sign has to be included every time that a creation operator crosses an annihilation operator. Overall, this gives the parity of the permutation needed to bring the creation operators on the left. Normal ordering of products of fields is then defined by expanding them in creation and annihilation operators and normal ordering all the products of these. Notice that contrary to the case of scalar fields, now the order of fields in the normal ordered product matters, although only for the overall sign, $: \psi(x)\psi(y) := - : \psi(y)\psi(x) :$. With this definition we finally obtain

$$H = \int d^3x\, : \bar{\psi}(x)\left(-i\vec{\nabla}\cdot\vec{\gamma} + m\right)\psi(x) := \int d\Omega_p\, p^0 \sum_s [b_s(\vec{p})^\dagger b_s(\vec{p}) + d_s(\vec{p})d_s(\vec{p})^\dagger]\,, \qquad (5.112)$$

as desired. Notice that if we had quantised by commutators, then Eq. (5.110) would not be positive-definite.[37]

---

[37]The microcausality condition for ψ and $\psi^\dagger$ would have been satisfied by choosing commutation relations if at the same time we had treated d_s as a creation instead of an annihilation operator. This would have led here to a non positive-definite Hamiltonian.

One can compute the generators of translations and Lorentz transformations from the Lagrangian making use of Noether's theorem.[38] The (canonical) energy-momentum tensor reads

$$\Theta^{\mu\nu} =: \frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi)}\partial^\nu \psi - \eta^{\mu\nu}\mathscr{L} := \tfrac{i}{2} : \bar{\psi}\gamma^\mu \overset{\leftrightarrow}{\partial^\nu}\psi : \tag{5.113}$$

where we have imposed normal ordering to avoid singularities due to products of fields at the same spacetime point. Notice that the Lagrangian vanishes on the EOM, so it does not contribute to Eq. (5.113). The four-momenta are

$$P^\mu = \tfrac{i}{2}\int d^3x : \bar{\psi}(x)\gamma^0 \overset{\leftrightarrow}{\partial^\mu}\psi(x) := i \int d^3x : \psi(x)^\dagger \partial^\mu \psi(x) : . \tag{5.114}$$

From the infinitesimal transformation law

$$\psi'_\alpha(x') = \psi_\alpha(x) + \frac{1}{2}\omega_{\rho\sigma}S^{(\rho\sigma)}_{\alpha\beta}\psi_\beta(x) = \psi_\alpha(x) + \frac{i}{4}\omega_{\rho\sigma}(\sigma^{\rho\sigma})_{\alpha\beta}\psi_\beta(x), \tag{5.115}$$

one obtains the Lorentz generators through the formula

$$J^{(\rho\sigma)} = -\int d^3x\, \mathcal{M}^{0,\rho\sigma},$$

$$\mathcal{M}^{\mu,\rho\sigma} = x^\rho \Theta^{\mu\sigma} - x^\sigma \Theta^{\mu\rho} + \mathcal{S}^{\mu,\rho\sigma} \tag{5.116}$$

$$\mathcal{S}^{\mu,\rho\sigma} = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi_\alpha)}S^{(\rho\sigma)}_{\alpha\beta}\psi_\beta = i(\bar{\psi}\gamma^\mu)_\alpha \frac{i}{2}(\sigma^{\rho\sigma})_{\alpha\beta}\psi_\beta = -\frac{1}{2}\bar{\psi}\gamma^\mu\sigma^{\rho\sigma}\psi,$$

and they read

$$J^{(\rho\sigma)} = -\int d^3x \left\{ x^\rho \Theta^{0\sigma} - x^\sigma \Theta^{0\rho} - \tfrac{1}{2}\psi^\dagger \sigma^{\rho\sigma}\psi \right\}. \tag{5.117}$$

One can compute explicitly the *commutators* to find

$$[P_\mu, \psi(x)] = -i\partial_\mu \psi(x),$$
$$[J^{(\rho\sigma)}, \psi(x)] = i\left(x^\rho \partial^\sigma - x^\sigma \partial^\rho + \tfrac{i}{2}\sigma^{\rho\sigma}\right)\psi(x) = i\left(x^\rho \partial^\sigma - x^\sigma \partial^\rho\right)\psi(x) - \tfrac{1}{2}\sigma^{\rho\sigma}\psi(x). \tag{5.118}$$

The Lagrangian Eq. (5.102) is also invariant under the U(1) transformation

$$\psi' = e^{i\alpha}\psi, \qquad \bar{\psi}' = e^{-i\alpha}\bar{\psi}. \tag{5.119}$$

The corresponding Noether current is

$$J^\mu =: \bar{\psi}\gamma^\mu \psi :, \tag{5.120}$$

and the corresponding charge is

$$Q = \int d^3x : \bar{\psi}\gamma^0 \psi := \int d^3x : \psi^\dagger \psi := \int d\Omega_p \sum_s \left\{ b_s(\vec{p})^\dagger b_s(\vec{p}) - d_s(\vec{p})^\dagger d_s(\vec{p}) \right\}. \tag{5.121}$$

---

[38]Here the Hermitian version of the Lagrangian is used to obtain manifestly Hermitian currents.

6 Interacting theories

So far we have discussed only the case of free fields, for spin 0 and spin $\frac{1}{2}$. This is of quite limited utility to describe the real world, where particles interact in various ways. The next task is then that of building a theory that describes interacting particles while complying with the requests of Poincaré invariance and locality. The use of local fields allows to easily keep track of these two issues, thanks to their loclaity and their simple transformation properties. In the framework of field theory, the method of *canonical quantisation* is a convenient approach to thie problem of building interacting theories: being based on the Lagrangian formalism, it exhibits manifestly the symmetries of the theory, and the requirement of microcausality is satisfied automatically by imposing the canonical commutation or anticommutation relations. Let us list here the steps of the canonical quantisation program.

- Take a classical Lagrangian $\mathscr{L}$ which is a real function of fields $\phi_a(x)$ and their derivatives, $\mathscr{L} = \mathscr{L}(\phi(x), \partial\phi(x))$. Poincaré invariance is manifest if the Lagrangian does not depend explicitly on x and is a scalar under Lorentz transformations, i.e., it transform like $\mathscr{L}(\phi(x), \partial\phi(x)) \to \mathscr{L}(\phi(\Lambda^{-1}x), \partial\phi(\Lambda^{-1}x))$. This is easy to achieve is the fields are assumed to transform in some representation of the Lorentz group.

- Derive the EOM by the usual action principle, and solve them (if you can...).

- Obtain the conjugate momenta as usual, and impose canonical commutation/anticommutation relations at equal time,

$$[\phi_a(x), \phi_b(y)]_{\mathrm{ET}} = [\pi_a(x), \pi_b(y)]_{\mathrm{ET}} = 0 , \qquad [\phi_a(x), \pi_b(y)]_{\mathrm{ET}} = i\delta_{ab}\delta^{(3)}(\vec{x} - \vec{y}) . \quad (6.1)$$

This request makes the fields into linear operators on some Hilbert space. Finding such a Hilbert space is part of the problem of solving the theory, together with solving the EOM. The imposition of CCR/CAR guarantees that observables built out of the field operators will automatically satisfy microcausality.

- Noether's theorem entails the existence of Hermitian generators of the continuous symmetries of the Lagrangian.[39] This leads to a unitary representation of the Poincaré group, under which the quantum field operators transform as their classical counterpart. This guarantees that our quantum system exhibits the desired symmetry.

Unfortunately, it is almost never possible to complete this program in practice: the EOM for interacting theories are usually nonlinear, and it is not known how to solve them. It is therefore necessary to find some approximation technique that allows us to extract something useful from the canonical quantisation program.

6.1 Interaction picture

In many cases of practical interest, the Hamiltonian of the system can be split into a free part and an interacting part, $H = H_0 + V$. Here the free part H_0 is the Hamiltonian of some system which we know how to solve explicitly, for example one of the free-field Hamiltonians discussed

---

[39]Reality of the Lagrangian at the classical level leads to Hermiticity at the quantum level, from which Hermiticity of the Noether charges follows, possibly after settling some operator-ordering issues.

in the previous sections. The interaction part V contains every other term appearing in the full Hamiltonian. The idea is that V can be looked at as a perturbation to the free Hamiltonian H_0, and its effect evaluated in successive steps. This is the *perturbative quantisation* approach, which we now discuss in detail.

Let $\phi(x)$ be the field that realises the canonical quantisation program, and $\pi(x)$ its conjugate momentum. These fields solve the equations of motion and obey the CCR/CAR, and therefore

$$\phi(x) = \phi(t, \vec{x}) = e^{iHt}\phi(0, \vec{x})e^{-iHt}, \qquad \dot{\phi}(t, \vec{x}) = i[H, \phi(t, \vec{x})],$$
$$\pi(x) = \pi(t, \vec{x}) = e^{iHt}\pi(0, \vec{x})e^{-iHt}, \qquad \dot{\pi}(t, \vec{x}) = i[H, \pi(t, \vec{x})]. \tag{6.2}$$

The full Hamiltonian is obtained from the Lagrangian as usual,

$$H = \int d^3x \left[\pi(t, \vec{x})\dot{\phi}(t, \vec{x}) - \mathscr{L}(\phi(t, \vec{x}), \partial\phi(t, \vec{x}))\right], \tag{6.3}$$

where it is understood that $\partial_0\phi$ has to be expressed as a function of ϕ and π. In the cases we will consider, the full Lagrangian can be written as a free Lagrangian plus an interaction term that depends on the fields but not on their derivatives, $\mathscr{L} = \mathscr{L}_0 + \mathscr{L}_I$, with $\mathscr{L}_I = \mathscr{L}_I(\phi)$. It then follows that as a function of fields and their derivatives, the canonical momenta in the full interacting theory satisfy

$$\pi(\phi, \partial\phi) = \frac{\partial\mathscr{L}}{\partial(\partial_0\phi)} = \frac{\partial\mathscr{L}_0}{\partial(\partial_0\phi)} = \pi_0(\phi, \partial\phi), \tag{6.4}$$

i.e., they have the same functional form as the conjugate momentum π_0 of the free theory; by the same token, $\partial_0\phi$ in the full theory will be the same function of ϕ and π as in the free theory. From Eq. (6.3) we then find in this case

$$H = H[\phi, \pi] = \int d^3x \left[\pi(t, \vec{x})\dot{\phi}(t, \vec{x}) - \mathscr{L}_0(\phi(t, \vec{x}), \partial\phi(t, \vec{x})) - \mathscr{L}_I(\phi(t, \vec{x}))\right]$$
$$= H_0[\phi, \pi] - \int d^3x\, \mathscr{L}_I(\phi(t, \vec{x})) = H_0[\phi, \pi] + V[\phi]. \tag{6.5}$$

In the cases of interest, H is time independent. On the other hand, after splitting it into H_0 and V, these will be separately time dependent. Let us do the splitting at $t = 0$,

$$H[\phi(t, \vec{x}), \pi(t, \vec{x})] = H[\phi(0, \vec{x}), \pi(0, \vec{x})] = H_0[\phi(0, \vec{x}), \pi(0, \vec{x})] + V[\phi(0, \vec{x})]. \tag{6.6}$$

From now on, H_0 and V will be those obtained using the interacting fields and momenta at $t = 0$. Let us now define the fields in the interaction picture as fields ewvolving in time with the free Hamiltonian, and coinciding with the full interacting fields (in the Heisenberg picture) at $t = 0$,

$$\phi_{\text{in}}(t, \vec{x}) \equiv e^{iH_0 t}\phi_{\text{in}}(0, \vec{x})e^{-iH_0 t}, \qquad \phi_{\text{in}}(0, \vec{x}) = \phi(0, \vec{x}),$$
$$\pi_{\text{in}}(t, \vec{x}) \equiv e^{iH_0 t}\pi_{\text{in}}(0, \vec{x})e^{-iH_0 t}, \qquad \pi_{\text{in}}(0, \vec{x}) = \pi(0, \vec{x}). \tag{6.7}$$

At $t = 0$ the fields in the interaction picture obey the CCR/CAR, and since their values at time t is obtained via a unitary transformation, they will obey the CCR/CAR at all times. Since they evolve in time with the free Hamiltonian and obey the CCR/CAR, they automatically obey the Hamilton equations of motion of the free theory:

$$\dot{\phi}_{\text{in}} = i[H_0, \phi_{\text{in}}] = \frac{\delta H_0}{\delta\pi}, \qquad \dot{\pi}_{\text{in}} = i[H_0, \pi_{\text{in}}] = -\frac{\delta H_0}{\delta\pi}. \tag{6.8}$$

The fields in the interaction picture are then nothing else but free fields, and we already know exactly what they look like. If, for example, H_0 is the free Hamiltonian for the charged scalar field, we will have that

$$\phi(t,\vec{x}) = \int d\Omega_p \left\{ a(\vec{p})e^{-ip\cdot x} + b(\vec{p})^\dagger e^{ip\cdot x} \right\}, \qquad \pi(t,\vec{x}) = \dot{\phi}(t,\vec{x}), \qquad (6.9)$$

where $a(\vec{p})$, $a(\vec{p})^\dagger$, $b(\vec{p})$, $b(\vec{p})^\dagger$, are the usual annihilation and creation operators. At this point, we define also the interaction Hamiltonian in the interaction picture, $V_I(t)$, as

$$V_I(t) \equiv e^{iH_0 t} V[\phi_{\text{in}}(0,\vec{x})]e^{-iH_0 t} = V[\phi_{\text{in}}(t,\vec{x})] = -\int d^3x \, \mathscr{L}_I(\phi_{\text{in}}(t,\vec{x})). \qquad (6.10)$$

Although it seems that we have made progress, in practice it is so only marginally: for example, if we want to determine the spectrum of the theory, we still have to solve the same eigenvalue problem $H|\psi\rangle = E|\psi\rangle$, and although we have expressed H as a functional of free fields, this does not make the eigenvalue problem any easier to solve. On the other hand, if we now assume that the interaction V is small, for example because it enters the Hamiltonian with some small numerical prefactor, then we can attack the eigenvalue problem perturbatively, and solve it by successive approximations.

6.2   The S-matrix

The interaction picture is particularly useful in the study of scattering problems. In a scattering experiment two bunches of particles are thrown against each other, and the outcome of the collision is observed. It is a fact of Nature that, first of all, a system of free particles can be prepared in a laboratory, i.e., a system of particles that do not interact appreciably with each other and just move on straight-line trajecotries at constant speed. It is another fact of Nature that if one waits enough after the collision, then what is observed is a gain a system of free particles. If the state of the scattering system is determined by some state vector ψ in the Hilbert space, so that at time t the state is $e^{-iHt}\psi$ (in the Schrödinger picture), what experience shows us is that the time evolution of the system looks like the free time evolution both at early times, i.e., at the beginning of the scattering experiment when particles are still far away from eachg other, and at late times, i.e., at the end of the scattering process after a sufficiently long time has elapsed after the collision and particles are not interacting with each other any more. Mathematically, we can express this by saying that as $t \to -\infty$ and $t \to +\infty$, the state vector $e^{-iHt}\psi$ of a scattering system, evolving with the full Hamiltonian H, can be approximated better and better by vectors $e^{-iH_0 t}\varphi_i$ and $e^{-iH_0 t}\varphi_f$, respectively, evolving in time with the free Hamiltonian H_0. Here φ_i and φ_f are vectors in the same Fock space describing free particles. More precisely, we expect that for a state vector ψ descrbing a sacttering state we have

$$\begin{aligned} \|e^{-iHt}\psi - e^{-iH_0 t}\varphi_i\| &\to 0 \text{ for } t \to -\infty, \\ \|e^{-iHt}\psi - e^{-iH_0 t}\varphi_f\| &\to 0 \text{ for } t \to +\infty, \end{aligned} \qquad (6.11)$$

for some $\varphi_{i,f}$. Here $\|\dots\|$ denotes the norm derived from the Hilbert space scalar product. Conversely, if we have a state that looks like $e^{-iH_0 t}\varphi$ in the far past we expect it to describe the early time evolution of some state $e^{-iHt}\psi_+$; similarly, if we have a state that looks like $e^{-iH_0 t}\varphi$

in the far future, we expect it to describe the late time evolution of some state $e^{-iHt}\psi_-$. We then expect that for every free-particle state vector φ there are vectors $\psi_\pm$ such that

$$\|e^{-iHt}\psi_+ - e^{-iH_0t}\varphi\| \to 0 \text{ for } t \to -\infty \,,$$
$$\|e^{-iHt}\psi_- - e^{-iH_0t}\varphi\| \to 0 \text{ for } t \to +\infty \,,$$
(6.12)

These last two equations can be recast as

$$\lim_{t\to\mp\infty} \|\psi_\pm - e^{iHt}e^{-iH_0t}\varphi\| = 0 \,.$$
(6.13)

This defines the *in states* ψ_+ and the *out states* ψ_- as the limits

$$\psi_\pm = \lim_{t\to\mp\infty} e^{iHt}e^{-iH_0t}\varphi \equiv \Omega_\pm\varphi \,,$$
(6.14)

The linear operators $\Omega_\pm$ are the *scattering (or Møller) operators.* Being the limits (in the strong sense) of unitary operators, they preserve the norm of vectors, and are therefore *isometries*, i.e., $\|\psi_\pm\| = \|\Omega_\pm\varphi\| = \|\varphi\|$, for all φ, which implies the property $\Omega_\pm^\dagger\Omega_\pm = \mathbf{1}$. The proof of these properties goes as follows. One has

$$\|\psi_\pm\| = \lim_{t\to\mp\infty} \|\psi_\pm - e^{iHt}e^{-iH_0t}\varphi + e^{iHt}e^{-iH_0t}\varphi\|$$
$$\leq \lim_{t\to\mp\infty} \|\psi_\pm - e^{iHt}e^{-iH_0t}\varphi\| + \|e^{iHt}e^{-iH_0t}\varphi\| = \|\varphi\| \,,$$
(6.15)

where we made use of Schwartz inequality and of $\|e^{iHt}e^{-iH_0t}\varphi\| = \|\varphi\|$. On the other hand,

$$\|\varphi\| = \lim_{t\to\mp\infty} \|e^{iHt}e^{-iH_0t}\| = \lim_{t\to\mp\infty} \|e^{iHt}e^{-iH_0t}\varphi - \psi_\pm + \psi_\pm\|$$
$$\leq \lim_{t\to\mp\infty} \|e^{iHt}e^{-iH_0t}\varphi - \psi_\pm\| + \|\psi_\pm\| = \|\psi_\pm\| \,,$$
(6.16)

and so the equality sign must hold. By definition of adjoint then, $\|\varphi\|^2 = \|\psi_\pm\|^2 = (\psi_\pm, \psi_\pm) = (\Omega_\pm\varphi, \Omega_\pm\varphi) = (\varphi, \Omega_\pm^\dagger\Omega_\pm\varphi)$. Since any scalar product can be written as a linear combination of squared norms,[40] this implies $(\varphi_1, \Omega_\pm^\dagger\Omega_\pm\varphi_2) = (\varphi_1, \varphi_2)$, and so $\Omega_\pm^\dagger\Omega_\pm = \mathbf{1}$. This does not mean that $\Omega_\pm$ are unitary: in fact, this requires that their image is the whole Hilbert space, or equivalently that $\Omega_\pm\Omega_\pm^\dagger = \mathbf{1}$. In principle, it is possible that the theory contains states (*bound states*)that are not accessible via scattering experiments. It is also possible in principle that the spaces of in and out states differ. From the physical standpoint, since the states ψ_- corresponding to the outcomes of scattering experiments can be used as initial states of other scattering experiments, we expect the two spaces to be the same; proving this mathematically is a hard task. For what concerns us, we will assume that the theory has no bound state unreachable by a scattering experiment, and that in and out states span the same space, which is then the whole Hilbert space of the system. With these extra assumptions, the scattering operators $\Omega_\pm$ are guaranteed to be unitary. In summary, these operators connect the freely evolving state vectors at early and late times with the exact state vector, evolving with the full Hamiltonian.

Suppose now that we have prepared the system in a state φ_i, which allows to identify the state of the system at early times as $e^{-iH_0t}\varphi_i$, for $t \to -\infty$. Long after the collision has taken place

---

[40]This is $4(\chi, \varphi) = 4\mathrm{Re}\,(\chi, \varphi) + 4\mathrm{Im}\,(\chi, \varphi) = \|\chi + \varphi\|^2 - \|\chi - \varphi\|^2 + \|\chi - i\varphi\|^2 - \|\chi + i\varphi\|^2$.

we measure the properties of the emerging particles. What is the probability that the final state that we observe is identified by the state vector φ_f, meaning that at late times the system evolves as $e^{-iH_0 t}\varphi_f$, for $t \to +\infty$? If we knew the exact state vectors $\psi_+ = \Omega_+ \varphi_i$ and $\psi_- = \Omega_- \varphi_f$ then this probability would simply be the absolute value square of the transition amplitude $\langle \psi_- | \psi_+ \rangle$, i.e., the scalar product of the two vectors; equivalently, this can be seen as the expectation value of the projection operator $|\psi_-\rangle\langle\psi_-|$, corresponding to our detectors identifying that specific state. No matter how we interpret it, the result is $P_{fi} = |\langle \psi_- | \psi_+ \rangle|^2 = \langle \psi_+ | \psi_- \rangle \langle \psi_- | \psi_+ \rangle$. This quantity is time-independent, so we can use the state vectors at any time to compute it. We can then comute the scalar product as follows,

$$\langle \psi_- | \psi_+ \rangle = \langle \psi_- | e^{iHT_f} e^{-iHT_f} | \psi_+ \rangle = \langle \psi_- | e^{iHT_f} e^{-iH(T_f - T_i)} e^{-iHT_i} | \psi_+ \rangle. \tag{6.17}$$

By the first equality we can use the state vectors at $t = T_f$ when the measurement is made to compute the transition amplitude; the evolved of the initial state is obtained by translating forward in the future by $T_f - T_i$ the initial state, prepared at $t = T_i$. This is all academic, since we do not have direct access to $|\psi_\pm\rangle$. However, in the limit of $T_f \to infty$ and $T_i \to -\infty$ we can replace the exactly evolving states with the freely evolving states, which are defined by our experimental setup, and are therefore under control. We then have

$$\begin{aligned}
\langle \psi_- | \psi_+ \rangle &= \lim_{T_f \to +\infty} \lim_{T_i \to -\infty} \langle \psi_- | e^{iHT_f} e^{-iH(T_f - T_i)} e^{-iHT_i} | \psi_+ \rangle \\
&= \lim_{T_f \to +\infty} \lim_{T_i \to -\infty} \langle \varphi_f | e^{iH_0 T_f} e^{-iHT_f} e^{iHT_i} e^{-iH_0 T_i} | \varphi_i \rangle \\
&= \langle \varphi_f | \Omega_-^\dagger \Omega_+ | \varphi_i \rangle = \langle \varphi_f | S | \varphi_i \rangle \equiv S_{fi},
\end{aligned} \tag{6.18}$$

where we have defined the *S-operator*,

$$S \equiv \Omega_-^\dagger \Omega_+. \tag{6.19}$$

Its matrix elements $S_{fi} = \langle \varphi_f | S | \varphi_i \rangle$ constitute the S-matrix. The S operator can be expressed in a very useful way by recalling the definition of the scattering operators, or by looking at the second line in Eq. (6.18). We have, recalling Eqs. (2.115) and (2.117),

$$\begin{aligned}
S &= \lim_{T_f \to +\infty} \lim_{T_i \to -\infty} e^{iH_0 T_f} e^{-iHT_f} e^{iHT_i} e^{-iH_0 T_i} = \lim_{T_f \to +\infty} \lim_{T_i \to -\infty} \mathcal{U}(T_f, 0)\mathcal{U}(0, T_i) \\
&= \lim_{T_f \to +\infty} \lim_{T_i \to -\infty} \mathcal{U}(T_f, T_i) = \mathcal{U}(+\infty, -\infty) = \text{Texp}\left\{ -i \int_{-\infty}^{+\infty} dt\, V_I(t) \right\},
\end{aligned} \tag{6.20}$$

where $V_I(t) = e^{iH_0 t} V e^{-iH_0 t}$ is the interaction Hamiltonian in the interaction picture.

The discussion above is rather general, and makes no reference to quantum field theory. If now we consider a relativistic quantum theory obtained by means of the canonical quantisation program, $V_I(t)$ is precisely the operator defined in Eq. (6.10), which is built in terms of fields in the interaction picture, so in practice it is built out of free fields, which we know how to manage. As for the initial and final states of freely evolving particles, they are built out of the creation and annihilation operators associated to the fields in the interaction picture, so again we know howe to deal with them. The relevant matrix elements are then

$$S_{fi} = \langle \vec{p}_1' s_1' \alpha_1', \vec{p}_2' s_2' \alpha_2', \dots, \vec{p}_{n'}' s_{n'}' \alpha_{n'}' | S | \vec{p}_1 s_1 \alpha_1, \vec{p}_2 s_2 \alpha_2, \dots, \vec{p}_n s_n \alpha_n \rangle, \tag{6.21}$$

where $\vec{p}_i$, s_i are the momenta and third component of the spin of the particles in the initial state, whose other quantum numbers are denoted collectively by α_i; primed quantities have the same meaning but for the particles in the final state.

As in the problem of finding the spectrum of the theory, finding the exact solution of the theory means computing exactly the S-operator, which is no easy task. On the other hand, if V, and so V_I, is a small perturbation, then we can power-expand S,

$$
\begin{aligned}
S &= \mathrm{Texp}\left\{-i \int_{-\infty}^{+\infty} dt\, V_I(t)\right\} \\
&= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{+\infty} d\tau_1 \ldots \int_{-\infty}^{+\infty} d\tau_n T\left\{V_I(\tau_1)\ldots V_I(\tau_n)\right\},
\end{aligned}
\tag{6.22}
$$

and compute the S-matrix elements order by order in the perturbation. We already know that products of fields at the same spacetime point lead to problems with infinities. To get rid of (part of) these, we take V_I to be normal-ordered: this does not change the symmetries of the S operator. Recalling again Eq. (6.10), and imposing normal ordering, we have that

$$
\begin{aligned}
S &= \mathrm{Texp}\left\{i \int d^4x\, :\mathscr{L}_I(\phi_{\mathrm{in}}(x)):\right\} \\
&= \sum_{n=0}^{\infty} \frac{i^n}{n!} \int d^4x_1 \ldots \int d^4x_n\, T\left\{:\mathscr{L}_I(\phi_{\mathrm{in}}(x_1)): \ldots :\mathscr{L}_I(\phi_{\mathrm{in}}(x_n)):\right\}.
\end{aligned}
\tag{6.23}
$$

6.3 Symmetries of the S-matrix

We discuss now the general properties of the S operator. The most important property is that it is a unitary operator, even when $\Omega_\pm$ are not unitary but simply isometric. Assume that a bound-state subspace exists in the Hilbert space of the system. Such states cannot be reached in a scattering experiment, and are orthogonal to the scattering states, which on the other hand can be obtained as $\psi_\pm = \Omega_\pm \varphi$ for some φ.[41] For a bound state χ we then have $0 = (\chi, \psi_\pm) = (\chi, \Omega_\pm \varphi) = (\Omega_\pm^\dagger \chi, \varphi)$, and so $\Omega_\pm^\dagger \chi = 0$. For a scattering state we have instead that $\psi_\pm = \Omega_\pm \varphi$ can be inverted to give[42] $\Omega_\pm^{-1} \psi_\pm = \varphi$, and so $(\psi_\pm, \Omega_\pm \varphi') = (\Omega_\pm \varphi, \Omega_\pm \varphi') = (\varphi, \varphi') = (\Omega_\pm^{-1}\psi_\pm, \varphi') = (\Omega_\pm^\dagger \psi_\pm, \varphi')$, so $\Omega_\pm^{-1} = \Omega_\pm^\dagger$. We then conclude that $\Omega_\pm \Omega_\pm^\dagger = \mathbf{1} - \Pi_B$ where Π_B is the projector on the bound state subspace. We then have for the S operator

$$
\begin{aligned}
S^\dagger S &= \Omega_+^\dagger \Omega_- \Omega_-^\dagger \Omega_+ = \Omega_+^\dagger(\mathbf{1} - \Pi_B)\Omega_+ = \Omega_+^\dagger \Omega_+ = \mathbf{1}, \\
SS^\dagger &= \Omega_-^\dagger \Omega_+ \Omega_+^\dagger \Omega_- = \Omega_-^\dagger(\mathbf{1} - \Pi_B)\Omega_- = \Omega_-^\dagger \Omega_- = \mathbf{1},
\end{aligned}
\tag{6.24}
$$

where we use the fact that $\Pi_B \Omega_\pm = 0$. The S operator is thus unitary. From now one we will assume fro simplicity that there are no bound states, so that $\Omega_\pm \Omega_\pm^\dagger = \mathbf{1}$.

