Introduction to Relativistic Quantum Fields

Jan Smit
Institute for Theoretical Physics
University of Amsterdam
Valckenierstraat 65, 1018XE Amsterdam
The Netherlands
jsmit@science.uva.nl

Lecture notes 2007

Contents

0.1 Why Field Theory?

Quantum Field Theory is our basic framework for the description of particles and their interactions. To the shortest distances which can be explored with current accelerators, which is about 10^{-25} cm, a theory called the 'Extended Standard Model' (ESM)[1] provides an accurate description of hadrons, leptons, gauge bosons. Field theory is also used in cosmology, e.g. combining general relativity ('geometrodynamics') with the ESM, with a scalar field added to incorporate inflation.

The world is evidently quantal, but why fields? Classical field theory is *local*. Interactions are described by differential equations at a point in space and time. They only refer to the immediate neighborhood of that point through derivatives of finite order, usually only up to second order. The equations referring to space-time points in Amsterdam do not refer to what goes on in Paris. A typical example is given by the the electromagnetic field interacting with electrons, which are described by the Maxwell equations and the equations for the Lorentz forces acting on the electrons. The particles create propagating electromagnetic fields which influence in turn the particles. In the quantum theory the electromagnetic field describes also particles, the photons, which can be created or annihilated by the electrons. Action at a distance (in space and time) can be avoided in this description. Locality is the space-time version of causality. The word causality suggests also an temporal order of cause and effect, and this is impletemented by retarded boundary conditions.

These ideas have been questioned from time to time, but alternative descriptions (such Feynman and Wheeler's absorber theory) have not been able to elicit the same intuitive appeal as field theory and have not been pursued very much.

As we will see, quantising fields leads to a description of arbitrarily many identical particles, bosons or fermions (in three spatial dimensions). This description is elegant and practical, which is another reason for the success of field theory. Moreover, there are phenomena which cannot be captured in terms of particles, for which the field formulation is essential. These are typically situations where strong fields prevail, which occur in classical electrodynamics but also in geometrodynamics (e.g. near black holes). Another example is quantum chromodynamcs, where quarks and gluons are confined into hadrons, the protons, neutrons, pions, glueballs, etc.

Field theory is based on the existence of space-time. The latter may perhaps be explainable in terms of an underlying theory, such as 'M theory'. It may take some time before such an extension can be tested by experimental results.

[1] By this we mean the renormalizable extension of the Standard Model (SM) to allow for non-zero neutrino masses; a.k.a. the νMSM.

0.2 These notes

These lecture notes present an introduction to relativistic quantum field theory, leading to quantum electrodynamics in covariant gauges and opening the road to the Standard Model. Our approach is based on the idea that fields are the basic variables, which upon quantization lead (or may not lead) to particles. Other authors (notably Veltman, Weinberg) follow an opposite line of thought: particles are basic and fields are to be constructed in accordance with their properties and general principles. This latter approach appears to be essentially perturbative.

We will briefly touch upon renormalization and the non-perturbative lattice formulation in the case the scalar field. For gauge fields and fermions our presentation of the path integral stays at the formal level,[2] which is sufficient for developing perturbation theory. In case of fermions, the complications one encounters in truly non-perturbative (lattice) formulations are remarkable and expose the 'cheating' implied by staying at the formal level. Progress here is steady and can be traced in the proceedings of the yearly International Symposium on Lattice Field Theory.

0.3 Books

Lecture notes are no substitute for a book. The following books are refered to in the text by name of authors:

Books on mathematical methods:

Jon Mathews and R.L. Walker, Mathematical Methods of Physics, Benjamin 1970.

R.B. Dingle, Asymptotic expansions, their derivation and interpretation, Academic Press 1973.

M.J. Lighthill, Introduction to Fourier analysis and generalised functions, Cambridge University Press 1958.

H.F. Jones, Groups, Representations and Physics (2nd edition), Institute of Physics 2003.

Books on classical electrodynamics and general relativity:

J.D. Jackson, Classical Electrodynamics, Wiley 1975/1998.

[2]'Formal' = jargon for 'pretty but mathematically imprecise'.

C.W. Misner, K.S. Thorn and J.A. Wheeler, Gravitation, Freeman 1973.

S. Weinberg, General Relativity and Cosmology, John Wiley and Sons 1972.

Books on relativistic QFT:

J.D. Bjorken and S.D. Drell, I: Relativistic Quantum Mechanics, McGraw-Hill (1964).

J.D. Bjorken and S.D. Drell, II: Relativistic Quantum Fields, McGraw-Hill (1965).

L.S. Brown, Quantum Field Theory, Cambridge University Press 1992.

Ta-Pei Cheng and Ling-Fong Li, Gauge Theory of Elementary Particle Physics, Oxford University Press (1984).

See also *ibid*, Problems and Solutions (2000).

R.P. Feynman, *The reason for antiparticles*, S. Weinberg, *Towards the final laws of physics*, in 'Elementary particles and the Laws of Physics', The 1986 Dirac Memorial Lectures, Cambridge University Press 1987.

C. Itzykson and J.-B. Zuber, Quantum Field Theory, McGraw-Hill (1980).

M.E. Peskin and D.V. Schroeder, An Introduction to Quantum Field Theory, Perseus 1995.

Stefan Pokorski, Gauge Field Theories, 2nd edition, Cambridge University Press 2000

P. Ramond, Field Theory: A Modern Primer (second edition), Addison Wesley (1989).

L. Ryder, Quantum Field Theory, Cambridge University Press 1996.

G. Sterman, Introduction to Quantum field Theory, Cambridge University Press 1993.

M. Veltman, Diagrammatica, Cambridge Lecture Notes in Physics, 1994.

S. Weinberg, The Quantum Theory of Fields, I: Fundamentals, Cambridge University Press 1995.

S. Weinberg, The Quantum Theory of Fields, II: Modern Applications, Cambridge University Press 1996.

S. Weinberg, The Quantum Theory of Fields, III, Cambridge University Press 1999.

B. de Wit and J. Smith, Field Theory in Particle Physics I, North-Holland (1986).

Books on quantum field theory and critical phenomena:

M. Le Bellac and G. Barton, Quantum and Statistical Field Theory, Clarendon 1992.

C. Itzykson and J-M. Drouffe, Statistical Field Theory I & II, Cambridge University Press 1989.

G. Parisi, Statistical Field Theory, Perseus 1998.

J. Zinn-Justin, Quantum Field Theory and Critical Phenomena, Clarendon 1996.

The following books are specifically on lattice field theory:

M. Creutz, Quarks, Gluons and Lattices, Cambridge University Press (1983).

G. Münster, I. Montvay, Quantum Fields on a Lattice, Cambridge University Press (1994).

H.J. Rothe, Introduction to Lattice Gauge Theories, World Scientific 1992 or later.

J. Smit, Introduction to Quantum Fields on a Lattice, Cambridge University Press, 2002.

See also the Proceedings of the yearly meetings Lattice 'XX.

Books on (specialized topics in) particle physics:

D.H. Perkins, Introduction to High Energy Physics, CUP (2000).

I.I. Bigi and A.I. Sanda, CP Violation, CUP (2000).

G. Branco, L. Lavoura and J. Silva, CP Violation, Oxford University Press (1999).

A book on the Casimir effect:

K.A. Milton, The Casimir effect, World Scientific 2001.

There are a number of useful lecture notes which are easily available, e.g. via the internet, for example:

P. van Baal, A Course in Field Theory, Instituut-Lorentz,
http://www-lorentz.leidenuniv.nl/vanbaal/FTcourse.html

P.J. Mulders, Quantum Field Theory, ITP, Vrije Universiteit,
http://www.nat.vu.nl/ mulders/lectures.html

J. Smit, Introduction to Quantum Field Theory 1994/95, ITFA 1995,
http://staff.science.uva.nl/ jsmit/
(The approach taken here differs substantially from the present notes.)

0.4 Conventions and notation

The following conventions will be used:

- $\hbar = c = 1$. The dimensions of various quantities are like [mass] = [energy] = [momentum] = $[\varphi(x)]$ = [length^{-1}] = [time^{-1}]. Actions are dimensionless. To convert to ordinary units use appropriate powers of $\hbar$ and c. A particularly useful combination is $\hbar c = 197.3$ MeV fm, where fm (femto meter or Fermi) denotes the unit of length 10^{-13} cm. For example a mass m of 200 MeV corresponds to a length $1/m$ of about 1 fm.

- Rationalized Gauss (Lorentz-Heaviside) units for electromgnetism. The unit of electromagnetic charge $e \approx 0.30$ ($\alpha = e^2/4\pi \approx 1/137$). The charge of the electron is $-e$.

- Minkowski metric

$$\begin{aligned}
x_\mu &= \eta_{\mu\nu} x^\nu, \quad \eta_{11} = \eta_{22} = \eta_{33} = -\eta_{00} = 1, & (1)\\
x^0 &= -x_0, \quad x_k = x^k, \quad k = 1, 2, 3, & (2)\\
\partial_\mu &= \frac{\partial}{\partial x^\mu}, & (3)\\
x^2 &= x_\mu x^\mu = \mathbf{x}^2 - x_0^2, \quad \partial^2 = \partial_\mu \partial^\mu = \nabla^2 - \partial_0^2, & (4)\\
px &= p_k x^k + p_0 x^0 = p_k x^k - p^0 x^0 = \mathbf{px} - p^0 x^0, & (5)\\
\epsilon_{0123} &= +1, & (6)\\
d^4x &= dx^0\, dx^1\, dx^2\, dx^3, \quad d^4p = dp^0\, dp^1\, dp^2\, dp^3. & (7)
\end{aligned}$$

  The same metric is used in the books by Brown, Weinberg, and Misner-Thorn-Wheeler. Many other authors (e.g. Bjorken and Drell, Peskin and Schroeder) use the metric with $\eta_{00} = +1$.

- Dirac matrices:

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}, \quad \alpha^\mu = i\beta\gamma^\mu, \quad \beta = i\gamma^0, \quad \gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{8}$$

  ($i\gamma^\mu = \gamma^\mu$[Bjorken&Drell], $\gamma_5 = \gamma_5$[Bjorken&Drell].) Left and right handed chiral projectors $P_L = (1 - \gamma_5)/2$, $P_R = (1 + \gamma_5)/2$.

- Lorentz invariant volume element for mass m,

$$d\omega_p = \frac{d^3p}{(2\pi)^3 2p^0}, \quad p^0 = E(\mathbf{p}) = \sqrt{\mathbf{p}^2 + m^2}. \tag{9}$$

  Depending on the context, p^0 can be an arbitrary variable (e.g. a dummy variable in $\int d^4p$), or 'on the energy shell' as in $d\omega_p$.

Chapter 1

Classical fields

We recall here some familiar classical fields, introduce the Lorentz group, the canonical formalism and the action functional, its symmetries and the corresponding conserved quantities such as energy and angular momentum.

1.1 Maxwell field

The electric and magnetic fields **E** and **B** constitute the Lorentz covariant anti-symmetric tensor field $F_{\mu\nu}(x)$, $F_{\mu\nu} = -F_{\nu\mu}$, such that

$$E_k(x) = -F_{0k} = F^{0k}, \quad B_k = \frac{1}{2}\epsilon_{klm}F_{lm}. \tag{1.1}$$

In terms of $F_{\mu\nu}$, Maxwell's equations can be written in Lorentz invariant form,

$$\epsilon^{\kappa\lambda\mu\nu}\partial_\lambda F_{\mu\nu} = 0, \tag{1.2}$$

$$\partial_\mu F^{\mu\nu} = -j^\nu, \tag{1.3}$$

where $j^\nu(x)$ is the electromagnetic four-current density, or current for short, which has to satisfy

$$\partial_\mu j^\mu = 0, \tag{1.4}$$

for consistency with (1.3) (check: $-\partial_\mu j^\mu = \partial_\mu\partial_\nu F^{\nu\mu} = 0$ because $F^{\mu\nu} = -F^{\nu\mu}$ whereas $\partial_\mu\partial_\nu = +\partial_\nu\partial_\mu$). In jargon we say that the current is conserved.[1] The homogeneous equations (1.2) can be satisfied identically by expressing $F_{\mu\nu}(x)$ in terms of the (four-)vector potential $A_\mu(x)$,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.5}$$

In terms of A_μ the inhomogenous Maxwell equations (1.3) take the form

$$\partial_\mu\partial^\mu A^\nu - \partial^\nu\partial_\mu A^\mu = -j^\nu. \tag{1.6}$$

[1]The terminology: a current $j^\mu(x)$ is 'conserved', simply means: $\partial_\mu j^\mu(x) = 0$. It is of course the total charge $Q = \int d^3x\, j^0(x)$ which is conserved.

An important aspect of this description is gauge invariance: $F_{\mu\nu}$ is invariant under the gauge transformations

$$A'_\mu(x) = A_\mu(x) + \partial_\mu \omega(x), \tag{1.7}$$
$$F'_{\mu\nu} = F_{\mu\nu}(x). \tag{1.8}$$

The equations for A_μ are gauge invariant, which implies that their solution is not unique. To obtain a unique solution of (1.6) one may impose a gauge condition. Two well-known gauge conditions are

$$\partial_\mu A^\mu = 0, \quad \text{Lorentz gauge (also called Landau gauge)}, \tag{1.9}$$
$$\partial_k A^k = 0, \quad \text{Coulomb gauge (also called radiation gauge)}. \tag{1.10}$$

In the Lorentz gauge (1.6) reduces to

$$\partial^2 A^\mu = -j^\mu, \tag{1.11}$$

which are four hyperbolic wave equations for A_μ (recall $\partial^2 = \nabla^2 - \partial_0^2$).

One has learned in the course of time that the A_μ are the basic variables for the description of the electromagnetic field. On the other hand, physically observable quantities have to be gauge invariant. However, this does not mean that everything physical is expressable in terms of the field strength $F_{\mu\nu}$. In the quantum theory this is spectacularly illustrated by the Aharonov-Bohm effect.

1.2 Einstein field

In General Relativity the gravitational field is described in terms of the metric tensor field $g_{\mu\nu}(x)$ while matter gives rise to an energy-momentum tensor field $T^{\mu\nu}(x)$. These have to satisfy the Einstein equations

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R + \Lambda g^{\mu\nu} = 8\pi G T^{\mu\nu}, \tag{1.12}$$

together with dynamical equations for matter. Here $R^{\mu\nu}$ and R are the Ricci tensor and scalar constructed out of $g_{\mu\nu}$, G is Newton's constant and Λ the cosmological constant. We shall not go into detail here as our working arena will be Minkowski space within special relativity. But note that the energy-momentum tensor will also play an important role in the following. It is good to keep in mind its role as a source of the gravitional field.

1.3 Scalar field

For illustrative purpose we will often use a scalar field $\varphi(x)$, which carries no vector or tensor indices. A typical equation for such a field is the Klein-Gordon equation with source $J(x)$,

$$(-\partial^2 + m^2)\varphi = J. \tag{1.13}$$

This equation is similar to (1.11), except for the parameter m, which has dimension of $(\text{length})^{-1}$ or frequency (recall that $c = 1$ in our units, $\hbar$ does not enter in classical field theory). To get a feeling for its meaning, consider a plane wave

$$\varphi(x) = e^{ikx} + \text{c.c.} = e^{i\mathbf{kx}-ik^0x^0} + \text{c.c.} \tag{1.14}$$

(c.c. denotes 'complex conjugate'). This is a solution of (1.13) for the case $J = 0$, provided that

$$k^2 + m^2 = 0, \quad k^0 = \pm\sqrt{\mathbf{k}^2 + m^2}. \tag{1.15}$$

So m is the frequency k^0 for zero wave vector, $\mathbf{k} = 0$. Another example is the solution for a static point source at the origin,

$$\varphi(\mathbf{x}) = \frac{e^{-m|\mathbf{x}|}}{4\pi|\mathbf{x}|}, \tag{1.16}$$

$$(-\nabla^2 + m^2)\varphi(\mathbf{x}) = \delta(\mathbf{x}). \tag{1.17}$$

The solution (1.16) is called the Yukawa potential. We see that it decays exponentially fast as $|\mathbf{x}| \to \infty$ with the scale set by $1/m$, the range of the potential.

As a classical field, the scalar field is not so familiar as the electromagnetic or or gravitational fields. The reason is that in applications to relativistic physics the particles described by quantized scalar fields are usually unstable with very short life times. Furthermore, we shall see that m is the mass of such particles; $1/m = \hbar/mc$ their Compton wavelength, is typically a very short distance, such that φ decays rapidly to zero away from its source. The minimal frequencies are also very large. For example, for pions $m^{-1} = 1.4 \times 10^{-13}$ cm or $m = 2 \times 10^{23}$ s^{-1}. In addition, wave packet solutions made out of superpositions of plane waves tend to spread rapidly because of the dispersion relation $k^0 = \pm\sqrt{m^2 + \mathbf{k}^2}$ (see e.g. the discussion in Jackson sect. 7.9).