From the definition of the scattering operators we can prove easily that

$$
\begin{aligned}
e^{iHs}\Omega_\pm &= \lim_{t\to\mp\infty} e^{iH(t+s)}e^{-iH_0 t} = \lim_{t\to\mp\infty} e^{iH(t+s)}e^{-iH_0(t+s)}e^{iH_0 s} \\
&= \lim_{t\to\mp\infty} e^{iHt}e^{-iH_0 t}e^{iH_0 s} = \Omega_\pm e^{iH_0 s},
\end{aligned}
\tag{6.25}
$$

---

[41]We still assume that the in and out states $\psi_\pm$ span the same subspace of the Hilbert space.

[42]The scattering operators are by definition onto the subspace of scattering states, and the are one-to-one since they are isometric, $\Omega_\pm \varphi_1 = \Omega_\pm \varphi_2 \Rightarrow \varphi_1 = \Omega_\pm^\dagger \Omega_\pm \varphi_2 = \varphi_2$, so there is an inverse map from the scattering state subspace to the space of free particle states.

for any s, and so

$$H\Omega_\pm = \Omega_\pm H_0\,. \tag{6.26}$$

Using this intertwining relation and its adjoint $H_0\Omega_\pm^\dagger = \Omega_\pm^\dagger H$ we can show that

$$H_0 S = H_0\Omega_-^\dagger \Omega_+ = \Omega_-^\dagger H\Omega_+ = \Omega_-^\dagger \Omega_+ H_0 = SH_0 \Longrightarrow [H_0, S] = 0\,. \tag{6.27}$$

Since the interaction Lagrangian is taken to be invariant under translations, both in time and space, this result can be obtained more directly from $e^{ia\cdot P_0} S e^{-ia\cdot P_0} = S$, from which we obtain

$$[P_{0\mu}, S] = 0\,, \tag{6.28}$$

of which Eq. (6.27) is the $\mu = 0$ case. Here the subscript 0 indicates that the generators of translations $P_{0\mu}$ are those of the free theory, builit out of th efileds in the interaction picture. If the interaction Lagrangian is invariant under any internal symmetry of the free Lagrangian, then S will be invariant under the corresponding unitary transformations, $U_{\text{int}}^\dagger S U_{\text{int}} = S$, and will therefore commute with the corresponding generators Q_0^a,

$$[Q_0^a, S] = 0\,. \tag{6.29}$$

Lorentz transformations, on the other hand, require a special treatment. Under a unitary tranformation $U_0(\Lambda^{-1}) = U(\Lambda)^\dagger$, where again the subscript 0 indicates that we are using the representation of the symmetry in the free theory, we find for the S operator

$$U_0(\Lambda) S U_0(\Lambda)^\dagger = \sum_{n=0}^\infty \frac{i^n}{n!} \int d_1^x \dots \int d^4x_n\, U_0(\Lambda) T_x \left\{:\mathscr{L}_I(\phi_{\text{in}}(x_1)): \dots :\mathscr{L}_I(\phi_{\text{in}}(x_n)):\right\} U_0(\Lambda)^\dagger$$

$$= \sum_{n=0}^\infty \frac{i^n}{n!} \int d^4x_1 \dots \int d^4x_n\, T_x \left\{:\mathscr{L}_I(\phi_{\text{in}}(\Lambda x_1)): \dots :\mathscr{L}_I(\phi_{\text{in}}(\Lambda x_n)):\right\}\,. \tag{6.30}$$

Here the subscript x in T_x stands for the fact that time-ordering is done according to the values of x_j^0 in the original reference frame. We made use of the fact that the interaction Lagrangian is a Lorentz scalar, $U_0(\Lambda)^\dagger \mathscr{L}_I(\phi_{\text{in}}(x)) U_0(\Lambda) = \mathscr{L}_I(\phi_{\text{in}}(\Lambda^{-1}x))$. An orthocronous Lorentz transformation will not change the relative ordering in time of timelike or lightlike-separated points x, y, so that $\theta(x^0 - y^0) = \theta((\Lambda x)^0 - (\Lambda y)^0)$, but nothing can be said in generale for spacelike-separated points. Here is where the requirement of microcausality plays a crucial role: for causal fields we are guaranteed that[43] $[\mathscr{L}_I(x), \mathscr{L}_I(y)] = 0$ if $(x - y)^2 < 0$, so the relative time-ordering of spacelike-separated points has no consequence. After changing frame via Λ, a point x_j can "cross" in time only those points that are spacelike-separated from it, but then the operator $\mathscr{L}(\Lambda x_j)$ can be brought to the position it should have in the new time ordering [the one with respect to the new time coordinates $(\Lambda x_i)^0$] without paying any price. Its relative position in time with respect to timelike-separated points will not change. We have then that

$$\begin{aligned} T_x &\left\{:\mathscr{L}_I(\phi_{\text{in}}(\Lambda x_1)): \dots :\mathscr{L}_I(\phi_{\text{in}}(\Lambda x_n)):\right\} \\ &= T_{\Lambda x}\left\{:\mathscr{L}_I(\phi_{\text{in}}(\Lambda x_1)): \dots :\mathscr{L}_I(\phi_{\text{in}}(\Lambda x_n)):\right\}\,, \end{aligned} \tag{6.31}$$

---

[43]This holds for the normal-ordered interaction Lagrangian as well.

either because the relative ordering has not change, or because it did not matter in the first place. We conclude that

$$
\begin{aligned}
U_0(\Lambda) S U_0(\Lambda)^\dagger &= \sum_{n=0}^\infty \frac{i^n}{n!} \int d^4 x_1 \dots \int d^4 x_n \, T_{\Lambda x} \{: \mathscr{L}_I(\phi_{\text{in}}(\Lambda x_1)) : \dots : \mathscr{L}_I(\phi_{\text{in}}(\Lambda x_n)) :\} \\
&= \sum_{n=0}^\infty \frac{i^n}{n!} \int d^4 (\Lambda^{-1} x)_1 \dots \int d^4 (\Lambda^{-1} x)_n \, T_{\Lambda x} \{: \mathscr{L}_I(\phi_{\text{in}}(x_1)) : \dots : \mathscr{L}_I(\phi_{\text{in}}(x_n)) :\} \\
&= \sum_{n=0}^\infty \frac{i^n}{n!} \int d^4 x_1 \dots \int d^4 x_n \, T_x \{: \mathscr{L}_I(\phi_{\text{in}}(x_1)) : \dots : \mathscr{L}_I(\phi_{\text{in}}(x_n)) :\} ,
\end{aligned}
\tag{6.32}
$$

since $|\det \Lambda^{-1}| = 1$. We then conclude that under the unitary tranformation $U_0(\Lambda)$ the S operator remains invariant,

$$
U_0(\Lambda)^\dagger S U_0(\Lambda) = S \implies [U_0(\Lambda), S] = 0 .
\tag{6.33}
$$

This proves the desired Lorentz invariance of the S-matrix, which includes of course invariance under rotations, soi in particualr we have

$$
[\vec{J}_0, S] = 0 .
\tag{6.34}
$$

Eqs. (6.28), (6.29) and (6.34) imply conservation of momentum, internal charges and angular momentum, respectively. Denoting with $\mathcal{O}$ any of these observables, the fact that $[\mathcal{O}, S] = 0$ implies that if $\varphi_{i,f}$ are eigenstates of $\mathcal{O}$ with eigenvalues $o_{i,f}$, then

$$
0 = \langle \varphi_f | [\mathcal{O}, S] | \varphi_i \rangle = \langle \varphi_f | \mathcal{O}S - S\mathcal{O} | \varphi_i \rangle = (o_f - o_i) \langle \varphi_f | S | \varphi_i \rangle ,
\tag{6.35}
$$

and so $\langle \varphi_f | S | \varphi_i \rangle$ must vanish unless $o_f = o_i$.

Invariance under Lorentz transformations and translations of the S operator, expressed by Eq. (6.33), allows us to construct a unitary representation of the Poincaré group on the in and out states. Recall that in general

$$
\begin{aligned}
U_0(\Lambda)|\vec{p}\,\sigma\rangle &= \sum_{\bar\sigma} \mathscr{D}_{\bar\sigma\sigma}(W(\Lambda, \vec{p}))|\Lambda\vec{p}\,\bar\sigma\rangle , \\
U_0(a)|\vec{p}\,\sigma\rangle &= e^{-ia\cdot p}|\vec{p}\,\sigma\rangle .
\end{aligned}
\tag{6.36}
$$

Here we discuss single-particle states; the generalisation to multiparticle states is straightforward. In and out momentum eigenstates are defined as

$$
|\vec{p}\,\sigma\pm\rangle = \Omega_\pm|\vec{p}\,\sigma\rangle .
\tag{6.37}
$$

Up to this point, calling them "momentum eigenstates" refers only to the fact that they are connected to the free momentum eigenstates by the scattering operators. From Eq. (6.36) we obtain

$$
\begin{aligned}
\Omega_\pm U_0(\Lambda)|\vec{p}\,\sigma\rangle &= \sum_{\bar\sigma} \mathscr{D}_{\bar\sigma\sigma}(W(\Lambda, \vec{p}))\Omega_\pm|\Lambda\vec{p}\,\bar\sigma\rangle = \sum_{\bar\sigma} \mathscr{D}_{\bar\sigma\sigma}(W(\Lambda, \vec{p}))|\Lambda\vec{p}\,\bar\sigma\pm\rangle \\
&= \Omega_\pm U_0(\Lambda)\Omega_\pm^\dagger \Omega_\pm|\vec{p}\,\sigma\rangle = \Omega_\pm U_0(\Lambda)\Omega_\pm^\dagger |\vec{p}\,\sigma\pm\rangle , \\
\Omega_\pm U_0(a)|\vec{p}\,\sigma\rangle &= e^{-ia\cdot p}\Omega_\pm|\vec{p}\,\sigma\rangle = e^{-ia\cdot p}|\vec{p}\,\sigma\pm\rangle = \Omega_\pm U_0(a)\Omega_\pm^\dagger \Omega_\pm|\vec{p}\,\sigma\rangle \\
&= \Omega_\pm U_0(a)\Omega_\pm^\dagger |\vec{p}\,\sigma\pm\rangle ,
\end{aligned}
\tag{6.38}
$$

which can be summarised as

$$U_\pm(\Lambda)|\vec{p}\,\sigma\pm\rangle = \sum_{\bar\sigma} \mathscr{D}_{\bar\sigma\sigma}(W(\Lambda,\vec{p}))|\Lambda\vec{p}\,\bar\sigma\pm\rangle\,,$$

$$U_\pm(a)|\vec{p}\,\sigma\pm\rangle = e^{-ia\cdot p}|\vec{p}\,\sigma\pm\rangle\,,$$

(6.39)

having defined

$$U_\pm(\Lambda) \equiv \Omega_\pm U_0(\Lambda)\Omega_\pm^\dagger\,, \qquad U_\pm(a) \equiv \Omega_\pm U_0(a)\Omega_\pm^\dagger\,. \tag{6.40}$$

The operators $U_\pm$ are indeed unitary[44], and provide a group representation:

$$U_\pm(\Lambda)^\dagger U_\pm(\Lambda) = \Omega_\pm U_0(\Lambda)^\dagger \Omega_\pm^\dagger \Omega_\pm U_0(\Lambda)\Omega_\pm^\dagger = \Omega_\pm\Omega_\pm^\dagger = \mathbf{1}\,,$$

$$U_\pm(\Lambda_1)U_\pm(\Lambda_2) = \Omega_\pm U_0(\Lambda_1)\Omega_\pm^\dagger \Omega_\pm U_0(\Lambda_2)\Omega_\pm^\dagger = \Omega_\pm U_0(\Lambda_1)U_0(\Lambda_2)\Omega_\pm^\dagger \Omega_\pm U_0(\Lambda_1\Lambda_2)\Omega_\pm^\dagger \quad (6.41)$$

$$= U_\pm(\Lambda_1\Lambda_2)\,,$$

and similarly for translations. On the other hand, Poincaré invariance requires that the physics be the same in any inertial reference frame, and so we need for the S-matrix elements that

$$\langle\Lambda\vec{p}_1', \Lambda\vec{p}_2', \ldots, \Lambda\vec{p}_{n'}' - |\Lambda\vec{p}_1, \Lambda\vec{p}_2, \ldots, \Lambda\vec{p}_n +\rangle = \langle\vec{p}_1', \vec{p}_2', \ldots, \vec{p}_{n'}' - |\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n +\rangle\,, \tag{6.42}$$

where for notational simplicity we considered the case of a single type of scalar particle. Here we wrote the S-matrix element as the scalar product of in and out momentum eigenstates, which by Eq. (6.26) are eigenstates of the full Hamiltonian. In and out momentum eigenstates form two complete sets in the space of scattering states, which we assumed to be the whole of the Hilbert space of the system. Eq. (6.39) tells us that the left-hand side of this equation, which is the basic request for Lorentz invariance, can be written as

$$\langle\vec{p}_1', \vec{p}_2', \ldots, \vec{p}_{n'}' - |U_-(\Lambda)^\dagger U_+(\Lambda)|\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n +\rangle = \langle\vec{p}_1', \vec{p}_2', \ldots, \vec{p}_{n'}' - |\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_n +\rangle\,, \quad (6.43)$$

so that Lorentz invariance is obtained if U_+ and U_- are in fact the *same* unitary representation of the Lorentz group. Here is where Eq. (6.33) comes into play: we have that

$$U_-(\Lambda)^\dagger U_+(\Lambda) = \Omega_- U_0(\Lambda)^\dagger \Omega_-^\dagger \Omega_+ U_0(\Lambda)\Omega_+^\dagger = \Omega_- U_0(\Lambda)^\dagger S U_0(\Lambda)\Omega_+^\dagger$$

$$= \Omega_- U_0(\Lambda)^\dagger U_0(\Lambda) S\Omega_+^\dagger = \Omega_- S\Omega_+^\dagger = \Omega_- \Omega_-^\dagger \Omega_+ \Omega_+^\dagger = \mathbf{1}\,,$$

(6.44)

which implies $U_-(\Lambda) = U_+(\Lambda) \equiv U(\Lambda)$. One similarly proves that $U_-(a) = U_+(a) \equiv U(a)$. From infinitesimal transformations we then obtain the exact symmetry generators P_μ, $\vec{J}$ and $\vec{K}$. We have then found one and the same unitary representation of the Poincaré group acting on the in and out states of the system, i.e., on the exact energy eigenstates, which transforms them according to the same transformation law as the corresponding free states. This implies then that the in and out momentum eigenstates are in fact eigenstates of the exact momentum, and also of the third component of the angular momentum:

$$U(a)|\vec{p}\,\sigma\pm\rangle = e^{-ia\cdot P}|\vec{p}\,\sigma\pm\rangle = e^{-ia\cdot p}|\vec{p}\,\sigma\pm\rangle \quad\Longrightarrow\quad P^\mu|\vec{p}\,\sigma\pm\rangle = p^\mu|\vec{p}\,\sigma\pm\rangle\,,$$

$$U(R_3(\theta))|\vec{p}\,\sigma\pm\rangle = e^{i\theta J_3}|\vec{p}\,\sigma\pm\rangle = e^{i\theta\sigma}|\vec{p}\,\sigma\pm\rangle \quad\Longrightarrow\quad J_3|\vec{p}\,\sigma\pm\rangle = \sigma|\vec{p}\,\sigma\pm\rangle\,.$$

(6.45)

It follows from the invariance of the free vacuum under $U_0(\Lambda)$ and $U_0(a)$ that the states $\Omega_\pm|0\rangle \equiv |0\pm\rangle$ are invariant under $U(\Lambda)$ and $U(a)$: $U|0\pm\rangle = U\Omega_\pm|0\rangle = \Omega_\pm U_0|0\rangle = \Omega_\pm|0\rangle = |0\pm\rangle$.

---

[44]At least on the subspace of scattering states.

These states are normalised to 1. If there were more than one invariant state under the exact transformations U, say $|\text{VAC}\rangle \equiv |0+\rangle$ and $|\text{VAC}'\rangle$ with $\langle \text{VAC}'|\text{VAC}\rangle = 0$, then besides $|0\rangle = \Omega_+^\dagger|\text{VAC}\rangle$ there would be another state $|0'\rangle = \Omega_+^\dagger|\text{VAC}'\rangle$ orthogonal to $|0\rangle$ and invariant under all the U's. But $|0\rangle$ is the only such state, and so there is no $|\text{VAC}'\rangle$. This means that $|0\pm\rangle$ must be proportional to each other, $|0-\rangle = e^{-i\phi}|0+\rangle$, with the proportionality constant being a phase due to the normalisation of these states. This implies that

$$\langle 0 - |0 +\rangle = \langle 0|S|0\rangle = e^{i\phi} \, . \tag{6.46}$$

This result could have been obtained more directly from the fact that $US|0\rangle = SU|0\rangle = S|0\rangle$ implies that $S|0\rangle$ is an invariant state, normalised to one due to unitarity of S, hence $S|0\rangle = e^{i\phi}|0\rangle$. Since only $|S_{fi}|^2$ is physically meaningful, we can ignore this phase, essentially redefining S so that $\langle 0|S|0\rangle = 1$.

6.4   The perturbative expansion: Wick's theorem

In experiments, the case of the final state coinciding with the initial state is practically impossible to observe, as it would interfere with the experimental setup. Even if it were possible to consider this case, it would be impossible to distinguish between the lack of any interaction between the colliding particle, and the presence of some interaction that however gives a final state coinciding with the initial one. For these reasons, what is studied in practice is the case of final and initial states being different. It is then customary to explicitly subtract the no-interaction contribution to the S-matrix. Furthermore, given that energy and momentum are conserved in a scattering process, it is convenient to factor out a four-momentum-conserving delta function. One then writes

$$S_{fi} = \delta_{fi} + i(2\pi)^4 \delta^{(4)}(P_f - P_i)\mathcal{M}_{fi} \, , \tag{6.47}$$

Here initial and final states are taken to be momentum eigenstates, $\delta_{fi} = \langle \varphi_f|\varphi_i\rangle$, and the i factor is conventional. The quantity $\mathcal{M}_{fi}$ is the one directly relevant for experiments, as it is the one entering the practically measurable quantity related to the transition probability, namely the *cross section* of the scattering processes. The purpose of the theory to is to provide predictions for the cross sections, a task which is essentially accomplished once that we have obtained $\mathcal{M}_{fi}$. In this subsection we discuss how this can be done in perturbation theory.

The basic assumption of perturbation theory is that the exact matrix element $\langle \varphi_f|S|\varphi_i\rangle$ can be well approximated with the lowest-order terms of the expansion of S, Eqs. (6.22) and (6.23). Subtracting the non-interaction term we have[45]

$$\langle \varphi_f|S - \mathbf{1}|\varphi_i\rangle = \langle \varphi_f|i \int d^4x : \mathscr{L}_I(x) : +\frac{i^2}{2} \int d^4x_1 \int d^4x_2 T\left\{: \mathscr{L}_I(x_1) :: \mathscr{L}_I(x_2) :\right\} + \dots |\varphi_i\rangle \, . \tag{6.48}$$

The basic object of interest is therefore the matrix element

$$\langle \varphi_f| \int d^4x_1 \int d^4x_2 \dots \int d^4x_n T\left\{: \mathscr{L}_I(x_1) :: \mathscr{L}_I(x_2) : \dots : \mathscr{L}_I(x_n) :\right\} |\varphi_i\rangle \, . \tag{6.49}$$

---

[45]We adopt the simplified notation $\mathscr{L}_I(x) = \mathscr{L}_I(\phi_{\text{in}}(x))$.

It is worth noting that thanks to translation invariance, if we denote with P_i and P_f the total initial and final four momentum, we find

$$
\begin{aligned}
&\langle \varphi_f | \int d^4 x_1 \ldots \int d^4 x_n T \left\{ : \mathscr{L}_I(x_1) : \ldots : \mathscr{L}_I(x_n) : \right\} | \varphi_i \rangle \\
&= \langle \varphi_f | e^{-ix_n \cdot P} \int d^4 x_1 \ldots \int d^4 x_n T \left\{ : \mathscr{L}_I(x_1 - x_n) : \ldots : \mathscr{L}_I(0) : \right\} e^{ix_n \cdot P} | \varphi_i \rangle \\
&= \int d^4 x_n e^{-ix_n \cdot (P_f - P_i)} \\
&\quad \times \langle \varphi_f | \int d^4 y_1 \ldots \int d^4 y_{n-1} T \left\{ : \mathscr{L}_I(y_1) : \ldots : \mathscr{L}_I(y_{n-1}) :: \mathscr{L}_I(0) : \right\} | \varphi_i \rangle \\
&= (2\pi)^4 \delta^{(4)}(P_f - P_i) \\
&\quad \times \langle \varphi_f | \int d^4 y_1 \ldots \int d^4 y_{n-1} T \left\{ : \mathscr{L}_I(y_1) : \ldots : \mathscr{L}_I(y_{n-1}) :: \mathscr{L}_I(0) : \right\} | \varphi_i \rangle,
\end{aligned}
\tag{6.50}
$$

i.e., the momentum-conserving delta function discussed above. The matrix elements Eq. (6.49) can be computed explicitly making use of *Wick's theorem*. In the case of a single Hermitian scalar field, this theorem states the following:

$$
\begin{aligned}
T\left(\varphi(x_1) \ldots \varphi(x_n)\right) = \sum_{m=0}^{\left[\frac{n}{2}\right]} &\left\{ : \varphi(x_1) \ldots \varphi(x_{n-2m}) : D(x_{n-2m+1}, x_{n-2m+2}) \ldots D(x_{n-1}, x_n) \right. \\
&\left. + \text{ other pairings} \right\}.
\end{aligned}
\tag{6.51}
$$

Here the sum is over the number m of pairings of coordinates $\{x_1, \ldots, x_n\}$, from 0 to the maximal possible value, i.e., the integer part $\left[\frac{n}{2}\right]$ of $\frac{n}{2}$, and over the all the possible such pairings. The quantity $D(x, y)$, which we will refer to as the *contraction* of two fields, or the *propagator*, is given by

$$
\begin{aligned}
D(x, y) &= \langle 0 | T\left(\varphi(x)\varphi(y)\right) | 0 \rangle = \langle 0 | e^{-iy \cdot P} T\left(\varphi(x - y)\varphi(0)\right) e^{iy \cdot P} | 0 \rangle \\
&= \langle 0 | T\left(\varphi(x - y)\varphi(0)\right) | 0 \rangle = D(x - y).
\end{aligned}
\tag{6.52}
$$

Notice also that for a real field

$$
D(x - y) = \langle 0 | T\left(\varphi(x)\varphi(y)\right) | 0 \rangle = \langle 0 | T\left(\varphi(y)\varphi(x)\right) | 0 \rangle = D(y - x).
\tag{6.53}
$$

To prove this theorem, and also to understand how it works in practice, we begin with the simplest case, $n = 2$. In this case, denoting with $\varphi_\pm$ the positive and negative frequency parts ofr the field, we have

$$
\begin{aligned}
\varphi(x_1)\varphi(x_2) &= [\varphi_+(x_1) + \varphi_-(x_1)][\varphi_+(x_2) + \varphi_-(x_2)] =: \varphi(x_1)\varphi(x_2) : +[\varphi_+(x_1), \varphi_-(x_2)] \\
&=: \varphi(x_1)\varphi(x_2) : +\langle 0 | [\varphi_+(x_1), \varphi_-(x_2)] | 0 \rangle =: \varphi(x_1)\varphi(x_2) : +\langle 0 | \varphi_+(x_1)\varphi_-(x_2) | 0 \rangle \\
&=: \varphi(x_1)\varphi(x_2) : +\langle 0 | \varphi(x_1)\varphi(x_2) | 0 \rangle,
\end{aligned}
\tag{6.54}
$$

where we have made use of the fact that $[\varphi_+(x_1), \varphi_-(x_2)]$ is a *c*-number and is thus identical to its vacuum expectation value, and further modified its expression by including terms that annihilate

the vacuum and therefore have no effect. Imposing now time ordering, since $:\varphi(x_1)\varphi(x_2):=:\varphi(x_2)\varphi(x_1):$ we find

$$T\left(\varphi(x_1)\varphi(x_2)\right) =:\varphi(x_1)\varphi(x_2): +\langle 0|T\left(\varphi(x_1)\varphi(x_2)\right)|0\rangle =:\varphi(x_1)\varphi(x_2): +D(x_1,x_2). \quad (6.55)$$

The proof for general n proceeds by induction. Assume that the theorem is true for n fields, and consider the time-ordered product of $n+1$ fields. Assume that $x_1^0 > x_2^0 > \ldots > x_n^0 > x_{n+1}^0$: if we prove that it is true in this case, it will be true for any other time ordering, since it would suffice to relabel the coordinates $x_j = x'_{Pj}$ to find the same situation, up to having primed coordinates. Then

$$T\left(\varphi(x_1)\ldots\varphi(x_n)\varphi(x_{n+1})\right) = \varphi(x_1)\ldots\varphi(x_n)\varphi(x_{n+1}) = T\left(\varphi(x_1)\ldots\varphi(x_n)\right)\varphi(x_{n+1})$$

$$= \sum_{m=0}^{\left[\frac{n}{2}\right]}\Bigg\{ :\varphi(x_1)\ldots\varphi(x_{n-2m}): D(x_{n-2m+1},x_{n-2m+2})\ldots D(x_{n-1},x_n) \quad (6.56)$$

$$+ \text{ other pairings}\Bigg\}\varphi(x_{n+1}).$$

We have then to compute

$$:\varphi(x_1)\ldots\varphi(x_{n-2m}):\varphi(x_{n+1}) =:\varphi(x_1)\ldots\varphi(x_{n-2m}):[\varphi_+(x_{n+1})+\varphi_-(x_{n+1})]$$

$$=:\varphi(x_1)\ldots\varphi(x_{n-2m})\varphi_+(x_{n+1}): + :\varphi(x_1)\ldots\varphi(x_{n-2m}):\varphi_-(x_{n+1}) \quad (6.57)$$

and variations thereof obtained by permuting $\{x_1,\ldots,x_n\}$.[46] The general contribution to the last term in Eq. (6.57), when we write down explicitly the normal ordering, reads

$$\varphi_-(x_{P(1)})\ldots\varphi_-(x_{P(k)})\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)})\varphi_-(x_{n+1})\,, \quad (6.58)$$

with P a permutation of $\{x_1,\ldots,x_{n-2m}\}$. To obtain a normal-ordered product we need to bring $\varphi_-(x_{n+1})$ to the left, and this is easily seen to yield

$$\varphi_-(x_{P(1)})\ldots\varphi_-(x_{P(k)})\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)})\varphi_-(x_{n+1})$$

$$= \varphi_-(x_{n+1})\varphi_-(x_{P(1)})\ldots\varphi_-(x_{P(k)})\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)})$$

$$+ \varphi_-(x_{P(1)})\ldots\varphi_-(x_{P(k)})[\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)}),\varphi_-(x_{n+1})]$$

$$= \varphi_-(x_{n+1})\varphi_-(x_{P(1)})\ldots\varphi_-(x_{P(k)})\Bigg\{\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)})$$

$$+ \sum_{j=k+1}^{n-2m}\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)})/\varphi_+(x_{P(j)})[\varphi_+(x_{P(j)}),\varphi_-(x_{n+1})]\Bigg\} \quad (6.59)$$

$$= \varphi_-(x_{n+1})\varphi_-(x_{P(1)})\ldots\varphi_-(x_{P(k)})\Bigg\{\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)})$$

$$+ \sum_{j=k+1}^{n-2m}\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)})/\varphi_+(x_{P(j)})D(x_{P(j)},x_{n+1})\Bigg\}\,,$$

---

[46]Notice that permutations that affect only $\{x_1,\ldots,x_{n-2m}\}$ do not give new terms.

where in the last passage we made use of the fact that $x^0_{P(j)} > x^0_{n+1}$. Here the notation $\ldots/\varphi_+(x_{P(j)})$ means that this field has to be removed from the product. Summing over k and over permutations P we find

$$\sum_{k,P} \varphi_-(x_{P(1)})\ldots\varphi_-(x_{P(k)})\varphi_+(x_{P(k+1)})\ldots\varphi_+(x_{P(n-2m)})\varphi_-(x_{n+1})$$

$$=: \varphi(x_1)\ldots\varphi(x_{n-2m}) : \varphi_-(x_{n+1}) \tag{6.60}$$

$$=: \varphi_-(x_{n+1})\varphi(x_1)\ldots\varphi(x_{n-2m}) : + \sum_{j=1}^{n-2m} : \varphi(x_1)\ldots\varphi(x_{n-2m})/\varphi(x_j) : D(x_j, x_{n+1}),$$

and including the term with positive-frequency contribution $\varphi_+(x_{n+1})$ appearing in Eq. (6.57) we obtain

$$: \varphi(x_1)\ldots\varphi(x_{n-2m}) : \varphi(x_{n+1}) =: \varphi(x_{n+1})\varphi(x_1)\ldots\varphi(x_{n-2m}) :$$

$$+ \sum_{j=1}^{n-2m} : \varphi(x_1)\ldots\varphi(x_{n-2m})/\varphi(x_j) : D(x_j, x_{n+1}), \tag{6.61}$$

which we plug in Eq. (6.56) to find

$$\varphi(x_1)\ldots\varphi(x_n)\varphi(x_{n+1})$$

$$= \sum_{m=0}^{\left[\frac{n}{2}\right]} \left\{ \left[: \varphi(x_1)\ldots\varphi(x_{n-2m})\varphi(x_{n+1}) : + \sum_{j=1}^{n-2m} : \varphi(x_1)\ldots\varphi(x_{n-2m})/\varphi(x_j) : D(x_j, x_{n+1}) \right] \right.$$

$$\left. \times D(x_{n-2m+1}, x_{n-2m+2})\ldots D(x_{n-1}, x_n) + \text{other pairings} \right\}$$

$$= \sum_{m=0}^{\left[\frac{n}{2}\right]} \left\{ : \varphi(x_1)\ldots\varphi(x_{n-2m})\varphi(x_{n+1}) : \right.$$

$$\left. \times D(x_{n-2m+1}, x_{n-2m+2})\ldots D(x_{n-1}, x_n) + \text{other pairings, } x_{n+1} \text{ not contracted} \right\}$$

$$+ \sum_{m=1}^{\left[\frac{n+1}{2}\right]} \left\{ : \varphi(x_1)\ldots\varphi(x_{n+1-2m}) : D(x_{n-2m+1}, x_{n-2m+2})\ldots D(x_n, x_{n+1}) \right.$$

$$\left. + \text{other pairings, } x_{n+1} \text{ contracted} \right\}$$

$$= \sum_{m=0}^{\left[\frac{n+1}{2}\right]} \left\{ : \varphi(x_1)\ldots\varphi(x_{n+1-2m}) : \right.$$

$$\left. \times D(x_{n-2m+1}, x_{n-2m+2})\ldots D(x_n, x_{n+1}) + \text{other pairings} \right\}, \tag{6.62}$$

which finally proves the theorem.

The result above is easily generalised to any number of real or complex scalar fields. Denoting them with φ_a, the only modification to the formula above is that contractions have to be

considered only for fields of the same type, since $[\varphi_a, \varphi_b] = 0$, and moreover only for φ_a and its adjoint (which is the field itself in the real case), since for complex fields $[\varphi_a, \varphi_a] = 0$ and $[\varphi_a, \varphi_b^\dagger] \neq 0$.

In the case of interest to us, the fields in the time ordered product are already partially normal ordered. In this case, contractions have to be considered only among fields belonging to different normal-ordered block. This is easily understood from the proof we gave above: a contraction appears only when we try to bring a negative-frequency field to the left of a positive-frequency field, but this never occurs for fields ina normal-ordered block.

Wick's theorem can also be generalised to fermions. In that case one has to be aware of the presence of signs, due to the anticommutation relations between fermion fields. A calculation similar to Eq. (6.54) shows that

$$
\begin{aligned}
\psi(x_1)\bar{\psi}(x_2) &= [\psi_+(x_1) + \psi_-(x_1)][\bar{\psi}_+(x_2) + \bar{\psi}_-(x_2)] \\
&= \psi_+(x_1)\bar{\psi}_+(x_2) + \psi_+(x_1)\bar{\psi}_-(x_2) + \psi_-(x_1)\bar{\psi}_+(x_2) + \psi_-(x_1)\bar{\psi}_-(x_2) \\
&= \psi_+(x_1)\bar{\psi}_+(x_2) - \bar{\psi}_-(x_2)\psi_+(x_1) + \psi_-(x_1)\bar{\psi}_+(x_2) + \psi_-(x_1)\bar{\psi}_-(x_2) + \{\psi_+(x_1), \bar{\psi}_-(x_2)\} \\
&=: \psi(x_1)\bar{\psi}(x_2) : + \langle 0|\psi(x_1)\bar{\psi}(x_2)|0\rangle \,.
\end{aligned}
$$
(6.63)

Time-ordering for fermions requires the introduction of minus signs, to be consistent with the anticommutation relations. For two fields,

$$
T\left(\psi_1(x_1)\psi_2(x_2)\right) = \theta(x_1^0 - x_2^0)\psi_1(x_1)\psi_2(x_2) - \theta(x_2^0 - x_1^0)\psi_2(x_2)\psi_1(x_1) \,,
$$
(6.64)

so that $T\left(\psi_1(x_1)\psi_2(x_2)\right) = -T\left(\psi_2(x_2)\psi_1(x_1)\right)$. This generalises to

$$
\begin{aligned}
&T\left(\psi_1(x_1)\ldots\psi_n(x_n)\right) = \\
&\sum_{\mathrm{P}}(-1)^{\sigma_{\mathrm{P}}}\theta(x_{\mathrm{P}(1)}^0 - x_{\mathrm{P}(2)}^0)\ldots\theta(x_{\mathrm{P}(n-1)}^0 - x_{\mathrm{P}(n)}^0)\psi_{\mathrm{P}(1)}(x_{\mathrm{P}(1)})\ldots\psi_{\mathrm{P}(n)}(x_{\mathrm{P}(n)}) \,,
\end{aligned}
$$
(6.65)

where σ_{P} is the parity of the permutation P, i.e., the number of transpositions required to obtain it modulo 2. Defining the contraction for fermion fields,

$$
S(x, y) = \langle 0|T\left(\psi(x)\bar{\psi}(y)\right)|0\rangle \,,
$$
(6.66)

we have then

$$
T\left(\psi(x_1)\bar{\psi}(x_2)\right) =: \psi(x_1)\bar{\psi}(x_2) : + S(x_1, x_2) \,.
$$
(6.67)

Notice that Eq. (6.66) is the only nontrivial contraction in this case. Notice furthermore that thanks to translation invariance we have

$$
S(x, y) = \langle 0|T\left(\psi(x - y)\bar{\psi}(0)\right)|0\rangle = S(x - y) \,,
$$
(6.68)

as in the scalar case. For a general number of fields, Wick's theorem reads

$$
\begin{aligned}
&T\left(\psi(x_1)\ldots\psi(x_n)\bar{\psi}(y_1)\ldots\bar{\psi}(y_{\bar{n}})\right) \\
&= \sum_{m=0}^{\min(n,\bar{n})} (-1)^\sigma : \psi(x_1)\ldots\psi(x_{n-m})\bar{\psi}(y_1)\ldots\bar{\psi}(y_{\bar{n}-m}) : S(x_{n-m+1}, y_{\bar{n}-m+1})\ldots S(x_n, y_{\bar{n}}) \\
&\quad + \text{other pairings} \,.
\end{aligned}
$$
(6.69)

Here σ is the parity of the permutation that reorders the fields according to

$$x_1, \ldots, x_n, y_1, \ldots, y_{\bar{n}} \to x_1, \ldots, x_{n-m}, y_1, \ldots, y_{\bar{n}-m}, x_{n-m+1}, y_{\bar{n}-m+1}, \ldots, x_n, y_{\bar{n}}. \qquad (6.70)$$

One can show that $\sigma = m(\bar{n} - \frac{m+1}{2})$. For the other pairings, the parity of the appropriate permutation is required, and any exchange of two fields with respect to the term displayed explicitly in Eq. (6.69) must be paid wiht en extra minus sign. An example will make the procedure clear. For $n = \bar{n} = 2$ one finds (here $\sigma = 2m - \frac{m(m+1)}{2}$, so the corresponding permutation is odd)

$$
\begin{aligned}
T\left(\psi(x_1)\psi(x_2)\bar{\psi}(y_1)\bar{\psi}(y_2)\right) = {} & : \psi(x_1)\psi(x_2)\bar{\psi}(y_1)\bar{\psi}(y_2) : \\
& - : \psi(x_1)\bar{\psi}(y_1) : S(x_2, y_2) - : \psi(x_2)\bar{\psi}(y_2) : S(x_1, y_1) \\
& + : \psi(x_1)\bar{\psi}(y_2) : S(x_2, y_1) + : \psi(x_2)\bar{\psi}(y_1) : S(x_1, y_2) \\
& - S(x_1, y_1)S(x_2, y_2) + S(x_1, y_2)S(x_2, y_1) \,.
\end{aligned}
\qquad (6.71)
$$

6.5 Feynman diagrams

Making use of Wick's theorem one can compute the S-matrix elements, since the matrix elements of normal-ordered products of fields are easily obtained: they are given by the products of the coefficients of the annihilation operators that destroy a particle in the initial state, and those of the creation operators that destroy a particle in the final state, summed over all the possible ways in which particles and creation/annihilation operators can be paired.

There is a convenient graphic method that allows to easily keep track of the various terms coming out of the perturbative expansion due to Wick's theorem. This is the method of Feynman diagrams, which we now discuss for two different theories.

6.5.1 ϕ^4 theory

Consider the theory of a self-interacting Hermitian scalar field ϕ, with interaction Lagrangian

$$\mathscr{L}_I = \frac{\lambda}{4!}\phi^4 \,, \qquad (6.72)$$

where the *coupling constant* λ is a real parameter that we assume to be small, so to justify the perturbative expansion of the S-matrix,

$$
\begin{aligned}
\langle\varphi_f|S - \mathbf{1}|\varphi_i\rangle &= \langle\varphi_f| \sum_{n=0}^{\infty} \frac{i^n}{n!} \int d^4x_1 \ldots \int d^4x_n \, T\left\{: \mathscr{L}_I(x_1): \ldots : \mathscr{L}_I(x_n):\right\} |\varphi_i\rangle \\
&= \sum_{n=0}^{\infty} \frac{1}{n!}\left(\frac{i\lambda}{4!}\right)^n \int d^4x_1 \ldots \int d^4x_n \, \langle\varphi_f|T\left(: \phi(x_1)^4 : \ldots : \phi(x_n)^4 :\right)|\varphi_i\rangle \,.
\end{aligned}
\qquad (6.73)
$$

The matrix elements in the second line of Eq. (6.73) are computed by means of Wick's theorem, expanding the time-ordered product in a sum of normal-ordered products. Among the various terms, only those will contribute that have precisely the same number of fields as particles in the initial and final states. In fact, writing the normal-ordered product in terms of positive and negative-frequency parts of the fields, we see that for the matrix element not to vanish. we need as many positive-frequency components as particles in the initial state, and as many

107

negative-frequency components in the final state. In this way, the annihilation operators in the positive-frequency components, acting on the right, will remove the particles from the initial state, turning it into the vacuum state. Similarly, the creation operators in negative-frequency components will act on the left turning the final state into the vacuum state. What remains to be determined is in how many ways this can be done: for a given normal-ordered product with the right amount of fields, we can choose in a number of ways from which fields we take the positive-frequency part and from which we take the negative-fequency part instead, or stated differently, which fields we will take to act on the initial state and which on the final state. On top of this, each of the positive-frequency fields can be associated with any of the initial-state particles, and simialrly for the negative-frequency fields and the final state particles.

The evaluation of the matrix elements has been reduced essentially to a problem in combinatorics, which is dealt with most efficiently by means of a graphic device. For each term $i : \mathscr{L}(x) :$ in the S-matrix expansion draw an *interaction vertex*,

$$\times\;x \qquad \longrightarrow \quad i\frac{\lambda}{4!}$$

where for the time being we are including the position x of the vertex. Each line correpsonds to one of the fields appearing in the interaction Lagrangian. To any such vertex it is associated a numerical factor $i\frac{\lambda}{4!}$, in correspondence with the numerical factor in front of $: \phi^4 :$. We can now represent each of the terms in Wick's theorem graphically, by connecting two lines when the corresponding fields are contracted. Such lines, running between vertices, are the *internal lines* of the diagram, and to each such line, running from vertex y to x, there is an associated factor $D(x, y)$ coming from Wick's theorem. Since the interaction Lagrangian is normal ordered, one should not contract two lines coming from the same vertex.[47] The remaining uncontracted lines correspond to the fields appearing in the normal-ordered products in Wick's theorem, and for them to contribute to a specific matrix element, they have to be as many as the particles involved in the process. These lines are the *external lines* of the diagram, and each of them has to be associated to an incoming or outgoing particle. To each of them is associated a numerical factor discussed below.

It is now easy to keep track of the combinatorics. Given a matrix element with specified incoming and outgoing particles, and given the term in the S-operator expansion with a specified number of interaction Lagrangians, the contribution of such term to such matrix elements is obtained by drawing a vertex for each interaction Lagrangian, associating the lines with particles or connecting them with one another in all the possible ways that satisfy the following two requirements:

- each particle must be associated with a line;

- all the lines must be either associated with a particle or connected with one another.

In other words, there must be an external line for each particle and viceversa, and each line must be either an external line or an internal line of the diagram. We may call these the admissible diagrams.

---

[47]This takes care of the problem of *tadpole diagrams*, which would yield a divergent contribution to the matrix elements, and would have to be subtracted anyway.

What is left at this point is evaluating the effect of fields on the particle states. To each diagram corresponds a specific way of associating the unpaired fields to the particles, which in turn corresponds to one of the contributions to the matrix element of the normal-ordered product between the initial and final particle states. Each of these contributions is obtained by letting the annihilation and creation operators act on the particle states, and selecting the term with the chosen particle/field association. One begins with

$$A = \langle \vec{p}_1', \ldots, \vec{p}_{n'}' | : \phi(x_1) \ldots \phi(x_m) : | \vec{p}_1, \ldots, \vec{p}_n \rangle, \tag{6.74}$$

where some of the $x_1, \ldots, x_m$ can be the same vertex, and where one must have $m = n + n'$, otherwise the matrix element vanishes. Choosing which fields act on the initial particles and which on the final particles corresponds to choosing for which x_j we take the positive-frequency part of the field, and for which we take the negative-frequency one. The matrix element A in Eq. (6.74) decomposes then in a sum of terms,

$$
\begin{aligned}
A &= \sum_{\mathrm{P}} \langle \vec{p}_1', \ldots, \vec{p}_{n'}' | \phi_-(x_{\mathrm{P}(1)}) \ldots \phi_-(x_{\mathrm{P}(n')}) \phi_+(x_{\mathrm{P}(n'+1)}) \ldots \phi_+(x_{\mathrm{P}(n'+n)}) | \vec{p}_1, \ldots, \vec{p}_n \rangle \\
&= \sum_{\mathrm{P}} \langle \vec{p}_1', \ldots, \vec{p}_{n'}' | \phi_-(x_{\mathrm{P}(1)}) \ldots \phi_-(x_{\mathrm{P}(n')}) |0\rangle \langle 0| \phi_+(x_{\mathrm{P}(n'+1)}) \ldots \phi_+(x_{\mathrm{P}(n'+n)}) | \vec{p}_1, \ldots, \vec{p}_n \rangle,
\end{aligned}
\tag{6.75}
$$

where in the second line we have included the projector on the vacuum state $|0\rangle\langle 0|$ at no cost, since no particle should remain after the annihilation and creation operators have done their job of removing particles from the initial and final states, respectively. We are then left with the task of evaluating matrix elements of the type

$$B = \langle 0| \phi_+(x_1) \ldots \phi_+(x_l) | \vec{p}_1, \ldots, \vec{p}_l \rangle, \tag{6.76}$$

Any factor in Eq. (6.75) is obtained either by an appropriate labelling of vertices, or by taking the complex conjugate $\langle 0| \phi_+(x_1) \ldots \phi_+(x_l) | \vec{p}_1, \ldots, \vec{p}_l \rangle^* = \langle \vec{p}_1, \ldots, \vec{p}_l | \phi_-(x_1) \ldots \phi_-(x_l) |0\rangle$.[48] Now,

$$
\begin{aligned}
B &= \int d\Omega_{q_1} \ldots \int d\Omega_{q_l} \, e^{-iq_1 \cdot x_1} \ldots e^{-iq_l \cdot x_l} \langle 0| a(\vec{q}_1) \ldots a(\vec{q}_l) a(\vec{p}_1)^\dagger \ldots a(\vec{p}_l)^\dagger |0\rangle \\
&= \int d\Omega_{q_1} \ldots \int d\Omega_{q_l} \, e^{-iq_1 \cdot x_1} \ldots e^{-iq_l \cdot x_l} \langle 0| [a(\vec{q}_1), [\ldots, [a(\vec{q}_l), a(\vec{p}_1)^\dagger \ldots a(\vec{p}_l)^\dagger]]] |0\rangle.
\end{aligned}
\tag{6.77}
$$

The multiple commutator can be computed to give

$$
\begin{aligned}
&[a(\vec{q}_1), [\ldots, [a(\vec{q}_l), a(\vec{p}_1)^\dagger \ldots a(\vec{p}_l)^\dagger]]] \\
&= \sum_j [a(\vec{q}_1), [\ldots, [a(\vec{q}_{l-1}), a(\vec{p}_1)^\dagger \ldots a(\vec{p}_{j-1})^\dagger a(\vec{p}_{j+1})^\dagger \ldots a(\vec{p}_l)^\dagger]]] (2\pi)^3 2p_j^0 \delta^{(3)}(\vec{q}_l - \vec{p}_j) \\
&= \sum_{\mathrm{P}} \prod_j (2\pi)^3 2p_j^0 \delta^{(3)}(\vec{q}_j - \vec{p}_{\mathrm{P}(j)}),
\end{aligned}
\tag{6.78}
$$

where the sum is over permutations of $\{1, \ldots, l\}$, corresponding to all the possible ways of pairing the annihilation and creation operators. The matrix element B becomes then

$$
\begin{aligned}
B &= \int d\Omega_{q_1} \ldots \int d\Omega_{q_l} \, e^{-iq_1 \cdot x_1} \ldots e^{-iq_l \cdot x_l} \sum_{\mathrm{P}} \prod_j (2\pi)^3 2p_j^0 \delta^{(3)}(\vec{q}_j - \vec{p}_{\mathrm{P}(j)}) \\
&= \sum_{\mathrm{P}} \prod_j e^{-ip_{\mathrm{P}(j)} \cdot x_j} = \sum_{\mathrm{P}} \prod_j \langle 0| \phi_+(x_j) | \vec{p}_{\mathrm{P}(1)} \rangle.
\end{aligned}
\tag{6.79}
$$

---

[48]We have used the commutativity of the negative-frequency fields to reorder them.

Each permutation corresponds to a specific way of associating the external lines with the colliding particles, so it corresponds to a specific diagram. In such a diagram we then have to include the following factor for an incoming particle with momentum $\vec{p}$ associated to a field $\phi(x)$ coming out of vertex x,

$$\langle 0|\phi(x)|\vec{p}\rangle = \langle 0|\phi_+(x)|\vec{p}\rangle = \int d\Omega_q\, e^{-iq\cdot x}\langle 0|a(\vec{q})|\vec{p}\rangle = \int d\Omega_q\, e^{-iq\cdot x}\langle 0|a(\vec{q})a(\vec{p})^\dagger|0\rangle$$
$$= \int d\Omega_q\, e^{-iq\cdot x}\langle 0|[a(\vec{q}), a(\vec{p})^\dagger]|0\rangle = \int d\Omega_q\, e^{-iq\cdot x}(2\pi)^3 2p^0\delta^{(3)}(\vec{q}-\vec{p})\langle 0|0\rangle = e^{-ip\cdot x}\,. \tag{6.80}$$

Similarly, we have to include the following factor for an outgoing particle with momentum $\vec{p}$ associated to a field $\phi(x)$ coming out of vertex x,

$$\langle\vec{p}|\phi(x)|0\rangle = \langle\vec{p}|\phi_-(x)|0\rangle = \int d\Omega_q\, e^{iq\cdot x}\langle\vec{p}|a(\vec{q})^\dagger|0\rangle = \int d\Omega_q\, e^{iq\cdot x}\langle 0|a(\vec{p})a(\vec{q})^\dagger|0\rangle$$
$$= \int d\Omega_q\, e^{iq\cdot x}\langle 0|[a(\vec{p}), a(\vec{q})^\dagger]|0\rangle = \int d\Omega_q\, e^{iq\cdot x}(2\pi)^3 2p^0\delta^{(3)}(\vec{q}-\vec{p})\langle 0|0\rangle = e^{ip\cdot x}\,. \tag{6.81}$$

Let us summarise our discussion, and give the *Feynman rules in coordinate space.* Given an S-matrix element with n incoming and n' outgoing particles of specified momenta, and a specified term in its perturbative expansion with v interaction Lagrangians, we proceed as follows for its evaluation.

- Draw all possible (and admissible) diagrams with v interaction vertices at $x_1,\ldots,x_v$ and $n+n'$ external lines, associated with the colliding particles;

- include a factor $i\frac{\lambda}{4!}$ for each vertex;

- include a factor $D(x_i, x_j)$ (propagator from x_j to x_i) for each internal line running from x_j to x_i, associated with the contraction $\langle 0|T(\phi(x_i)\phi(x_j))|0\rangle$;

- for each incoming particle, include a factor $e^{-ip_i\cdot x_j}$ if particle $\vec{p}_i$ is attached to an external line coming from the vertex x_j;

- for each outogoing particle, include a factor $e^{ip_i\cdot x_j}$ if particle $\vec{p}_i$ is attached to an external line coming from the vertex x_j;

- integrate over the position of the vertices;

- sum up all the contributions including the factor $1/v!$.

The last two passages complete the calculation since the previous ones, discussed in detial above, only evaluate the matrix elements of the time-ordered product of interaction Lagrangians.

When considering all the possible contributions at a given perturbative order, one typically encounters many identical terms. A source of degeneracy is the association between fields and external particles: for a given topology of the diagram it does not matter which of the four fields in a vertex is actually associated to a given particle, since one always obtains the same contribution. Similarly, when contracting fields from two vertices to form internal lines, one obtains the same factor now matter which of the fields are used. Finally, diagrams which before integration differ only by the labelling of the vertices will yield the same contribution to the

matrix element, since the position of the vertices is integrated over. Unless a diagram has special symmetries, the first two sources of degeneracy yield a factor 4! for each vertex, since the fields in each vertex can be freely permuted yielding the same contribution. This cancels out the factor $\frac{1}{4!}$ associated to each vertex. The last source of degeneracy instead, again in the case of absence of special symmetries, yields a factor $v!$ corresponding to all possible relabellings, which cancels out the factor $\frac{1}{v!}$.

It is possible, nonetheless, that permuting fields or vertices (before integration) does *not* yield a new and equivalent diagram, but just the same diagram. This is the case, for example, if two pairs of fields from two vertices are contracted: of the four ways of ordering the two pairs of fields, only two correspond to different contractions. For the sake of clarity label these fields as $\phi_1(x), \phi_2(x)$ and $\phi_1(y), \phi_2(y)$. In general one expects a degeneracy factor 2×2 from independent permutations of the fields from the two vertices. When contracting the two pairs together, however, one can only form the combinations $[(\phi_1(x), \phi_1(y)), (\phi_2(x), \phi_2(y))]$ and $[(\phi_1(x), \phi_2(y)), (\phi_2(x), \phi_1(y))]$, since exchanging the fields in both vertices does not give a new combination.[49] Similarly, this happens if two vertices in a diagram are not only equivalent but topologically identical. For example, one may have two vertices with all the four fields contracted together. In this case there is no difference between the two vertices, and exchanging their labels gives rise to exactly the same diagram. The bottom line of this discussion is that symmetry factors should be evaluated by carefully counting the possible ways to obtain topologically equivalent diagrams. We will see how to to this in practice in a couple of examples.

To proceed further in the evaluation of matrix elements, recall from the previous subsection that for the contraction we have $D(x, y) = D(x - y)$ thanks to translation invariance. We can then write it as a Fourier transform with respect to a single momentum,

$$D(x - y) = \int \frac{d^4 p}{(2\pi)^4} \, e^{-ip \cdot (x-y)} \tilde{D}(p) \,. \tag{6.82}$$

To each internal line $D(x, y)$ we can then associate a momentum p flowing from y to x. Using this representation we can carry out explicitly the integrations over the position of vertices. Suppose a vertex x is attached to incoming particles $\vec{p}_1, \ldots \vec{p}_n$ and outgoing particles $\vec{p}_1', \ldots \vec{p}_n'$, and furthermore to internal lines $D(x, y_1), \ldots, D(x, y_m)$ and $D(z_1, x), \ldots, D(z_{m'}, x)$. Using Eq. (6.82), we can write the internal line contributions as

$$D(x - y_j) = \int \frac{d^4 q_j}{(2\pi)^4} \, e^{-iq_j \cdot (x-y_j)} \tilde{D}(q_j) \,,$$

$$D(z_j - x) = \int \frac{d^4 k_j}{(2\pi)^4} \, e^{-ik_j \cdot (z_j-x)} \tilde{D}(k_j) \,. \tag{6.83}$$

Exchanging the order of integrations over internal momenta and positions of the vertices, we can isolated the factors that depend on x and perform the corresponding integration. there are factors $e^{-ip_j \cdot x}$ for the incoming particles, $e^{-ip_j' \cdot x}$ for the outgoing particles, $e^{-iq_j \cdot x}$ for incoming internal lines and $e^{ik_j \cdot x}$ for outgoing internal lines.[50] The integration is straightforward, and yields

$$\int d^4 x \, e^{-i \sum_j p_j \cdot x} e^{i \sum_j p_j' \cdot x} e^{-i \sum_j q_j \cdot x} e^{i \sum_j k_j \cdot x} = (2\pi)^4 \delta^{(4)} \left(\sum p_j - \sum p_j' + \sum q_j - \sum k_j \right) \,. \tag{6.84}$$

---

[49] One obtains $[(\phi_2(x), \phi_2(y)), (\phi_1(x), \phi_1(y))]$ $(\phi_2(x), \phi_1(y)), [(\phi_1(x), \phi_2(y))]$, but the square brackets correspond to *unordered pairs*, so these do not differ from the previous ones.

[50] The distinction between incoming and outgoing lines is actually irrelevant for the Hermitian scalar field, but it becomes important for the charged scalar field and the Dirac field.

Having assigned momenta to the internal lines, this relation simply establishes the *conservation of four-momentum at each interaction vertex*. What is left to do now is to collect all the momentum-conserving delta functions, include all the factors $\tilde{D}(q)$ from each contraction, and integrate over the interanl momenta. Together with the sum over topologically distinct diagrams multiplied by the appropriate symmetry factors, this completes the evaluation of a given matrix element at the desired perturbative order.

We can at this point give a new set of rules, the *Feynman rules in momentum space*, to evaluate S-matrix elements. Before doing this, it is high time we defined more precisely the concept of *topologically equivalent diagrams*. Two diagrams are topologically equivalent if they can be deformed into one another continuously, i.e., without cutting or glueing parts of the diagram. For example, exchanging two external lines coming from the same vertex gives a topologically equivalent diagram; exchanging the position of two vertices while keeping the same connections ot external particles and other vertices gives a topologically equivalent diagram; rotating or reflecting a diagram gives an equivalent one; and so on. Let now the colliding particles and the perturbative order be specified as before.

- Draw all possible, admissible and topologically distinct diagrams with v interaction vertices and $n+n'$ external lines associated with the colliding particles. Vertices have now no labels;

- associate four-momenta to external and internal lines: for the external lines, associate a four-momentum equal to that of the corresponding particle, flowing in the vertex if the particle is in the initial state and flowing out of the vertex if the particle is in the final state; for internal lines, let it flow along the direction of the internal line (which in the case at hand is completely arbitrary; see however footnote 50);

- for each vertex, include a factor $i\frac{\lambda}{4!}$ and a a momentum-conserving delta function,

$$(2\pi)^4\delta^{(4)}(p_i - p_o + q_i - q_o)\,,$$

  where p_i and p_o are the sum of the four-momenta of incoming and outgoing particles attached to the vertex, respectively, and q_i and q_o are the sum of momenta of internal lines flowing respectively in or out of the vertex;

- for each internal line of momentum q, include a factor $\tilde{D}(q)$ (propagator in momentum space);

- for each external line associate a factor 1;

- integrate over all internal momenta;

- sum up all the contributions including the factor $1/v!$.