However, in the nonrelativistic domain such scalar fields *do* occur classically in systems showing superfluidity or (normal) superconductivity. Suppose φ is slowly varying in space compared to m^{-1}. Then it is useful to derive a nonrelativistic form of the Klein-Gordon equation by separating out the high frequency oscillations, writing

$$\varphi(\mathbf{x}, t) = \frac{1}{2}e^{-imt}\,\psi(\mathbf{x}, t) + \frac{1}{2}e^{imt}\,\psi^*(\mathbf{x}, t), \tag{1.18}$$

$$\partial_t\varphi(\mathbf{x}, t) = \frac{-im}{2}e^{-imt}\,\psi(\mathbf{x}, t) + \frac{im}{2}e^{imt}\,\psi^*(\mathbf{x}, t), \tag{1.19}$$

$$\psi(\mathbf{x}, t) = e^{imt}\left[\varphi(\mathbf{x}, t) + \frac{i}{m}\partial_t\varphi(\mathbf{x}, t)\right], \tag{1.20}$$

with the complex conjugate equation for ψ^*. This is a change of variables in which φ and $\partial_t\varphi$ are represented by two new independent fields, the real and imaginary parts of ψ. The equation $(\partial^2 - m^2)\varphi = 0$ is equivalent to

$$i\partial_t\psi = \frac{-\nabla^2}{2m}\left(\psi + e^{2imt}\psi^*\right). \tag{1.21}$$

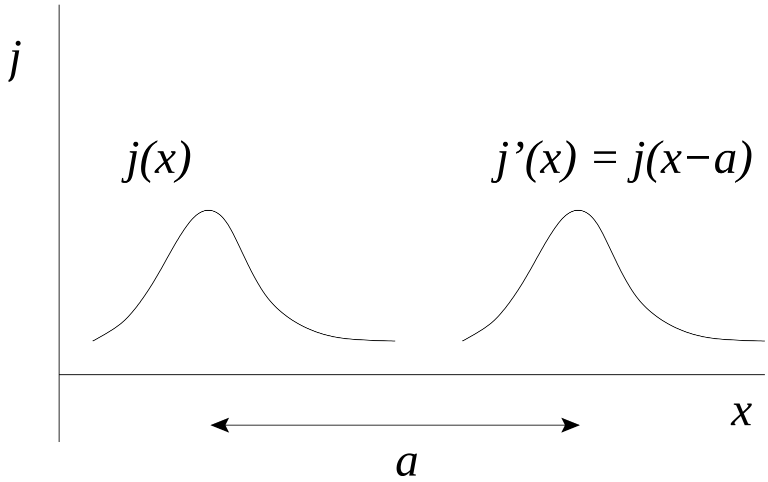

Figure 1.1: Translation of a field.

Assuming that measurements involve time scales much larger than m we can neglect the rapidly oscillating term $\propto \exp(i2mt)$. The resulting equation is identical to the Schrödinger equation, but this should not confuse us. The above ψ has nothing to do with this quantum mechanical wave function: it is simply a classical field which happens to be complex. Usually there are also additional source terms like $\psi^*\psi^2$ in the nonrelativistic field equations (the field providing its own source), in which case one sometimes uses the even more confusing terminology 'nonlinear Schrödinger equation'. Observable quantities should of course be real. Simple examples are given by $\psi^*\psi$ and $(\psi^*\nabla\psi - \nabla\psi * \psi)/2im$.

In the application to superfluid ^{4}He, m is of the order of the mass of the helium atom. In the application to superconductivity m is the mass of a Cooper pair (about twice the electron mass) and ψ is called the Landau-Ginsberg field ('the Cooper-pair field').

1.4 Poincaré group

The field equations of the previous sections are invariant under translations and Lorentz transformations. Such transformations form a group, the Poincaré group, which contains translations and Lorentz transformations as subgroups. This will now be discussed in more detail[2].

We start with translations. Let $j^\mu(x)$ be the electromagnetic current of the system. Suppose we translate the system over a distance a^μ in spacetime.[3] Then the corresponding current is $j'^\mu(x) = j^\mu(x-a)$, see figure 1.1. So under transla-

[2]Our review is brief, see e.g. Jones or courses on group theory for more information.

[3]This is a so-called 'active' transformation. In the equivalent 'passive' viewpoint the system is unchanged but we make a coordinate transformation $x \to x' = x + a$.

tions:

$$\begin{aligned} x' &= x + a, && (1.22)\\ \varphi'(x') &= \varphi(x), \quad \text{or } \varphi'(x) = \varphi(x-a), && (1.23)\\ A'_\mu(x') &= A_\mu(x), \quad \text{or } A'_\mu(x) = A_\mu(x-a), && (1.24) \end{aligned}$$

and similar for other fields. The field equations are obviously invariant, e.g. $\partial'_\mu F'^{\mu\nu}(x') + j'^\nu(x') = 0$ implies $\partial_\mu F^{\mu\nu}(x) + j^\nu(x) = 0$, because $\partial'_\mu \equiv \partial/\partial x'^\mu = \partial/\partial x^\mu = \partial_\mu$.

For many derivations it is useful to consider infinitesimal transformations, written as $1 - ia^\mu P_\mu + O(a^2)$. Here 1 denotes the abstract unit element (no translation) and the coefficients P_μ of the infinitesimal parameters a^μ are the generators of the translation group. A finite translation can be abstractly written as $\exp(-ia^\mu P_\mu)$. For a representation of P_μ, consider an infinitesimal translation of a scalar field:

$$\varphi'(x) = \varphi(x-a) = \varphi(x) - a^\mu \partial_\mu \varphi(x) + \cdots. \tag{1.25}$$

Writing this as

$$(1 - ia^\mu P_\mu + \cdots)\varphi(x), \tag{1.26}$$

we see that

$$P_\mu \to -i\partial_\mu \tag{1.27}$$

is a representation of the generators P_μ.

Next we turn to Lorentz transformations. Recall that we use a metric tensor of the form

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 \end{pmatrix}_{\mu\nu} . \tag{1.28}$$

Lorentz transformations

$$x'^\mu = \ell^\mu{}_\nu x^\nu, \tag{1.29}$$

leave invariant the quadratic form $x^2 = x^\mu x_\mu = \eta_{\mu\nu} x^\mu x^\nu$,

$$x'^2 = x^2, \tag{1.30}$$

which implies

$$\eta_{\mu\nu} \ell^\mu{}_\rho \ell^\nu{}_\sigma = \eta_{\rho\sigma}, \quad \text{or} \quad \ell_\nu{}^\rho \ell^\nu{}_\sigma = \delta_{\rho\sigma}. \tag{1.31}$$

We can write the transformations also in matrix form,

$$\begin{aligned} x &= \begin{pmatrix} x^0\\ x^1\\ x^2\\ x^3 \end{pmatrix}, \quad x' = \ell x, && (1.32)\\ (x')_\mu &= (\ell)_{\mu\nu}(x)_\nu, \quad (x)_\mu \equiv x^\mu, \quad (\ell)_{\mu\nu} \equiv \ell^\mu{}_\nu. && (1.33) \end{aligned}$$

Eq. (1.31) reads in matrix notation

$$\ell^T \eta \ell = \eta, \quad \text{or } \ell^{-1} = \eta \ell^T \eta, \quad (\ell^{-1})_{\mu\nu} = \ell_\nu{}^\mu, \tag{1.34}$$

where T denotes transposition. Note the subtlety in the connection between matrix and tensor notation: $(\eta)_{\mu\nu} = \eta_{\mu\nu} = \eta^{\mu\nu}$, but $(\ell)_{\mu\nu} = \ell^\mu{}_\nu$.

For a special Lorentz transformation (boost) in the 3-direction,

$$\ell = \begin{pmatrix} \cosh\chi & 0 & 0 & \sinh\chi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\chi & 0 & 0 & \cosh\chi \end{pmatrix}, \tag{1.35}$$

where $\cosh\chi = \gamma$, $\gamma = 1/\sqrt{1-v^2}$, $\sinh\chi = \gamma v$, with v the velocity of the transformation in units $c = 1$. For a rotation about the three axis over an angle ϕ,

$$\ell = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi & 0 \\ 0 & \sin\phi & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.36}$$

The transformations form a group, which includes the rotation group, parity (P) and timereversal (T). The latter two are defined by

$$(\ell_P x)^0 = x^0, \ (\ell_P x)^k = -x^k, \tag{1.37}$$

$$(\ell_T x)^0 = -x^0, \ (\ell_T x)^k = x^k, \tag{1.38}$$

where $k = 1, 2, 3$. The rotations and boosts have $\det \ell = 1$, while ℓ_P and ℓ_T have determinant -1. The transformations with $\det \ell = +1$ form a subgroup, the proper Lorentz group. In the following we shall mean this group when referring to Lorentz transformations; parity and time reversal will be mentioned separately.

We denote matrix representations of the Lorentz group by D,

$$\ell \to D(\ell). \tag{1.39}$$

Fields $\chi(x)$ transform as

$$\chi'(x') = D\chi(x). \tag{1.40}$$

Simple examples are given by

$$D = 1, \quad \chi'(x') = \chi(x), \quad \text{scalar field}, \tag{1.41}$$

$$(D)_{\mu\rho} = \ell^\mu{}_\rho, \quad \chi'^\mu(x') = \ell^\mu{}_\rho \chi^\rho(x), \quad \text{vector field}, \tag{1.42}$$

$$(D)_{\mu\nu,\rho\sigma} = \ell^\mu{}_\rho \ell^\nu{}_\sigma, \quad \chi'^{\mu\nu}(x') = \ell^\mu{}_\rho \ell^\nu{}_\sigma \chi^{\rho\sigma}(x), \quad \text{tensor field}. \tag{1.43}$$

See figure 1.2 for an illustration of the transformation rule of the vector field for the case of rotations in the active formulation.

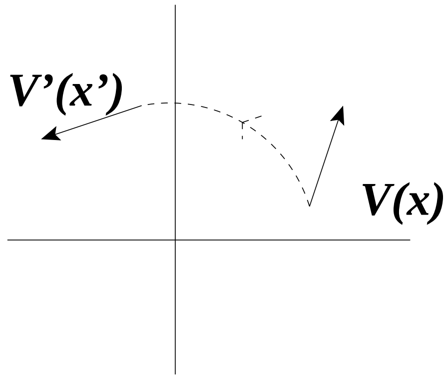

Figure 1.2: Rotation of a vector field in space: the vector $\mathbf{V}$ is rotated into $\mathbf{V}'$ and its base point $\mathbf{x}$ is transformed into $\mathbf{x}'$, $V'_k(\mathbf{x}') = \ell_{kl}V_l(\mathbf{x})$, $x'_k = \ell_{kl}x_l$, just as would follow from a rotation of the coordinate frame.

The vector representation is the defining representation of the Lorentz group. An example of a vector field is the electromagnetic current $j^\mu(x)$. Under rotations it has a spin zero component, the scalar (under rotations) $j^0(x)$, and a spin one component, the vector (under rotations) $\mathbf{j}(x)$. An example of a tensor field is the electromagnetic field $F^{\mu\nu}(x)$. It is an antisymmetric tensor under Lorentz transformations, while under rotations it consists of two vectors $\mathbf{E}(x)$ and $\mathbf{B}(x)$, $E^k = F^{0k}$, $B^k = (1/2)\epsilon_{klm}F^{lm}$, $k, l, m = 1, 2, 3$.

To verify the invariance of field equations under Lorentz tranformations we check that ∂_μ transforms as a (covariant) vector. For example

$$\begin{aligned}\partial'_\mu\varphi'(x') &= \frac{\partial}{\partial x'^\mu}\varphi'(x') = \frac{\partial}{\partial x'^\mu}\varphi(x) = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu}\varphi(x) = (\ell^{-1})_{\nu\mu}\partial_\nu\varphi(x)\\ &= \ell_\mu{}^\nu\partial_\nu\varphi(x),\end{aligned} \tag{1.44}$$

where we used (1.34). It is now straightforward to check the Lorentz invariance of field equations, e.g. $\partial'_\mu F'^{\mu\nu}(x') + j'^\nu(x') = 0$ implies $\partial_\mu F^{\mu\nu}(x) + j^\nu(x) = 0$.

Poincaré transformations

$$x' = \ell x + a \tag{1.45}$$

leave invariant $(x - y)^2$ and form therefore also a group. The field equations are of course also invariant under this combined group. For more information see Weinberg I or Ryder.

Consider next infinitesimal Lorentz transformations,

$$\ell^\mu{}_\rho = \eta^\mu{}_\rho + \omega^\mu{}_\rho, \tag{1.46}$$

where $\omega^\mu{}_\rho$ is infinitesimal (note that $\eta^\mu{}_\rho = \delta_{\mu\rho}$). The relation (1.31) is satisfied if

$$\omega_{\mu\nu} = -\omega_{\nu\mu} \tag{1.47}$$

(also the indices of ω are raised and lowered with the metric tensor). For a rotation in the 1-2 plane about the 3-axis over a positive angle ϕ we have

$$\omega^1{}_2 = \omega_{12} = -\omega_{21} = -\phi, \tag{1.48}$$

(cf. (1.36) and Problem 10). Similarly, for special Lorentz transformation in the 0-3 plane of Minkowski space with hyperbolic angle χ, a boost in the positive 3-direction with velocity v,

$$\omega^0{}_3 = -\omega_{03} = \omega_{30} = \chi = \tanh^{-1} v \tag{1.49}$$

(cf. (1.35) and Problem 10; infinitesimally $\chi = v$).

For a representation D we write

$$D = 1 + i\frac{1}{2}\omega^{\alpha\beta} S_{\alpha\beta} + O(\omega^2), \quad S_{\alpha\beta} = -S_{\beta\alpha}, \tag{1.50}$$

where the $S_{\alpha\beta}$ are matrices specifying the representation. They are called the generators in the representation D. In the defining (vector) representation,

$$(D)_{\mu\nu} = \ell^\mu{}_\nu = \delta_{\mu\nu} + i\frac{1}{2}\omega^{\alpha\beta}\,(S_{\alpha\beta})_{\mu\nu}. \tag{1.51}$$

It follows from this by comparison with (1.46), that in the defining representation $S_{\alpha\beta}$ is given by

$$(S_{\alpha\beta})_{\mu\nu} = -i(\eta^\mu_\alpha \eta_{\beta\nu} - \eta^\mu_\beta \eta_{\alpha\nu}), \quad \text{defining representation} \tag{1.52}$$

(just substitute and check). It is now straightforward to verify that the generators satisfy the commutation relations

$$[S_{\alpha\beta}, S_{\gamma\delta}] = i(\eta_{\alpha\gamma} S_{\beta\delta} - \eta_{\beta\gamma} S_{\alpha\delta} - \eta_{\alpha\delta} S_{\beta\gamma} + \eta_{\beta\delta} S_{\alpha\gamma}). \tag{1.53}$$

A finite transformation can be written as

$$D = \exp\left(i\frac{1}{2}\omega^{\alpha\beta} S_{\alpha\beta}\right). \tag{1.54}$$

If we have a set of matrices $S_{\alpha\beta}$ satisfying the commutation relations (1.53) then we have a representation of the Lorentz group. This follows from the Baker-Cambell-Haussdorff theorem: if $D = \exp(M) = \exp(i\omega_{\alpha\beta} S_{\alpha\beta}/2)$ and similar for $D' = \exp(M')$, then $D'' \equiv DD'$ can be expressed as the exponential of a series in multiple commutators of M and M', $D'' = \exp(M + M' + [M, M']/2 + \cdots)$, which is completely determined by the commutation relations (1.53).

Note that the boosts do not form a group. This can be seen from the fact that the set of boost generators S_{0k} is not closed under commutation. For example $[S_{02}, S_{03}] = -iS_{23}$, which generates rotations about the three axis. So the rotations are needed to complement the boosts into the (proper) Lorentz group.

We end this introduction with the commutation relations of the generators of the Poincaré group. Let us denote for the moment the abstract generators of the Lorentz group by $J_{\alpha\beta}$, so $S_{\alpha\beta}$ is a representation of $J_{\alpha\beta}$. A representation is provided by the transformation rule for the spacetime argument of fields; e.g. for an infinitesimal transformation of a scalar field:

$$\begin{aligned}\varphi'(x) &= \varphi(\ell^{-1}x) = \varphi(x) - \omega^{\mu\nu}x_\nu\partial_\mu\varphi(x) \\ &= \varphi(x) + i\frac{1}{2}\omega^{\mu\nu}(-ix_\mu\partial_\nu + ix_\nu\partial_\mu)\varphi(x). \end{aligned} \tag{1.55}$$

Writing this as

$$\left(1 + i\frac{1}{2}\omega^{\alpha\beta}J_{\alpha\beta} + \cdots\right)\varphi(x), \tag{1.56}$$

we see that we have the representation

$$J_{\alpha\beta} \to L_{\alpha\beta} \equiv -ix_\alpha\partial_\beta + ix_\beta\partial_\alpha. \tag{1.57}$$

For a general field $\chi(x)$ transforming as $\chi'(x) = D(\ell)\chi(\ell^{-1}x)$ we have the representation

$$J_{\alpha\beta} \to L_{\alpha\beta} + S_{\alpha\beta}. \tag{1.58}$$

The commutation relations of the generators of the Poincaré group now follow straightforwardly from (1.27) and (1.57):

$$\begin{aligned}[J_{\alpha\beta}, J_{\gamma\delta}] &= i(\eta_{\alpha\gamma}J_{\beta\delta} - \eta_{\beta\gamma}J_{\alpha\delta} - \eta_{\alpha\delta}J_{\beta\gamma} + \eta_{\beta\delta}J_{\alpha\gamma}), \\ [J_{\alpha\beta}, P_\mu] &= i(\eta_{\mu\alpha}P_\beta - \eta_{\mu\beta}P_\alpha), \\ [P_\mu, P_\nu] &= 0. \end{aligned} \tag{1.59}$$

1.5 Action

A powerful tool in our considerations will be the action S, from which the equations of motion can be derived by requiring it to be stationary under small variations of the dynamical variables. Symmetries of S lead to symmetries of the equations of motion and to quantities which are conserved in time: Noether's theorem. Furthermore, the action plays a crucial role in the path integral formulation of quantum theory.