Since four-momentum is conserved at every vertex, it follows that the four-momentum flowing in from the incoming particles is equal to the four-momentum flowing out with the outgoing particles. We can then trade one of the delta functions at the vertices with an overall, momentum-conserving delta function $(2\pi)^4\delta^{(4)}(P_f - P_i)$. If we also factor out one i, what we obtain is the matrix element $\mathcal{M}_{fi}$ introduced in Eq. (6.47). At this point we have all the tools we need to evalute matrix elements, except for the explciit expression of the propagator. This will be discussed later, as it is needed only when the explicit value of the matrix element will be needed.

It is now time to discuss a couple of examples to make the general discussion given above more concrete. Consider the matrix element $\langle \vec{p}'|S - \mathbf{1}|\vec{p}\rangle$.[51] This vanishes to first order in the coupling constant, as $\langle \vec{p}'| : \phi(x)^4 : |\vec{p}\rangle = 0$, since there are too many fields and too few particles. To second order we have to evaluate

$$\frac{1}{2}\left(\frac{i\lambda}{4!}\right)^2 \int d^4x \int d^4y \ \langle \vec{p}'|T\left(:\phi(x)^4 :: \phi(y)^4 :\right)|\vec{p}\rangle . \qquad (6.85)$$

Of all the terms obtained via Wick's theorem, only $: \phi(x)\phi(y) : D(x-y)^3$ has a nonzero matrix element here; such a term can be obtained in $4 \times 4!$ ways. Moreover, one sees that

$$\langle \vec{p}'| : \phi(x)\phi(y) : |\vec{p}\rangle = \langle \vec{p}'|\phi_-(x)\phi_+(y) + \phi_-(y)\phi_+(x)|\vec{p}\rangle = e^{i(p'\cdot x - py)} + e^{i(p'\cdot y - px)} , \qquad (6.86)$$

and the two terms give the same contribution after integration, so that all in all

$$\frac{1}{2}\left(\frac{i\lambda}{4!}\right)^2 \int d^4x \int d^4y \ \langle \vec{p}'|T\left(:\phi(x)^4 :: \phi(y)^4 :\right)|\vec{p}\rangle$$

$$= \frac{(i\lambda)^2}{3!} \int d^4x \int d^4y e^{i(p'\cdot x - py)} D(x-y)^3$$

$$= \frac{(i\lambda)^2}{3!} \int d^4x \int d^4y \int d\Omega_{q_1} \int d\Omega_{q_2} \int d\Omega_{q_3} \ e^{i(p'\cdot x - py)} e^{-i(q_1+q_2+q_3)\cdot(x-y)} \tilde{D}(q_1)\tilde{D}(q_2)\tilde{D}(q_3)$$

$$= \frac{(i\lambda)^2}{3!} \int d\Omega_{q_1} \int d\Omega_{q_2} \int d\Omega_{q_3} \ (2\pi)^4\delta^{(4)}(p' - q_1 - q_2 - q_3)$$

$$\times (2\pi)^4\delta^{(4)}(p - q_1 - q_2 - q_3)\tilde{D}(q_1)\tilde{D}(q_2)\tilde{D}(q_3)$$

$$= i(2\pi)^4\delta^{(4)}(p' - p)\left[\frac{i\lambda^2}{3!} \int d\Omega_{q_1} \int d\Omega_{q_2} \tilde{D}(q_1)\tilde{D}(q_2)\tilde{D}(p - q_1 - q_2)\right] ,$$

$$(6.87)$$

where the term in square brackets is the contribution to $\mathcal{M}_{fi}$.

Let us now do the evaluation via Feynman diagrams, beginning with the coordinate-space formulation. There is only one diagram with two vertices and two external lines, up to relabelling the vertices, see Fig. 1 (left). To count the degeneracy of the diagram, we proceed as depicted in Fig. 2. We first choose to attach $\vec{p}$ to vertex y, and then associate it with one of its four lines: this can be done in 4 ways. We now have to attach $\vec{p}'$ to vertex x, assocating it with one of its four lines: this can also be done in 4 ways. We then have to contract the remaining lines os that none is left out. We pick one of the three lines left from y and contract it with one of the three coming from x: this can be done in 3 ways. We next pick one of the two remaining lines in y and attach it to one of the two remaining lines in x: this can be done in 2 ways. The remaining lines from y and x must be contracted together. We can then repeat the procedure choosing instead to attach $\vec{p}$ to x and $\vec{p}^{prime}$ to y. This is a different diagram since the external particles are attached to different vertices in the two cases; it is however an equivalent diagram since vertex labels do not matter for the final result. All in all we obtain a factor of $2 \times 44!$. We then associate $e^{-ip\cdot y}$ to the external line attached to the incoming particle, $e^{ip'\cdot x}$ to the external line attached to the outgoing particle, and $D(x-y)$ to each of the internal lines connecting y to

---

[51]Since $S|\vec{p}\rangle$ transforms under Lorentz transformations as a one-particle state due to $[U_0(\Lambda), S] = 0$, and since four-momentum is conserved, $S|\vec{p}\rangle$ must equal $|\vec{p}\rangle$ up to a multiplicative phase factor. Here we are interested in the Feynman diagrams obtained at a given perturbative order.
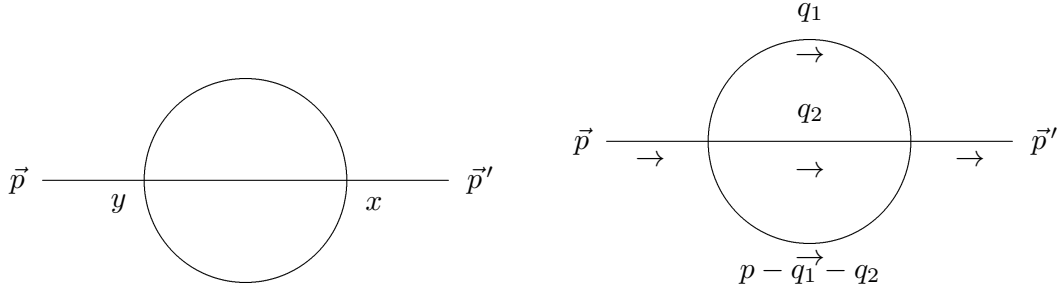
Figure 1: Diagram for the lowest-order nonzero $1 \to 1$ S-matrix element.
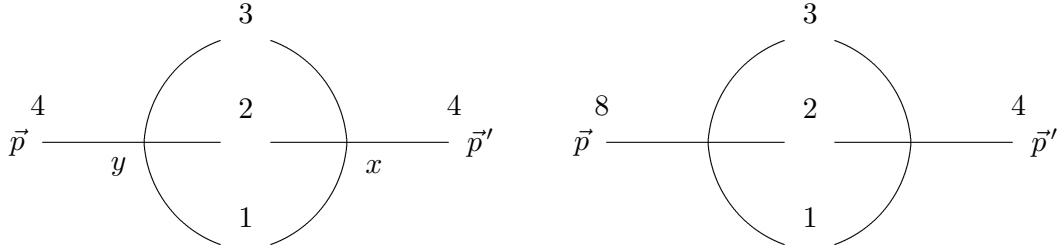

Figure 2: Degeneracy counting for the diagram.

x. Together with the degeneracy factor and integrating over the position of vertices we obtain the second line in Eq. (6.87).

In the momentum space formulation the relevant diagram is given in Fig. 1 (right), differing from the coordinate-space one by the absence of vertex labels and by the association of momenta to the internal lines. This must be done so that momentum is conserved at each vertex. The degeneracy counting is similar to the one in coordinate space, the only difference being that since we have no labels for the vertices, the assignement of an external line to the incoming particle $\vec{p}$ can be done in 8 ways, while the assignement of an external line to the outgoing particle can now be done only in 4 ways, since it has to be attached to a different vertex. The rest of the argument is as before. We then write down all the factors associate to the diagram, integrate over the internal momenta which are not fixed by the delta functions, and obtain the last line of Eq. (6.87).

A more interesting matrix element is that for the $2 \to 2$ elastic scattering process,

$$\langle \vec{p}_1', \vec{p}_2' | S - \mathbf{1} | \vec{p}_1, \vec{p}_2 \rangle . \tag{6.88}$$

To lowest order one easily finds by direct computation

$$
\begin{aligned}
\langle \vec{p}_1', \vec{p}_2' | S - \mathbf{1} | \vec{p}_1, \vec{p}_2 \rangle &= \frac{i\lambda}{4!} \int d^4x \, \langle \vec{p}_1', \vec{p}_2' | : \phi(x)^4 : | \vec{p}_1, \vec{p}_2 \rangle \\
&= i(2\pi)^4 \delta^{(4)}(p_1' + p_2' - p_1 - p_2) \lambda .
\end{aligned}
\tag{6.89}
$$

To lowest order, $\mathcal{M}_{fi} = \lambda$. The calculation via Feynman rules is straighforward: in coordinate space we have to draw one vertex and associate its lines with the colliding particles; no line remains for contractions (which are not allowed anyway due to normal ordering). There are 4! ways of attaching the lines to the external particles, and each of them give the same contribution,
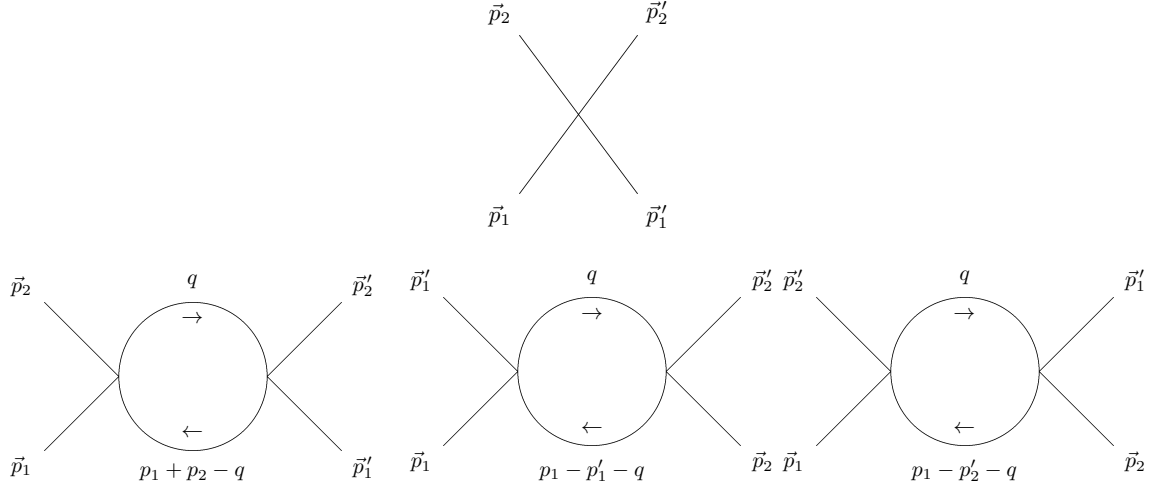
Figure 3: Diagrams for the lowest and next-to-lowest order $2 \to 2$ S-matrix element.

so

$$\langle \vec{p}_1', \vec{p}_2' | S - \mathbf{1} | \vec{p}_1, \vec{p}_2 \rangle = \frac{i\lambda}{4!} 4! \int d^4 x \, e^{i(p_1' + p_2' - p_1 - p_2) \cdot x} = i(2\pi)^4 \delta^{(4)}(p_1' + p_2' - p_1 - p_2)\lambda \,. \quad (6.90)$$

The momentum space calculation is even faster: no internal lines means no propagators and no internal momenta to integrate over, the only delta function associated to a vertex is exactly the one giving the overall conservation of four-momentum, and the result $\mathcal{M}_{fi} = \lambda$ is obtained by just taking into account the combinatorics. The next perturbative order is less trivial, and we compute it directly by making use of Feynman rules in momentum space. There are three topologically distinct ways of attaching four particles ot the lines coming out of two interaction vertices,[52] depicted in Fig. The degeneracy factor is evaluted as outlined above, and it is $8 \times 4 \times 3 \times 2 = (4!)^2$ for each of the diagrams, which multiplied by the numerical prefactor $\frac{1}{2}\frac{1}{(4!)^2}$ gives $\frac{1}{2}$. Conservation of momentum at each vertex leaves only one internal momentum to integrate, and the final expression for $\mathcal{M}_{fi}$ is thus

$$\mathcal{M}_{fi}|_{\mathcal{O}(\lambda^2)} = \frac{i\lambda^2}{2} \int d^4 q \left\{ \tilde{D}(q)\tilde{D}(p_1 + p_2 - q) + \tilde{D}(q)\tilde{D}(p_1 - p_1' - q) + \tilde{D}(q)\tilde{D}(p_1 - p_2' - q) \right\} \,. \quad (6.91)$$

6.5.2 Yukawa coupling

Consider next a theory containing a real scalar field and a complex Dirac field, of masses M and m respectively, coupled via the following interaction vertex (*Yukawa coupling*),

$$\mathcal{L}_I(x) = g\bar{\psi}(x)\psi(x)\phi(x) \,. \quad (6.92)$$

The Feynman rules discussed in the previous subsection can be extended to include fermionic fields. In this case, however, we have to distinguish between the two fermionic lines coming out of a vertex, since one corresponds to ψ and the other to its Dirac conjugate $\bar{\psi}$. Making use of the

---

[52]Here we consider only connected diagrams, i.e., diagrams that have no disjoint subdiagrams.

perturbative expansion of the S-matrix and of Wick's theorem, we reduce the problem to the evaluation of matrix elements of normal-ordered products of fields between the initial and final particle states. The only differences with the previous subsection is that now we have to keep track of the minus signs appearing when exchanging two fermionic fields, and that the action of a field on a one-particle state brings about the appropriate Dirac spinor corresponding to the particle. It is convenient to discuss these issues in a practical case. Consider the lowest-order contribution to the S-matrix element for elastic scattering of two fermions,[53]

$$
\begin{aligned}
F &= \langle \vec{p}_1' s_1', \vec{p}_2' s_2' | S - \mathbf{1} | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle \\
&= \frac{(ig)^2}{2} \int d^4 x \int d^4 y \, \langle \vec{p}_1' s_1', \vec{p}_2' s_2' | T \left(: \bar{\psi}(x) \psi(x) \phi(x) :: \bar{\psi}(y) \psi(y) \phi(y) : \right) | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle .
\end{aligned}
\tag{6.93}
$$

All the fermion fields have to be used to destroy particles in the initial and final state, so the only term that survives in the Wick expansion is

$$
F = \frac{(ig)^2}{2} \int d^4 x \int d^4 y \, D(x - y) \langle \vec{p}_1' s_1', \vec{p}_2' s_2' | : \bar{\psi}(x) \psi(x) \bar{\psi}(y) \psi(y) : | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle .
\tag{6.94}
$$

No minus sign has appeared yet, since the ordering of the fields is exactly as in the time-ordered product. Writing now the normal-ordered product in Eq. (6.94) in terms of positive and negative frequency components, we realise that the only term that contribute is that in which we take the positive-frequency component of $\psi(x)$ and $\psi(y)$, to destroy the particles in the initial state, and the negative-frequency component of $\bar{\psi}(x)$ and $\bar{\psi}(y)$, to destroy the particles in the final state,

$$
\begin{aligned}
&\langle \vec{p}_1' s_1', \vec{p}_2' s_2' | : \bar{\psi}(x) \psi(x) \bar{\psi}(y) \psi(y) : | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle \\
&= \langle \vec{p}_1' s_1', \vec{p}_2' s_2' | : \bar{\psi}_-(x) \psi_+(x) \bar{\psi}_-(y) \psi_+(y) : | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle \\
&= \langle \vec{p}_1' s_1', \vec{p}_2' s_2' | : \bar{\psi}_{-\alpha}(x) \psi_{+\alpha}(x) \bar{\psi}_{-\beta}(y) \psi_{+\beta}(y) : | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle \\
&= \langle \vec{p}_1' s_1', \vec{p}_2' s_2' | \bar{\psi}_{-\beta}(y) \bar{\psi}_{-\alpha}(x) \psi_{+\alpha}(x) \psi_{+\beta}(y) | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle \\
&= \langle \vec{p}_1' s_1', \vec{p}_2' s_2' | \bar{\psi}_{-\beta}(y) \bar{\psi}_{-\alpha}(x) | 0 \rangle \langle 0 | \psi_{+\alpha}(x) \psi_{+\beta}(y) | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle ,
\end{aligned}
\tag{6.95}
$$

where in the third line we have made explciit the Dirac index of the fermionic fields, in the fourth line we have reordered the fields, with no extra sign since $\bar{\psi}_-(y)$ has been moved left past two fields, and in the last passage we have conveniently inserted the vacuum projector at no cost. We now have to evaluate the simple matrix elements

$$
\langle 0 | \psi_+(x) \psi_+(y) | \vec{p}_1 s_1, \vec{p}_2 s_2 \rangle , \qquad \langle \vec{p}_1' s_1', \vec{p}_2' s_2' | \bar{\psi}_-(y) \bar{\psi}_-(x) | 0 \rangle .
\tag{6.96}
$$

---

[53]There is no $\mathcal{O}(g)$ contribution since there are four external particles and only two fermionic fields.

For the first one we find

$$\langle 0|\psi_+(x)\psi_+(y)|\vec{p}_1 s_1, \vec{p}_2 s_2\rangle = \langle 0|\psi_+(x)\psi_+(y)b_{s_1}(\vec{p}_1)^\dagger b_{s_2}(\vec{p}_2)^\dagger|0\rangle$$

$$= \int d\Omega_{q_1}\int d\Omega_{q_2}\sum_{t_1}\sum_{t_2}u_{t_2}(\vec{q}_2)u_{t_1}(\vec{q}_1)e^{-i(q_2\cdot x+q_1\cdot y)}\langle 0|b_{t_2}(\vec{q}_2)b_{t_1}(\vec{q}_1)b_{s_1}(\vec{p}_1)^\dagger b_{s_2}(\vec{p}_2)^\dagger|0\rangle$$

$$= \int d\Omega_{q_1}\int d\Omega_{q_2}\sum_{t_1}\sum_{t_2}u_{t_2}(\vec{q}_2)u_{t_1}(\vec{q}_1)e^{-i(q_2\cdot x+q_1\cdot y)}\langle \vec{q}_1 t_1, \vec{q}_2 t_2|\vec{p}_1 s_1, \vec{p}_2 s_2\rangle$$

$$= \int d\Omega_{q_1}\int d\Omega_{q_2}\sum_{t_1}\sum_{t_2}u_{t_2}(\vec{q}_2)u_{t_1}(\vec{q}_1)e^{-i(q_2\cdot x+q_1\cdot y)}(2\pi)^3 2p_1^0(2\pi)^3 2p_2^0$$

$$\times\left[\delta_{s_1 t_1}\delta_{s_2 t_2}\delta^{(3)}(\vec{p}_1-\vec{q}_1)\delta^{(3)}(\vec{p}_2-\vec{q}_2)-\delta_{s_1 t_2}\delta_{s_2 t_1}\delta^{(3)}(\vec{p}_1-\vec{q}_2)\delta^{(3)}(\vec{p}_2-\vec{q}_1)\right]$$

$$= u_{s_2}(\vec{p}_2)u_{s_1}(\vec{p}_1)e^{-i(p_2\cdot x+p_1\cdot y)}-u_{s_1}(\vec{p}_1)u_{s_2}(\vec{p}_2)e^{-i(p_1\cdot x+p_2\cdot y)}.$$

(6.97)

Recall that $u_s(\vec{p})$ carries a Dirac index α, $u_{s\alpha}(\vec{p})$ so that the spinorial factors above are not the same. Explicitly,

$$\langle 0|\psi_{+\alpha}(x)\psi_{+\beta}(y)|\vec{p}_1 s_1, \vec{p}_2 s_2\rangle = u_{s_2\alpha}(\vec{p}_2)u_{s_1\beta}(\vec{p}_1)e^{-i(p_2\cdot x+p_1\cdot y)}-u_{s_1\alpha}(\vec{p}_1)u_{s_2\beta}(\vec{p}_2)e^{-i(p_1\cdot x+p_2\cdot y)}.$$

(6.98)

A similar calculation yields

$$\langle \vec{p}_1's_1', \vec{p}_2's_2'|\bar{\psi}_{-\beta}(y)\bar{\psi}_{-\alpha}(x)|0\rangle = \bar{u}_{s_2'\alpha}(\vec{p}_2')\bar{u}_{s_1'\beta}(\vec{p}_1')e^{i(p_2'\cdot x+p_1'\cdot y)}-\bar{u}_{s_1'\alpha}(\vec{p}_1')\bar{u}_{s_2'\beta}(\vec{p}_2')e^{i(p_1'\cdot x+p_2'\cdot y)}.$$

(6.99)

Putting the two pieces together we find

$$\langle \vec{p}_1's_1', \vec{p}_2's_2'|:\bar{\psi}(x)\psi(x)\bar{\psi}(y)\psi(y):|\vec{p}_1 s_1, \vec{p}_2 s_2\rangle$$

$$= \bar{u}_{s_2'}(\vec{p}_2')u_{s_2}(\vec{p}_2)\bar{u}_{s_1'}(\vec{p}_1')u_{s_1}(\vec{p}_1)e^{i[(p_2'-p_2)\cdot x+(p_1'-p_1)\cdot y]}$$

$$- \bar{u}_{s_2'}(\vec{p}_2')u_{s_1}(\vec{p}_1)\bar{u}_{s_1'}(\vec{p}_1')u_{s_2}(\vec{p}_2)e^{i[(p_2'-p_1)\cdot x+(p_1'-p_2)\cdot y]}+x\leftrightarrow y.$$

(6.100)

The contribution to the S-matrix element of the term with x and y exchanged is obviously the same, since $D(x)=D(-x)$. We can then write

$$F = (ig)^2\int d^4x\int d^4y\, D(x-y)\bigg\{\bar{u}_{s_1'}(\vec{p}_1')u_{s_1}(\vec{p}_1)\bar{u}_{s_2'}(\vec{p}_2')u_{s_2}(\vec{p}_2)e^{i(p_2'-p_2)\cdot x}e^{i(p_1'-p_1)\cdot y}$$

$$- \bar{u}_{s_2'}(\vec{p}_2')u_{s_1}(\vec{p}_1)\bar{u}_{s_1'}(\vec{p}_1')u_{s_2}(\vec{p}_2)e^{i(p_2'-p_1)\cdot x}e^{i(p_1'-p_2)\cdot y}\bigg\}.$$

(6.101)

Using the momentum representation for the scalar propagator we can integrate over the position of the vertices and get finally

$$F = (ig)^2(2\pi)^4\delta^{(4)}(p_1'+p_2'-p_1-p_2)\bigg\{\bar{u}_{s_1'}(\vec{p}_1')u_{s_1}(\vec{p}_1)\bar{u}_{s_2'}(\vec{p}_2')u_{s_2}(\vec{p}_2)\tilde{D}(p_1-p_1')$$

$$- \bar{u}_{s_2'}(\vec{p}_2')u_{s_1}(\vec{p}_1)\bar{u}_{s_1'}(\vec{p}_1')u_{s_2}(\vec{p}_2)\tilde{D}(p_1-p_2')\bigg\}.$$

(6.102)

We are now ready to formulate the first part of the Feynman rules for theories with fermions.

- An interaction vertex consists of one dashed line, one line with an arrow pointing towards the vertex and one line with an arrow pointing away from the vertex: the dashed line corresponds to the scalar field,[54] the ingoing directed line to the field ψ and the outgoing

---

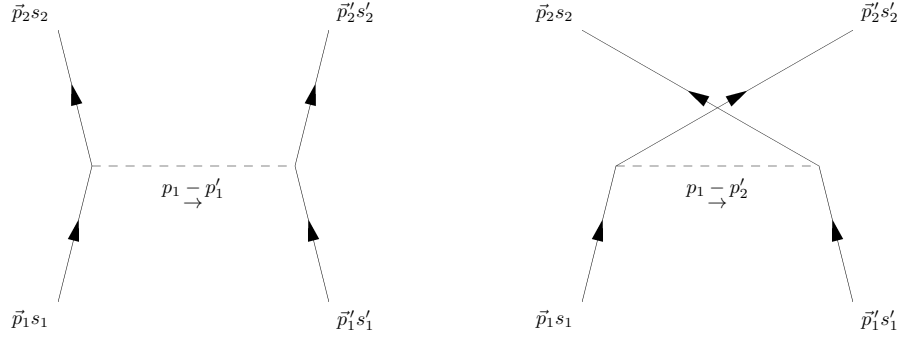[54]We use a dashed rather than a solid line to make more clear the difference between scalr and fermionic lines.

Figure 4: Diagrams for the lowest order $2 \to 2$ fermion scattering S-matrix element.

on to the field $\bar{\psi}$. It comes with the numerical factor ig.

- In coordinate space, associate a factor $u_s(\vec{p})e^{-ip\cdot x}$ to an incoming particle attached to vertex x, which corresponds to an fermion line entering the vertex, and associate a factor $\bar{u}_{s'}(\vec{p}')e^{ip'\cdot x}$ to an outgoing particle attached to the same vertex x, which corresponds to a fermion line leaving the vertex. The Dirac index of the spinors is contracted between $\bar{u}_{s'}(\vec{p}')$ and $u_s(\vec{p})$, so it can be thought to "flow" along the fermionic lines in the direction opposite to that of the arrows. In momentum space, associate only the spinors to the particles, and a momentum-conserving delta function to each vertex.

- A relative minus sign has to be included for diagrams that differ by the exchange of two external lines, which corresponds to the exchange of two fermions in a multiparticle state.

The scalar field is treated as above in the ϕ^4 theory. The Feynman diagrams relevant to this scattering matrix element are shown in Fig. 4.

Consider next the elastic scattering of a fermion and an antifermion. To lowest order,

$$
\begin{aligned}
F &= \langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}|S - \mathbf{1}|\vec{p}_1 s_1, \overline{\vec{p}_2 s_2}\rangle \\
&= \frac{(ig)^2}{2}\int d^4x \int d^4y \, \langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}|T\left(:\bar{\psi}(x)\psi(x)\phi(x)::\bar{\psi}(y)\psi(y)\phi(y):\right)|\vec{p}_1 s_1, \overline{\vec{p}_2 s_2}\rangle
\end{aligned}
\tag{6.103}
$$

$$
= \frac{(ig)^2}{2}\int d^4x \int d^4y \, D(x-y)\langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}|:\bar{\psi}(x)\psi(x)\bar{\psi}(y)\psi(y):|\vec{p}_1 s_1, \overline{\vec{p}_2 s_2}\rangle,
$$

where we have already made use of Wick's theorem. In the normal ordered product we need a positive-frequency field ψ_+ for the initial fermion, a positive-frequency field $\bar{\psi}_+$ for the initial antifermion, a negative-frequency field $\bar{\psi}_-$ for the final fermion, and a negative-frequency field

ψ_- for the final antifermion. The terms that contribute are then

$$\langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}| : \bar{\psi}(x)\psi(x)\bar{\psi}(y)\psi(y) : |\vec{p}_1s_1, \overline{\vec{p}_2s_2}\rangle$$

$$= \langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}|\{ : \bar{\psi}_+(x)\psi(x)_+\bar{\psi}_-(y)\psi_-(y) : + : \bar{\psi}_+(x)\psi(x)_-\bar{\psi}_-(y)\psi_+(y) :$$

$$+ : \bar{\psi}_-(x)\psi(x)_+\bar{\psi}_+(y)\psi_-(y) : + : \bar{\psi}_-(x)\psi(x)_-\bar{\psi}_+(y)\psi_+(y) : \}|\vec{p}_1s_1, \overline{\vec{p}_2s_2}\rangle$$

$$= \langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}|\{\bar{\psi}_-(y)\psi_-(y)\bar{\psi}_+(x)\psi(x)_+ - \bar{\psi}_-(y)\psi(x)_-\bar{\psi}_+(x)\psi_+(y) \tag{6.104}$$

$$- \bar{\psi}_-(x)\psi_-(y)\bar{\psi}_+(y)\psi(x)_+ + \bar{\psi}_-(x)\psi(x)_-\bar{\psi}_+(y)\psi_+(y)\}|\vec{p}_1s_1, \overline{\vec{p}_2s_2}\rangle$$

$$= \langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}|\bar{\psi}_-(x)\psi_-(x)|0\rangle\langle 0|\bar{\psi}_+(y)\psi(y)_+|\vec{p}_1s_1, \overline{\vec{p}_2s_2}\rangle$$

$$- \langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}|\bar{\psi}_-(y)\psi(x)_-|0\rangle\langle 0|\bar{\psi}_+(x)\psi_+(y)|\vec{p}_1s_1, \overline{\vec{p}_2s_2}\rangle + \ x \leftrightarrow y\,.$$

Terms differing only by the exchange of x and y give the same result after integration, so we need consider only the two terms written explicitly here. It is straightforward to show that

$$\langle 0|\bar{\psi}_{+\alpha_2}(x_2)\psi(x_1)_{+\alpha_1}|\vec{p}_1s_1, \overline{\vec{p}_2s_2}\rangle$$

$$= \int d\Omega_{q_1} \int d\Omega_{q_2} \sum_{t_1}\sum_{t_2} \bar{v}_{t_2\alpha_2}(\vec{q}_2)u_{t_1\alpha_1}(\vec{q}_1)e^{-i(q_2\cdot x_2 + q_1\cdot x_1)}\langle \vec{q}_1t_1, \overline{\vec{q}_2t_2}|\vec{p}_1s_1, \overline{\vec{p}_2s_2}\rangle$$

$$= \int d\Omega_{q_1} \int d\Omega_{q_2} \sum_{t_1}\sum_{t_2} \bar{v}_{t_2\alpha_2}(\vec{q}_2)u_{t_1\alpha_1}(\vec{q}_1)e^{-i(q_2\cdot x_2 + q_1\cdot x_1)} \tag{6.105}$$

$$\times \delta_{t_1s_1}(2\pi)^3 2p_1^0\delta^{(3)}(\vec{p}_1 - \vec{q}_1)\delta_{t_2s_2}(2\pi)^3 2p_2^0\delta^{(3)}(\vec{p}_2 - \vec{q}_2)$$

$$= \bar{v}_{s_2\alpha_2}(\vec{p}_2)u_{s_1\alpha_1}(\vec{p}_1)e^{-i(p_2\cdot x_2 + p_1\cdot x_1)}\,,$$

and similarly that

$$\langle \vec{p}_1's_1', \overline{\vec{p}_2's_2'}|\bar{\psi}_{-\alpha_1}(x_1)\psi_{-\alpha_2}(x_2)|0\rangle = \bar{u}_{s_1'\alpha_1}(\vec{p}_1')v_{s_2'\alpha_2}(\vec{p}_2')e^{i(p_2'\cdot x_2 + p_1'\cdot x_1)}\,. \tag{6.106}$$

Putting all the pieces together we find

$$F = (ig)^2 \int d^4x \int d^4y\, D(x-y)\Big\{ \bar{u}_{s_1'}(\vec{p}_1')v_{s_2'}(\vec{p}_2')\bar{v}_{s_2}(\vec{p}_2)u_{s_1}(\vec{p}_1)e^{i[(p_1'+p_2')\cdot x - (p_1+p_2)\cdot y]}$$

$$- \bar{u}_{s_1'}(\vec{p}_1')u_{s_1}(\vec{p}_1)\bar{v}_{s_2\alpha_2}(\vec{p}_2)v_{s_2'\alpha_2}(\vec{p}_2')e^{i[(p_2'-p_2)\cdot x + (p_1'-p_1)\cdot y]}\Big\}\,. \tag{6.107}$$

It is now easy to go over to momentum space and write

$$F = i(2\pi)^4\delta^{(4)}(p_1'+p_2'-p_1-p_2)ig^2\Big\{ \bar{u}_{s_1'}(\vec{p}_1')v_{s_2'}(\vec{p}_2')\bar{v}_{s_2}(\vec{p}_2)u_{s_1}(\vec{p}_1)\tilde{D}(p_1+p_2)$$

$$- \bar{u}_{s_1'}(\vec{p}_1')u_{s_1}(\vec{p}_1)\bar{v}_{s_2\alpha_2}(\vec{p}_2)v_{s_2'\alpha_2}(\vec{p}_2')\tilde{D}(p_1-p_1')\Big\}\,. \tag{6.108}$$

We have then learnt a new Feynman rule.

- In coordinate space, associate a factor $\bar{v}_s(\vec{p})e^{-ip\cdot x}$ to an incoming antiparticle attached to vertex x, which corresponds to a fermion line leaving the vertex, and associate a factor $v_{s'}(\vec{p}')e^{ip'\cdot x}$ to an outgoing antiparticle attached to vertex x, which corresponds to a fermion line entering the vertex. The Dirac index is contracted between spinors attached
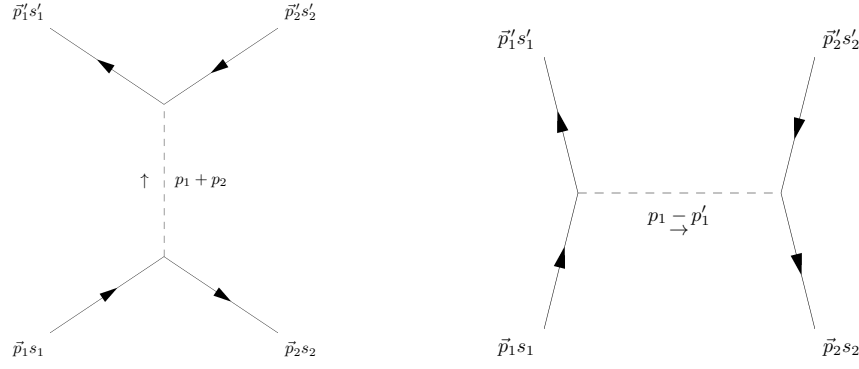
119

Figure 5: Diagrams for the lowest order $2 \to 2$ fermion-antifermion scattering S-matrix element.

to the same vertex, no matter if they correspond to particles or antiparticles, and "flows" along the fermionic lines in the direction opposite to that of the arrows. In momentum space, associate only the spinors to the antiparticles, and a momentum-conserving delta function to each vertex.

The diagrams corresponding to Eq. (6.108) are shown in Fig. 5, and applying the rules described so far one finds straighforwardly all the pieces of the formula, except for the relative minus sign between the diagrams. To understand why there is such a sign, let us go back to the case of fermio-fermion scattering. There, the diagram with the plus sign is that in which the first particle in the initial state and the first particle in the final state are connected to the first vertex,[55] and similarly the other two particles are connected to the second vertex. In this case the negative-frequency component of the field $\bar{psi}$ in the first vertex has to go past both fields in the second vertex, so that one gets a factor $(-1)^2 = 1$, to be the first operator acting on the final state to annihilate the first particle, which gives no further minus signs. The positive-frequency component of ψ at the same vertex is the first operator acting on the initial state, so no extra sign appears when it annihilates the first particle. The negative-frequency component of the field $\bar{psi}$ in the second vertex is now in the right position to annihilate the second particle in the final state, and the positive-frequency component of ψ in the same vertex is in the right position to annihilate the second particle in the initial state: no extra sign appears, and overall we have a plus sign. In the other diagram the two $\bar{\psi}_-$ are exchanged, and a relative minus sign appears. In fermion-antifermion scattering, the fields are in the right position when ψ_+ and $\bar{psi}_+$ from the first vertex act on the initial state and ψ_- and $\bar{psi}_-$ from the second vertex act on the final state. When we bring the negative-frequency component of $\bar{\psi}$ in the first vertex to act on the particle in the final state, we have to move it past both fields in the other vertex, and so there is no minus sign. Instead, when we bring the negative-frequency component of ψ in the second vertex to act on the antiparticle in the final state, it has to move past the field $\bar{\psi}$ in the same vertex, and this must be paid with a minus sign. In the case of antifermion-antifermion scattering this happens twice, so the diagrams are the same as in Fig. (4) with the orientation of the arrows reversed, but with the same signs for the two diagrams.

We still have to discuss internal fermion lines. To this end, let us consider fermion-antifermion annihilation into two scalars. The lowest-order contribution to the corresponding S-matrix

---

[55]The choice of which of the two vertices is the first is conventional and irrelevant, since exchanging completely two vertices does not bring about any minus signs.

element is

$$F = \langle \vec{q}_1, \vec{q}_2 | S - \mathbf{1} | \vec{p}_1 s_1, \overline{\vec{p}_2 s_2} \rangle$$

$$= \frac{(ig)^2}{2} \int d^4x \int d^4y \, \langle \vec{q}_1, \vec{q}_2 | T \left(: \bar{\psi}(x)\psi(x)\phi(x) :: \bar{\psi}(y)\psi(y)\phi(y) : \right) | \vec{p}_1 s_1, \overline{\vec{p}_2 s_2} \rangle$$

$$= \frac{(ig)^2}{2} \int d^4x \int d^4y \, \langle \vec{q}_1, \vec{q}_2 | : \bar{\psi}_\alpha(x)\psi_\beta(y)\phi(x)\phi(y) : S_{\alpha\beta}(x - y)$$

$$+ : \bar{\psi}_\beta(y)\psi_\alpha(x)\phi(x)\phi(y) : S_{\beta\alpha}(y - x) | \vec{p}_1 s_1, \overline{\vec{p}_2 s_2} \rangle$$

$$= (ig)^2 \int d^4x \int d^4y \, \langle \vec{q}_1, \vec{q}_2 | : \phi(x)\phi(y) : |0\rangle\langle 0| : \bar{\psi}_\alpha(x)\psi_\beta(y) : | \vec{p}_1 s_1, \overline{\vec{p}_2 s_2} \rangle S_{\alpha\beta}(x - y)$$

$$= (ig)^2 \int d^4x \int d^4y \left(e^{i(q_1 \cdot y + q_2 \cdot x)} + e^{i(q_1 \cdot x + q_2 \cdot y)} \right) e^{-i(p_1 \cdot y + p_2 \cdot x)} \bar{v}_{s_2}(\vec{p}_2) S(x - y) u_{s_1}(\vec{p}_1).$$

(6.109)

Going over to momentum space

$$F = i(2\pi)^4 \delta^{(4)}(q_1 + q_2 - p_1 - p_2) ig^2 \bar{v}_{s_2}(\vec{p}_2) \left(\tilde{S}(p_1 - q_1) - \tilde{S}(p_1 - q_2) \right) u_{s_1}(\vec{p}_1). \qquad (6.110)$$

We can now add another rule to our set.

- In coordinate space, internal fermion lines are oriented from the vertex y where the contracted $\bar{\psi}(y)$ field is to the vertex where the contracted $\psi(x)$ field is, consistently with the fact that for each vertex there is one fermion line that enters and one that exits. To such a line is associated a fermion propagator $S(x - y)$. The two Dirac indices of the propagator are contracted again following the flow of the lines in the opposite direction. In momentum space, internal lines are oriented, and must appear in diagrams consistently with the fact that one fermion line enters a vertex and one exits. A momentum q is associated to the line, flowing in the direction of the arrow, and a propagator $\tilde{S}(q)$ is included, with Dirac indices contracted as explained above.

We can now understand a general rule for minus signs in the presence of antifermionic lines crossing a diagram from side to side. Supppose we have a diagram with a fermionic line crossing the diagram from side to side. Collecting all the factors for this line, one finds from the Feynman rules

$$\bar{u}_{s'}(\vec{p}') e^{ip' \cdot x} S(x - x_1) S(x_1 - x_2) \ldots S(x_{n-1} - y) e^{-ip \cdot y} u_s(\vec{p}). \qquad (6.111)$$

This corresponds to the matrix element of the following term in the Wick expansion,

$$: \bar{\psi}(x)\psi(y) : S(x - x_1) S(x_1 - x_2) \ldots S(x_{n-1} - y) \qquad (6.112)$$

taken between one-particle states, $\langle 1' | : \bar{\psi}(x)\psi(y) : |1\rangle = \bar{u}_{1'} e^{ip' \cdot x} e^{-ip \cdot y} u_1$ (in shorthand notation). If we had instead antiparticles in the initial and final state, we would find instead

$$\bar{v}_s(\vec{p}) e^{-ip \cdot x} S(x - x_1) S(x_1 - x_2) \ldots S(x_{n-1} - x) e^{ip \cdot y} v_{s'}(\vec{p}') \qquad (6.113)$$

from the Feynman rules, possibly up to a sign. This term comes from the matrix element of the same operator as in Eq.(6.112), but now

$$\langle \bar{1}' | : \bar{\psi}(x)\psi(y) : |\bar{1}\rangle = -\langle \bar{1}' | : \psi(y)\bar{\psi}(x) : |\bar{1}\rangle = -\bar{v}_{1'} e^{ip' \cdot y} e^{-ip \cdot x} v_1, \qquad (6.114)$$

and so a minus sign must be included. This is a general rule:

- For each antifermionic line crossing the diagram from side to side (i.e., from the final to the initial state), include a minus sign.

A similar minus sign has to be included for fermionic loops: when there is a closed fermionic line, the ψ from the last vertex is contracted with the $\bar{\psi}$ from the first vertex, and has to cross an odd number of fields to get in the right position. In coordinate space one then finds $-\text{tr}\, S(x_1 - x_2)S(x_2 - x_3)\ldots S(x_n - x_1)$, where the trace takes into account the fact that the flow of the Dirac index goes back to the first vertex after passing the last one.

We have now concluded our derivation of Feynman rules. In the next subsection we discuss the missing bit, namely the evaluation of propagators.

6.6 Propagators

We begin the discussion of propagators starting from the theory of partial differential equations. The Klein-Gordon equation $(\Box + m^2)\varphi = 0$ can be derived by an action principle from the Klein-Gordon Lagrangian $\mathscr{L}_{\text{KG}} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{m^2}{2}\varphi^2$, describing a free real massive scalar field. If we now add an interaction part $\mathscr{L}(\varphi) = -V(\varphi)$, the EOM would change to

$$(\Box + m^2)\varphi(x) = -V'(\varphi(x)) \equiv -j(x), \tag{6.115}$$

i.e., they would be modified by the presence of a source term. In order to solve Eq. (6.115), it is convenient to solve first the following equation,

$$(\Box + m^2)G(x) = -\delta^{(4)}(x). \tag{6.116}$$

After finding $G(x)$, it is straightforward to show that

$$\varphi(x) = \varphi_0(x) + \int d^4y\, G(x-y)j(y), \tag{6.117}$$

where φ_0 solves the free KG equation, $(\Box + m^2)\varphi_0 = 0$, is a solution of Eq. (6.115):

$$(\Box + m^2)\varphi(x) = (\Box + m^2)\varphi_0(x) + \int d^4y\, (\Box_x + m^2)G(x-y)j(y)$$
$$= -\int d^4y\, \delta^{(4)}(x-y)j(y) = -j(x). \tag{6.118}$$

The function G is called *propagator*, as it propagates the effect of the source through spacetime. Mathematically, it is the inverse of the differntial operator $(\Box + m^2)$. The solution of Eq. (6.116) is most easily obtained in momentum space, i.e., taking a Fourier transform. Setting

$$G(x) = \int \frac{d^4p}{(2\pi)^4}\, e^{-ip\cdot x}\tilde{G}(p), \tag{6.119}$$

we find

$$(p^2 - m^2)\tilde{G}(p) = 1, \qquad \tilde{G}(p) = \frac{1}{p^2 - m^2}. \tag{6.120}$$

Now we should go back to coordinate space, but here we run into a problem: the solution is singular at $p^2 = m^2$, i.e., for $p^0 = \pm\sqrt{\vec{p}^2 + m^2}$, and we cannot perform the integral Eq. (6.119) without a prescription to avoid the singularities on the integration path. Such a prescription is
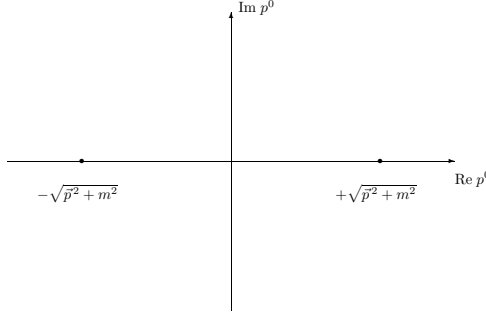
Figure 6: Singularities of the Klein-Gordon propagator in the complex plane.

far from unique. On the other hand, this just reflects the fact that there is not a unique choice of boundary conditions, which have to be imposed on φ to single out one out of the whole family of solutions in Eq. (6.117). The ambiguity in the choice of prescription thus reflects a physical ambiguity, and can be settled by means of physical arguments.

In the classical theory of waves, we want to impose that causality is reflected in our solution, in the sense that the state of the source at the spacetime point x can only affect the state of the solution at spacetime points in the forward lightcone of x. This leads to ask that the *retarded propagator* $G_{\mathrm{ret}}(x)$ vanishes at negative times. Eq. (6.115) must then be supplemented by the boundary condition $G_{\mathrm{ret}}(x) = 0$ $x^0 < 0$. This corresponds to a specific prescription in handling the singularity. Recalling the Fourier transform of the Heaviside theta function,

$$\theta(x^0) = \int \frac{d\omega}{2\pi i} e^{-i\omega x^0} \frac{-1}{\omega + i\epsilon}, \tag{6.121}$$

we see that the right prescription is

$$\tilde{G}_{\mathrm{ret}}(p) = \frac{1}{(p^0 + i\epsilon)^2 - \vec{p}^2 - m^2}. \tag{6.122}$$

In this way the poles are pushed into the lower half of the complex plane. Indeed, using the residue theorem to compute the integral over p^0,

$$\begin{aligned}
G_{\mathrm{ret}}(x) &= \int \frac{d^4 p}{(2\pi)^4} e^{-ip\cdot x} \frac{1}{(p^0 + i\epsilon)^2 - \vec{p}^2 - m^2} \\
&= \int \frac{d^3 p}{(2\pi)^3} \int \frac{dp^0}{2\pi} e^{-ip\cdot x} \frac{1}{p^0 + i\epsilon - \sqrt{\vec{p}^2 + m^2}} \frac{1}{p^0 + i\epsilon + \sqrt{\vec{p}^2 + m^2}},
\end{aligned} \tag{6.123}$$

we should close the contour in the lower half of the complex plane if $x^0 > 0$, and in the upper half of the complex plane if $x^0 < 0$. In the first case we pick $-2\pi i$ times the residue of the two poles, while in the second case we find no pole and so no contribution:

$$\begin{aligned}
G_{\mathrm{ret}}(x) &= \int \frac{d^4 p}{(2\pi)^4} e^{-ip\cdot x} \frac{1}{(p^0 + i\epsilon)^2 - \vec{p}^2 - m^2} \\
&= -i\theta(x^0) \int \frac{d^3 p}{(2\pi)^3 2\sqrt{\vec{p}^2 + m^2}} \left(e^{-i(\sqrt{\vec{p}^2+m^2}x^0 - \vec{p}\cdot\vec{x})} - e^{-i(-\sqrt{\vec{p}^2+m^2}x^0 - \vec{p}\cdot\vec{x})} \right) \\
&\quad - i\theta(x^0) \int \frac{d^3 p}{(2\pi)^3 2p^0} \left(e^{-ip\cdot x} - e^{ip\cdot x} \right).
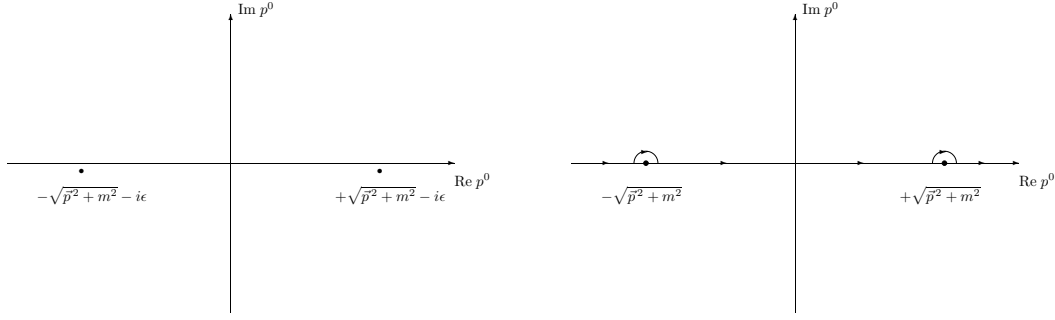\end{aligned} \tag{6.124}$$

123

Figure 7: Prescription to deal with the singluaritaies to obtain the retarded propagator (left). This corresponds in practice to computing the integral over p^0 using a modified integration path in the complex plane (right).

In the last pasage we have written $p^0 = \sqrt{\vec{p}^2 + m^2}$, and changed integration variable $\vec{p} \to -\vec{p}$ in the second term in brackets.

Let us consider now a slight modification of the problem. Instead of Eq. (6.116) consider

$$(\Box + m^2)D(x) = -i\delta^{(4)}(x)\,. \tag{6.125}$$

Setting

$$D(x) = \int \frac{d^4p}{(2\pi)^4}\, e^{-ip\cdot x}\tilde{D}(p)\,, \tag{6.126}$$

the general solution in momentum space is

$$\tilde{D}(p) = \frac{i}{p^2 - m^2}\,. \tag{6.127}$$

Consider now the following prescrtiption for the poles: we shift the one with positive real part down in the lower half of the complex plane, and the one with negative real part up in the upper half of the complex plane. This means that we take

$$\tilde{D}(p) = \frac{i}{p^2 - m^2 + i\epsilon}\,, \tag{6.128}$$

so that the poles are at $p^0 = \pm\sqrt{\vec{p}^2 + m^2} \mp i\epsilon$. Going back to coordinate space, making use of the residue theorem, we find now

$$
\begin{aligned}
D(x) &= \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot x}\frac{i}{p^2 - m^2 + i\epsilon}\\
&= \int \frac{d^3p}{(2\pi)^3}\int \frac{dp^0}{2\pi}\, ie^{-ip\cdot x}\frac{1}{p^0 - \sqrt{\vec{p}^2 + m^2} + i\epsilon}\frac{1}{p^0 + \sqrt{\vec{p}^2 + m^2} - i\epsilon}\\
&= \int \frac{d^3p}{(2\pi)^3 2\sqrt{\vec{p}^2 + m^2}}\left[(-i)\theta(x^0)e^{-i\sqrt{\vec{p}^2+m^2}x^0} + i(-1)\theta(-x^0)e^{i\sqrt{\vec{p}^2+m^2}x^0}\right]ie^{i\vec{p}\cdot\vec{x}}\\
&= \int \frac{d^3p}{(2\pi)^3 2p^0}\left[\theta(x^0)e^{-ip\cdot x} + \theta(-x^0)e^{-ip\cdot x}\right]\,.
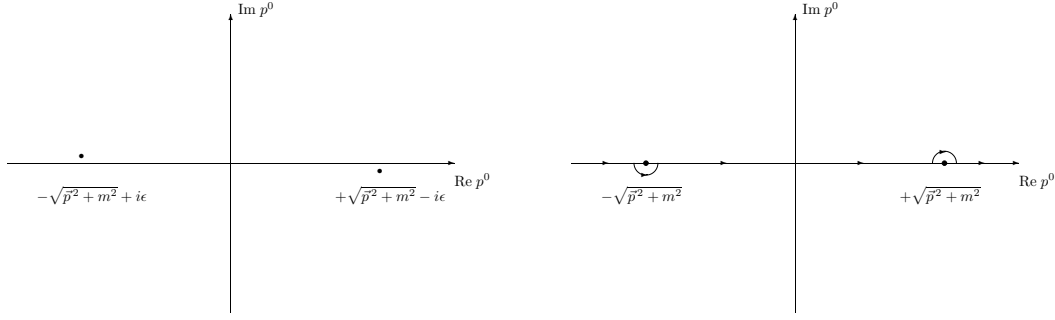\end{aligned}
\tag{6.129}
$$

Figure 8: Prescription to deal with the singluaritaies to obtain the causal propagator (left). This corresponds in practice to computing the integral over p^0 using a modified integration path in the complex plane (right).

This is the *causal* or *Feynman propagator*. What is it good for? Consider the vacuum expectation value of the the product of two free real scalar quantum fields,

$$\langle 0|\varphi(x)\varphi(y)|0\rangle = \int d\Omega_p \int d\Omega_q \, \langle 0|a(\vec{p})a(\vec{q})^\dagger|0\rangle e^{-i(p\cdot x - q\cdot y)} = \int d\Omega_p \, e^{-ip\cdot(x-y)} = \Delta(x) \,. \quad (6.130)$$

If we consider the contraction of two scalar fields, i.e., the vacuum expectation value of the time-ordered product, we then find

$$\langle 0|T\left(\varphi(x)\varphi(y)\right)|0\rangle = \int d\Omega_p \left(\theta(x^0 - y^0)e^{-ip\cdot(x-y)} + \theta(y^0 - x^0)e^{-ip\cdot(y-x)}\right) = D(x-y) \,, \quad (6.131)$$

which is precisely the propagator found in Eq. (6.129). This justifies the use of the term "propagator" for this object. One can prove directly, by making use of the EOM and the CCR, that the contraction $D(x)$ is a propagator, i.e., it solve the KG equation with a delta term. Indeed, after showing that

$$\begin{aligned}
\partial_{x0}^2 T\left(\varphi(x)\varphi(y)\right) &= \partial_{x0}\left\{T\left(\partial_{x0}\varphi(x)\varphi(y)\right) + \delta(x^0 - y^0)[\varphi(x), \varphi(y)]_{\text{ET}}\right\} \\
&= \partial_{x0}T\left(\partial_{x0}\varphi(x)\varphi(y)\right) = T\left(\partial_{x0}^2\varphi(x)\varphi(y)\right) + \delta(x^0 - y^0)[\partial_{x0}\varphi(x), \varphi(y)]_{\text{ET}} \\
&= T\left(\partial_{x0}^2\varphi(x)\varphi(y)\right) + \delta(x^0 - y^0)[\pi(x), \varphi(y)]_{\text{ET}} = T\left(\partial_{x0}^2\varphi(x)\varphi(y)\right) = -i\delta^{(4)}(x-y) \,,
\end{aligned} \quad (6.132)$$

we find

$$(\Box_x + m^2)T\left(\varphi(x)\varphi(y)\right) = T\left((\Box_x + m^2)\varphi(x)\varphi(y)\right) - i\delta^{(4)}(x-y) = -i\delta^{(4)}(x-y) \,. \quad (6.133)$$

however, in this way we cannot show that the "$-i\epsilon$" prescription is the correct one to obtain $D(x)$.

It is easy to show that for a complex (charged) scalar field the following results hold,

$$\langle 0|T\left(\varphi(x)\varphi(y)\right)|0\rangle = \langle 0|T\left(\varphi(x)^\dagger\varphi(y)^\dagger\right)|0\rangle = 0 \,, \quad (6.134)$$

and

$$\langle 0|T\left(\varphi(x)\varphi(y)^\dagger\right)|0\rangle = D(x-y) \,. \quad (6.135)$$

125

This expectation value can be recast as

$$\langle 0|T\left(\varphi(x)\varphi(y)^\dagger\right)|0\rangle = \theta(x^0 - y^0)\langle 0|\varphi(x)\varphi(y)^\dagger|0\rangle + \theta(y^0 - x^0)\langle 0|\varphi(y)^\dagger\varphi(x)|0\rangle$$
$$= \theta(x^0 - y^0)\langle\varphi(x)^\dagger|\varphi(y)^\dagger\rangle + \theta(y^0 - x^0)\langle\varphi(y)|\varphi(x)\rangle\,, \tag{6.136}$$

where the states $|\varphi(x)^\dagger\rangle$ and $|\varphi(x)\rangle$ correspond to a particle or an antiparticle, respectively, being created at x and destroyed at y. Let us now recast Eq. (6.131) as follows,

$$D(x - y) = \theta(x^0 - y^0)\int d^3p\,\phi_{(+)p}(x)\phi_{(+)p}(x)^* + \theta(y^0 - x^0)\int d^3p\,\phi_{(-)p}(x)\phi_{(-)p}(x)^*\,, \tag{6.137}$$

where

$$\phi_{(\pm)p}(x) = \frac{e^{\mp ip\cdot x}}{\sqrt{(2\pi)^3 2p^0}} \tag{6.138}$$

are positive-energy and negative-energy solutions of the KG equation. As the propagator carries the field from y to x, we see that first term in Eq. (6.131) propagates a positive-energy solution forward in time, while the second one propagates a negative-energy solution backward in time. The latter can be interpreted as a positive-energy state travelling actually forward in time: we then obtain the reinterpretation of negative-energy particle states as positive-energy antiparticle state, since all the quantum numbers of the particle have to be changed for the interpretation to be consistent.