Consider first an action for the scalar field,

$$S = \int d^4x \left(-\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 + J\varphi\right), \tag{1.60}$$

where the integration is over some compact domain M of spacetime. The dynamical variables are here $\varphi(x)$, whereas $J(x)$ is considered to be a given function of

x, it is an *external source*. Consider the variation of S under a small variation $\delta\varphi$ of φ which vanishes on the boundary ∂M of the domain M:

$$\begin{aligned}\delta S \equiv S[\varphi+\delta\varphi]-S[\varphi] &= \int d^4x\,(-\partial_\mu\varphi\partial^\mu\delta\varphi - m^2\varphi\delta\varphi + J\delta\varphi)\\ &= \int d^4x\,(\partial_\mu\partial^\mu\varphi - m^2\varphi + J)\delta\varphi, \qquad (1.61)\end{aligned}$$

where we made a partial integration in the second line and used the fact that $\delta\varphi = 0$ on ∂M. Requiring $\delta S = 0$ for arbitrary $\delta\varphi$ gives the Klein-Gordon equation with source J,

$$(\partial^2 - m^2)\varphi + J = 0, \qquad (1.62)$$

in the interior of M. (Allowing $\delta\varphi$ not to vanish on ∂M would in addition lead to boundary conditions for φ.)

A suitable action for the nonrelativistic scalar field (cf. eq. (1.21)) is given by

$$S = \int d^3x\,dt\,\left(\frac{1}{2}\psi^* i\partial_t\psi - \frac{1}{2}i\partial_t\psi^*\,\psi - \frac{1}{2m}\nabla\psi^*\nabla\psi + \eta^*\psi + \psi^*\eta\right), \qquad (1.63)$$

where η is the nonrelativistic analog of J.

Consider next the following action for the electromagnetic field,

$$S = \int d^4x\,\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + J^\mu A_\mu\right), \qquad (1.64)$$

where J^μ is an external current and the integral is again over M. Under a variation δA_μ of A_μ we have

$$\begin{aligned}\delta S &\equiv S[A+\delta A] - S[A]\\ &= \int d^4x\,\delta\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + J^\mu A_\mu\right), \qquad (1.65)\\ \delta F_{\mu\nu} &= \partial_\mu(A_\nu+\delta A_\nu) - \partial_\nu(A_\mu+\delta A_\mu) - F_{\mu\nu}\\ &= \partial_\mu\delta A_\nu - \partial_\nu\delta A_\mu, \qquad (1.66)\\ \delta(F_{\mu\nu}F^{\mu\nu}) &= 2F^{\mu\nu}\delta F_{\mu\nu} = 4F^{\mu\nu}\partial_\mu\delta A_\nu, \qquad (1.67)\\ \delta S &= \int d^4x\,(-F^{\mu\nu}\partial_\mu\delta A_\nu + J^\mu\delta A_\mu)\\ &= \int d^4x\,(\partial_\mu F^{\mu\nu} + J^\nu)\delta A_\nu. \qquad (1.68)\end{aligned}$$

We made a partial integration in the last step and assumed that the surface term is zero, which is correct if we impose that δA_μ vanishes on ∂M. Requiring $\delta S = 0$ for arbitrary variations in M gives Maxwell's equations

$$\partial_\mu F^{\mu\nu} + J^\nu = 0. \qquad (1.69)$$

The action for the gravitational field without matter is given by

$$S_g = \frac{1}{16\pi G}\int d^4x\,\sqrt{-\det g}\,(R - 2\Lambda). \qquad (1.70)$$

Under a variation $\delta g_{\mu\nu}$ we have[4]

$$\delta S_g = \frac{1}{16\pi G}\int d^4x\,\sqrt{-g}\,\left(-R^{\mu\nu} + \frac{1}{2}g^{\mu\nu}R - \Lambda g^{\mu\nu}\right)\delta g_{\mu\nu}. \tag{1.71}$$

The demonstration of this is quite involved, see for example Weinberg's book on general relativity. Setting δS_g to zero gives the Einstein equations without the $T^{\mu\nu}$ term. To get the full Einstein equations including $T^{\mu\nu}$ we have to add to S_g a term representing the contribution of matter, which we denote by S_m. With the total action

$$S = S_g + S_m, \tag{1.72}$$

and

$$\delta S_m = \int d^4x\,\sqrt{-g}\,\frac{1}{2}T^{\mu\nu}\delta g_{\mu\nu}, \tag{1.73}$$

we get the full Einstein equations including $T^{\mu\nu}$ by setting $\delta S = 0$. Note that 'matter' is just a name for anything not composed of $g_{\mu\nu}$ only. For example, it could be a bunch of point particles, but is may also be the electromagnetic field or a scalar field. We shall now derive the form of $T^{\mu\nu}$ for the scalar field and the electromagnetic field.

The minimal prescription for constructing an action that is invariant under general coordinate transformations from a Lorentz invariant action is:

a) make the volume element invariant under general coordinate transformations

$$d^4x \to d^4x\,\sqrt{-g(x)}; \tag{1.74}$$

b) replace derivatives by covariant derivatives.

With this prescription the scalar field action becomes

$$S_m[\varphi, g] = -\int d^4x\,\sqrt{-g}\,\left(\frac{1}{2}g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi + \frac{1}{2}m^2\varphi^2\right) \tag{1.75}$$

(for a scalar field the covariant derivative is just the ordinary derivative). Similarly the generally covariant generalization of the Maxwell action is

$$S_m[A, g] = -\int d^4x\,\sqrt{-g}\,\left(\frac{1}{4}g^{\kappa\lambda}g^{\mu\nu}F_{\kappa\mu}F_{\lambda\nu}\right). \tag{1.76}$$

The calculation of the variation of these matter actions with respect to $g_{\mu\nu}$ is not difficult (Problem 5) and leads to

$$T^{\mu\nu} = g^{\mu\alpha}g^{\nu\beta}\partial_\alpha\varphi\partial_\beta\varphi - g^{\mu\nu}\left(\frac{1}{2}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi + \frac{1}{2}m^2\varphi^2\right) \tag{1.77}$$

[4] Acccording to standard notation, $g \equiv \det g$ (the determinant of the matrix g), and $g^{\mu\nu} \equiv (g^{-1})_{\mu\nu}$, so $g^{\kappa\lambda}g_{\lambda\mu} = \delta_{\kappa\mu}$.

for the scalar field and

$$T^{\mu\nu} = F^{\mu\rho} F^{\nu}{}_{\rho} - \frac{1}{4} g^{\mu\nu} F^{\rho\sigma} F_{\rho\sigma}, \tag{1.78}$$

for the electromagnetic field, where $F^{\mu\rho} = g^{\mu\alpha} g^{\rho\beta} F_{\alpha\beta}$, etc. Specializing to Minkowski space is easy: $g_{\mu\nu} \to \eta_{\mu\nu}$.

1.6 Canonical formalism

Up to now we have used a manifestly covariant formalism for Lorentz invariant theories. For the relation to quantum theory, the canonical formalism in terms of a hamiltonian, which is a function of 'p's and q's', is appealing because of the correspondence between Poisson brackets and commutators $(q_k, p_l) = \delta_{kl} \leftrightarrow [\hat{q}_k, \hat{p}_l]/i\hbar = \delta_{kl}$. Because time is singled out as special (e.g. $p_k = \partial_t q_k$) this formalism breaks manifest covariance and it is complicated for gauge fields. However, it is important for the proper statement of the initial value problem and useful for an introduction to quantum field theory. We give the basics here for a scalar field theory described by the action

$$S = -\int d^4x \left[\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi + V(\varphi)\right], \tag{1.79}$$

$$V(\varphi) = \frac{1}{2}\kappa\varphi^2 + \frac{1}{4}\lambda\varphi^4. \tag{1.80}$$

We have added a φ^4 term to the action, which makes theory more interesting because the field equation is now nonlinear.[5] Its importance is parametrized by λ. The action can be written as ($x^0 = t$, $\dot{\varphi} = \partial_0\varphi$)

$$S = \int dt\, L[\varphi, \dot{\varphi}], \tag{1.81}$$

$$L = \int d^3x\, \frac{1}{2}\,\dot{\varphi}^2 - U[\varphi], \tag{1.82}$$

$$U[\varphi] = \int d^3x \left[V(\varphi) + \frac{1}{2}\,\partial_k\varphi\partial_k\varphi\right], \tag{1.83}$$

where the dot denotes $\partial/\partial t$. This looks like the lagrangian for a sum of systems, one for each $\mathbf{x}$, which are coupled by the spatial gradient term. The canonical momentum conjugate to φ is defined as[6]

$$\pi(\mathbf{x}) = \frac{\delta L}{\delta\dot{\varphi}(\mathbf{x})} = \dot{\varphi}(\mathbf{x}), \tag{1.84}$$

[5] This so-called φ^4 theory is a very useful model for illustration.

[6] The time dependence is implicit in the following. See the appendix for the definition of the functional derivative.

where we have used a functional derivative. As discussed in the appendix this is just a generalization of a partial derivative from discrete indices to continuous indices (here $\mathbf{x}$). In the same generalization the hamiltonian is given by the Legendre transform

$$H[\varphi, \pi] = \int d^3x\, \pi\dot{\varphi} - L[\varphi, \dot{\varphi}], \tag{1.85}$$

$$= \int d^3x\, \frac{1}{2}\pi^2 + U[\varphi]. \tag{1.86}$$

The Poisson bracket (A, B) of $A[\varphi, \pi]$ and $B[\varphi, \pi]$ generalizes to

$$(A, B) = \int d^3x \left[\frac{\delta A}{\delta\varphi(\mathbf{x})}\frac{\delta B}{\delta\pi(\mathbf{x})} - \frac{\delta B}{\delta\varphi(\mathbf{x})}\frac{\delta A}{\delta\pi(\mathbf{x})}\right]. \tag{1.87}$$

With the Poisson brackets

$$(\varphi(\mathbf{x}), \pi(\mathbf{y})) = \delta(\mathbf{x} - \mathbf{y}), \quad (\varphi(\mathbf{x}), \varphi(\mathbf{y})) = (\pi(\mathbf{x}, \pi(\mathbf{y})) = 0, \tag{1.88}$$

we now expect the equations of motion to follow from the Hamilton equations:

$$\dot{\varphi}(\mathbf{x}) = (\varphi(\mathbf{x}), H) = \pi(\mathbf{x}), \quad \dot{\pi}(\mathbf{x}) = (\pi(\mathbf{x}), H) = -\frac{\delta}{\delta\varphi(\mathbf{x})}U. \tag{1.89}$$

This is the case indeed (cf. Problem 6). The canonical form of the action principle is

$$S = \int dt \left(\int d^3x\, \pi\dot{\varphi} - H[\varphi, \pi]\right), \tag{1.90}$$

$$\delta_\pi S = \int d^4x\, (\dot{\varphi} - \pi)\delta\pi = 0 \Rightarrow \dot{\varphi} = \pi, \tag{1.91}$$

$$\delta_\varphi S = \int d^4x \left(-\dot{\pi} + \nabla^2\varphi - \frac{\partial V}{\partial\varphi}\right)\delta\varphi = 0 \Rightarrow \dot{\pi} = -\frac{\delta U}{\delta\varphi}. \tag{1.92}$$

We finally note that the hamiltonian has the nice form of an integral over an energy density $\mathcal{H}$. This $\mathcal{H}$ is equal to T^{00} obtained from the variation (1.73) of the matter action $S_m[\varphi, g]$, in the limit of Minkowski spacetime $g_{\mu\nu} \to \eta_{\mu\nu}$. In more general cases there may be a difference between $\mathcal{H}$ and T^{00}, but the total energy is unambiguous, $\int d^3x\, \mathcal{H} = \int d^3x\, T^{00}$.

1.7 Symmetries and Noether's theorem

The invariances of a theory are summarized by the symmetries of its action. This is a nice idea, although it turns out that there are exceptions, so-called *anomalies* where the classical action has more symmetry than the corresponding quantum theory. For now this does not concern us and we shall explore the spacetime symmetries of the action and their consequences according to Noether's theorem:

To each one parameter symmetry group of the action corresponds a conserved quantity.

We shall illustrate how these conserved quantities can be found in a few examples which will be useful later. Consider the action for the scalar field

$$S = \int d^4x \, \mathcal{L}, \quad \mathcal{L} = -\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2. \tag{1.93}$$

The lagrangian (density) $\mathcal{L}$ transforms as a scalar field under Poincaré transformations. If the integration domain M is taken infinite, the action is symmetric (invariant), $S[\varphi'] = S[\varphi]$, with $\varphi'(x) = \varphi(\ell^{-1}x - a)$. This follows from a simple change of integration variables, the jacobian is 1. However, we prefer to keep the integration domain finite so as not to have to discuss convergence of the action integral.

Consider translations. According to Noether there should be four conserved quantities, one for each parameter a^μ, $\mu = 0, 1, 2, 3$. A convenient way to find these is to perform infinitesimal spacetime dependent translations $\epsilon^\mu(x)$ which vanish on the boundary ∂M.[7] Then $\varphi'(x) = \varphi(x - \epsilon(x)) = \varphi(x) - \epsilon^\mu(x)\partial_\mu\varphi(x)$, or

$$\delta\varphi(x) \equiv \varphi'(x) - \varphi(x) = -\epsilon^\mu(x)\partial_\mu\varphi(x). \tag{1.94}$$

Because ϵ depends on x we no longer expect S to be invariant: the derivatives act nontrivially on ϵ. So we may expect the form

$$\delta S = \int d^4x \, T^{\mu\nu}(x)\partial_\mu\epsilon_\nu(x). \tag{1.95}$$

Now, when the equations of motion (field equations) for φ are satisfied, $\delta S = 0$ for any $\delta\varphi$ vanishing on ∂M, hence, also for the variation (1.94). Making a partial integration and using the arbitrariness of $\epsilon_\mu(x)$, we conclude that

$$\begin{aligned} 0 &= \delta S = -\int d^4x \, \partial_\mu T^{\mu\nu}\epsilon_\nu, \\ \Rightarrow & \quad \partial_\mu T^{\mu\nu} = 0, \end{aligned} \tag{1.96}$$

in the interior of M. So we expect four conserved currents, $T^{\mu\nu}$, $\nu = 0, \cdots, 3$, to which correspond four conserved 'charges',

$$P^\nu = \int d^3x \, T^{0\nu}. \tag{1.97}$$

As will be verified below, $T^{\mu\nu}$ is the energy-momentum tensor of the scalar field and P^ν is its total energy-momentum four-vector. Note that Noether's conserved

[7] This method has the advantage that it can be taken over directly in the quantum theory in the path integral formulation.

quantities are not necessarily normalized in standard ways: the way we identified $T^{\mu\nu}$ depends on the normalization of the infinitesimal parameters ϵ^ν. Note furthermore that for P^ν to transform as a four-vector it is crucial that $\partial_\mu T^{\mu\nu} = 0$, see Problem 8.

We now calculate $T^{\mu\nu}$:

$$\begin{aligned}
\delta S &= -\int d^4x \, (\partial^\mu \varphi \partial_\mu \delta\varphi + m^2 \varphi \delta\varphi) \\
&= \int d^4x \, [\partial^\mu \varphi \partial_\mu (\epsilon^\nu \partial_\nu \varphi) + m^2 \varphi \epsilon^\nu \partial_\nu \varphi] \\
&= \int d^4x \, (-\epsilon^\nu \partial_\nu \mathcal{L} + \partial^\mu \varphi \partial_\nu \varphi \partial_\mu \epsilon^\nu) \\
&= \int d^4x \, [\partial_\nu(-\epsilon^\nu \mathcal{L}) + (\partial^\mu \varphi \partial_\nu \varphi + \eta^\mu{}_\nu \mathcal{L}) \partial_\mu \epsilon^\nu] \, .
\end{aligned} \tag{1.98}$$

The integral over the total divergence in (1.98) vanishes because $\epsilon = 0$ on ∂M and comparing with (1.95) we find

$$T^{\mu\nu} = \partial^\mu \varphi \partial^\nu \varphi + \eta^{\mu\nu} \mathcal{L}. \tag{1.99}$$

This is identical to (1.77) obtained from the coupling to the gravitational field, in the limit $g_{\mu\nu} \to \eta_{\mu\nu}$.

The Noether form for the energy-momentum tensor of the electromagnetic field turns out *not* to be equal to the standard form following from (1.73). The integrated P^ν are however identical. The $T^{\mu\nu}$ from (1.73) is more physical as it plays a dynamical role in gravitation (but its derivation is not so easy for spinor fields). The Noether form can be modified by adding a divergence-less quantity which integrates to zero in the total P^ν, such that the energy-momentum tensor satisfies standard criteria such as symmetry,

$$T^{\nu\mu} = T^{\mu\nu}, \tag{1.100}$$

and gauge invariance. For more information, see Ryder, Itzykson & Zuber, Weinberg I, and Problem 9.