Let us now discuss the fermion contraction. First we show that it satisfies the following inhomogeneous Dirac equation,

$$(i\slashed{\partial} - m)S(x) = i\delta^{(4)}(x)\,, \tag{6.139}$$

so that we may call it as well fermion propagator. Indeed,[56]

$$(i\slashed{\partial}_x - m)\langle 0|T\left(\psi(x)\bar\psi(y)\right)|0\rangle$$
$$= \langle 0|T\left(i(\slashed{\partial}_x - m)\psi(x)\bar\psi(y)\right)|0\rangle + i\gamma^0\delta(x^0 - y^0)\langle 0|\{\psi(x), \psi(y)^\dagger\gamma^0\}_{\mathrm{ET}}|0\rangle \tag{6.140}$$
$$= i\gamma^0\delta^{(4)}(x - y)\gamma^0 = i\delta^{(4)}(x - y)\,.$$

We can then solve the equation in Fourier transform,

$$(\slashed{p} - m)\tilde{S}(p) = i\,, \tag{6.141}$$

where

$$S(x) = \int\frac{d^4p}{(2\pi)^4}\,e^{-ip\cdot x}\tilde{S}(p)\,. \tag{6.142}$$

Multiplying by $\slashed{p} + m$ we find

$$(p^2 - m^2)\tilde{S}(p) = i(\slashed{p} + m)\,, \tag{6.143}$$

---

[56]If you are confused by how the Dirac indices are contracted, write them out explicitly:

$$(i\slashed{\partial}_x - m)_{\alpha\beta}\langle 0|T\left(\psi_\beta(x)\bar\psi_\gamma(y)\right)|0\rangle = i\gamma^0_{\alpha\beta}\delta(x^0 - y^0)\langle 0|\{\psi_\beta(x), \psi_\delta(y)^\dagger\gamma^0_{\delta\gamma}\}_{\mathrm{ET}}|0\rangle = i\gamma^0_{\alpha\beta}\delta^{(4)}(x - y)\delta_{\beta\delta}\gamma^0_{\delta\gamma}\,,$$

which equals $i\delta^{(4)}(x - y)\delta_{\alpha\gamma}$.

which can be inverted to give

$$\tilde{S}(p) = \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \, .\tag{6.144}$$

here the choice of prescription is inspired by the scalar case. Let us verify that it is the correct one. Integrating over p^0 and using the residue theorem we find ($E_p \equiv \sqrt{\vec{p}^2 + m^2}$)

$$\begin{aligned}
S(x) &= \int \frac{d^4p}{(2\pi)^4} e^{-ip\cdot x} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \\
&= i\Bigg\{ \theta(x^0 - y^0)(-i) \int \frac{d^3p}{(2\pi)^3} e^{-i(E_p x^0 - \vec{p}\cdot\vec{x})} \frac{E_p\gamma^0 - \vec{p}\cdot\vec{\gamma} + m}{2E_p} \\
&\quad + \theta(x^0 - y^0)(+i) \int \frac{d^3p}{(2\pi)^3} e^{-i(-E_p x^0 - \vec{p}\cdot\vec{x})} \frac{-E_p\gamma^0 - \vec{p}\cdot\vec{\gamma} + m}{-2E_p} \Bigg\} \\
&= \Bigg\{ \theta(x^0 - y^0) \int \frac{d^3p}{(2\pi)^3} e^{-i(E_p x^0 - \vec{p}\cdot\vec{x})} \frac{E_p\gamma^0 - \vec{p}\cdot\vec{\gamma} + m}{2E_p} \\
&\quad + \theta(x^0 - y^0) \int \frac{d^3p}{(2\pi)^3} e^{-i(-E_p x^0 + \vec{p}\cdot\vec{x})} \frac{-E_p\gamma^0 + \vec{p}\cdot\vec{\gamma} + m}{2E_p} \Bigg\} \\
&= \int d\Omega_p \left[\theta(x^0 - y^0) e^{-ip\cdot x}(\not{p} + m) + \theta(y^0 - x^0) e^{ip\cdot x}(-\not{p} + m) \right] .
\end{aligned}\tag{6.145}$$

On the other hand,

$$\begin{aligned}
\langle 0 | T\left(\psi_\beta(x)\bar{\psi}_\gamma(0)\right) | 0 \rangle &= \int d\Omega_p \int d\Omega_q \sum_{s,s'} \langle 0 | \theta(x^0) e^{-ip\cdot x} u_s(\vec{p})\bar{u}_{s'}(\vec{q}) b_s(\vec{p}) b_{s'}(\vec{q})^\dagger \\
&\qquad\qquad - \theta(-x^0) e^{ip\cdot x} v_s(\vec{q})\bar{v}_{s'}(\vec{p}) d_s(\vec{q}) d_{s'}(\vec{p})^\dagger | 0 \rangle \\
&= \int d\Omega_p \sum_s \left[\theta(x^0) e^{-ip\cdot x} u_s(\vec{p})\bar{u}_s(\vec{p}) - \theta(-x^0) e^{ip\cdot x} v_s(\vec{p})\bar{v}_s(\vec{p}) \right] \\
&= \int d\Omega_p \sum_s \left[\theta(x^0) e^{-ip\cdot x}(\not{p} + m) - \theta(-x^0) e^{ip\cdot x}(\not{p} - m) \right] ,
\end{aligned}\tag{6.146}$$

which shows that the prescription is correct.

7 Quantisation of the electromagnetic field

There is one more thing to do before building QED, namely quantising the electromagnetic field. As anticipated in Section 2.6, this is made complicated by the presence of gauge symmetry. Recall that Maxwell equations can be derived from the following Lagrangian,

$$\mathscr{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - J_\mu A^\mu \, ,\tag{7.1}$$

and read

$$\Box A^\nu - \partial^\nu \partial_\mu A^\mu = J^\nu \, .\tag{7.2}$$

As discussed in Section 2.6, the $\mu = 0$ equation is not a dynamical equation for A^0 but rather a constraint,

$$\vec{\nabla}^2 A^0 + \partial_0 \vec{\nabla} \cdot \vec{A} = -J^0 \,, \tag{7.3}$$

which can be solved for A^0 by inverting the Laplacian,

$$A^0 = -\frac{1}{\Delta} \left(\partial_0 \vec{\nabla} \cdot \vec{A} + J^0 \right) \,. \tag{7.4}$$

This can be done easily in Fourier transform. Transforming only with respect to the spatial coordinates,

$$A_\mu(x) = \int \frac{d^3 p}{(2\pi)^3} e^{i\vec{p} \cdot \vec{x}} \tilde{A}_\mu(t, \vec{p}) \,, \tag{7.5}$$

we find

$$\tilde{A}^0 = \frac{1}{\vec{p}^2} \left(i\vec{p} \cdot \partial_0 \vec{\tilde{A}} + \tilde{J}^0 \right) \,. \tag{7.6}$$

One has now to transform back to coordinate space. This yields

$$A^0(t, \vec{x}) = \int \frac{d^3 p}{(2\pi)^3} e^{i\vec{p} \cdot \vec{x}} \frac{1}{\vec{p}^2} \left(i\vec{p} \cdot \partial_0 \vec{\tilde{A}}(t, \vec{p}) + \tilde{J}^0(t, \vec{p}) \right) = \int d^3 y \frac{\partial_0 \vec{\nabla} \cdot \vec{A}(t, \vec{y}) + J^0(t, \vec{y})}{4\pi |\vec{x} - \vec{y}|} \,. \tag{7.7}$$

Plugging now A^0 in the remaining equations one finds

$$\Box \vec{A}_\perp = \vec{j}_\perp \,, \tag{7.8}$$

where the transverse fields $\vec{A}_\perp$ are given by $\vec{A}_\perp = \vec{A} - \vec{\nabla} \frac{1}{\Delta} \vec{\nabla} \cdot \vec{A}$, and simnilarly the transverse part of the current is $\vec{j}_\perp = \vec{j} - \vec{\nabla} \frac{1}{\Delta} \vec{\nabla} \cdot \vec{j}$. Since $\vec{\nabla} \cdot \vec{A}_\perp = 0$, the divergence of $\vec{A}$ and the temporal component A^0 only enter the constraint equation Eq. (7.3), and are not determined dynamically. The dynamical degrees of freedom are only the transverse components of the four-potential, which reflects the physical fact that photons happen to show only two independent polarisations. This is also in agreement with our general discussion of irreducible representations of the Poincaré group for massless particles, where we saw that in the case of nonzero spin only two states of opposite helicity are present. Photons are indeed expected to be massless, since they mediate a long-range force.

The equation of motion and the constraint can be satisfied by infinitely many choices for A_μ: if we replace $A_\mu \to A_\mu + \partial_\mu \Lambda$, with an arbitrary function of spacetime Λ, both the dynamical equation and the constraint are unchanged:

$$\delta \vec{A}_\perp = \vec{\nabla} \Lambda - \vec{\nabla} \frac{1}{\Delta} \vec{\nabla} \cdot \vec{\nabla} \Lambda = 0 \,,$$
$$\delta(\text{constraint}) = \vec{\nabla}^2 \partial_0 \Lambda - \partial_0 \vec{\nabla} \cdot \vec{\nabla} \Lambda = 0 \,. \tag{7.9}$$

This invariance of the EOM is called *gauge invariance*, and can actually be seen already at the level of the Lagrangian: the field-strength tensor $F_{\mu\nu}$ is manifestly invariant under a gauge transformation $A_\mu \to A_\mu + \partial_\mu \Lambda$, while the coupling to matter changes by a total derivative since the current is conserved, $\delta(J^\mu A_\mu) = J^\mu \partial_\mu \Lambda = \partial_\mu (J^\mu \Lambda)$. The reduced number of actual degrees of freedom and invariance under gauge transformations require that we proceed carefully

in quantising the theory. In fact, if we compute the conjugate momenta from the Lagrangian, we find that

$$\pi^\mu = \frac{\partial \mathscr{L}}{\partial(\partial_0 A_\mu)} = -F^{0\mu}\,, \tag{7.10}$$

so that $\pi^0 = 0$ identically, and we cannot use the canonical quantisation approach out of the box.

Before discussing how quantisation can be done, let us mention how it *cannot* be done. One could try to use the physical electric and magnetic fields $\vec{E}$ and $\vec{B}$: these are all dynamical degrees of freedom, and there are no problems due to the redundancy of the description. Although this is perfectly viable classically, there are quantum phenomena which can be explained only if we couple matter to the four-potential: this means that the four-potential is actually physical (to some extent). One could then try to build this field as we did with the scalar and Dirac fields, starting from the photon states and combining the corresponding creation and annihilation operators into a local field which transforms irreducibly under the Lorentz group. Unfortunately, it is not possible to build such a field for massless, spin 1 particles (see Weinberg I for details).

Let us now discuss the ways to go. One is to isolate the physical degrees of freedom and qunatise them only. The idea is that the redundancy of gauge invariance can be removed by imposing some extra condition on the fields, i.e., *fixing the gauge*, so that the solution to the constraint becomes unique. Here we briefly sketch this approach. If we impose that the vector potential is divergenceless, $\vec{\nabla} \cdot \vec{A} = 0$ (*Coulomb* or *radiation gauge*), the constraint becomes $\vec{\nabla}^2 A^0 = -J^0$, which in the absence of sources implies $A^0 = 0$. Let us focus on this case. The gauge field in Coulomb gauge only shows the two physical, transverse degrees of freedom. We can then impose canonical commutation relations on these,[57] and so obtain a quantised field describing only the two physical photons. The disadvantage of this approach is that Lorentz invariance is not manifest, which becomes cumbersome when including interactions. This is a disadvantage of quantisation in any non-covariant gauge, like the *axial gauge* $A^3 = 0$ or the *temporal gauge* $A^0 = 0$. The three gauge conditions discussed here are all legitimate: if we have a field A_μ that does not satisfy them, we can always find a new, gauge-transformed one that will. For example, in Coulomb gauge, if $\vec{\nabla} \cdot \vec{A} \neq 0$, then $\vec{A}' = \vec{A} - \vec{\nabla}\Lambda$ will if we choose Λ to satisfy $\Delta\Lambda = \vec{\nabla} \cdot \vec{A}$, which can always be done. After a Lorentz transformation, though, the new field will not satisfy the Coulomb condition, and a further gauge transformation is needed to re-establish it.

A Lorentz-invariant choice of gauge is the Loren(t)z gauge[58] $\partial_\mu A^\mu = 0$. This is also a legitimate gauge, since a nonzero $\partial_\mu A^\mu$ can be gauged away by transforming $A_\mu \to A_\mu + \partial_\mu \Lambda$ with $\Box\Lambda = -\partial_\mu A^\mu$. However, the solution is not unique: if $\partial_\mu A^\mu = 0$, we can make a gauge transformation with $\Box\Lambda = 0$ (that has nontrivial solutions) and still have the gauge condition satisfied. In this gauge the EOM simplify to $\Box A^\mu = J^\mu$; for a free electromagnetic field, $J^\mu = 0$, we simply find four massless KG equations $\Box A_\mu = 0$, on top of which the gauge condition has to be imposed.

---

[57]This has to be done carefully, since the commutation relations must be consistent with the gauge-fixing condition and the constraint equation.

[58]This gauge was first introduced by *L. V. Lorenz*, and not by *H. A. Lorentz* as it is generally believed, cf. the paper J. D. Jackson and L. B. Okun, Rev. Mod. Phys. 73 (2001) 663, on the history of gauge theories.

7.1 Gupta-Bleuler quantisation

From now on we will focus on the free case, and try to quantise the theory. The procedure that we will follow is different from the one used in Coulomb gauge: instead of starting from Maxwell's Lagrangian $\mathscr{L}_M = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, fix the gauge and solve the equations of motion, and then quantise the solution, we start from a modified Lagrangian which has the gauge-fixed EOM as its EOM, solve them and quantise the field, and then impose somehow the gauge condition on the quantised field. This procedure is called *Gupta-Bleuler quantisation*. We start from the following modified Lagrangian,[59]

$$\mathscr{L}' \equiv -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\partial_\mu A^\mu)^2 = \mathscr{L}_F + \text{total divergence}\,,$$

$$\mathscr{L}_F \equiv -\frac{1}{2}\partial_\mu A_\nu \partial^\mu A^\nu\,.$$

$$(7.11)$$

The extra term breaks gauge invariance, so it should be possible to quantise the theory by canonical methods. The equations of motions are easily seen to be

$$\Box A_\mu = 0\,,\tag{7.12}$$

and they are of course the same irrespectively of us using $\mathscr{L}'$ or $\mathscr{L}_F$. These are precisely the gauge-fixed Maxwell equations in Lorenz gauge, so if we were dealing with classical fields, using this Lagrangian would be exactly equivalent to using Maxwell Lagrangian and using the Lorenz gauge condition. We now proceed with the canonical program, and impose CCR on the field solving Eq. (7.12) and its conjugate momenta. We use $\mathscr{L}_F$ to find

$$\pi^\mu = \frac{\partial \mathscr{L}_F}{\partial(\partial_0 A_\mu)} = -\partial^0 A^\mu = -\partial_0 A^\mu\,.\tag{7.13}$$

Now all $\pi^\mu \neq 0$, and the canonical approach is viable. Notice that π^μ is the momentum conjugate to A_μ, not to A^μ. The CCR read

$$[A_\mu(x), A_\nu(y)]_{\mathrm{ET}} = [\pi_\mu(x), \pi_\nu(y)]_{\mathrm{ET}} = 0\,,$$

$$[A_\mu(x), \pi_\nu(y)]_{\mathrm{ET}} = i\eta_{\mu\nu}\delta^{(3)}(\vec{x}-\vec{y})\,.$$

$$(7.14)$$

The EOM can be solved easily: they are just four KG equations for a real field (as A_μ must be Hermitian for the Lagrangian to be Hermitian), and so we can write down the solution right away,

$$A_\mu(x) = \int d\Omega_p \left\{ a_\mu(\vec{p})e^{-ip\cdot x} + a_\mu(\vec{p})^\dagger e^{ip\cdot x} \right\}\,,\tag{7.15}$$

for operator-valued coefficients $a_\mu(\vec{p})$ and $a_\mu(\vec{p})^\dagger$, whose commutation relations are determined by the CCR. We can exploit again our knowledge of the scalar case to write

$$a_\mu(\vec{p}) = \int d^3x\, e^{ip\cdot x}i\overleftrightarrow{\partial_0}A_\mu(x)\,,$$

$$a_\mu(\vec{p})^\dagger = \int d^3x\, e^{-ip\cdot x}(-i\overleftrightarrow{\partial_0})A_\mu(x)\,,$$

$$(7.16)$$

---

[59]The subscript F stands for Fermi, who first proposed this Lagrangian to quantise the electromagnetic field.

to find that

$$[a_\mu(\vec{p}), a_\nu(\vec{q})] = [a_\mu(\vec{p})^\dagger, a_\nu(\vec{q})^\dagger] = 0$$
$$[a_\mu(\vec{p}), a_\nu(\vec{q})^\dagger] = -\eta_{\mu\nu} 2p^0 (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) \,. \tag{7.17}$$

From these relations we can proceed to build Fock space in the usual way. Everything looks nice here, with one troubling exception: for $\mu = 0$ we find that $[a_0(\vec{p}), a_0(\vec{q})^\dagger] = -2p^0(2\pi)^3\delta^{(3)}(\vec{p}-\vec{q})$ which implies negative norm for the corresponding states. This is physically unacceptable, since the presence of states with negative norm ("*ghosts*") imply negative probabilities, and thus a breakdown of unitarity. Before looking ofr a way out of this trouble, let us disentangle the Lorentz index μ from the actual polarisation of the photons. Let us set

$$a_\mu(\vec{p}) = \sum_{\lambda=0}^{3} \varepsilon_\mu^{(\lambda)} a_{(\lambda)}(\vec{p}) \,, \tag{7.18}$$

where $\varepsilon_\mu^{(\lambda)}$ are four orthonormal vectors forming a complete set, in the following sense,

$$\varepsilon^{(\lambda)} \cdot \varepsilon^{(\lambda')*} = \eta^{\mu\nu} \varepsilon_\mu^{(\lambda)} \varepsilon_\nu^{(\lambda')*} = \eta^{\lambda\lambda'} \,, \qquad \eta_{\lambda\lambda'} \varepsilon_\mu^{(\lambda)} \varepsilon_\nu^{(\lambda')*} = \eta_{\mu\nu} \,. \tag{7.19}$$

The field operator reads

$$A_\mu(x) = \int d\Omega_p \sum_{\lambda=0}^{3} \left\{ \varepsilon_\mu^{(\lambda)} a_{(\lambda)}(\vec{p}) e^{-ip\cdot x} + \varepsilon_\mu^{(\lambda)*} a_{(\lambda)}(\vec{p})^\dagger e^{ip\cdot x} \right\} \,. \tag{7.20}$$

An explicit choice is the following: assuming that $\vec{p}$ is along direction 3, for $\lambda = 0, 3$ we take

$$\varepsilon_\mu^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \,, \qquad \varepsilon_\mu^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \,, \tag{7.21}$$

while for $\lambda = 1, 2$ we either take

$$\varepsilon_\mu^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ i \\ 0 \end{pmatrix} \,, \qquad \varepsilon_\mu^{(2)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} \,, \tag{7.22}$$

corresponding to *circular polarisations* and thus helicity eigenstates, or

$$\varepsilon_\mu^{(1)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \,, \qquad \varepsilon_\mu^{(2)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \,, \tag{7.23}$$

corresponding to *linear polarisations*. These correspond to the physical states of a massless, spin-1 particle. In general,

$$\varepsilon_\mu^{(3)} = \begin{pmatrix} 0 \\ \frac{\vec{p}}{|\vec{p}|} \end{pmatrix} \,, \tag{7.24}$$

and $\varepsilon_\mu^{(1,2)}$ are taken to be in the spatial plane orthogonal to $\vec{p}$. Now, since $\eta^{\lambda\lambda} a_{(\lambda)}(\vec{p}) = \varepsilon^{(\lambda)\mu*} a_\mu(\vec{p})$ (no sum over λ), we find that

$$
\begin{aligned}
[a_{(\lambda)}(\vec{p}), a_{(\lambda')}(\vec{q})] &= [a_{(\lambda)}(\vec{p})^\dagger, a_{(\lambda')}(\vec{q})^\dagger] = 0 \\
[a_{(\lambda)}(\vec{p}), a_{(\lambda')}(\vec{q})^\dagger] &= -\eta_{\lambda\lambda'} 2p^0 (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) \,.
\end{aligned}
\tag{7.25}
$$

The states of *temporal photons*, created by $a_{(0)}(\vec{p})^\dagger$, have negative norms, and are therefore clearly unphysical. *Transverse photons* ($\lambda = 1, 2$) and *longitudinal photons* $\lambda = 3$) have states with positive norm; longitudinal photons are however unphysical as well.

We would like to get rid somehow of the negative norm states and of the unphysical polarisations. This is where the gauge condition, which we have not imposed yet, comes to our rescue. How can we impose $\partial_\mu A^\mu = 0$? We might try to impose it as an operator identity, but this is bound to fail: in fact, since

$$
\begin{aligned}
0 &= [\partial_\mu A^\mu(x), A^0(y)]_{\mathrm{ET}} = [\partial_0 A_0(x), A_0(y)]_{\mathrm{ET}} + [\vec{\nabla} \cdot \vec{A}(x), A_0(y)]_{\mathrm{ET}} \\
&= [\partial_0 A_0(x), A_0(y)]_{\mathrm{ET}} = [A_0(y), -\partial_0 A_0(x)]_{\mathrm{ET}} = [A_0(y), \pi_0(x)]_{\mathrm{ET}} = i\delta^{(3)}(\vec{x} - \vec{y}) \neq 0 \,,
\end{aligned}
\tag{7.26}
$$

this would be inconsistent. We might then try to impose it on the Hilbert space of states as a condition selecting the physical states $|\Psi\rangle$, $\partial_\mu A^\mu(x)|\Psi\rangle = 0$, but this would leave the vacuum state out of the physical subspace. Indeed, given that $a_\mu(\vec{p})|0\rangle = 0$, we would have

$$
\partial_\mu A^\mu(x)|0\rangle = \int d\Omega_p \, e^{ip\cdot x} a_\mu(\vec{p})^\dagger |0\rangle \neq 0 \,.
\tag{7.27}
$$

A viable possibility is to ask for the positive-frequency component of $\partial_\mu A^\mu(x)$ to annihilate the physical states,

$$
\partial_\mu A_+^\mu(x)|\Psi\rangle = 0 \,.
\tag{7.28}
$$

The gauge condition is then realised on average,

$$
\langle\Psi|\partial_\mu A^\mu(x)|\Psi\rangle = \langle\Psi| \left(\partial_\mu A_+^\mu(x)|\Psi\rangle\right) + \left(\langle\Psi|\partial_\mu A_-^\mu(x)\right) |\Psi\rangle = 0 \,,
\tag{7.29}
$$

and more generally on has that $\langle\Psi'|\partial_\mu A^\mu(x)|\Psi\rangle = 0$ for physical states $|\Psi\rangle$, $|\Psi'\rangle$. The first two steps in constructing the physical Hilbert space of the system is thus to first construct the usual Fock space, which in the case at hand has indefinite norm, and to select then a subspace $\mathcal{H}_{\mathrm{phys}}$ by imposing the condition Eq. (7.28), $\mathcal{H}_{\mathrm{phys}} = \{\,|\Psi\rangle \in \mathcal{H} \,|\, \partial_\mu A_+^\mu(x)|\Psi\rangle = 0\}$. This eliminates negative-norm states. Indeed, consider $\mathcal{H}_{\mathrm{phys}}$ with the scalar product (and thus the norm) inherited from the Fock space $\mathcal{H}$. The most general state in $\mathcal{H}$ can be written as a superposition of basis vectors $|\psi\rangle = |\psi_T\rangle|\phi\rangle$, where $|\psi_T\rangle$ contains only transverse modes and $|\phi\rangle$ contains temporal and longitudinal modes. The operator $\partial_\mu A^\mu(x)$ reads

$$
\begin{aligned}
\partial_\mu A^\mu(x) &= -i \int d\Omega_p \, \left\{ p^\mu a_\mu(\vec{p}) e^{-ip\cdot x} - p^\mu a_\mu(\vec{p})^\dagger e^{ip\cdot x} \right\} \\
&= -i \int d\Omega_p \sum_{\lambda=0}^3 \left\{ p^\mu \varepsilon_\mu^{(\lambda)} a_{(\lambda)}(\vec{p}) e^{-ip\cdot x} + p^\mu \varepsilon_\mu^{(\lambda)*} a_{(\lambda)}(\vec{p})^\dagger e^{ip\cdot x} \right\} \\
&= -i \int d\Omega_p \, \left\{ L(\vec{p}) e^{-ip\cdot x} - L(\vec{p})^\dagger e^{ip\cdot x} \right\} \,,
\end{aligned}
\tag{7.30}
$$

where

$$L(\vec{p}) \equiv p^\mu a_\mu(\vec{p}) = p \cdot \varepsilon^{(\lambda)} a_{(\lambda)}(\vec{p}) = p^0 a_{(0)}(\vec{p}) - |\vec{p}| a_{(3)}(\vec{p}) = p^0 [a_{(0)}(\vec{p}) - a_{(3)}(\vec{p})] \,,$$
$$L(\vec{p})^\dagger \equiv p^\mu a_\mu(\vec{p})^\dagger = p \cdot \varepsilon^{(\lambda)*} a_{(\lambda)}(\vec{p})^\dagger = p^0 a_{(0)}(\vec{p})^\dagger - |\vec{p}| a_{(3)}(\vec{p})^\dagger = p^0 [a_{(0)}(\vec{p})^\dagger - a_{(3)}(\vec{p})^\dagger] \,.$$
$$(7.31)$$

Imposing Eq. (7.28) for all x is equivalent to impose

$$L(\vec{p})|\Psi\rangle = 0 \,. \tag{7.32}$$

Since $L(\vec{p})$ acts only on temporal and longitudinal photons, Eq. (7.32) for vectors of the form $|\psi_T\rangle|\phi\rangle$ amounts to

$$L(\vec{p})|\phi\rangle = 0 \implies [a_{(0)}(\vec{p}) - a_{(3)}(\vec{p})]|\phi\rangle = 0 \quad \forall \vec{p}. \tag{7.33}$$

This implies that for physical states $a_{(0)}(\vec{p})|\phi\rangle = a_{(3)}(\vec{p})]|\phi\rangle$ and so

$$\langle\phi|a_{(0)}(\vec{p})^\dagger a_{(0)}(\vec{p})|\phi\rangle = \langle\phi|a_{(3)}(\vec{p})^\dagger a_{(3)}(\vec{p})]|\phi\rangle \,, \tag{7.34}$$

i.e., the occupation number of temporal and longitudinal photons with momentum $\vec{p}$ are the same. Equivalently, we have that

$$\langle\phi|a_0(\vec{p})^\dagger a_0(\vec{p})|\phi\rangle = \langle\phi|a_3(\vec{p})^\dagger a_3(\vec{p})]|\phi\rangle \,, \tag{7.35}$$

for the operators a_μ and their adjoint. What does this imply for the norm of the states? The most general vector satisfying Eq. (7.32) for a given content in transverse photons is of the form

$$|\Psi\rangle = |\psi_T\rangle|\phi\rangle \,, \qquad |\phi\rangle = \sum_n |\phi_n\rangle \,,$$
$$|\phi_n\rangle = \int d\Omega_{p_1} \ldots \int d\Omega_{p_n}\, C^{(n)}(\vec{p}_1, \ldots, \vec{p}_n\,) L(\vec{p}_1)^\dagger \ldots L(\vec{p}_n)^\dagger |0\rangle \,.$$
$$(7.36)$$

Such states are certainly physical states: since

$$[L(\vec{p}), L(\vec{q})^\dagger] = p^\mu q^\nu [a_\mu(\vec{p}), a_\nu(\vec{q})] = -p \cdot q\, 2p^0 (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) = 0 \,, \tag{7.37}$$

where we used $p^2 = 0$, we have that $L(\vec{p})$ can be commuted with all the $L(\vec{p_j}\,)^\dagger$ to reach the vacuum state vector and annihilate it. But the states $L(\vec{p}_1)^\dagger \ldots L(\vec{p}_n)^\dagger |0\rangle$ are precisely the basis vectors containing only temporal and longitudinal photons that satisfy the condition Eq. (7.33). The states $|\phi_n\rangle$ are all orthogonal to each other, and have zero norm for $n \neq 0$:

$$\langle\phi'_{n'}|\phi_n\rangle = \int d\Omega_{p_1} \ldots \int d\Omega_{p_n} \int d\Omega_{q_1} \ldots \int d\Omega_{q_{n'}}\, C^{(n)}(\vec{p}_1, \ldots, \vec{p}_n\,) C'^{(n')}(\vec{q}_1, \ldots, \vec{q}_{n'}\,)$$
$$\times \langle 0|L(\vec{q}_{n'}) \ldots L(\vec{q}_1) L(\vec{p}_1)^\dagger \ldots L(\vec{p}_n)^\dagger |0\rangle = \delta_{n'0}\delta_{n0} \,,$$
$$(7.38)$$

where we have used Eq. (7.37). Up to an irrelevant rescaling, there are only two interesting choices for $C^{(0)}$, namely $C^{(0)} = 1$ or $C^{(0)} = 0$. Consider first the case $C^{(0)} = 1$. We find that

$$\langle\Psi|\Psi\rangle = \langle\psi_T|\psi_T\rangle \,, \tag{7.39}$$

and more generally for two vectors of the type in Eq. (7.36),

$$\langle \Psi' | \Psi \rangle = \langle \psi_T' | \psi_T \rangle \,, \tag{7.40}$$

independently of the functions $C^{(n)}$ for $n \geq 1$, i.e., independently of the content in unphysical photons. On the other hand, if we take $C^{(0)} = 0$ then $\langle \Psi | \Psi \rangle = 0$: zero-norm states are still present in $\mathcal{H}_{\text{phys}}$. However, any of the vectors in Eq. (7.36) represents the same physical state if $|\psi_T\rangle$ is the same, the difference between them being a zero-norm state. Even more generally, any vector of the form

$$|\Psi_Z\rangle = |\psi_T\rangle |0\rangle + \sum_{n=1}^{\infty} |\chi_{T\,n}\rangle |\phi_n\rangle = |\Psi\rangle + |Z\rangle \,, \tag{7.41}$$

for arbitrary choice of the transverse photon state vectors $|\chi_{T\,n}\rangle$, represents the same physical state. In fact, they all have the same norm, since $|Z\rangle$ is orthogonal to $\Psi\rangle$ and has zero norm, and moreover the scalar product between two such vectors is independent of the zero-norm part,

$$\langle \Psi'_{Z'} | \Psi_Z \rangle = (\langle \Psi' | + \langle Z' |)(|\Psi\rangle + |Z\rangle) = \langle \Psi' | \Psi \rangle \,. \tag{7.42}$$

We can then define an equivalence relation $\sim$ by

$$|\Psi'\rangle \sim |\Psi\rangle \quad \text{if} \quad |\Psi'\rangle = |\Psi\rangle + |Z\rangle \tag{7.43}$$

for some zero-norm vector $|Z\rangle$ in $\mathcal{H}_{\text{phys}}$. The true physical space is then the quotient space $\mathcal{H}_{\text{phys}}/\sim$, i.e., the space formed by the equivalence classes with respect to $\sim$, and with the scalar product inherited from $\mathcal{H}$, which can be consistently computed using any pair of representatives of the equivalence classes, since the result is always the same.