The Noether consequence of Lorentz invariance now follows. The scalar field action is invariant under the infinitesimal Lorentz transformation $\ell^\mu{}_\nu = \eta^\mu_\nu + \omega^\mu{}_\nu$, $\varphi'(x) = \varphi(\ell^{-1}x) = \varphi(x) - \omega^{\mu\nu} x_\nu \partial_\mu \varphi(x)$. Making the transformation spacetime dependent (and vanishing on the spacetime boundary) we get

$$\delta\varphi(x) = -\omega^{\mu\nu}(x) x_\nu \partial_\mu \varphi(x). \tag{1.101}$$

This has the form (1.94) with $\epsilon^\mu(x) = \omega^{\mu\nu}(x) x_\nu$. So we can take over eq. (1.95) to get

$$\begin{aligned}
\delta S &= \int d^4x \, T^{\mu\nu} \partial_\mu (\omega_{\nu\rho} x^\rho) \\
&= \int d^4x \, (T^{\mu\nu} x^\rho \partial_\mu \omega_{\nu\rho} + T^{\mu\nu} \omega_{\nu\mu}) \\
&= \frac{1}{2} \int d^4x \, (T^{\mu\nu} x^\rho - T^{\mu\rho} x^\nu) \partial_\mu \omega_{\nu\rho}.
\end{aligned} \tag{1.102}$$

In the last line we used the antisymmetry of $\omega_{\mu\nu}$ and the symmetry of $T^{\mu\nu}$ found in (1.99). Making a partial integration, ∂_μ acts on the expression in parenthesis and because the $\omega_{\mu\nu}$ are arbitrary, the currents

$$J^{\mu\alpha\beta} = x^\alpha T^{\mu\beta} - x^\beta T^{\mu\alpha} \tag{1.103}$$

are divergence-free,

$$\partial_\mu J^{\mu\alpha\beta} = 0, \tag{1.104}$$

and the Noether invariants

$$J^{\alpha\beta} = \int d^3x\, (x^\alpha T^{0\beta} - x^\beta T^{0\alpha}) \tag{1.105}$$

are time-independent. These are the generalized (from rotations to Lorentz transformations) angular momenta of the field. The angular momentum density of the field, J^{0ab}, is easiest to interpret: it has the form of an outer product of x^a with the momentum density T^{0a}, as in $l_{ab} = x_a p_b - x_b p_a$, or $l_a \equiv \frac{1}{2}\epsilon_{abc} l_{ab} = \epsilon_{abc} x_b p_c$ for a single particle.

Note that $J^{0k} = x^0 P^k - \int d^3x\, x^k T^{00}$ depends explicitly on time. This is just right to ensure time-independence in the canonical formalism, according to $dJ^{0k}/dt = \partial J^{0k}/\partial t + (J^{0k}, H) = P^k + (J^{0k}, H)$, where $\partial J^{0k}/\partial t$ denotes the explicit time derivative.

A straightforward calculation (check!) shows that the Poisson brackets of the Noether invariants P_μ and $J_{\mu\nu}$ with the canonical variables φ and π are given by

$$(\varphi(x), P_k) = -\partial_k \varphi(x), \quad (\varphi(x), P_0) = -\pi(x), \tag{1.106}$$

$$(\pi(x), P_k) = -\partial_k \pi(x), \quad (\pi(x), P_0) = -\nabla^2 \varphi(x) + \frac{\partial V(\varphi(x))}{\partial \varphi(x)}, \tag{1.107}$$

which can also be written as

$$(\varphi(x), P_\mu) = -\partial_\mu \varphi(x), \quad (\pi(x), P_\mu) = -\partial_\mu \pi(x). \tag{1.108}$$

Similarly on can write

$$(\varphi(x), J_{\mu\nu}) = -(x_\mu \partial_\nu - x_\nu \partial_\mu)\varphi(x), \tag{1.109}$$

etc. So the P_μ and $J_{\mu\nu}$ generate Poincaré transformations via the Poisson algebra, and indeed, their Poisson brackets are just given by (1.59) with a commutator replaced by i times the Poisson bracket, or $-i[\ ,] \to (\ ,)$:

$$\begin{aligned} (J_{\alpha\beta}, J_{\gamma\delta}) &= \eta_{\alpha\gamma} J_{\beta\delta} - \eta_{\beta\gamma} J_{\alpha\delta} - \eta_{\alpha\delta} J_{\beta\gamma} + \eta_{\beta\delta} J_{\alpha\gamma}, \\ (J_{\alpha\beta}, P_\mu) &= \eta_{\mu\alpha} P_\beta - \eta_{\mu\beta} P_\alpha, \\ (P_\mu, P_\nu) &= 0. \end{aligned} \tag{1.110}$$

1.8 Summary

In field theory the interactions are local. The electromagnetic and gravitational fields are familiar classical fields, the somewhat less familiar relativistic scalar field is often used for simplicity of presentation. We can formulate an action, which is stationary (extremal) when the equations of motion (field equations) are satisfied. The canonical formalism leads to field equations in Hamilton form. Lorentz transformations (including rotations) and translations form a group called the Poincaré group. These transformations are *symmetries* of the action. The Noether invariants corresponding to Poincaré transformations are the total four-momentum $P^\nu = \int d^3x\, T^{0\nu}$ and generalized angular momenta $J^{\alpha\beta} = \int d^3x\, (x^\alpha T^{0\beta} - x^\beta T^{0\alpha})$, with $T^{\mu\nu}$ the energy-momentum tensor. The conservation laws have a local expression in the form of divergence equations, $\partial_\mu T^{\mu\nu} = 0$, $\partial_\mu(x^\alpha T^{\mu\beta} - x^\beta T^{\mu\alpha}) = 0$. The Noether invariants in turn act as generators of the symmetry group and their Poisson brackets satisfy the corresponding Poincaré algebra.

1.9 Appendix: Functional derivative

Consider the action for the scalar field (1.60). It is a function of infinitely many variables, namely $\varphi(x)$ for each $x \in M$. This is the continuum analog of a function of many variables labeled by a discrete index k, say $f(\alpha) = f(\{\alpha_k\})$. In the case of continuous indices like x we speak of a functional.[8] The functional derivative is the analog of the partial derivative. In the function case we can write the change of $f(\alpha)$ under an infinitesimal variation $\delta\alpha_k$ as

$$\delta f(\alpha) = \sum_k g_k(\alpha)\delta\alpha_k, \tag{1.111}$$

with

$$g_k(\alpha) = \frac{\partial f(\alpha)}{\partial \alpha_k} \tag{1.112}$$

the partial derivative. Similarly, in the functional case we can work the variation in the form

$$\delta F[\varphi] = \int d^4x\, G[x,\varphi]\,\delta\varphi(x) \tag{1.113}$$

(we have seen how to do this in the case of the various actions using partial integration), and the functional derivative is defined as

$$\frac{\delta F[\varphi]}{\delta\varphi(x)} = G[x,\varphi]. \tag{1.114}$$

For example, according to (1.61) the functional derivative of the scalar field action is given by

$$\frac{\delta S}{\delta\varphi(x)} = (\partial^2 - m^2)\varphi(x) + J(x), \tag{1.115}$$

[8] We indicate functionals with square brackets.

while the variation (1.68) of the Maxwell field action tells us that

$$\frac{\delta S}{\delta A_\mu(x)} = \partial_\lambda F^{\lambda\mu}(x) + J^\mu(x). \tag{1.116}$$

Let us give some more examples for the case of one dimensional indices x and k:

FUNCTIONAL $F[\varphi]$	FUNCTION $f(\alpha)$
$F = a + \int dx\, b(x)\varphi(x)$	$f = a + \sum_k b_k$
$+ \int dx\, dy\, c(x,y)\varphi(x)\varphi(y) + \cdots$	$+ \sum_{kl} c_{kl}\alpha_k\alpha_l + \cdots$
$\Rightarrow \frac{\delta F}{\delta\varphi(x)} = b(x) + 2\int dy\, c(x,y)\varphi(y) + \cdots$	$\Rightarrow \frac{\partial f}{\partial \alpha_k} = b_k + 2\sum_l c_{kl}\alpha_l + \cdots$
$F = \varphi(x) \Rightarrow \frac{\delta F}{\delta\varphi(y)} = \delta(x-y)$	$f = \alpha_k \Rightarrow \frac{\partial f}{\partial\alpha_l} = \delta_{kl}$
$\frac{\delta\varphi(x)^n}{\delta\varphi(y)} = n\varphi(x)^{n-1}\delta(x-y)$	$\frac{\partial\alpha_k^n}{\partial\alpha_l} = n\alpha_k^{n-1}\delta_{kl}$
$\frac{\delta}{\delta\varphi(y)}\left[\frac{\partial}{\partial x}\varphi(x)\right] = \frac{\partial}{\partial x}\,\delta(x-y)$	$\frac{\partial}{\partial\alpha_l}(\alpha_{k+1} - \alpha_k) = \delta_{k+1,l} - \delta_{k,l},$

where $\delta(x-y)$ and δ_{kl} are the Dirac and Kronecker delta functions.

1.10 Appendix: Rudiments of representations of continuous groups

A group G is a collection of elements $g \in G$ with the properties

- there is a multiplication rule: $g_1 \in G$, $g_2 \in G \rightarrow g_1 g_2 \in G$
- there is a unit element e (often denoted by 1): $ge = eg = g$
- there is a inverse g^{-1}: $gg^{-1} = g^{-1}g = e$

Examples:

discrete groups of 90° rotations, groups of continuous rotations in 1,2,... dimensions

abelian group (named after Abel) satisfy $g_1 g_2 = g_2 g_1$ (for all $g_{1,2} \in G$)
non-abelian group $g_1 g_2 \neq g_2 g_1$ (not necessarily for all $g_{1,2} \in G$)

Elements of continuous groups, such as the rotation group in three dimensions, or the Poincaré group, can be written in exponential form

$$g = e^{i\omega_p S_p} \equiv e + \sum_{n=1}^{\infty} \frac{1}{n!}\,(i\omega_p S_p)^n \tag{1.117}$$

(summation over repeated indices), where ω_p are parameters (real numbers) and S_p are called generators of the group. For example, for rotations, ω_p, $p = 1, 2, 3$

are angles of rotation and S_p are hermitian 3×3 matrices. The commutator of two generators is a superposition of generators

$$[S_p, S_q] = i f_{pqr} \, S_r, \tag{1.118}$$

in which the coefficients (real numbers) f_{pqr} are called the structure constants of the group.

A representation of the group is a mapping $g \to D(g)$ which is itself a group, with

$$g_1 g_2 = g_3 \to D(g_1)D(g_2) = D(g_3). \tag{1.119}$$

Elements $D^{(j)}(g)$ of a representation j of G can also be written in exponential form,

$$g = e^{i\omega_p S_p} \to D^{(j)}(g) = e^{i\omega_p S_p^{(j)}}, \tag{1.120}$$

in which the $S_p^{(j)}$ are called the generators in the representation j. They satisfy the same commutation relations as the original S_p:

$$[S_p^{(j)}, S_q^{(j)}] = i f_{pqr} \, S_r^{(j)}. \tag{1.121}$$

Conversely, if we are able to construct generators $S_p^{(j)}$ satisfying (1.121), the we have also constructed a representation of the group G:

$$g = e^{i\omega_p S_p}, \; g' = e^{i\omega'_p S_p}, \; g'' = gg' = e^{i\omega''_p S_p} \to D^{(j)}(g'') = e^{i\omega''_p S_p^{(j)}}. \tag{1.122}$$

This follows from (1.117), (1.120), (1.118), (1.121) and the Baker-Campbell-Haussdorff formula

$$e^M \, e^{M'} = e^{M+M'+(1/2)[M,M']+\cdots},$$

where the ... consist of multiple commutators of M and M'.

1.11 Problems

1. *Mathematical methods*

   Asymptotic expansions occur frequently.

   a. Familiarize your selve with the stationary phase approximation and the saddle point approximation. See for example Mathews and Walker, or Dingle.

   b. We will freely consider Fourier integrals which do not converge. Such integrals have meaning within the theory of generalized functions (distributions). See for example Lighthill. As an important example, consider

$$\tilde{\theta}(x) = \int_{-\infty}^{\infty} dt \, e^{ixt} \, \theta(t), \tag{1.123}$$

where $\theta(t)$ is the Heavyside step function,

$$\begin{aligned} \theta(t) &= 1, \quad t > 0, \\ &= 0, \quad t < 0. \end{aligned} \tag{1.124}$$

This Fourier integral may be evaluated by introducing a convergence factor $\exp(-\epsilon t)$, where ϵ is arbitrarily small postive. Then

$$\tilde{\theta}(x) = \frac{i}{x + i\epsilon}. \tag{1.125}$$

We have the following identity between distributions

$$\frac{1}{x + i\epsilon} = P\frac{1}{x} - i\pi\delta(x). \tag{1.126}$$

Here $\delta(x)$ is the Dirac delta function and P denotes the Cauchy principal value: if $f(x)$ is a test function, then

$$\int dx\, P\frac{1}{x}\, f(x) \equiv \lim_{\delta \downarrow 0} \left[\int_{\delta}^{\infty} dx\, \frac{f(x)}{x} + \int_{-\infty}^{-\delta} dx\, \frac{f(x)}{x} \right]. \tag{1.127}$$

For finite ϵ, make a plot of the real and imaginary parts of $1/(x + i\epsilon)$ and convince yourself that (1.126) is correct in the limit $\epsilon \downarrow 0$.

Using the identity $\theta(t) + \theta(-t) = 1$, verify that

$$\tilde{1}(x) \equiv \int_{-\infty}^{\infty} dt\, e^{ixt} = 2\pi\delta(x)$$

c. Consider the integral

$$I(t) = \int_{-\infty}^{\infty} dx\, f(x) e^{ixt}, \tag{1.128}$$

which we assume to be convergent. Let C be the closed contour in the complex z plane from $(-R, 0)$ to $(+R, 0)$ on the real axis and then back along a semicircle in the upper half plane with radius R.

Show that for $t > 0$ the above integral can be evaluated as

$$I(t) = \lim_{R \to \infty} \int_C dz\, f(z) e^{izt} \tag{1.129}$$

(assume for example that $f(z) \to 0$ like $1/|z|$ as $|z| \to \infty$). Verify that for $t < 0$ the integral along the semicircle in the upper half plane does *not* converge, and that the analog contour integral for this case is along a semicircle in the *lower* half plane.

2. *Fourier and Cauchy methods*

To prove (1.17) we may use the Fourier representation

$$\varphi(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{kx}}\, \tilde{\varphi}(\mathbf{k}), \tag{1.130}$$

together with

$$\delta(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{kx}}. \tag{1.131}$$

Then $\tilde{\varphi}$ has to satisfy

$$(\mathbf{k}^2 + m^2)\tilde{\varphi}(\mathbf{k}) = 1 \to \varphi(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}\, \frac{e^{i\mathbf{kx}}}{m^2 + \mathbf{k}^2}. \tag{1.132}$$

This integral may be done in spherical coordinates with $\mathbf{kx} = kr\cos\theta$, integrating first over angles,

$$\begin{aligned} \varphi(\mathbf{x}) &= \frac{1}{(2\pi)^2}\int_0^\infty dk\, \frac{k^2}{m^2+k^2}\, \frac{2\sin kr}{kr}, & (1.133)\\ &= \frac{1}{4\pi r}\mathrm{Re}\int_{-\infty}^\infty \frac{dk}{2\pi i}\, \frac{2k\, e^{ikr}}{m^2+k^2}, & (1.134)\end{aligned}$$

and then over k using contour integration, by closing the contour in the upper half of the complex k-plane and picking up the pole at $k = im$ (cf. Problem 1.c).

Verify this.

3. *Green functions*

Solutions of the Klein-Gordon equation with source can be obtained in terms of Green functions,

$$\begin{aligned} \varphi(x) &= \varphi_{\mathrm{in}}(x) + \int d^4x'\, G^R(x-x')J(x'),\\ \varphi(x) &= \varphi_{\mathrm{out}}(x) + \int d^4x'\, G^A(x-x')J(x'),\end{aligned}$$

where φ_{in} and φ_{out} are solutions of the homogenous equation $(-\partial^2+m^2)\varphi_{\mathrm{in,out}} = 0$, and G^R and G^A are retarded and advanced Green functions, which satisfy

$$(-\partial^2 + m^2)G^{R,A}(x-x') = \delta^4(x-x'). \tag{1.135}$$

We assume here infinite spacetime.

The names retarted and advanced indicate that

$$\begin{aligned} G^R(x-x') &= 0, \quad x^0 < x^{0\prime},\\ G^A(x-x') &= 0, \quad x^0 > x^{0\prime}. \end{aligned} \tag{1.136}$$

In fact, $G^{R/A}(x-x')$ is nonzero only in the future/past light cone of x' (cf. Problem 4). Assuming that the source $J(x)$ is nonzero only in a bounded region of spacetime, it follows that $\varphi(x) \to \varphi_{\substack{\text{out}\\ \text{in}}}(x)$ for $x^0 \to \pm\infty$.

The Green functions may be represented as Fourier transforms

$$G^{\overset{A}{R}}(x) = \int \frac{d^4p}{(2\pi)^4}\, e^{ipx}\, \frac{1}{m^2 + \mathbf{p}^2 - (p^0 \mp i\epsilon)^2}, \tag{1.137}$$

where $\epsilon > 0$ is infinitesimal, i.e. it is arbitrarily small, positive, and will be set to zero where its presence is not needed anymore. In the representation above it is needed to prescribe how the poles at $p^0 = \pm\sqrt{m^2 + \mathbf{p}^2}$ are to be avoided in the integral.

Given (1.137), verify eqs. (1.135).

By evaluating the integral over p^0 in (1.137), verify that

$$G^{\overset{A}{R}}(x) = \mp\theta(\mp x^0)\Delta(x), \tag{1.138}$$

where $\theta(x^0)$ is the step function (1.124) and

$$\Delta(x) = i\int \frac{d^3p}{(2\pi)^3 2p^0}\left(e^{ipx} - e^{-ipx}\right), \quad p^0 = \sqrt{\mathbf{p}^2 + m^2}. \tag{1.139}$$

Eqs. (1.136) follow from (1.138).

4. *Lorentz invariance of* $G^{A/R}$

The advanced and retarted Green functions are Lorentz invariant. This is suggested by the representation (1.137) if we ignore the infinitesimal ϵ. More precisely it follows from:

i) the Lorentz invariance $\Delta(x)$ in (1.139),
ii) the fact that this function vanishes for spacelike separations,

$$\Delta(x) = 0, \quad x^2 > 0. \tag{1.140}$$

Under these circumstances the distinction $x^0 > 0$ or < 0 in (1.138) is a Lorentz invariant property. The advanced/retarted Green functions are nonzero only inside the past/future light cones.

Property i) relies on the Lorentz invariance of the integration volume element

$$d\omega_p \equiv \frac{d^3p}{(2\pi)^2 2p^0}, \tag{1.141}$$

i.e.

$$d\omega_{\ell p} = d\omega_p. \tag{1.142}$$

Verify this for a Lorentz tranformation along the 3-axis with velocity $v < 1$:

$$p'^0 = \gamma p^0 + \gamma v p^3, \quad p'^3 = \gamma p^3 + \gamma v p^0, \quad p'^1 = p^1, \quad p'^2 = p^2, \tag{1.143}$$

where $\gamma = 1/\sqrt{1-v^2}$ is the relativistic dilatation factor.

Property ii) follows from the fact that a spacelike interval $(x^0, \mathbf{x})$ can be transformed into one with time separation zero, $(0, \mathbf{x}')$, by a Lorentz transformation, and the fact that $\Delta(0, \mathbf{x}) = 0$.
Verify.

5. *$T^{\mu\nu}$ from coupling to $g_{\mu\nu}$*

Verify the expressions (1.77) and (1.78) for the energy-momentum tensors of the scalar and electromagnetic fields, using (1.73) as the definition of $T^{\mu\nu}$.

For the variations the folowing formulas are useful. Let $\hat{g}$ denote the matrix with matrix elements $g_{\mu\nu}$. Then ($g^{\mu\nu} \equiv (\hat{g}^{-1})_{\mu\nu}$):

$$\hat{g}\hat{g}^{-1} = 1 \rightarrow \delta\hat{g}\,\hat{g}^{-1} + \hat{g}\,\delta\hat{g}^{-1} = 0 \rightarrow \delta\hat{g}^{-1} = -\hat{g}^{-1}\delta\hat{g}\hat{g}^{-1}. \tag{1.144}$$

$$\det\hat{g} = \frac{1}{4!}\epsilon^{\kappa\lambda\mu\nu}\epsilon^{\alpha\beta\gamma\delta}g_{\kappa\alpha}g_{\lambda\beta}g_{\mu\gamma}g_{\nu\delta}, \tag{1.145}$$

so

$$\begin{aligned}\delta\det\hat{g} &= \frac{1}{3!}\epsilon^{\kappa\lambda\mu\nu}\epsilon^{\alpha\beta\gamma\delta}g_{\lambda\beta}g_{\mu\gamma}g_{\nu\delta}\delta g_{\kappa\alpha} \\ &= \det\hat{g}\,g^{\kappa\alpha}\delta g_{\kappa\alpha}.\end{aligned} \tag{1.146}$$

(We implicitly derived Cramer's rule for the inverse of a matrix.)

6. *Hamilton equations*

Verify the statement following eq. (1.89) in steps:

a) Derive the field equations (equations of motion) form the action principle for the φ^4 theory given by (1.80).

b) Calculate the Poisson brackets of $\varphi(\mathbf{x})$ and $\pi(\mathbf{x})$ with the hamiltonian $H[\varphi, \pi]$ and verify in this way that Hamilton's equations are equivalent to the field equations found under a).

7. *Functional derivatives may give distributions*

From eq. (1.61) follows that

$$\frac{\delta S}{\delta\varphi(x)} = (\partial^2 - m^2)\varphi(x) + J(x). \tag{1.147}$$

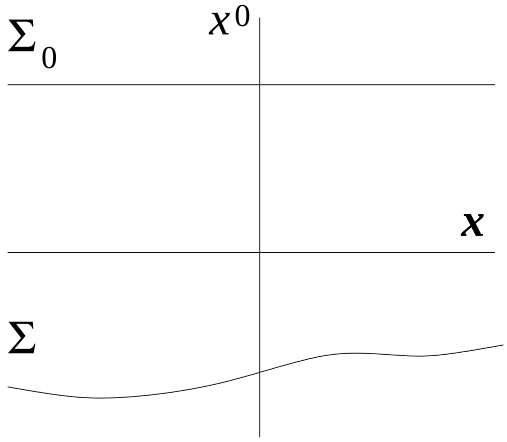

Figure 1.3: The constant-time hyperplane Σ_0 and an arbitrary spacelike hypersurface Σ in Minskowski space.

Verify that differentiating this once again gives

$$\frac{\delta^2 S}{\delta\varphi(x)\delta\varphi(x')} = (\partial^2 - m^2)\delta^4(x - x'). \tag{1.148}$$

This can be viewed as a 'matrix' with continuous indices x and x'. Calculate the action of this matrix on the 'vector' $\varphi(x')$, i.e.
$\int d^4x' \, [\delta^2 S/\delta\varphi(x)\delta\varphi(x')] \, \varphi(x')$.

8. *Conserved charges and Lorentz invariance*

   Let $j^\mu(x)$ be a conserved current (jargon for 'divergence-free current density'), i.e. $\partial_\mu j^\mu = 0$. Assume it drops to zero faster than $1/x^2$ in spacelike directions. Show that the corresponding charge

$$Q = \int d^3x \, j^0(x) \tag{1.149}$$

   is conserved (time independent).

   By integrating $\partial_\mu j^\mu$ over the four-volume between the spacelike hypersurfaces Σ and Σ_0 and using Gauss's law, verify that

$$Q = \int_\Sigma d\sigma_\mu(x) \, j^\mu(x). \tag{1.150}$$

   Show that Q is Lorentz invariant, both from the active and passive point of view.

   Hint for the active case: Let $j'^\mu(x) = \ell^\mu{}_\nu j^\nu(\ell^{-1}x)$ be the current of the transformed system. To be shown is $Q' = \int d^3x \, j'^0(x) = \int d^3x \, j^0(x) = Q$. Make a transformation of variables $x = \ell y$, and use the identity

$$\epsilon_{\mu\alpha\beta\gamma}\ell^\alpha{}_{\alpha'}\ell^\beta{}_{\beta'}\ell^\gamma{}_{\gamma'} = \ell_\mu{}^{\mu'}\epsilon_{\mu'\alpha'\beta'\gamma'} \det(\ell) \tag{1.151}$$

(verify!). Then $d\sigma_\mu(x) = \ell_\mu{}^\rho d\sigma_\rho(y)$, and it follows that

$$\begin{aligned} Q' = \int_{\Sigma_0} d\sigma_\mu(x) j'^\mu(x) &= \int_{\Sigma_1} d\sigma_\nu(y) j^\nu(y) \\ &= \int_{\Sigma_0} d\sigma_\mu(y) j^\mu(y) = Q, \end{aligned} \tag{1.152}$$

where Σ_1 is the hypersurface for the integral over the (dummy) y-variables corresponding to the hypersurface x^0 = constant. In the last step we used the previous result (1.150). Make a sketch of Σ_0 and σ_1 similar to figure 1.3, for the case of a boost in the 3-direction.

Similarly, show that $P^\mu = \int d^3x\, T^{0\mu}$ transforms as a four-vector.

9. *Noether and gravitational energy-momentum tensor*

Consider the action (1.75) for a scalar field coupled minimally to the Einstein field, $S = S[\varphi, g]$. Let $T^{\mu\nu}$ be the energy-momentum tensor that is the source in the Einstein field equations, which is given by (1.73), and let $T_{\rm N}^{\mu\nu}$ the one found by the Noether method. We have seen that $T^{\mu\nu} = T_{\rm N}^{\mu\nu}$ in the Minkowski limit $g_{\mu\nu} = \eta_{\mu\nu}$. In this problem we want to shed some light on this equality. The action S is invariant under the general coordinate transformations

$$x'^\mu = f^\mu(x), \quad \varphi'(x') = \varphi(x), \quad g'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x^\sigma}\frac{\partial x'^\mu}{\partial x'^\nu}\, g_{\rho\sigma}(x), \tag{1.153}$$

i.e. $S[\varphi', g'] = S[\varphi, g]$. Consider an infinitesimal transformation

$$x'^\mu = x^\mu + \epsilon^\mu(x). \tag{1.154}$$

Then, to first order in ϵ^μ,

$$\begin{aligned} \delta\varphi(x) &\equiv \varphi'(x) - \varphi(x) = -\epsilon^\mu(x)\partial_\mu\varphi(x), \qquad (1.155) \\ \delta g_{\mu\nu(x)} &\equiv g'_{\mu\nu}(x) - g_{\mu\nu}(x) = -\epsilon^\rho(x)\partial_\rho g_{\mu\nu}(x) - g_{\mu\rho}(x)\partial_\nu\epsilon^\rho(x) - g_{\nu\rho}(x)\partial_\mu\epsilon^\rho(x). \end{aligned}$$

The invariance of S has the consequence

$$0 = \delta S = \int d^4x \left(\frac{\delta S}{\delta g_{\mu\nu}}\delta g_{\mu\nu} + \frac{\delta S}{\delta\varphi}\delta\varphi\right). \tag{1.156}$$

The first term on the right hand side of (1.156) gives

$$\int d^4x\, \sqrt{-g}\, \frac{1}{2} T^{\mu\nu}\delta g_{\mu\nu} = \int d^4x\, T^{\mu\nu}(-\partial_\mu\epsilon_\nu), \tag{1.157}$$

in the Minkowski limit.

The infinitesimal coordinate transformation looks like a spacetime dependent translation, and (1.155) is identical to the variation (1.94) used in the derivation of $T_{\rm N}^{\mu\nu}$. So the second term on the right hand side of (1.156) produces, as in (1.95),

$$\int d^4x\, T_{\rm N}^{\mu\nu}\partial_\mu\epsilon_\nu. \tag{1.158}$$

Since the sum of the two terms is zero we conclude that

$$\partial_\mu T^{\mu\nu} = \partial_\mu T_{\rm N}^{\mu\nu}. \tag{1.159}$$

Note that we have not assumed that the equations of motion are valid: the above identity holds for arbitrary φ.

For our minimally coupled action we actually found the stronger identity $T^{\mu\nu} = T_{\rm N}^{\mu\nu}$, but this does not have to be the case. For example, adding to the minimally coupled action the term

$$-\int d^4x\,\sqrt{-g}\,\frac{1}{2}\xi R\varphi^2, \tag{1.160}$$

leads to a different $T^{\mu\nu}$. In the Minkowski limit it is given by

$$T_\xi^{\mu\nu} = T^{\mu\nu} - \xi(\partial^\mu\partial^\nu - \eta^{\mu\nu}\partial^2)\varphi^2, \tag{1.161}$$

where $T^{\mu\nu}$ is the old tensor corresponding to $\xi = 0$.

Verify that $T_\xi^{\mu\nu}$ and $T^{\mu\nu}$ lead to the same total energy-momentum, $\int d^3x\, T_\xi^{0\nu} = \int d^3x\, T^{0\nu}$.

For $\xi = 1/6$ this new energy-momentum tensor is sometimes called the 'improved energy-momentum tensor'. Verify, using the equations of motion, that the improved tensor satisfies

$$\eta_{\mu\nu}T_{1/6}^{\mu\nu} = -m^2\varphi^2. \tag{1.162}$$

So the trace of $T_{1/6}^{\mu\nu}$ vanishes for $m^2 = 0$. This property is relevant for the description of scale invariance.

10. *The exponential parametrization*

    Let $S_{\alpha\beta}$ be the generators (1.52) in the defining representation. Verify by expansion that $\ell = \exp(-i\phi S_{12})$ is equal to the rotation matrix (1.36). Verify that $\ell = \exp(i\chi S_{03})$ is the boost matrix (1.35).

11. *Vector-matrices*

    Let $R = \exp(-i\phi^k S_k)$ be a general rotation matrix, where $S_k \equiv (1/2)\epsilon_{klm}\, S_{lm}$. The boost generators S_{0k} are vector-matrices (vectors for short) under rotations in the sense that $R^{-1}S_{0k}R = R^k{}_l S_{0l}$. Verify this for infinitesimal rotations (you need to check the commutation relations $[S_k, S_{0l}] = i\epsilon_{klm}\, S_{0m}$).

Chapter 2

Quantized scalar fields

This chapter introduces the basics of the theory of the canonically quantized scalar field. We introduce the particle interpretation, touch briefly on applications to scattering and decay processes, and discuss Lorentz symmetry and locality in the quantum theory.

2.1 Canonical quantization

Consider the φ^4 model described by the action

$$S = -\int d^4x \left[\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi + V(\varphi)\right], \tag{2.1}$$

$$V(\varphi) = \epsilon + \frac{1}{2}\kappa\varphi^2 + \frac{1}{4}\lambda\varphi^4. \tag{2.2}$$

For later use we have added a constant ϵ to (1.80). It may be interpreted as representing the cosmological constant. In a world in which the above action would describe all existing matter, shifting Λ from the graviational field action S_g to the matter action S_m would give $\epsilon = \Lambda/8\pi G$ (cf. (1.70)).

For interpretation of the quantized field the expressions for the energy, momentum etc. (the Noether invariants) are important, in particular

$$P^0 = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\varphi)^2 + V(\varphi)\right] = H(\varphi, \pi), \tag{2.3}$$

$$P_k = -\int d^3x\, \pi\partial_k\varphi, \tag{2.4}$$

$$J_k = -\int d^3x\, \pi(\epsilon_{klm}x_l\partial_m)\varphi. \tag{2.5}$$

We now quantize the theory by replacing the classical fields φ and π by operators φ and π in Hilbert space, such that their commutators go over into their Poisson brackets in the formal classical correspondence limit $\hbar \to 0$: $[\ ,\ ]/i\hbar \to (\ ,\ )$. So at say, $t = 0$, we put (keeping Planck's constant $\hbar$ explicit for the moment)

$$[\varphi(\mathbf{x}), \pi(\mathbf{y})] = i\hbar\delta^3(\mathbf{x} - \mathbf{y}), \quad [\varphi(\mathbf{x}), \varphi(\mathbf{y})] = [\pi(\mathbf{x}), \pi(\mathbf{y})] = 0. \tag{2.6}$$

These are called the canonical commutation relations. In the Heisenberg picture (where the operators are time dependent) they are supposed to hold at equal times. The above relations are a straightforward generalization of the case of discretely many variables. One realization of the commutation relations is the coordinate representation:

$$\varphi(\mathbf{x}) \to \text{multiplication by } \varphi(\mathbf{x}), \quad \pi(\mathbf{x}) \to \frac{\hbar\delta}{i\delta\varphi(\mathbf{x})}, \tag{2.7}$$

acting on Schrödinger wave *functionals* $\psi[\varphi]$. This realization is basic to the path integral approach which will be introduced later. In this chapter we follow another approach which is geared to the particle interpretation of the quantized field.

2.2 Free field

For $\lambda = 0$ the hamiltonian is at most quadratic in the canonical variables. For reasons that will become clear later on we change the notation in the quantum theory by adding a subscript zero to the parameters in the action, so the hamiltonian is given by

$$H = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}\kappa_0\varphi^2 + \epsilon_0\right]. \tag{2.8}$$

This model with hamiltonian quadratic in the fields is called the free theory, because it is equivalent to a collection of uncoupled harmonic oscillators, as will now be shown by going over to 'momentum space'. To simplify the presentation we first assume only one spatial dimension. Afterwards, we can easily generalize to three spatial dimensions. We furthermore assume space to be a circle with circumference L, i.e. $0 \leq x \leq L$ with periodic boundary conditions at $0, L$ and $\int dx = \int_0^L dx$. We expand the fields at time $t = 0$ in Fourier modes,

$$\varphi(x) = \frac{1}{\sqrt{L}}\sum_p e^{ipx}\,\tilde{\varphi}_p, \quad \pi(x) = \frac{1}{\sqrt{L}}\sum_p e^{ipx}\,\tilde{\pi}_p, \tag{2.9}$$

$$\tilde{\varphi}_p = \frac{1}{\sqrt{L}}\int_0^L dx\, e^{-ipx}\,\varphi(x), \quad \tilde{\pi}_p = \frac{1}{\sqrt{L}}\int_0^L dx\, e^{-ipx}\,\pi(x), \tag{2.10}$$

where $p = 2\pi n/L$, $n = 0, \pm 1, \pm 2, \cdots$. The modes are eigenfunctions of the gradient operator $\partial/\partial x$ with periodic boundary conditions. Since the fields are hermitian, $\varphi^\dagger(x) = \varphi(x)$, the Fourier components satisfy the relations

$$\tilde{\varphi}_p^\dagger = \tilde{\varphi}_{-p}, \quad \tilde{\pi}_p^\dagger = \tilde{\pi}_{-p}. \tag{2.11}$$

The hamiltonian and the momentum operator are diagonal in this representation:

$$H = \sum_p \frac{1}{2}[\tilde{\pi}_p^\dagger \tilde{\pi}_p + (p^2+\kappa_0)\tilde{\varphi}_p^\dagger \tilde{\varphi}_p] + \epsilon_0 L, \tag{2.12}$$

$$P = -\sum_p \tilde{\pi}_p^\dagger \tilde{\varphi}_p ip. \tag{2.13}$$

The hamiltonian looks like that of a sum of harmonic oscillators with frequencies

$$\omega_p = \sqrt{p^2+m^2}, \quad m^2 = \kappa_0, \tag{2.14}$$

where we have chosen $\kappa_0 > 0$. As in the case of the harmonic oscillator, it is very useful to introduce creation and annihilation operators, $a_p^\dagger$ and a_p, one for each mode:

$$a_p = \frac{1}{\sqrt{2\omega_p}}(\omega_p \tilde{\varphi}_p + i\tilde{\pi}_p), \quad a_p^\dagger = \frac{1}{\sqrt{2\omega_p}}(\omega_p \tilde{\varphi}_{-p} - i\tilde{\pi}_{-p}), \tag{2.15}$$

$$\tilde{\varphi}_p = \frac{1}{\sqrt{2\omega_p}}(a_p + a_{-p}^\dagger), \quad \tilde{\pi}_p = \frac{1}{\sqrt{2\omega_p}}(-i\omega_p a_p + i\omega_p a_{-p}^\dagger), \tag{2.16}$$

where we used (2.11). The creation and annihilation operators satisfy the commutation relations

$$[a_p, a_q^\dagger] = \delta_{pq}, \quad [a_p, a_q] = [a_p^\dagger, a_q^\dagger] = 0. \tag{2.17}$$