There is more to physics than state vectors though: observable quantities are represented as linear Hermitian operators $\mathcal{O} = \mathcal{O}^\dagger$ on the Hilbert space of states, and they have to be built out of fields. They are therefore defined first of all on the whole Fock space, and then restricted to $\mathcal{H}_{\text{phys}}$. For the final construction of the true physical Hilbert space to be acceptable, we need that the matrix elements of physical observables be independent of the representative vector that one uses for their evaluation: this identifies which observables are good observables. This can only happens if their matrix elements between zero-norm states, and between positive-norm and zero-norm states all vanish. This is equivalent to ask that the matrix elements of the commutators $[L(\vec{p}), \mathcal{O}]$ and $[L(\vec{p})^\dagger, \mathcal{O}]$ vanish between physical states. A sufficient condition for this to happen is

$$[L(\vec{p}), \mathcal{O}] = ML(\vec{p}) \,, \qquad [L(\vec{p})^\dagger, \mathcal{O}] = -L(\vec{p})^\dagger M^\dagger \,, \tag{7.44}$$

for some operator M. This is the case for observables of the form

$$\mathcal{O} = \int d\Omega_p \, O(\vec{p}) N(\vec{p}) \,, \tag{7.45}$$

where

$$N(\vec{p}) \equiv -\eta^{\mu\nu} a_\mu(\vec{p}) a_\nu(\vec{p})^\dagger = -a_0(\vec{p}) a_0(\vec{p})^\dagger + \sum_{j=1}^{3} a_j(\vec{p}) a_j(\vec{p})^\dagger \,, \tag{7.46}$$

134

which equals

$$N(\vec{p}) = -\eta^{\mu\nu} \sum_{\lambda,\lambda'} \varepsilon_\mu^{(\lambda)}(\vec{p}) a_{(\lambda)}(\vec{p}) \varepsilon_\nu^{(\lambda')}(\vec{p})^* a_{(\lambda')}(\vec{p})^\dagger = -\eta_{\lambda,\lambda'} a_{(\lambda)}(\vec{p}) a_{(\lambda')}(\vec{p})^\dagger$$

$$= -a_{(0)}(\vec{p}) a_{(0)}(\vec{p})^\dagger + \sum_{j=1}^{3} a_{(j)}(\vec{p}) a_{(j)}(\vec{p})^\dagger \,. \tag{7.47}$$

Indeed, since

$$[L(\vec{p}), N(\vec{q})] = -p^\mu \eta^{\nu\rho}[a_\mu(\vec{p}), a_\nu(\vec{q}) a_\rho(\vec{q})^\dagger] = -p^\mu \eta^{\nu\rho} a_\nu(\vec{q})[a_\mu(\vec{p}), a_\rho(\vec{q})^\dagger]$$
$$= -p^\mu \eta^{\nu\rho} a_\nu(\vec{q})(-\eta_{\mu\rho}) 2q^0 \delta^{(3)}(\vec{p} - \vec{q}) = p^\nu a_\nu(\vec{p}) 2q^0 \delta^{(3)}(\vec{p} - \vec{q}) \tag{7.48}$$
$$= L(\vec{p}) 2q^0 \delta^{(3)}(\vec{p} - \vec{q}) \,,$$

we have

$$[L(\vec{p}), \mathcal{O}] = \int d\Omega_q\, O(\vec{q})[L(\vec{p}), N(\vec{q})] = L(\vec{p}) \int d\Omega_q\, O(\vec{q}) 2q^0 \delta^{(3)}(\vec{p} - \vec{q}) = L(\vec{p}) O(\vec{p}) \,. \tag{7.49}$$

For these operators the contribution of the unphysical polarisations drops from the expecation values: recalling Eq. (7.34), one finds that

$$\langle \Psi | \mathcal{O} | \Psi \rangle = \int d\Omega_p\, O(\vec{p}) \langle \Psi | \sum_{j=1}^{2} a_j(\vec{p}) a_j(\vec{p})^\dagger | \Psi \rangle \,. \tag{7.50}$$

The form Eq. (7.45) is actually that of the four-momentum operators. These can be obtained via Noether's theorem from the Lagrangian $\mathscr{L}_F$, which we write here explcitly:

$$\mathscr{L}_F = -\frac{1}{2} \partial_\mu A_\nu \partial^\mu A^\nu = -\frac{1}{2} \partial_\mu A_0 \partial^\mu A_0 + \frac{1}{2} \partial_\mu A_j \partial^\mu A^j \,. \tag{7.51}$$

This is the Lagrangian of four real scalar fields, one of which comes with the wrong sign, $\mathscr{L}_F = -\mathscr{L}_0 + \sum_{j=1}^{3} \mathscr{L}_j$. It follows that the Noether charges will be the corresponding sum of Noether charges for scalar fields, i.e. $Q = -Q_0 + \sum_{j=1}^{3} Q_j$. Then

$$P^\mu = \int d\Omega_p\, p^\mu N(\vec{p}) = \int d\Omega_p\, p^\mu \left[-a_0(\vec{p}) a_0(\vec{p})^\dagger + \sum_{j=1}^{3} a_j(\vec{p}) a_j(\vec{p})^\dagger \right]$$

$$= \int d\Omega_p\, p^\mu \left[-a_{(0)}(\vec{p}) a_{(0)}(\vec{p})^\dagger + \sum_{j=1}^{3} a_{(j)}(\vec{p}) a_{(j)}(\vec{p})^\dagger \right] \,. \tag{7.52}$$

One can show that also the Lorentz generators are of the "good" type, satisfying Eq. (7.44), so they are good observables as well. Any gauge-invariant functional of A_μ is also a good observable. In fact, one has

$$[L(\vec{p}), A_\mu(x)] = p^\nu \int d\Omega_q\, e^{iq\cdot x}[a_\nu(\vec{p}), a_\mu(\vec{q})^\dagger] = p^\nu \int d\Omega_q\, e^{iq\cdot x}(-\eta_{\nu\mu}) 2p^0 \delta^{(3)}(\vec{p} - \vec{q})$$
$$= -p_\mu e^{ip\cdot x} = \partial_\mu \left(i e^{ip\cdot x} \right) \equiv \delta_{L(\vec{p})} A_\mu(x) \,, \tag{7.53}$$

which is essentially a gauge transformation. For a functional $\mathcal{F}[A]$ one then finds

$$[L(\vec{p}), \mathcal{F}[A]] = \int d^4x \, \frac{\delta \mathcal{F}[A]}{\delta A_\mu(x)} \, \delta_{L(\vec{p})} A_\mu(x) = 0 \,, \tag{7.54}$$

since $\mathcal{F}$ is gauge-invariant.[60] Similarly, one shows that

$$[L(\vec{p})^\dagger, A_\mu(x)] == p_\mu e^{-ip\cdot x} = \partial_\mu \left(ie^{-ip\cdot x} \right) \equiv \delta_{L(\vec{p})^\dagger} A_\mu(x) \,, \tag{7.55}$$

and so

$$[L(\vec{p})^\dagger, \mathcal{F}[A]] = \int d^4x \, \frac{\delta \mathcal{F}[A]}{\delta A_\mu(x)} \, \delta_{L(\vec{p})^\dagger} A_\mu(x) = 0 \,. \tag{7.56}$$

Gauge invariance of an observable therefore implies that it commutes with all $L(\vec{p})$ and $L(\vec{p})^\dagger$.

A few final comments are now in order. Instead of the Lagrangians $\mathscr{L}'$ or $\mathscr{L}_F$ of Eq. (7.11), one could use

$$\mathscr{L}_\alpha \equiv -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\alpha}(\partial_\mu A^\mu)^2 \,, \tag{7.57}$$

which leads to the following EOM,

$$\partial_\mu F^{\mu\nu} + \frac{1}{\alpha} \partial^\nu (\partial_\mu A^\mu) = \Box A^\nu - \left(1 - \frac{1}{\alpha} \right) \partial^\nu (\partial_\mu A^\mu) = 0 \,. \tag{7.58}$$

Taking the divergence with respect to ν we find

$$\frac{1}{\alpha} \Box (\partial_\mu A^\mu) = 0 \,. \tag{7.59}$$

If we impose $\partial_\mu A_\mu = 0$ everywhere at $t = -\infty$, it will remain zero at all times, and the solution of Eq. (7.58) will be a solution of Maxwell's equations in Lorenz gauge. Quantisation can be done starting from this Lagrangian, solving the EOM and imposing $\partial_\mu A_\mu = 0$ on physical states in the same way we have done above. To see that the physics will be the same, it suffices to check that the symmetry generators (from which all the observables can be built out) change by terms that vanish when acting on physical states. We then have to check that the change δQ^a of a Noether charge due to the change $\delta \mathscr{L} = (\frac{1}{\alpha} - 1)(\partial_\mu A^\mu)^2$ of the Lagrangian does not affect physical states. We have

$$\begin{aligned}
\delta Q^a &= \int d^3x \left\{ \frac{\partial \delta \mathscr{L}}{\partial(\partial_0 A_\mu)} \left[\mathcal{M}^a_{\mu\nu} A^\nu + \mathcal{A}^{a\nu} \partial_\nu A_\mu \right] - \mathcal{A}^{a0} \delta \mathscr{L} \right\} \\
&= \left(1 - \frac{1}{\alpha} \right) \int d^3x \left\{ \eta^{\mu 0}(\partial_\beta A^\beta) \left[\mathcal{M}^a_{\mu\nu} A^\nu + \mathcal{A}^{a\nu} \partial_\nu A_\mu \right] - \frac{1}{2} \mathcal{A}^{a0}(\partial_\beta A^\beta)^2 \right\} \,,
\end{aligned} \tag{7.60}$$

and when imposing normal ordering we always have either $\partial_\beta A^\beta_+$ on the right or $\partial_\beta A^\beta_-$ on the left, so that there is no contribution of $: \delta Q^a :$ to matrix elements between physical states.

Let us work out the implications of the consequences Eqs. (7.54) and (7.56) of gauge invariance of an observable imply, in the case of a scalar normal-ordered observable. The most general

---

[60]Gauge-invariance means that $\mathcal{F}[A+\partial\Lambda] - \mathcal{F}[A] = 0$. For infinitesimal transformations $0 = \mathcal{F}[A+\epsilon\partial\Lambda] - \mathcal{F}[A] = \epsilon \int d^4x \, \frac{\delta \mathcal{F}[A]}{\delta A_\mu(x)} \, \partial_\mu \Lambda(x) + \mathcal{O}(\epsilon^2)$, so $\int d^4x \, \frac{\delta \mathcal{F}[A]}{\delta A_\mu(x)} \, \partial_\mu \Lambda(x) = 0$.

such observable is of the form

$$\mathcal{O} = \sum_{n,m} \mathcal{O}^{(n,m)} \,,$$

$$\mathcal{O}^{(n,m)} = \int d\Omega_{p_1} \ldots \int d\Omega_{p_n} \int d\Omega_{q_1} \ldots \int d\Omega_{q_m} \, O^{\mu_1 \ldots \mu_n \nu_1 \ldots \nu_m}(\vec{p}_1, \ldots, \vec{p}_n, \vec{q}_1, \ldots, \vec{q}_m) \qquad (7.61)$$

$$\times \, a_{\mu_1}(\vec{p}_1)^\dagger \ldots a_{\mu_n}(\vec{p}_n)^\dagger a_{\nu_1}(\vec{q}_1) \ldots a_{\nu_m}(\vec{q}_n) \,,$$

where Hermiticity imposes that

$$O^{\mu_1 \ldots \mu_n \nu_1 \ldots \nu_m}(\vec{p}_1, \ldots, \vec{p}_n, \vec{q}_1, \ldots, \vec{q}_m) = O^{\nu_1 \ldots \nu_m \mu_1 \ldots \mu_n}(\vec{q}_1, \ldots, \vec{q}_m, \vec{p}_1, \ldots, \vec{p}_n)^* \,, \qquad (7.62)$$

and Bose symmetry allows to take $O^{\mu_1 \ldots \mu_n \nu_1 \ldots \nu_m}(\vec{p}_1, \ldots, \vec{p}_n, \vec{q}_1, \ldots, \vec{q}_m)$ symmetric under exchange of $\mu_j, \vec{p}_j$ with $\mu_k, \vec{p}_k$, and similarly under exchange of $\nu_j, \vec{q}_j$ with $\nu_k, \vec{q}_k$. Imposing gauge invariance in the form $[L(\vec{k}), \mathcal{O}] = [L(\vec{k})^\dagger, \mathcal{O}] = 0$, we arrive at the conditions

$$p_{j\,\mu_j} O^{\mu_1 \ldots \mu_n \nu_1 \ldots \nu_m}(\vec{p}_1, \ldots, \vec{p}_n, \vec{q}_1, \ldots, \vec{q}_m) = q_{j\,\nu_j} O^{\mu_1 \ldots \mu_n \nu_1 \ldots \nu_m}(\vec{p}_1, \ldots, \vec{p}_n, \vec{q}_1, \ldots, \vec{q}_m) = 0 \,, \quad (7.63)$$

for all j. The coefficient functions should then be divergenceless in all indices.[61]

As a final comment, notice that in the presence of a nonzero current, Eqs. (7.57) and (7.58) are modified to

$$\mathscr{L}_\alpha = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\alpha}(\partial_\mu A^\mu)^2 - J^\mu A_\mu \,, \qquad (7.64)$$

and

$$\partial_\mu F^{\mu\nu} + \frac{1}{\alpha} \partial^\nu (\partial_\mu A^\mu) = \Box A^\nu - \left(1 - \frac{1}{\alpha}\right) \partial^\nu(\partial_\mu A^\mu) = J^\nu \,. \qquad (7.65)$$

Taking the divergence with respect to ν we find now

$$\frac{1}{\alpha} \Box (\partial_\mu A^\mu) = \partial_\nu J^\nu = 0 \,. \qquad (7.66)$$

The field $\partial_\mu A^\mu$ is thus a free field also in the presence of interaction with matter, if the coupling is with a conserved current. This means that also in the interacting case we can identify the positive and negative-frequency components of $\partial_\mu A^\mu$. We can proceed á la Gupta-Bleuler also in the interacting case, quantising canonically the theory in Lorenz gauge and then imposing the subsidiary condition $(\partial_\mu A^\mu)_+ |\Psi\rangle = 0$ to select physical states out of the full Hilbert space.

7.2 Photon propagator

We now compute the photon field contraction, or photon propagator. This quantity depends on the particular gauge used in the quantisation procedure: it is therefore crucial to show that physical quantities are independent of the choice of gauge. Having quantised the theory á la Gupta-Bleuler, the propoagator that we will obtain is the one in Lorenz gauge. The calculation

---

[61]The condition $p_\mu \tilde{O}(p) = 0$ is the Fourier transform of the divergenceless condition $\partial_\mu O(x) = 0$, hence the nomenclature.

is rather easy, and as one can expect from the discussion of the previous subsection, the result amounts essentially to four scalar propagators, one of which with the wrong sign. We have

$$
\begin{aligned}
D_{\mu\nu}(x - y) &\equiv \langle 0|T\left(A_\mu(x)A_\nu(y)\right)|0\rangle \\
&= \int d\Omega_p \int d\Omega_q \Big\{ \theta(x^0 - y^0)e^{-i(p\cdot x - q\cdot y)}\langle 0|a_\mu(\vec{p})a_\nu(\vec{q})^\dagger|0\rangle \\
&\qquad\qquad + \theta(y^0 - x^0)e^{-i(p\cdot y - q\cdot x)}\langle 0|a_\nu(\vec{p})a_\mu(\vec{q})^\dagger|0\rangle \Big\} \\
&= -\eta_{\mu\nu}\int d\Omega_p \left[\theta(x^0 - y^0)e^{-ip\cdot(x-y)} + \theta(y^0 - x^0)e^{ip\cdot(x-y)} \right] = -\eta_{\mu\nu}D(x - y).
\end{aligned}
$$
(7.67)

In momentum space,

$$
D_{\mu\nu}(x - y) = \frac{d^4p}{(2\pi)^4}\, e^{-ip\cdot x}\, \tilde{D}_{\mu\nu}(p),
$$
(7.68)

we find

$$
\tilde{D}_{\mu\nu}(p) = \frac{-i\eta_{\mu\nu}}{p^2 + i\epsilon}.
$$
(7.69)

The momentum-space expression follows (up to the prescription at the poles) from the fact that $D_{\mu\nu}$ solves the inhomogeneous equations of motion. In fact, since

$$
\begin{aligned}
\partial_{x\,0}^2 D_{\mu\nu}(x - y) &= \partial_{x\,0}\left\{\langle 0|T\left(\partial_{x\,0}A_\mu(x)A_\nu(y)\right)|0\rangle + \delta(x^0 - y^0)\langle 0|[A_\mu(x), A_\nu(y)]_{\mathrm{ET}}|0\rangle\right\} \\
&= \langle 0|T\left(\partial_{x\,0}^2 A_\mu(x)A_\nu(y)\right)|0\rangle + \delta(x^0 - y^0)\langle 0|[\partial_{x\,0}A_\mu(x), A_\nu(y)]_{\mathrm{ET}}|0\rangle \\
&= \langle 0|T\left(\partial_{x\,0}^2 A_\mu(x)A_\nu(y)\right)|0\rangle - \delta(x^0 - y^0)\langle 0|[\pi_\mu(x), A_\nu(y)]_{\mathrm{ET}}|0\rangle \\
&= \langle 0|T\left(\partial_{x\,0}^2 A_\mu(x)A_\nu(y)\right)|0\rangle + i\eta_{\mu\nu}\delta^{(4)}(x - y),
\end{aligned}
$$
(7.70)

we find that

$$
\begin{aligned}
\Box_x D_{\mu\nu}(x - y) &= \langle 0|T\left(\partial_{x\,0}^2 A_\mu(x)A_\nu(y)\right)|0\rangle + i\eta_{\mu\nu}\delta^{(4)}(x - y) - \langle 0|T\left(\Delta A_\mu(x)A_\nu(y)\right) \\
&= \langle 0|T\left(\Box A_\mu(x)A_\nu(y)\right)|0\rangle + i\eta_{\mu\nu}\delta^{(4)}(x - y) = i\eta_{\mu\nu}\delta^{(4)}(x - y),
\end{aligned}
$$
(7.71)

which in momentum space reads $-p^2\tilde{D}_{\mu\nu} = i\eta_{\mu\nu}$.

One can show that when quantising the photon field making use of one of the Lagrangians $\mathscr{L}_\alpha$ of Eq. (7.57), the photon propagator is found to solve the following equation,

$$
\left[\Box\eta_{\mu\rho} - \left(1 - \frac{1}{\alpha}\right)\partial_\mu\partial_\rho\right]D^\rho{}_\nu(x) = i\eta_{\mu\nu}\delta^{(4)}(x).
$$
(7.72)

In momentum space, this reads

$$
\left[p^2\eta_{\mu\rho} - \left(1 - \frac{1}{\alpha}\right)p_\mu p_\rho\right]\tilde{D}^\rho{}_\nu(p) = -i\eta_{\mu\nu}.
$$
(7.73)

The propagator is then obtained by inverting the matrix $M_{\mu\nu}^{(\alpha^{-1})}(p) = p^2\eta_{\mu\rho} - \left(1 - \frac{1}{\alpha}\right)p_\mu p_\rho$. Notice that in the absence of the extra term in the Lagrangian (corresponding to taking $\alpha \to \infty$), this matrix is not invertible: this reflects the gauge invariance of the action in that case, which leads to the presence of a zero mode for $M_{\mu\nu}^{(0)}(p)$, namely the longitudinal mode, $M_{\mu\nu}^{(0)}(p)p^\nu = 0$.

To invert $M_{\mu\nu}^{(\alpha^{-1})}(p)$ we can make use of this fact, and notice that $M_{\mu\nu}^{(\alpha^{-1})}(p) = M_{\mu\nu}^{(0)}(p) + \frac{1}{\alpha}p_\mu p_\nu$. The two terms are nonzero in orthogonal subspaces, and actually proportional to the corresponding projectors $\Pi_\perp$ and $\Pi_\parallel$. We can write

$$M_{\mu\nu}^{(\alpha^{-1})}(p) = p^2\left(\eta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}\right) + \frac{p^2}{\alpha}\frac{p_\mu p_\nu}{p^2} = p^2\Pi_\perp + \frac{p^2}{\alpha}\Pi_\parallel,\tag{7.74}$$

and obtain the propagator right away as

$$\tilde{D}_{\mu\nu}(p) = -i\left(\frac{1}{p^2}\Pi_\perp + \frac{\alpha}{p^2}\Pi_\parallel\right) = \frac{-i}{p^2}\left(\eta_{\mu\nu} - (1-\alpha)\frac{p_\mu p_\nu}{p^2}\right)\tag{7.75}$$

A detailed calculation allows to determine the correct prescription to obtain the field contraction. The momentum propagator in the gauge α reads

$$\tilde{D}_{\mu\nu}(p) = \frac{-i}{p^2 + i\epsilon}\left(\eta_{\mu\nu} - (1-\alpha)\frac{p_\mu p_\nu}{p^2 + i\epsilon}\right).\tag{7.76}$$

It is common to refer to the various choices of α as different gaauge choice, although it is always the Lorenz gauge that is being used. The choice $\alpha = 1$ is called *Feynman* (or *Fermi*) *gauge*. The choice $\alpha \to 0$ is called *Landau gauge*: notice that here the limit is taken after quantising the theory. Another choice of α with a name attached is Yennie gauge, $\alpha = 3$. Of course, physics must not depend on α, so we must ensure that wherever the photon propagator gets inserted, the longitudinal term gives no contribution.

8 Quantum Electrodynamics

We now have all the tools to formulate a relativistic quantum theory of the electromagnetic interaction. At the classical level, if we couple matter to the electromagnetic field A_μ via a current J^μ,

$$\mathscr{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - J_\mu A^\mu,\tag{8.1}$$

then the current must be conserved. This can be reinterpreted geometrically as requiring that gauge invariance be preserved also in the interacting case. In fact, on the one side, conservation of the current implies gauge invariance of the action, in the sense that this chages by the integral of a total divergence, which does not affect the EOM. On the other side, gauge invariance of the action (again, up to irrelevant boundary terms) implies that

$$\delta S = -\int d^4x\, J^\mu \partial_\mu \Lambda = -\int d^4x\, \partial_\mu(J^\mu\Lambda) - \int d^4x\, (\partial_\mu J^\mu)\Lambda \Rightarrow \partial_\mu J^\mu = 0,\tag{8.2}$$

since this must hold for arbitrary Λ.

Let us consider spinorial electrodynamics, in which spin-$\frac{1}{2}$ get coupled to the electromagnetic field. We know that from the Dirac Lagrangian,

$$\mathscr{L}_D = \bar{\psi}(i\slashed{\partial} - m)\psi,\tag{8.3}$$

we get a conserved current J^μ,

$$J^\mu = \bar{\psi}\gamma^\mu\psi,\tag{8.4}$$

with the associated conserved charge

$$Q = \int d^3x\, J^0 = \int d^3x\, :\psi^\dagger\psi := \int d\Omega_p \sum_s \left\{ b_s(\vec{p})^\dagger b_s(\vec{p}) + d_s(\vec{p})^\dagger d_s(\vec{p}) \right\}. \tag{8.5}$$

This current is associated to invariance under the U(1) transformation

$$\psi \to e^{i\alpha}\psi, \qquad \bar{\psi} \to \bar{\psi}e^{-i\alpha}. \tag{8.6}$$

If we couple A_μ to J^μ with some coupling constant e, the resulting Lagrangian,

$$\begin{aligned}
\mathscr{L}_{\text{QED}} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\slashed{\partial} - m)\psi - e\bar{\psi}\gamma^\mu\psi A_\mu \\
&= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\slashed{\partial} - m)\psi - e\bar{\psi}\slashed{A}\psi \\
&= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}[i\gamma^\mu(\partial_\mu + ieA_\mu) - m]\psi,
\end{aligned} \tag{8.7}$$

is still invariant under the U(1) transformation Eq. (8.6), so J^μ is conserved and gauge invariance still holds. Here

$$J^\mu = e\bar{\psi}\gamma^\mu\psi, \tag{8.8}$$

and the associated charge is

$$Q = e \int d^3x\, \psi^\dagger\psi. \tag{8.9}$$

The coupling constant e is the *electric charge* of the b-type particles.[62] The U(1) transformation Eq. (8.6) wiht x-independent α is called a *global gauge transformation*, for reasons that will become clear in a moment. If we make a gauge transformation $A_\mu \to A_\mu - \partial_\mu\Lambda$, the Lagrangian changes by the term

$$\delta_A\mathscr{L}_{\text{QED}} = e\bar{\psi}\gamma^\mu\psi\partial_\mu\Lambda = -e\bar{\psi}e^{-ie\Lambda}i\slashed{\partial}(e^{ie\Lambda})\psi. \tag{8.10}$$

This can be compensated by changing also $\psi \to e^{ie\Lambda}\psi$, $\bar{\psi} \to \bar{\psi}e^{-ie\Lambda}$, where Λ is now x-dependent, since in this case

$$\begin{aligned}
\delta_\psi\mathscr{L}_{\text{QED}} &= e\bar{\psi}e^{-ie\Lambda}i\slashed{\partial}(e^{ie\Lambda}\psi) - e\bar{\psi}i\slashed{\partial}\psi \\
&= e\bar{\psi}e^{-ie\Lambda}i\slashed{\partial}(e^{ie\Lambda})\psi + e\bar{\psi}e^{-ie\Lambda}e^{ie\Lambda}i\slashed{\partial}\psi - e\bar{\psi}\gamma^\mu i\slashed{\partial}\psi \\
&= e\bar{\psi}e^{-ie\Lambda}i\slashed{\partial}(e^{ie\Lambda})\psi = -\delta_A\mathscr{L}_{\text{QED}}.
\end{aligned} \tag{8.11}$$

This is a *local gauge transformation*, since the U(1) rotation performed on the Dirac field depends on the spacetime point. The Lagrangian is then invariant under the combined guage transformation

$$\begin{aligned}
A_\mu &\to A_\mu - \partial_\mu\Lambda, \\
\psi &\to e^{ie\Lambda}\psi, \\
\bar{\psi} &\to \bar{\psi}e^{-ie\Lambda},
\end{aligned} \tag{8.12}$$

---

[62]Here we are anticipating the passage to the interaction picture, where Q will be written in terms of free fields for which the particle interpretation is clear.

for an arbitrary function $\Lambda(x)$. Conversely, asking that the Lagrangian be invariant under such a transformation forces us to couple the electromagnetic and the Dirac fields as in Eq. (8.7), if we want to use derivatives of the fields. To see this, let us denote with D_μ the *covariant derivative*,

$$D_\mu \equiv \partial_\mu + ieA_\mu \,, \tag{8.13}$$

so that we can write

$$\mathscr{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\slashed{D} - m)\psi \,. \tag{8.14}$$

The covariant derivative of ψ transforms like ψ itself under a gauge tranformation:

$$D_\mu\psi \to (\partial_\mu + ieA_\mu - ie\partial_\mu\Lambda)e^{ie\Lambda}\psi = e^{ie\Lambda}(\partial_\mu + ieA_\mu)\psi = e^{ie\Lambda}D_\mu\psi \,. \tag{8.15}$$

To obtain a Lorentz-invariant and gauge-invariant action we are then force to use scalar combinations of $F_{\mu\nu}$, ψ and $D_\mu\psi$. The simplest possibility is precisely Eq. (8.14). This explains the origin of the so-called *minimal coupling* or *minimal substitution* in terms of gauge invariance.