The hamiltonian and the momentum operator can now be written in the form

$$H = \sum_p \left(a_p^\dagger a_p + \frac{1}{2}\right)\omega_p + \epsilon_0 L, \tag{2.18}$$

$$P = \sum_p a_p^\dagger a_p\, p. \tag{2.19}$$

We see that the hamiltonian is just a sum of independent harmonic oscillators. The ground state, i.e. the state with lowest energy, is given by (as usual, up to a phase factor)

$$a_p|0\rangle = 0, \quad \text{for all } p, \quad \langle 0|0\rangle = 1. \tag{2.20}$$

It is also an eigenstate of P with eigenvalue zero. The other simultaneous eigenstates of H and P are obtained from the ground state $|0\rangle$ by application of the creation operators,

$$|\{n_p\}\rangle = \prod_p \frac{(a_p^\dagger)^{n_p}}{\sqrt{n_p!}}|0\rangle, \tag{2.21}$$

with occupation numbers $n_p = 0, 1, \cdots$. All eigenstates are normalized to 1. The eigenvalues of H and P are given by

$$H|\{n_p\}\rangle = \left(E_0 + \sum_p n_p \omega_p\right)|\{n_p\}\rangle, \quad E_0 = \epsilon_0 L + \sum_p \frac{1}{2}\omega_p, \tag{2.22}$$

$$P|\{n_p\}\rangle = \left(\sum_p n_p\, p\right)|\{n_p\}\rangle. \tag{2.23}$$

Consider now the ground state energy density:

$$\epsilon \equiv \frac{E_0}{L} = \epsilon_0 + \frac{1}{L}\sum_p \frac{1}{2}\omega_p \quad (2.24)$$

$$\to \epsilon_0 + \int_{-\infty}^{\infty} \frac{dp}{2\pi}\,\frac{1}{2}\sqrt{p^2+m^2}, \quad L \to \infty. \quad (2.25)$$

The integral in the last line is the limit of a Riemann sum for $L \to \infty$:

$$\frac{1}{L}\sum_p F(p) = \frac{1}{2\pi}\sum_p \Delta p\, F(p) \to \int_{-\infty}^{\infty} \frac{dp}{2\pi}\, F(p), \quad \Delta p = \frac{2\pi}{L}. \quad (2.26)$$

The ground state energy as written is infinite, because the integral diverges at large p. The reason is that we are dealing with an infinite number of degrees of freedom. However, we can absorb this infinity in ϵ_0, such that ϵ is finite. We come back to this shortly.

We now generalize to three spatial dimensions. Let us choose ϵ_0 such that $\epsilon = 0$. Then we can summarize as follows:

$$\tilde{\varphi}_{\mathbf{p}} = \int d^3x\, \frac{e^{-i\mathbf{px}}}{\sqrt{L^3}}\, \varphi(\mathbf{x}), \quad \text{etc.}, \quad (2.27)$$

$$a_{\mathbf{p}} = \frac{1}{\sqrt{2\omega_{\mathbf{p}}}}\left(\omega_{\mathbf{p}}\tilde{\varphi}_{\mathbf{p}} + i\tilde{\pi}_{\mathbf{p}}\right), \quad (2.28)$$

$$\varphi(\mathbf{x}) = \sum_{\mathbf{p}}\left[a_{\mathbf{p}}\,\frac{e^{i\mathbf{px}}}{\sqrt{2\omega_{\mathbf{p}}L^3}} + a_{\mathbf{p}}^\dagger\,\frac{e^{-i\mathbf{px}}}{\sqrt{2\omega_{\mathbf{p}}L^3}}\right], \quad (2.29)$$

$$\pi(\mathbf{x}) = \sum_{\mathbf{p}}\left[-i\omega_{\mathbf{p}}a_{\mathbf{p}}\,\frac{e^{i\mathbf{px}}}{\sqrt{2\omega_{\mathbf{p}}L^3}} + i\omega_{\mathbf{p}}a_{\mathbf{p}}^\dagger\,\frac{e^{-i\mathbf{px}}}{\sqrt{2\omega_{\mathbf{p}}L^3}}\right], \quad (2.30)$$

$$[a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger] = \delta_{\mathbf{p},\mathbf{q}}, \quad [a_{\mathbf{p}}, a_{\mathbf{q}}] = [a_{\mathbf{p}}^\dagger, a_{\mathbf{q}}^\dagger] = 0, \quad (2.31)$$

$$P^\mu = \sum_{\mathbf{p}} a_{\mathbf{p}}^\dagger a_{\mathbf{p}}\, p^\mu, \quad P^0 = H, \quad p^0 = \omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}, \quad (2.32)$$

$$P^\mu|0\rangle = 0, \quad P^\mu|\mathbf{p}\rangle = p^\mu|\mathbf{p}\rangle, \quad |\mathbf{p}\rangle \equiv a_{\mathbf{p}}^\dagger|0\rangle = |1_{\mathbf{p}}\rangle, \quad (2.33)$$

$$P^\mu|\mathbf{p}_1\mathbf{p}_2\rangle = (p_1^\mu + p_2^\mu)|\mathbf{p}_1\mathbf{p}_2\rangle, \quad |\mathbf{p}_1\mathbf{p}_2\rangle \equiv a_{\mathbf{p}_1}^\dagger a_{\mathbf{p}_2}^\dagger|0\rangle, \quad (2.34)$$

etc. In (2.33) we used the convention that only non-zero occupation numbers are shown in the ket.

The interpretation of the scalar field model in terms of a collection of free particles is very suggestive. The ground state $|0\rangle$ is interpreted as representing the vacuum. The one particle state $|\mathbf{p}\rangle$ is the state with $n_{\mathbf{p}} = 1$ and all other $n_{\mathbf{q}} = 0$, $\mathbf{q} \neq \mathbf{p}$. The *mass* of the particles is $m = \sqrt{\kappa_0}$. Their *spin* is zero since their is no further index besides $\mathbf{p}$ to indicate a spin degree of freedom. More formally, it follows from (2.119) below that a particle state at rest ($\mathbf{p} = 0$) is invariant under rotations, so its total angular momentum is identically zero and the particles are spinless.

The two particle state[1] $|\mathbf{p}_1\mathbf{p}_2\rangle$ is symmetric in the interchange of the labels $\mathbf{p}_1$ and $\mathbf{p}_2$: the particles are *bosons*.

2.3 Renormalization of the cosmological constant

We now return to the energy density in the groundstate, ϵ. It is the vacuum expectation value of T^{00}. Calculating the expectation value of the full energy-momentum tensor gives in the infinite volume limit

$$\langle 0|T^{\mu\nu}(x)|0\rangle = -\epsilon_0\eta^{\mu\nu} + I^{\mu\nu}, \quad I^{\mu\nu} = \int d\omega_p\, p^\mu p^\nu \tag{2.35}$$

(independent of x in accordance to translation invariance), where we introduced the notation

$$d\omega_p \equiv \frac{d^3p}{(2\pi)^3 2p^0}. \tag{2.36}$$

Apart from conveniently absorbing numerical factors, this volume element of integration has the important property that it is Lorentz invariant (cf. Problem 1.4):

$$d\omega_{\ell p} = d\omega_p. \tag{2.37}$$

It follows that the integral is an invariant tensor under Lorentz transformations,

$$\ell^\mu_\alpha \ell^\nu_\beta I^{\alpha\beta} = \ell^\mu_\alpha \ell^\nu_\beta \int d\omega_p\, p^\alpha p^\beta = \int d\omega_{\ell p}\, (\ell p)^\mu (\ell p)^\nu = \int d\omega_p\, p^\mu p^\nu = I^{\mu\nu}. \tag{2.38}$$

There are only two independent Lorentz-invariant tensors, $\eta_{\mu\nu}$ and $\epsilon_{\kappa\lambda\mu\nu}$ (check!). Hence, the vacuum expectation value of $T^{\mu\nu}$ is proportional to $\eta^{\mu\nu}$:

$$\langle 0|T^{\mu\nu}(x)|0\rangle = -\epsilon\eta^{\mu\nu}. \tag{2.39}$$

We can now interpret ϵ as the true contribution to the cosmological constant, while ϵ_0 is just a parameter in the action. In standard jargon, $8\pi G\epsilon$ is the renormalized (or dressed) cosmological constant, and $8\pi G\epsilon_0$ the bare cosmological constant.

However, the integral (2.35) is badly divergent at large momenta. To make sense of it we should regularize it. Even better, we can start with a regularized formulation of the theory such that at every stage we have well defined expressions. This can be done, e.g. by replacing the spacetime continuum by a lattice, but it is cumbersome and we have learned that in many cases it is sufficient to

[1]This state can also be written as $|1_{\mathbf{p}_1}1_{\mathbf{p}_2}\rangle$, or $\sqrt{2}|2_{\mathbf{p}}\rangle$ if $\mathbf{p}_1 = \mathbf{p}_2 = \mathbf{p}$.

deal with the problem 'on the fly', by regulating divergent integrals in a consistent manner. We could simply cut off the momentum integration at $|\mathbf{p}| = \Lambda$. Using spherical coordinates this gives for T^{00}

$$\langle 0|T^{00}|0\rangle = \epsilon = \epsilon_0 + \frac{4\pi}{2(2\pi)^3}\int_0^\Lambda dp\, p^2\sqrt{p^2+m^2}. \tag{2.40}$$

For T^{kl}, rotational invariance tells us that it has the form

$$T^{kl} = p\,\delta_{kl}, \tag{2.41}$$

since δ_{kl} is the only relevant invariant tensor under rotations. In fact, p is the pressure of the ground state. It follows that $3p = \delta_{kl}T_{kl}$, and

$$p = -\epsilon_0 + \frac{1}{3}\frac{4\pi}{2(2\pi)^3}\int_0^\Lambda dp\,\frac{p^4}{\sqrt{p^2+m^2}}. \tag{2.42}$$

The problem with this regularization is, that it is not consistent with Lorentz invariance: we are treating space and time differently and $\langle 0|T^{\mu\nu}|0\rangle$ will not be proportional to $\eta^{\mu\nu}$ this way and $p \neq -\epsilon$. Inserting the identity $1 = (p^{-2}/3)\partial p^3/\partial p$ into (2.40) and making a partial integration gives

$$\epsilon = -p + \frac{1}{3}\frac{4\pi}{(2\pi)^3}\Lambda^3\sqrt{m^2+\Lambda^2}, \tag{2.43}$$

where the second term on the r.h.s. is the surface term. This also shows that if the unregularized integral were convergent and the surface term absent, $p = -\epsilon$ would follow indeed.

There are Lorentz covariant regularizations, for example dimensional regularization or Pauli-Villars regularization. The latter is simplest here to present and is as follows. Define $\langle 0|T^{\mu\nu}|0\rangle$ as

$$\langle 0|T^{\mu\nu}|0\rangle = -\epsilon_0\eta^{\mu\nu} + \int \frac{d^3p}{(2\pi)^3}\frac{1}{2}\sum_i c_i\,\frac{p_i^\mu p_i^\nu}{p_i^0}, \tag{2.44}$$

where $p_i^0 = \sqrt{m_i^2+\mathbf{p}^2}$, $\mathbf{p}_i = \mathbf{p}$, and the coefficients c_i and the masses m_i are chosen such that the integral converges, whith $c_1 \equiv 1$ and $m_1 \equiv m$. This regularization is Lorentz invariant because the c_i and m_i are invariant. When the masses m_i, $i > 1$ are sent to infinity the result diverges again, but we cancel this by a suitable choice of ϵ_0. See Problem 9 for more details.

Having set the vacuum energy density equal to zero we can now ask meaningful questions about the energy of the ground state in a finite volume. A famous example is the Casimir effect. This was originally discovered in QED but it applies also to our scalar field *mutatis mutandis* (two free massless scalar fields to represent the two spin states of the photon, Dirichlet boundary conditions).

However, let us use the language of QED anyway as it is more intuitive. Consider two parallel plates of a conductor a distance a apart, with a much smaller than the linear size L of the plates. The presence of the plates is taken into account by imposing boundary conditions corresponding to a perfect conductor. This shifts the ground state energy inside and outside the plates relative to the vacuum, and the result is (see e.g. the book by Milton, Itzykson and Zuber sect. 3-2-4, Van Baal sect. 2)

$$\Delta E = \frac{-\hbar\pi^2 L^2}{720 a^3}. \tag{2.45}$$

It corresponds to a tiny attractive force which has been verified by experiment.

2.4 Perturbation theory

In the general case that the action is of higher than second order in the fields the theory is said to be interacting, because there is then no Fourier or other representation in which the harmonic oscillators are uncoupled. In our scalar field model with hamiltonian[2]

$$H = \int d^3x \left[\frac{1}{2}\pi_0^2 + \frac{1}{2}(\nabla\varphi_0)^2 + \frac{1}{2}\kappa_0\varphi_0^2 + \frac{1}{4}\lambda_0\varphi_0^4 + \epsilon_0\right], \tag{2.46}$$

the strength of the anharmonic φ_0^4 term is monitored by λ_0, the coupling constant. Its presence changes the eigenvalues and eigenvectors of P^μ, and we have to recalculate the ground state and the single and multiparticle states. A useful tool is perturbation theory, e.g. making an expansion in powers of λ_0.

However, it is much better to make an expansion in a renormalized coupling constant λ that is closely related to experimental observation. In QED the corresponding coupling constant is the elementary charge e or the fine-structure constant $\alpha = e^2/4\pi$. The starting parameters in the quantum theory are not simply related to familiar quantities like charge and mass of a particle, and we make this explicit by giving them the subscript 0: ϵ_0, κ_0, λ_0. They are called the unrenormalized, or bare, parameters. The analogue physical quantities are denoted by ϵ, κ and λ, and are called the renormalized, or dressed, parameters. Furthermore, the 'strength' of the field turns out to be changed by the interaction and one introduces a renormalized field φ which differs from φ_0 by a factor that is traditionally written as

$$\varphi_0 = \sqrt{Z}\,\varphi. \tag{2.47}$$

The constant Z ($Z > 0$) is called the 'wave function renormalization constant'.

The fact that fields represent an infinite number of degrees of freedom easily leads to divergent integrals in perturbation theory. It turns out that such divergencies can be absorbed in the bare parameters, such that the renormalized

[2]The reason for the subscript 0 will be explained shortly.

ones come out finite. We have seen an example of this in ϵ_0 and its relation to ϵ. Similarly, expectation values of products of φ_0 may contain divergencies that are absorbed in Z such that the expectation values of φ are finite, when expressed in terms of the renormalized parameters. In perturbation theory, formulated as an expansion in λ, one finds

$$Z = 1 + O(\lambda^2), \quad \lambda_0 = \lambda + O(\lambda^2), \quad \kappa_0 = \kappa + O(\lambda), \quad \epsilon_0 = \epsilon + O(1). \tag{2.48}$$

This is in accordance with the results in the previous section, where we found for $\lambda_0 = 0 \equiv \lambda$ that $\epsilon_0 = \epsilon + \text{constant}$, and $\kappa_0 = m^2 \equiv \kappa$. We shall assume the ground-state energy to be zero, $\epsilon = 0$.

Consider now the time-evolution operator

$$U(t,0) = e^{-iHt}. \tag{2.49}$$

We want to expand this is powers of λ. At $t = 0$, the hamiltonian can be written as the sum of a free part[3]

$$H_0 = \int d^3x \left[\frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}m^2\varphi^2 + \epsilon_0^{\text{free}}\right] \tag{2.50}$$

$$= \sum_{\mathbf{p}} a_{\mathbf{p}}^\dagger a_{\mathbf{p}} \sqrt{m^2 + \mathbf{p}^2}, \tag{2.51}$$

and an interaction part $H_1 = H - H_0$. Note that in H_0, π is defined to have the usual commutation relation with φ, hence

$$\pi = \sqrt{Z}\,\pi_0. \tag{2.52}$$

Furthermore, m is the physical mass and ϵ_0^{free} is chosen such that the vacuum energy is zero in the free theory. The interaction hamiltonian has the form

$$H_1 = \int d^3x\, \frac{1}{4}\lambda\varphi^4 + \Delta H_1, \tag{2.53}$$

$$\Delta H_1 = \int d^3x \left[(Z^{-1} - 1)\frac{\pi^2}{2} + (Z-1)\frac{(\nabla\varphi)^2}{2} + \delta m^2 \frac{\varphi^2}{2} + \delta\lambda\frac{\varphi^4}{4} + \delta\epsilon\right] \tag{2.54}$$

$$\delta m^2 = Z\kappa_0 - m^2, \quad \delta\lambda = Z^2\,\lambda_0 - \lambda, \quad \delta\epsilon = \epsilon_0 - \epsilon_0^{\text{free}}. \tag{2.55}$$

The terms in ΔH_1 are called *counterterms*. In lowest nontrivial order of perturbation theory the counterterms are usually zero, and in these lecture notes we shall usually ignore them. The relation between bare and renormalized parameter will be studied in more detail in chapter 4. See, e.g. Peskin & Schroeder ch. 10 for more information on the counterterm method (also known as 'renormalized perturbation theory').

[3]The subscript 0 in H_0 means 'free' here, not 'bare'.

We want to expand the evolution operator in powers of H_1. It is wrong to simply expand the exponential because H_1 and H_0 do not commute. This is a standard problem in time-dependent perturbation theory, which is solved by introducing[4]

$$U_1(t,0) = e^{iH_0 t}\, e^{-iHt}, \tag{2.56}$$

differentiating this with respect to t,

$$\begin{aligned} i\partial_t U_1(t,0) &= e^{iH_0 t}\, H_1\, e^{-iH_0 t}\, U_1(t,0) = H_1(t)\, U_1(t,0), && (2.57)\\ H_1(t) &= e^{iH_0 t}\, H_1\, e^{-iH_0 t}, && (2.58)\end{aligned}$$

and integrating this from 0 to t. The result can be written as a *time-ordered product*, for example by discretizing time in small steps $a = t/N$, $t_n = na$,

$$\begin{aligned} U_1(t,0) &= \lim_{N\to\infty} e^{-iH_1(t_{N-1})a}\, e^{-iH_1(t_{N-2})a} \cdots e^{-iH_1(0)a} && (2.59)\\ &\equiv T\, e^{-i\int_0^t dt'\, H_1(t')} && (2.60)\\ &= 1 - i\int_0^t dt'\, H_1(t') + \frac{(-i)^2}{2!}\int_0^t dt'\, dt''\, T\, H_1(t')H_1(t'') + \cdots && (2.61)\end{aligned}$$

where T is the *time-ordering 'operator'*, defined as an instruction to order operators by increasing time from right to left:

$$\begin{aligned} T\, H_1(t_1)H_1(t_2) &= \theta(t_1-t_2)H_1(t_1)H_1(t_2) + \theta(t_2-t_1)H_1(t_2)H_1(t_1), && (2.62)\\ T\, H_1(t_1)\cdots H_1(t_k) &= H_1(t_{i_1})H_1(t_{i_2})\cdots H_1(t_{i_k}), \quad t_{i_1} > t_{i_2} > \cdots > t_{i_k} && (2.63)\end{aligned}$$

A derivation of (2.60) by iteration is given in most text books on quantum field theory. Multiplying the expansion for $U_1(t,0)$ by $e^{-iH_0 t}$ from the left gives the expansion for $U(t,0)$:

$$e^{-iHt} = e^{-iH_0 t} - ie^{-iH_0 t}\int_0^t dt'\, e^{iH_0 t'}\, H_1\, e^{-iH_0 t'} + O(H_1^2). \tag{2.64}$$

2.5 Scattering

One of the most interesting new possible effects due to the interaction is scattering of particles. Fortunately, to lowest non-trivial order we only need to know the particle states in zeroth order, i.e. the free states, and we shall not need to renormalize κ and λ.

[4]This is the evolution operator in the *interaction picture*, the formalism in which states evolve in time according to $|\psi,t\rangle = U_1(t,0)|\psi,0\rangle$, and operators $O(t) = e^{iH_0 t}\, O(0)\, e^{-iH_0 t}$. In the *Heisenberg picture* the states are time-independent and the operators evolve as $O(t) = e^{iHt}\, O(0)\, e^{-iHt}$; in the *Schrödinger picture* the operators are time-independent and the states evolve according to $|\psi,t\rangle = U(t,0)|\psi,0\rangle$. Arbitrary matrix elements of operators are identical in the three 'pictures'. Note that in general the splitting $H = H_0 + H_1$ depends on time (here chosen $t = 0$).

Consider the scattering $1 + 2 \to 3 + 4$. We start with a free two-particle state $|\mathbf{p}_1\mathbf{p}_2\rangle$ at time $t = 0$ and wish to calculate the probability amplitude for the transition to another such state $|\mathbf{p}_3\mathbf{p}_4\rangle$ at a later time t,

$$\langle\mathbf{p}_3\mathbf{p}_4|U(t,0)|\mathbf{p}_1\mathbf{p}_2\rangle, \quad U(t,0) = e^{-iHt}, \tag{2.65}$$

where $U(t,0)$ is the evolution operator. The hamiltonian H has the form

$$H = H_0 + H_1, \quad H_1 = \int d^3x\, \frac{1}{4}\lambda\varphi^4 + \Delta H_1. \tag{2.66}$$

with H_0 the free hamiltonian of the previous sections:

$$H_0|\mathbf{p}_1\mathbf{p}_2\rangle = (E_1 + E_2)|\mathbf{p}_1\mathbf{p}_2\rangle, \quad E_1 = E(\mathbf{p}_1) = \sqrt{\mathbf{p}_1^2 + m^2}, \tag{2.67}$$

etc, and the ΔH_1 is given in (2.55). For non-trivial scattering the final state is different from the initial state and the result would then be zero if H_1 were zero. Hence the scattering amplitude is at least of order H_1 (order λ). The expansion of the evolution operator in powers of H_1 is given in the previous section. We only need the first order expression (2.64), which leads to

$$\begin{aligned}\langle\mathbf{p}_3\mathbf{p}_4|U(t,0)|\mathbf{p}_1\mathbf{p}_2\rangle &= e^{-i(E_3+E_4)t}\,\frac{1-e^{i\Delta Et}}{\Delta E}\langle\mathbf{p}_3\mathbf{p}_4|H_1|\mathbf{p}_1\mathbf{p}_2\rangle + O(\lambda^2), &(2.68)\\ \Delta E &= E_3 + E_4 - E_1 - E_2, &(2.69)\end{aligned}$$

and

$$|\langle\mathbf{p}_3\mathbf{p}_4|U(t,0)|\mathbf{p}_1\mathbf{p}_2\rangle|^2 = \frac{2 - 2\cos(\Delta Et)}{(\Delta E)^2}\,|\langle\mathbf{p}_3\mathbf{p}_4|H_1|\mathbf{p}_1\mathbf{p}_2\rangle|^2. \tag{2.70}$$

We now turn to the matrix element of H_1. Using

$$\varphi(\mathbf{x}) = \sum_{\mathbf{q}}\left[\frac{e^{i\mathbf{q}\mathbf{x}}}{\sqrt{2E(\mathbf{q})L^3}}\,a_{\mathbf{q}} + \frac{e^{-i\mathbf{q}\mathbf{x}}}{\sqrt{2E(\mathbf{q})L^3}}\,a_{\mathbf{q}}^\dagger\right], \tag{2.71}$$

and using the fact that only terms contribute which do not change the number of particles (i.e. same number of annihilation and creation operators), we get terms of the form

$$\begin{aligned}\langle\mathbf{p}_3\mathbf{p}_4|a_{\mathbf{q}_3}^\dagger a_{\mathbf{q}_4}^\dagger a_{\mathbf{q}_1}a_{\mathbf{q}_2}|\mathbf{p}_1\mathbf{p}_2\rangle &= (\delta_{\mathbf{q}_1,\mathbf{p}_1}\delta_{\mathbf{q}_2,\mathbf{p}_2} + \delta_{\mathbf{q}_1,\mathbf{p}_2}\delta_{\mathbf{q}_2,\mathbf{p}_1})\langle\mathbf{p}_3\mathbf{p}_4|\mathbf{q}_3\mathbf{q}_4\rangle &(2.72)\\ &\to 2\delta_{\mathbf{q}_1,\mathbf{p}_1}\delta_{\mathbf{q}_2,\mathbf{p}_2}\langle\mathbf{p}_3\mathbf{p}_4|\mathbf{q}_3\mathbf{q}_4\rangle\\ &= 2\delta_{\mathbf{q}_1,\mathbf{p}_1}\delta_{\mathbf{q}_2,\mathbf{p}_2}(\delta_{\mathbf{q}_3,\mathbf{p}_3}\delta_{\mathbf{q}_4,\mathbf{p}_4} + \delta_{\mathbf{q}_3,\mathbf{p}_4}\delta_{\mathbf{q}_4,\mathbf{p}_3})\\ &\to 4\delta_{\mathbf{q}_1,\mathbf{p}_1}\delta_{\mathbf{q}_2,\mathbf{p}_2}\delta_{\mathbf{q}_3,\mathbf{p}_3}\delta_{\mathbf{q}_4,\mathbf{p}_4}, &(2.73)\end{aligned}$$

where the arrows indicate equivalence under relabeling of the dummy $\mathbf{q}$'s which are to be summed over. (In (2.72) we worked the $a_{\mathbf{q}}$'s to the right using the commutation relations until we got $a_{\mathbf{q}}|0\rangle = 0$.) There are five more such contributions, differering in the order of the operators ($aa^\dagger a^\dagger a, \ldots, aaa^\dagger a^\dagger$), which each give equivalent results (terms like $\delta_{\mathbf{q}_i,\mathbf{q}_j}$ do not contribute because the initial and final states differ). The result is then

$$\langle \mathbf{p}_3\mathbf{p}_4|H_1|\mathbf{p}_1\mathbf{p}_2\rangle = \frac{6\lambda}{\prod_i \sqrt{2E_i L^3}} \int d^3x\, e^{i(-\mathbf{p}_3-\mathbf{p}_4+\mathbf{p}_1+\mathbf{p}_2)\mathbf{x}} = \frac{6\lambda L^3}{\prod_i \sqrt{2E_i L^3}}\, \delta_{\mathbf{p}_3+\mathbf{p}_4,\mathbf{p}_1+\mathbf{p}_2}. \tag{2.74}$$

This gives for the probability

$$|\langle \mathbf{p}_3\mathbf{p}_4|U(t,0)|\mathbf{p}_1\mathbf{p}_2\rangle|^2 = \frac{(6\lambda)^2 L^6}{L^{12}\prod_i 2E_i}\, \frac{2-2\cos(\Delta E t)}{(\Delta E)^2}\, \delta_{\mathbf{p}_3+\mathbf{p}_4,\mathbf{p}_1+\mathbf{p}_2}. \tag{2.75}$$

We are interested in scattering into a domain Δ of final momenta,

$$\sum_{(\mathbf{p}_3,\mathbf{p}_4)\in\Delta} |\langle \mathbf{p}_3\mathbf{p}_4|U(t,0)|\mathbf{p}_1\mathbf{p}_2\rangle|^2 \tag{2.76}$$

$$\to \quad \frac{L^{-3}(6\lambda)^2}{4E_1E_2} \int_\Delta d\omega_3\, d\omega_4\, \frac{2-2\cos(\Delta E t)}{(\Delta E)^2}\, (2\pi)^3\delta^3(\mathbf{p}_3+\mathbf{p}_4-\mathbf{p}_1-\mathbf{p}_2) \tag{2.77}$$

$$d\omega_i = \frac{d^3p_i}{(2\pi)^3 2E_i}, \tag{2.78}$$

where the arrow indicates the infinite volume limit (2.26), which also implies

$$L^3\delta_{\mathbf{p},\mathbf{q}} \to (2\pi)^3\delta^3(\mathbf{p}-\mathbf{q}). \tag{2.79}$$

For large times t (on the scale of the typical inverse energies E^{-1}) we have the identity

$$\frac{2-2\cos(\Delta E t)}{(\Delta E)^2} = t\, 2\pi\delta(\Delta E) + O(1/t). \tag{2.80}$$

This can be shown by integration with a test function $F(E)$:

$$\int_{-\infty}^{\infty} dE\, F(E)\, \frac{2-2\cos Et}{E^2} = t\int_{-\infty}^{\infty} du\, F\left(\frac{u}{t}\right)\, \frac{2-2\cos u}{u^2} \tag{2.81}$$

$$= t\left[F(0)\int_{-\infty}^{\infty} du\, \frac{2-2\cos u}{u^2} + O(t^{-2})\right] \tag{2.82}$$

$$= tF(0)\, 2\pi + O(t^{-1}), \tag{2.83}$$

where we used $F(u/t) = F(0) + F'(0)u/t + O(t^{-2})$; the $F'(0)$ term drops out by symmetry.

Summarizing, we have the following result for the probability *rate*:

$$\Gamma_\Delta = \frac{\partial}{\partial t} \sum_{(\mathbf{p}_3,\mathbf{p}_4)\in\Delta} |\langle \mathbf{p}_3\mathbf{p}_4|U(t,0)|\mathbf{p}_1\mathbf{p}_2\rangle|^2 \tag{2.84}$$

$$= \frac{L^{-3}}{4E_1E_2} \int_\Delta d\omega_3\, d\omega_4\, (2\pi)^4\delta^4(p_3+p_4-p_1-p_2)\, (6\lambda)^2. \tag{2.85}$$

The probablity rate implies an *event rate*, which is expected to be proportional to the overlap of particle densities $\int d^3x\, n_1 n_2$. As will be discussed in greater detail in chapter 9, the results of scattering experiments are expressed in terms of the *cross section* σ_Δ. In a reference frame where the initial particle momenta are aligned it is defined by

$$\Gamma_\Delta^{\text{event}} = \sigma_\Delta v_{12} \int d^3x\, n_1 n_2, \tag{2.86}$$

with

$$v_{12} = |\mathbf{p}_1/E_1 - \mathbf{p}_2/E_2| \tag{2.87}$$

the relative velocity. Realizing that $\Gamma_\Delta^{\text{event}} = \Gamma_\Delta$ if we normalize to unit initial particle number, $\int d^3x\, n_{1,2} = 1$, and that the density of our initial particles is $n_{1,2} = 1/L^3$, we have the result

$$\sigma_\Delta = \frac{1}{4E_1 E_2 v_{12}} \int_\Delta d\omega_3\, d\omega_4\, (2\pi)^4 \delta^4(p_3 + p_4 - p_1 - p_2)\, |T|^2. \tag{2.88}$$

where T is the scattering amplitude (also called the invariant amplitude $\mathcal{M}$) for this case,

$$|T|^2 = (6\lambda)^2. \tag{2.89}$$

The prefactor can be expressed as a Lorentz scalar,

$$E_1 E_2 v_{12} = \sqrt{(p_1 p_2)^2 - m_1^2 m_2^2}, \tag{2.90}$$

and we see that if the integration domain Δ is invariantly specified, the cross section is a Lorentz scalar. For example, integrating over all momenta gives the total cross section (cf. Problem 3)

$$\sigma = \frac{1}{32\pi s}\,(6\lambda)^2, \quad s \equiv -(p_1 + p_2)^2, \tag{2.91}$$

where the Lorentz invariant s is equal to the total energy squared in the center of mass frame. In a more detailed specification of Δ we can fix the invariant momentum transfer t. The corresponding cross section is conventially written $d\sigma/dt$ (cf. Problem 3):

$$\frac{d\sigma}{dt} = \frac{1}{16\pi s(s - 4m^2)}\,(6\lambda)^2, \quad t \equiv -(p_1 - p_3)^2. \tag{2.92}$$

In the center of mass frame defined by $\mathbf{p}_1 + \mathbf{p}_2 = 0$, we have $t = -2|\mathbf{p}_1|^2(1 - \cos\theta)$, with θ the scattering angle (cf. Fig. 2.1), the angle between $\mathbf{p}_1$ and $\mathbf{p}_3$, and $|\mathbf{p}_1|^2 = (s - 4m^2)/4$. So we see that the differential cross section

$$\left[\frac{d\sigma}{d\Omega}\right]_{\text{cm}} = \frac{1}{64\pi^2 s}(6\lambda)^2 \tag{2.93}$$

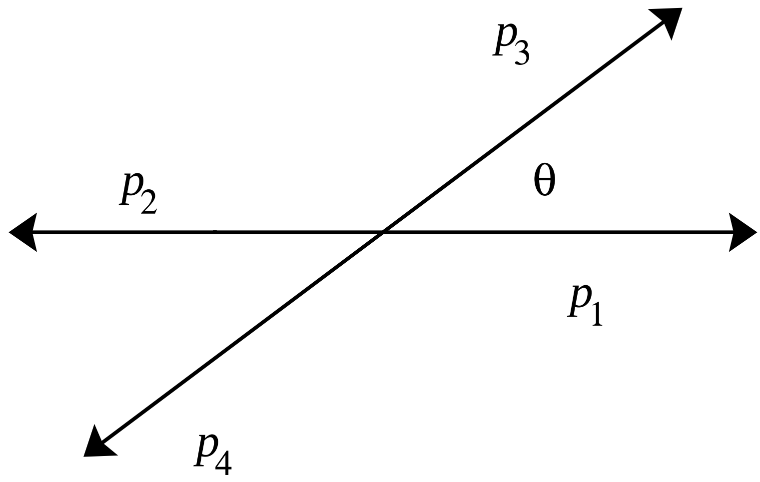

Figure 2.1: Three-momenta for scattering in the center of mass frame.

is isotropic. This is special to the φ^4 theory, later we will encounter more interesting differential cross sections.

The above derivation of the scattering amplitude has the benefit that it is short. In higher orders it gets complicated because it lacks manifest Lorentz covariance. Only the end results are covariant or invariant. Later we will develop more sophisticated calculational techniques which are manifestly covariant. Conceptually the above derivation can be improved by considering wave packet states which are localized in space (unlike the plane wave states used here which correspond to uniform density). This we will do in chapter 9.

2.6 Decay

Apart from leading to scattering, interactions may cause particles to be unstable, transforming them into two or more particles of a different species. For example, neutral pions are unstable and decay predominantly into two photons, $\pi^0 \to \gamma+\gamma$, with a mean life time $\tau = 8.8 \times 10^{-17}$ sec. The mean life time is the inverse of the decay rate Γ.

The possibility of decay can be illustrated by the following simple model involving only spinless particles. The model is specified by the action

$$S[\chi,\varphi] = -\int d^4x \left(\frac{1}{2}\partial_\mu\chi\partial^\mu\chi + \frac{1}{2}M^2\chi^2 + \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi + \frac{1}{2}m^2\varphi^2 + \frac{\lambda}{4}\varphi^4 + \frac{g}{2}\varphi^2\chi\right), \tag{2.94}$$

which describes two types of particles "χ" and "φ", with masses M and m, respectively. There are two interaction terms,

$$H_{\rm int} = \int d^3x \left(\frac{1}{2}g\varphi^2\chi + \frac{1}{4}\lambda\varphi^4\right), \tag{2.95}$$

with strengths parametrized by the coupling constants g and λ (g has dimension of mass). Apart from new types of scattering, the $g\varphi^2\chi$ term also allows for transitions $\chi \leftrightarrow \varphi + \varphi$.