Let us summarise our discussion so far. We have argued that a good Lagrangian for the description of spin-$\frac{1}{2}$ particles coupled to the electromagnetic field is the one given in Eq. (8.14). This can be motivated by the request of invariance under Lorentz transformations and under the local gauge transformation Eq. (8.12). The theory obtained starting from this Lagrangian has the Poincaré group and the U(1) group as its global symmetries, to which correspond conserved currents and charges according to Noether's theorem. The latter are the four-momentum generators P_μ, the Lorentz generators $J^{(\rho\sigma)}$ and the electric charge Q given in Eq. (8.9).

The EOM are obtained as usual via an action principle, and read

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= e\bar{\psi}\gamma^\nu\psi \,, \\ (i\slashed{D} - m)\psi &= 0 \,. \end{aligned} \tag{8.16}$$

In particular, the second equation and its adjoint imply by themselves the conservation of the *electric current* $J^\mu = e\bar{\psi}\gamma^\mu\psi$. Indeed, taking the adjoint of this equation in the form $i\slashed{\partial}\psi = (m + \slashed{A})\psi$, we find $(\partial_\mu\bar{\psi})\gamma = -\bar{\psi}(m + \slashed{A})$, so that

$$\partial_\mu(e\bar{\psi}\gamma^\mu\psi) = e[(\partial_\mu\bar{\psi})\gamma^\mu\psi + \bar{\psi}\gamma^\mu\partial_\mu\psi] = e\bar{\psi}[-(m + \slashed{A}) + (m + \slashed{A})]\psi = 0 \,. \tag{8.17}$$

Our next task is to quantise this theory. We already know that gauge invariance does not allow us to proceed directly via the canonical method, so we try to proceed again á la Gupta-Bleuler. We then modify the Lagrangian to

$$\mathscr{L}_{\text{QED|GB}} = -\frac{1}{4}\partial_\mu A_\nu \partial^\mu A^\nu + \bar{\psi}(i\slashed{D} - m)\psi \,, \tag{8.18}$$

having added the gauge non-invariant term $-\frac{1}{2}(\partial_\mu A^\mu)^2$ to the original Lagrangian (and having dropped irrelevant total derivative terms). We then proceed to solve the modified EOM

$$\begin{aligned} \Box A^\nu &= e\bar{\psi}\gamma^\nu\psi \,, \\ (i\slashed{D} - m)\psi &= 0 \,, \end{aligned} \tag{8.19}$$

after which we will impose the subsidiary condition on the Hilbert space,

$$(\partial_\mu A^\mu)_+(x)|\Psi\rangle = 0 \,. \tag{8.20}$$

As discussed at the end of Section 7.1, $\partial_\mu A^\mu$ is a free field also in the interacting case, so that it can be separated into positive and negative frequency parts. The Gupta-Bleuler procedure requires that we impose the usual CCR/CAR on the fields and their conjugate momenta,

$$\pi_A^\mu = \frac{\partial \mathscr{L}_{\mathrm{QED|GB}}}{\partial_0 A^\mu} = -\partial^0 A^\mu \,, \qquad \pi_\psi = \frac{\partial \mathscr{L}_{\mathrm{QED|GB}}}{\partial_0 \psi} = i\psi^\dagger \,, \tag{8.21}$$

and after that select physical states in the Hilbert space by imposing Eq. (8.20).

The procedure outlined above cannot be fully carried out in practice, so we have to resort to some approximation technique. We will then go over to the interaction picture and compute physical quantities by means of perturbastion theory. Since the interaction part of the Lagrangian does not depend on time derivatives of the field, we have that

$$V_I = -\int d^3x \, \mathscr{L}_I(\psi_{\mathrm{int}}, \bar\psi_{\mathrm{int}}, A_{\mathrm{int}\,\mu}) \,, \tag{8.22}$$
$$\mathscr{L}_I(\psi_{\mathrm{int}}, \bar\psi_{\mathrm{int}}, A_{\mathrm{int}\,\mu}) = -e\bar\psi_{\mathrm{int}} \slashed{A}_{\mathrm{int}} \bar\psi_{\mathrm{int}} \,.$$

From this we can develop perturbation theory as discussed in Section 6.

Before discussing in detail perturbation theory, we still have to clarify what happens to the subsidiary condition when going over to the interaction picture. Since $\partial \cdot A = \partial_\mu A^\mu$ is a Hermitian free field, $\Box \partial \cdot A = 0$, we can write it as

$$\partial \cdot A(x) = \int d\Omega_p \left\{ e^{-ip\cdot x} c(\vec{p}) + e^{ip\cdot x} c(\vec{p})^\dagger \right\} \,, \tag{8.23}$$

for some operators $c(\vec{p})$. All we need to know about them is that they can be extracted from $\partial \cdot A$ as usual,

$$c(\vec{p}) = \int d^3x \, e^{ip\cdot x} i\overleftrightarrow{\partial_0} \partial \cdot A(x) \,, \tag{8.24}$$

for arbitrary x^0. Since $A_\mu(x) = e^{iP\cdot x} A_\mu(0) e^{-iP\cdot x}$, thanks to translation invariance of the theory, we also have

$$\partial \cdot A(x) = \partial_\mu \left(e^{iP\cdot x} A^\mu(0) e^{-iP\cdot x} \right) = [iP_\mu, A^\mu(x)] = e^{iP\cdot x} [iP_\mu, A^\mu(0)] e^{-iP\cdot x}$$
$$= e^{iP\cdot x} \partial \cdot A(0) e^{-iP\cdot x} \,. \tag{8.25}$$

This implies

$$e^{iP\cdot a} c(\vec{p}) e^{-iP\cdot a} = \int d^3x \, e^{ip\cdot x} i\overleftrightarrow{\partial_0} \partial \cdot A(x+a) = e^{-ip\cdot a} \int d^3x \, e^{ip\cdot x} i\overleftrightarrow{\partial_0} \partial \cdot A(x) = e^{-ip\cdot a} c(\vec{p}) \,, \tag{8.26}$$

where we have made a change of variables in $\vec{x}$, and we have used the arbitrariness in the choice of the time variable in Eq. (8.24). We conclude that

$$[\partial \cdot A(x)]_+ = \int d\Omega_p \, e^{-ip\cdot x} c(\vec{p}) = e^{iP\cdot x} \left[\int d\Omega_p \, c(\vec{p}) \right] e^{-iP\cdot x} = e^{iP\cdot x} [\partial \cdot A(0)]_+ e^{-iP\cdot x} \,. \tag{8.27}$$

On the other hand, at $t = 0$ the fields and momenta in the Heisenberg and interaction pictures coincide,

$$A_{\mathrm{int}\,\mu}(0, \vec{x}) = A_\mu(0, \vec{x}) \,, \qquad \pi_{\mathrm{int}\,\mu}(0, \vec{x}) = -\partial_0 A_{\mathrm{int}\,\mu}(0, \vec{x}) \pi_\mu(0, \vec{x}) = -\partial_0 A_\mu(0, \vec{x}) \,, \tag{8.28}$$

142

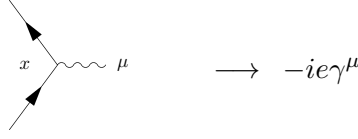$$x \;\rangle\!\!\!\sim\!\!\sim\!\!\sim\, \mu \qquad\qquad \longrightarrow \quad -ie\gamma^\mu$$

Figure 9: Interaction vertex in QED.

from which it follows that $[\partial \cdot A(0,\vec{x})]_+ = [\partial \cdot A_{\mathrm{int}}(0,\vec{x})]_+ = \partial \cdot A_{\mathrm{int}+}(0,\vec{x})$, where in the last passage we used our knowledge of free fields to write the positive-frequency part of the divergence of the free photon field as the diverge of the positive-frequency part of the field (this is not true for the interacting field). We now use the fact that the subsidiary condition Eq. (8.20) holds for all x. Since Eq. (8.20) implies also that $e^{iH_0 t}(\partial_\mu A^\mu)_+(0,\vec{x})e^{-iHt}|\Psi\rangle = 0$, for all t, we can write

$$
\begin{aligned}
0 &= \lim_{t\to-\infty} e^{iH_0 t}\int d^3x\, e^{ip\cdot x}[\partial\cdot A(0,\vec{x})]_+ e^{-iHt}|\Psi\rangle \\
&= \lim_{t\to-\infty} e^{iH_0 t}\int d^3x\, e^{ip\cdot x}\partial\cdot A_{\mathrm{int}+}(0,\vec{x})e^{-iHt}|\Psi\rangle \\
&= \lim_{t\to-\infty} e^{iH_0 t}\int d^3x\, e^{ip\cdot x}\partial\cdot A_{\mathrm{int}+}(0,\vec{x})e^{-iH_0 t}|\Psi_0\rangle \qquad\qquad (8.29)\\
&= \lim_{t\to-\infty} \int d^3x\, e^{ip\cdot x}\partial\cdot A_{\mathrm{int}+}(t,\vec{x})|\Psi_0\rangle \\
&= \int d^3x\, e^{ip\cdot x}\partial\cdot A_{\mathrm{int}+}(t,\vec{x})|\Psi_0\rangle = -ip^\mu a_\mu(\vec{p})|\Psi_0\rangle = -iL(\vec{p})|\Psi_0\rangle\,,
\end{aligned}
$$

where $|\Psi_0\rangle$ is defined as the free state leading to the in state $|\Psi\rangle$, $|\Psi\rangle = \Omega_+|\Psi_0\rangle$, and $a_\mu(\vec{p})$ are the annihilation operators associated to the free photon field in Lorenz gauge. We thus recover the Gupta-Bleuler subsidiary condition for free particle states. The same argument can be made for out states. We can then build the free physical Hilbert space as described in the previous Section, and then obtain the interacting physical Hilbert space through the scattering operators $\Omega_\pm$. Being unitary operators (or at least isometric), we are guaranteed that all negative and zero norm states are eliminated from the theory.

At this point we construct the S operator in the usual way, but we have to make sure that it has no matrix elements between physical and unphysical states, so that the identification of state vectors differing by zero-norm vectors does not affect the physics. We will show later that S is actually gauge invariant, in the sense that $[L(\vec{p}), S] = [L(\vec{p})^\dagger, S] = 0$, so that we end up having a well-defined unitary scattering matrix acting on a positive-norm Hilbert space. Since Lorentz invariance is manifest at all stages, we are guaranteed that the S matrix will also be Lorentz invariant. All that is left to do (it may seem...) is to actually compute the S-matrix. We now outline how this is done.

8.1 Feynman rules for spinorial electrodynamics

We now give the extra rules required to compute the perturbative series in QED. In general, the construction of the perturbative series proceeds as discussed in Section 6, expanding Dyson formula in powers of the coupling (here the electric charge e), and dealing with the time-ordered products of fields by means of Wick's theorem. All we need to add is how the interaction vertex
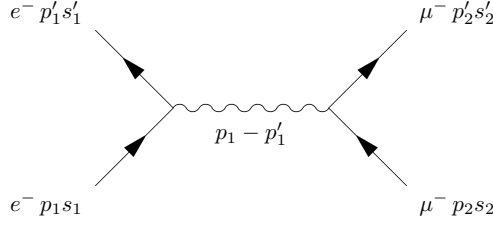
Figure 10: Lowest-order diagram for elastic $e^-\mu^-$ scattering.

behaves in the theory at hand, and how to deal with photon fields, both when they appear in contractions and when they act on the incoming and outgoing particles.

- The interaction vertex is represented by two fermionic lines, one ingoing and one outgoing, and a wiggly line corresponding to the photon field (see Fig. 9). The corresponding factor reads $-ie\gamma^\mu$, where the Dirac indices are contracted as usual by going against the flow of the fermion lines, while the Lorentz index μ is contracted with the contribution corresponding to the photon field.

- A photon field acting on an incoming or outgoing photon gives rise to an external photon line in the diagram. Photon states are denoted by $|\vec{p}\,\lambda\rangle$, where $\lambda = 1, 2$ is one of the two physical polarisations. To determine the contribution of an external line, it suffices to evaluate the following matrix elements,

$$\langle 0|A_\mu(x)|\vec{p}\,\lambda\rangle = \int d\Omega_q \sum_{\lambda'} e^{-iq\cdot x}\varepsilon_\mu^{(\lambda')}(\vec{q})\langle 0|a_{(\lambda')}(\vec{q})|\vec{p}\,\lambda\rangle = e^{-ip\cdot x}\varepsilon_\mu^{(\lambda)}(\vec{p})\,,$$
$$\langle \vec{p}\,\lambda|A_\mu(x)|0\rangle = \langle 0|A_\mu(x)|\vec{p}\,\lambda\rangle^* = e^{ip\cdot x}\varepsilon_\mu^{(\lambda)}(\vec{p})^*\,. \tag{8.30}$$

We have then the following rules: in coordinate space, associate a factor $e^{-ip\cdot x}\varepsilon_\mu^{(\lambda)}(\vec{p})$ to an incoming particle attached to the vertex x, and a factor factor $e^{+ip\cdot x}\varepsilon_\mu^{(\lambda)}(\vec{p})^*$ to an outogoing particle attached to the vertex x; in momentum space, drop the phase factors.

- The only nontrivial contraction for the photon field is $D_{\mu\nu}(x-y) = \langle 0|T(A_\mu(x)A_\nu(y))|0\rangle$. In a Feynman diagram, such a contraction corresponds to an internal line running from vertex y to vertex x,[63] and the corresponding factor $D_{\mu\nu}(x-y)$ must be included in coordinate space. In momentum space, a momentum p is associated to each internal line, and a factor $\tilde{D}_{\mu\nu}(p)$ must be included.

Let us apply these rules to the calculation of the simplest process, namely the elastic scattering of two different fermions. To fix the ideas, let them be an electron and a muon, both of charge $e = -|e|$. In this case there is a single Feynman diagram contributing to lowest order, the one shown in Fig. There are no special symmetries in this diagram, so the usual degeneracy counting applies, and we find

$$\mathcal{M}_{fi}^{(e\mu\to e\mu)} = \frac{1}{i}\bar{u}_{s_1'}^{(e)}(\vec{p}_1')(-ie)\gamma^\mu u_{s_1}^{(e)}(\vec{p}_1)\frac{(-i)\eta_{\mu\nu}}{(p_1-p_1')^2+i\epsilon}\bar{u}_{s_2'}^{(\mu)}(\vec{p}_2')(-ie)\gamma^\nu u_{s_2}^{(\mu)}(\vec{p}_2)$$
$$= \frac{e^2}{t}\bar{u}_{s_1'}^{(e)}(\vec{p}_1')\gamma^\mu u_{s_1}^{(e)}(\vec{p}_1)\bar{u}_{s_2'}^{(\mu)}(\vec{p}_2')\gamma_\mu u_{s_2}^{(\mu)}(\vec{p}_2)\,, \tag{8.31}$$

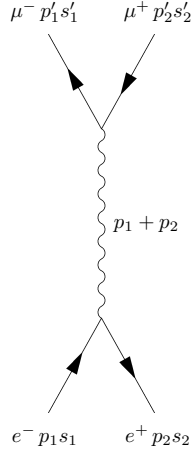[63]Photon lines are usually drawn without an orientation, since the photon is a neutral particle.

Figure 11: Lowest-order diagram for $e^- e^+ \to \mu^- \mu^+$ scattering.

where $t = (p_1 - p_1')^2$ is the usual Mandelstam variable. We can verify explicitly that the same result is obtained in any other α gauge. In fact, replacing the Feynman gauge propagator with the one for $\alpha \neq 1$ we find the extra contribution

$$\frac{1}{i}\bar{u}_{s_1'}^{(e)}(\vec{p}_1')(-ie)\gamma^\mu u_{s_1}^{(e)}(\vec{p}_1)\frac{i(1-\alpha)(p_1 - p_1')_\mu (p_1 - p_1')_\nu}{[(p_1 - p_1')^2 + i\epsilon]^2}\bar{u}_{s_2'}^{(\mu)}(\vec{p}_2')(-ie)\gamma^\nu u_{s_2}^{(\mu)}(\vec{p}_2), \qquad (8.32)$$

which vanishes since

$$\begin{aligned}
\bar{u}_{s_1'}^{(e)}(\vec{p}_1')\gamma^\mu u_{s_1}^{(e)}(\vec{p}_1)(p_1 - p_1')_\mu &= \bar{u}_{s_1'}^{(e)}(\vec{p}_1')(\slashed{p}_1 - \slashed{p}_1')u_{s_1}^{(e)}(\vec{p}_1) \\
&= \bar{u}_{s_1'}^{(e)}(\vec{p}_1')(m - m)u_{s_1}^{(e)}(\vec{p}_1) = 0,
\end{aligned} \qquad (8.33)$$

the other factor is seen to vanish as well since $p_1 - p_1' = p_2' - p_2$ due to momentum conservation. This reflects a general result: due to current conservation, the factor $O^{\mu_1 \cdots \mu_n}$ that in a Feynman diagram is contracted with internal and external photon lines satisfies $q_{j\,\mu_j}O^{\mu_1 \cdots \mu_n} = 0$ for all j, where q_j is the momentum flowing in the j internal or externl photon line. These are the so-called *Ward identities*, that guarantee the gauge invariance of the S-matrix.

Another simple process is the inelastic scattering process $e^- e^+ \to \mu^- \mu^+$. Also in this case one finds a single Feynman diagram to lowest order, which is evaluated to be

$$\begin{aligned}
\mathcal{M}_{fi}^{(e^- e^+ \to \mu^- \mu^+)} &= \frac{1}{i}\bar{u}_{s_1'}^{(e)}(\vec{p}_1')(-ie)\gamma^\mu v_{s_2'}^{(\mu)}(\vec{p}_2')\frac{(-i)\eta_{\mu\nu}}{(p_1 + p_2)^2 + i\epsilon}\bar{v}_{s_2}^{(\mu)}(\vec{p}_2)(-ie)\gamma^\nu u_{s_1}^{(e)}(\vec{p}_1) \\
&= \frac{e^2}{s}\bar{u}_{s_1'}^{(e)}(\vec{p}_1')\gamma^\mu v_{s_2'}^{(\mu)}(\vec{p}_2')\bar{v}_{s_2}^{(\mu)}(\vec{p}_2)\gamma_\mu u_{s_2}^{(e)}(\vec{p}_1),
\end{aligned} \qquad (8.34)$$

where $s = (p_1 + p_2)^2$ is the usual Mandelstam variable. For the processes $e^- e^- \to e^- e^-$ and $e^- e^+ \to e^- e^+$ there are instead two diagrams, similar to those in Figs. 4 and 5, except for the replacement of the internal scalar line with a photon line. Their value is of course different.

The physical quantity measured in experiments is not the transition amplitude $\mathcal{M}_{fi}$ but the cross section, which is proportional to $|\mathcal{M}_{fi}|^2$. Often the particles in the initial state are not polarised, and the spins of those in the final state are not observed. What is measured is then

the cross section averaged over initial spins and summed over final spins, in which enters the following averaged quantity,

$$\langle |\mathcal{M}_{fi}|^2 \rangle \equiv \frac{1}{(2s_1+1)(2s_2+1)} \sum_{s_1,s_2,s_f} |\mathcal{M}_{fi}(s_1,s_2 \to s_f)|^2 \,. \tag{8.35}$$

For processes involving fermions one can exploit the identities of Eq. (5.86) to simplify the final expressions. Let us see how this works in the case of elastic $e^-\mu^-$ scattering. From Eq. (8.31) we find

$$\langle |\mathcal{M}_{fi}^{(e\mu \to e\mu)}|^2 \rangle = \frac{1}{4} \sum_{s_1,s_2,s_1',s_2'} \left(\frac{e^2}{t}\right)^2 \bar{u}_{1'}^{(e)} \gamma^\mu u_1^{(e)} \bar{u}_{2'}^{(\mu)} \gamma_\mu u_2^{(\mu)} \bar{u}_1^{(e)} \gamma^\nu u_{1'}^{(e)} \bar{u}_2^{(\mu)} \gamma_\nu u_{2'}^{(\mu)}$$
$$= \frac{1}{4} \frac{e^4}{t^2} \mathrm{tr}\, \gamma^\mu (\not{p}_1 + m_e) \gamma^\nu (\not{p}_1' + m_e) \gamma_\mu (\not{p}_2 + m_\mu) \gamma_\nu (\not{p}_2' + m_\mu) \,. \tag{8.36}$$

The trace can be evaluated by making use of the following identities,

$$\mathrm{tr}\, \gamma^\mu \gamma^\nu = 4\eta^{\mu\nu} \,,$$
$$\mathrm{tr}\, \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma = 4(\eta^{\mu\nu}\eta^{\rho\sigma} - \eta^{\mu\rho}\eta^{\nu\sigma} + \eta^{\mu\sigma}\eta^{\nu\rho}) \,, \tag{8.37}$$

and the fact that the trace of an odd number of gamma matrices vanishes. We will see below how to proved these identities, for the time being we just use them to get

$$\langle |\mathcal{M}_{fi}^{(e\mu \to e\mu)}|^2 \rangle = 4\frac{e^4}{t^2} \left(\eta^{\mu\nu}(m_e^2 - p_1 \cdot p_1') + p_1^\mu p_1'^\nu + p_1^\nu p_1'^\mu\right) \left(\eta_{\mu\nu}(m_\mu^2 - p_2 \cdot p_2') + p_{2\mu} p_{2\nu}' + p_{2\nu} p_{2\mu}'\right) \,. \tag{8.38}$$

Notice now that $t = 2(m_e^2 - p_1 \cdot p_1') = 2(m_\mu^2 - p_2 \cdot p_2')$, $s = m_e^2 + m_\mu^2 + p_1 \cdot p_2 = m_e^2 + m_\mu^2 + p_1' \cdot p_2'$ and $u = (p_1 - p_2')^2 = m_e^2 + m_\mu^2 - p_1 \cdot p_2' = m_e^2 + m_\mu^2 - p_2 \cdot p_1'$, to find

$$\langle |\mathcal{M}_{fi}^{(e\mu \to e\mu)}|^2 \rangle = 4\frac{e^4}{t^2} \left(4\left(\frac{t}{2}\right)^2 + 2\frac{t}{2}(p_1 \cdot p_1' + p_2 \cdot p_2') + 2((p_1 \cdot p_2)(p_1' \cdot p_2') + (p_1 \cdot p_2')(p_1' \cdot p_2))\right)$$
$$= 4\frac{e^4}{t^2} \left(t^2 + t(m_e^2 + m_\mu^2 - t) + \frac{1}{2}((s - m_e^2 - m_\mu^2)^2 + (u - m_e^2 - m_\mu^2)^2)\right)$$
$$= 4\frac{e^4}{t^2} \left(t(m_e^2 + m_\mu^2) + \frac{1}{2}((s - m_e^2 - m_\mu^2)^2 + (u - m_e^2 - m_\mu^2)^2)\right) \,. \tag{8.39}$$

Using the identity $s + t + u = 2(m_e^2 + m_\mu^2)$, this can be written as

$$\langle |\mathcal{M}_{fi}^{(e\mu \to e\mu)}|^2 \rangle = 2\frac{e^4}{t^2} \left(6(m_e^2 + m_\mu^2)^2 - 4(s+u)(m_e^2 + m_\mu^2) + s^2 + u^2\right) \,. \tag{8.40}$$

In the limit of high energy at fixed t this simplifies to

$$\langle |\mathcal{M}_{fi}^{(e\mu \to e\mu)}|^2 \rangle \to 2e^4 \frac{s^2 + u^2}{t^2} \simeq 4e^4 \frac{s^2}{t^2} \,. \tag{8.41}$$

The first expression above coincides with the one found in the limit of massless fermions.

Let us now discuss how one can prove the trace identities Eq. (8.37). To this end, let us introduce the following matrix,

$$\gamma^5 = \frac{1}{4!i}\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma = \frac{1}{i}\epsilon_{0123}\gamma^0\gamma^1\gamma^2\gamma^3 = -\frac{1}{i}\gamma^0\gamma^1\gamma^2\gamma^3 = i\gamma^0\gamma^1\gamma^2\gamma^3 \,, \tag{8.42}$$

which in the spinor representation that we are using reads

$$\gamma^5 = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} . \tag{8.43}$$

It is easy to see that γ^5 anticommutes with all the γ^μ,

$$\{\gamma^5, \gamma^\mu\} = 0 \,, \tag{8.44}$$

and that $(\gamma^5)^2 = \mathbf{1}$. If we now compute the trace of an odd number of gamma matrices, we find

$$\begin{aligned} \mathrm{tr}\,\{\gamma^{\mu_1}\ldots\gamma^{\mu_{2k+1}}\} &= \mathrm{tr}\,\{\gamma^{\mu_1}\ldots\gamma^{\mu_{2k+1}}(\gamma^5)^2\} = \mathrm{tr}\,\{\gamma^5\gamma^{\mu_1}\ldots\gamma^{\mu_{2k+1}}\gamma^5\} \\ &= \mathrm{tr}\,\{\gamma^5(-1)^{2k+1}\gamma^5\gamma^{\mu_1}\ldots\gamma^{\mu_{2k+1}}\gamma^5\} = -\mathrm{tr}\,\{\gamma^{\mu_1}\ldots\gamma^{\mu_{2k+1}}\} = 0 \,. \end{aligned} \tag{8.45}$$

For the trace of the product of two gamma matrices, we can use the basic anticommutators to show

$$\mathrm{tr}\,\gamma^\mu\gamma^\nu = \mathrm{tr}\,\gamma^\nu\gamma^\mu = \frac{1}{2}\mathrm{tr}\,\{\gamma^\mu, \gamma^\nu\} = 4\eta^{\mu\nu} \,. \tag{8.46}$$

Finally, to prove the four-gamma identity we make repeated use of $\gamma^\mu\gamma^\nu = 2\eta^{\mu\nu} - \gamma^\nu\gamma^\mu$ to show that

$$\begin{aligned} \mathrm{tr}\,\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma &= 2\eta^{\mu\nu}\mathrm{tr}\,\gamma^\rho\gamma^\sigma - \mathrm{tr}\,\gamma^\nu\gamma^\mu\gamma^\rho\gamma^\sigma = 8\eta^{\mu\nu}\eta^{\rho\sigma} - \mathrm{tr}\,\gamma^\mu\gamma^\rho\gamma^\sigma\gamma^\nu \\ &= 8(\eta^{\mu\nu}\eta^{\rho\sigma} - \eta^{\mu\rho}\eta^{\nu\sigma}) + \mathrm{tr}\,\gamma^\mu\gamma^\sigma\gamma^\nu\gamma^\rho \\ &= 8(\eta^{\mu\nu}\eta^{\rho\sigma} - \eta^{\mu\rho}\eta^{\nu\sigma} + \eta^{\mu\sigma}\eta^{\nu\rho}) - \mathrm{tr}\,\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma \,, \end{aligned} \tag{8.47}$$

from which the identity follows.