Suppose at time zero the initial state contains only one χ-particle with four-momentum p. The probability at a later time t for the decay into two φ-particles with momenta q_1 and q_2 is then $|\langle \mathbf{q}_1(\varphi)\mathbf{q}_2(\varphi)|U(t,0)|\mathbf{p}(\chi)\rangle|^2$ (we use the same notation as in the scattering case). Going through similar steps as in the derivation of the rate for scattering, gives for the decay rate

$$\Gamma = \frac{1}{\tau} = \frac{\partial}{\partial t}\frac{1}{2}\sum_{\mathbf{q}_1\mathbf{q}_2} |\langle \mathbf{q}_1(\varphi)\mathbf{q}_2(\varphi)|U(t,0)|\mathbf{p}(\chi)\rangle|^2 \tag{2.96}$$

$$= \frac{1}{2p^0}\frac{1}{2}\int d\omega_1 d\omega_2 (2\pi)^4 \delta(q_1+q_2-p)\, g^2, \tag{2.97}$$

where we used

$$\langle \mathbf{q}_1(\varphi)\mathbf{q}_2(\varphi)|H_{\text{int}}|\mathbf{p}(\chi)\rangle = \frac{g}{2}\int d^3x \sum_{\mathbf{p}'\mathbf{q}_1'\mathbf{q}_2'} \frac{1}{\sqrt{8q_1'^0 q_2'^0 p'^0 L^9}} e^{i(\mathbf{p}'-\mathbf{q}_1'-\mathbf{q}_2')\cdot\mathbf{x}}$$

$$\langle \mathbf{q}_1(\varphi)\mathbf{q}_2(\varphi)|a^\dagger_{\mathbf{q}_1'}(\varphi)a^\dagger_{\mathbf{q}_1'}(\varphi)a_{\mathbf{p}'}(\chi)|\mathbf{p}(\chi)\rangle \tag{2.98}$$

$$= g\frac{1}{\sqrt{8q_1^0 q_2^0 p^0 L^3}}\,\delta_{\mathbf{q}_1+\mathbf{q}_2,\mathbf{p}}. \tag{2.99}$$

The explicit factor 1/2 in (2.96) avoids double counting the two identical particles in the final state.

This example illustrates that the transition at relatively large times on the scale of m^{-1}, M^{-1} (i.e. 'the decay'), is only possible if energy-momentum is conserved: $q_1 + q_2 = p$. Examining this for the case of a χ-particle at rest one finds that this leads to the condition that there has to be enough energy to create the two particles in the final state,

$$M \geq 2m. \tag{2.100}$$

The integral in (2.97) is Lorentz invariant. It depends only on g^2, M and m (cf. Problem 3),

$$\Gamma = \frac{|\mathbf{q}|}{16\pi M p^0}\, g^2, \quad |\mathbf{q}| = \frac{1}{2}\sqrt{M^2-4m^2}. \tag{2.101}$$

For a moving χ-particle the factor $1/p^0$ in (2.97) expresses the expected time dilatation.

The unstable particles can be produced in scattering, e.g. $\varphi(\mathbf{q}_1) + \varphi(\mathbf{q}_2) \to \chi(\mathbf{p})$, which is just the inverse of the decay process.

2.7 Representation of the Poincaré group

We have seen in a few examples that Lorentz invariance emerges in the infinite volume limit. We now go into somewhat more detail of the formal aspects of symmetry in the quantum theory. In addition to the momentum operators P^μ, also the angular momentum operators and generators of special Lorentz transformations

$$J^{\mu\nu} = \int d^3x \, (x^\mu T^{0\nu} - x^\nu T^{0\mu}) \tag{2.102}$$

are time independent. The commutation relations with the field operators can be calculated from the canonical commutation relation (Problem 4),

$$\begin{aligned} [\varphi(x), P_\mu] &= -i\partial_\mu \varphi(x), & (2.103) \\ [\varphi(x), J_{\mu\nu}] &= -i(x_\mu \partial_\nu - x_\nu \partial_\mu)\varphi(x), & (2.104) \end{aligned}$$

and these lead to the Poincaré algebra (1.59).

The transformations are represented in Hilbert space by unitary operators $U(a)$, $U(\ell)$. For translations we have

$$|\psi'\rangle = U(a)|\psi\rangle, \quad U(a) = e^{-ia^\mu P_\mu}. \tag{2.105}$$

Here $|\psi\rangle$ represents a state of the system and $|\psi'\rangle$ represents this state actively translated over a spacetime distance a^μ. To see that (2.105) is correct we calculate the expectation value of $\varphi(x)$:

$$\langle\psi|\varphi(x)|\psi\rangle = f(x) \Rightarrow \langle\psi'|\varphi(x)|\psi'\rangle = f(x-a), \tag{2.106}$$

where we used (cf. (1.25), (2.103) and Problem 6)

$$\begin{aligned} U(a)^\dagger \varphi(x) U(a) &= \varphi(x) - ia^\mu[\varphi(x), P_\mu] + \cdots \\ &= \varphi(x-a). & (2.107) \end{aligned}$$

Note that in the quantum theory $\varphi(x)$ corresponds to the *observables*, whereas classically its transformation properties were treated as representing the *state* of the system. Similarly we have for Lorentz transformations[5]

$$|\psi'\rangle = U(\ell)|\psi\rangle, \quad U(\ell) = e^{i\frac{1}{2}\omega^{\mu\nu}J_{\mu\nu}} \leftrightarrow \ell = e^{i\frac{1}{2}\omega^{\mu\nu}S_{\mu\nu}}, \tag{2.108}$$

and (cf. (1.55), (2.104))

$$\begin{aligned} U(\ell)^\dagger \varphi(x) U(\ell) &= \varphi(x) + i\frac{1}{2}\omega^{\mu\nu}[\varphi(x), J_{\mu\nu}] + \cdots \\ &= \varphi(\ell^{-1}x). & (2.109) \end{aligned}$$

[5]The sign in the exponent is here '+' by convention, but recall (cf. (1.48)) that $U(\ell) = \exp(-i\phi J_3)$ for an active rotation about the 3-axis over an angle ϕ.

The fact that the unitary operators exist guarantees invariance of transition amplitudes, $\langle\psi_1'|\psi_2'\rangle = \langle\psi_1|\psi_2\rangle$. If Lorentz invariance is broken, then $U(\ell)$ does not exist or is time dependent. Despite its non-covariant features, canonical quantization gives a Lorentz invariant theory.

In the following we shall use a convenient covariant normalization of particle states in the infinite volume limit

$$\langle p'|p\rangle = 2p^0(2\pi)^3\delta(\mathbf{p}' - \mathbf{p}), \quad p^0 = \sqrt{\mathbf{p}^2 + m^2}. \tag{2.110}$$

This has the property (cf. (2.37))

$$\int d\omega_p\, f(p)\, \langle p'|p\rangle = f(p'). \tag{2.111}$$

For the argument of ket and bra we use the four-momentum p, but note that here p^0 is not an independent variable. Comparing with our finite volume normalization we have

$$|p\rangle = \sqrt{2p^0L^3}|\mathbf{p}\rangle \tag{2.112}$$

(recall (2.79)). In infinite volume we expand the free scalar field in terms of covariant $a(p)$ and $a^\dagger(p)$,

$$\varphi(x) = \int d\omega_p\, \left[a(p)e^{ipx} + a^\dagger(p)e^{-ipx}\right]. \tag{2.113}$$

Comparison with the previous finite volume expansion at time zero

$$\varphi(\mathbf{x}) = \sum_{\mathbf{p}} \left[\frac{e^{i\mathbf{p}\mathbf{x}}}{\sqrt{2p^0L^3}}\, a_{\mathbf{p}} + \frac{e^{-i\mathbf{p}\mathbf{x}}}{\sqrt{2p^0L^3}}\, a_{\mathbf{p}}^\dagger\right] \tag{2.114}$$

shows that[6]

$$a(p) = \sqrt{2p^0L^3}a_{\mathbf{p}}, \tag{2.115}$$

$$[a(p), a^\dagger(p')] = 2p^0(2\pi)^3\delta(\mathbf{p}' - \mathbf{p}), \tag{2.116}$$

$$|p\rangle = a^\dagger(p)|0\rangle, \quad \text{etc.}, \tag{2.117}$$

$$P^\mu = \int d\omega_p\, a^\dagger(p)a(p)\, p^\mu. \tag{2.118}$$

From the transformation behavior (2.109) for the scalar field we infer that under Lorentz transformations the covariantly normalized objects transform simply,

$$U(\ell)a(p)U^\dagger(\ell) = a(\ell p), \quad U(\ell)|p\rangle = |\ell p\rangle. \tag{2.119}$$

Note that the *non*-covariantly normalized $a_{\mathbf{p}}$ and $|\mathbf{p}\rangle$ have a factor $\sqrt{(\ell p)^0/p^0}$ in their transformation rules.

[6]For a free field $a(p)$ is time-independent, it is the value of the Heisenberg operator at time zero: $a(p,t) = a(p)\exp(-ip^0t)$.

2.8 Where is the particle?

In the following we assume the scalar field to be free. Consider a normalized one particle state

$$|\psi\rangle = \int d\omega_p\, f(p)|p\rangle, \qquad \int d\omega_p\, |f(p)|^2 = 1. \tag{2.120}$$

Where is the particle? The answer to this question seems simply: measure its position. However, there is no natural particle position operator in quantum field theory. All operators are supposed to be made of the canonical variables $\varphi(\mathbf{x})$ and $\pi(\mathbf{x})$, for which $\mathbf{x}$ is just a label enumerating degrees of freedom. We can of course use our intuition from non-relativistic quantum mechanics and define localized states by

$$|\mathbf{x}\rangle = \int \frac{d^3p}{(2\pi)^3}\, e^{-i\mathbf{px}}\, \frac{1}{\sqrt{2p^0}}\, |p\rangle, \tag{2.121}$$

which satisfy

$$\langle\mathbf{x}|\mathbf{y}\rangle = \delta(\mathbf{x}-\mathbf{y}), \qquad \int d^3x\, |\mathbf{x}\rangle\langle\mathbf{x}| = \hat{1}_1, \tag{2.122}$$

where $\hat{1}_1$ is the unit operator in the one particle subspace. The $\sqrt{2p^0}$ is needed to convert to non-relativistic normalization. In terms of these localized states

$$\psi(\mathbf{x},t) \equiv \langle\mathbf{x}|\psi,t\rangle = \langle\mathbf{x}|e^{-iHt}|\psi\rangle = \int d\omega_p\, f(p)\sqrt{2p^0}\, e^{i\mathbf{px}-ip^0t} \tag{2.123}$$

is the candidate probability amplitude and

$$|\psi(\mathbf{x},t)|^2 \tag{2.124}$$

the probability density for finding the particle at $\mathbf{x}$ at time t. It is non-negative and the total probability is conserved in time,

$$\int d^3x\, |\psi(\mathbf{x},t)|^2 = 1. \tag{2.125}$$

Yet, there are some puzzling aspects to this interpretation: $|\psi(\mathbf{x},t)|^2$ is not the time component of a conserved probability current. Going through the usual steps we have

$$\frac{\partial}{\partial t}|\psi(\mathbf{x},t)|^2 = i\langle\psi,t|H|\mathbf{x}\rangle\psi(\mathbf{x},t) - i\psi(\mathbf{x},t)^*\langle\mathbf{x}|H|\psi,t\rangle. \tag{2.126}$$

Now the one particle hamiltonian is given by

$$\begin{aligned}
\langle\mathbf{x}|H|\psi,t\rangle &= \int d\omega_p\, \langle\mathbf{x}|p\rangle\langle p|H|\psi,t\rangle = \int d\omega_p\, \sqrt{m^2+\mathbf{p}^2}\, \langle\mathbf{x}|p\rangle\langle p|\psi,t\rangle \\
&= \sqrt{m^2-\nabla^2}\int d\omega_p\, \langle\mathbf{x}|p\rangle\langle p|\psi,t\rangle \\
&= \sqrt{m^2-\nabla^2}\,\psi(\mathbf{x},t) \equiv \int d^3y\, h(\mathbf{x}-\mathbf{y})\psi(\mathbf{y},t).
\end{aligned} \tag{2.127}$$

This is a non-local operator. One way to see this is by expansion of the square root in powers of ∇^2. This produces an infinite number of derivatives, which result in 'shifting $\psi(\mathbf{x})$ around the point $\mathbf{x}$', as in a Taylor series. The function $h(\mathbf{x}-\mathbf{y})$ is singular at $\mathbf{x}=\mathbf{y}$ (it is a distribution), but at large distances it decays like $\exp(-m|\mathbf{x}-\mathbf{y}|)$ (cf. Problem 7). So the right hand side of the equation of probability conservation (2.126) does *not* look like the divergence of a current. In fact, there is no such current that satisfies the requirements one would associate with a probability current.

There is a density which does have an associated current which is sometimes used as a probability density instead of $\psi(\mathbf{x},t)$. The function[7]

$$f(x) = \langle 0|\varphi(x)|\psi\rangle = \int d\omega_p\, f(p)e^{ipx} \tag{2.128}$$

satisfies the Klein-Gordon equation because $\varphi(x)$ does so (Heisenberg picture). In general, $f(x) \neq f^*(x)$ and then the current

$$j^\mu(x) = -f^*(x)i\partial^\mu f(x) + i\partial^\mu f^*(x)\, f(x) \tag{2.129}$$

does not vanish. It is easy to verify that it is a conserved current,

$$(\partial^2 - m^2)f(x) = 0 \Rightarrow \partial_\mu j^\mu(x) = 0. \tag{2.130}$$

Furthermore,

$$\int d^3x\, j^0(x) = 1 \tag{2.131}$$

(verify). Because of these attractive properties j^0 has been used as a probability density. However, as a probability density j^0 should be non-negative, but there is no guarantee that this is the case (not even for purely positive frequency solutions of the Klein-Gordon equation). So we are really stuck with the non-covariant $|\psi(\mathbf{x},t)|^2$.

Consider wave packet functions $f(p)$ that are sufficiently concentrated in a narrow region of size $\Delta\mathbf{p}$ around an average momentum $\bar{\mathbf{p}}$, such that the variation of p^0 in $\sqrt{2p^0}$ and $d\omega_p$ under the integrals in (2.120), (2.121), (2.123) and (2.128) can be neglected. In these circumstances

$$j^\mu(x) \approx 2\bar{p}^\mu|f(x)|^2, \quad |\psi(\mathbf{x},t)|^2 \approx j^0(\mathbf{x},t), \tag{2.132}$$

and we may think of $\mathbf{j}(x)$ as the probability current for $|\psi(\mathbf{x},t)|^2$. The condition $|\Delta\mathbf{p}| \ll \bar{p}^0$ corresponds in position space to $|\Delta\mathbf{x}| >> 1/\bar{p}^0 \geq 1/m$, i.e. to scales much larger than the Compton wavelength.

We have to accept that position of particles is not a natural concept in relativistic field theory, and that the momentum basis is favoured. However, non-relativistic intuition applies as long as we do not wish to localize particles on

[7]For simplicity we use the same symbol f for $f(x)$ as for $f(p)$, but the two functions ($f(p)$ and $f(x)$) are very different and should not be confused.

the scale of the Compton wavelength $1/m$. This is a very small scale for the usual particles, except for neutrinos which are nearly massless, and of course, the massless photon.

At a deeper level we ought to take into account *how* a position measurement is performed. Typically, this is through electromagnetic or gravitational interactions, for which we do have a relativistic formulation – the one we are developing here in terms of quantum fields. For example, as will become clearer after introducing QED, the electromagnetic current of a spinless charged particle has the form of (2.129) in the limit of vanishing charge: it is the current associated with the complex scalar field considered in Problem 8. The 'problem' of the absence of a covariant position-probability current appears to be a red herring.

Another way to measure the position of a particle is through its gravitational interaction. The energy momentum tensor is the source of the gravitational field, and for our one particle state it can be shown that[8]

$$\begin{aligned} \langle\psi|T^{\mu\nu}(x)|\psi\rangle &= \partial^\mu f^*(x)\partial^\nu f(x) + \partial^\nu f^*(x)\partial^\mu f(x) \\ &\quad - \eta^{\mu\nu}\left[\partial_\rho f^*(x)\partial^\rho f(x) + m^2 f^*(x) f(x)\right], \qquad (2.133) \end{aligned}$$

which resembles the classical expression. Similar expressions hold for massless particles.

2.9 Causality and locality

According to standard lore nothing can go faster than the speed of light. Yet, for a free relativistic particle, the amplitude

$$\langle\mathbf{x}|e^{-iH(t-t')}|\mathbf{x}'\rangle \equiv A(\mathbf{x},t;\mathbf{x}',t') \qquad (2.134)$$

is nonzero for spacelike separations $(\mathbf{x}-\mathbf{x}')^2 - (t-t')^2 > 0$. To see this, use the representation (2.121) to get

$$\begin{aligned} A(\mathbf{x},t;\mathbf{x}',t') &= \int \frac{d^3p}{(2\pi)^3}\, e^{i\mathbf{p}(\mathbf{x}-\mathbf{x}')-ip^0(t-t')}, \quad p^0 = \sqrt{m^2+p^2}, \\ &= \int d\omega_p\, 2p^0\, e^{ip(x-x')}, \\ &= 2\partial_0 \Delta^{(+)}(x-x'), \qquad (2.135) \\ \Delta^{(+)}(x-x') &\equiv i\int d\omega_p\, e^{ip(x-x')}. \qquad (2.136) \end{aligned}$$

The Lorentz invariant function $\Delta^{(+)}$ is nonzero for spacelike separations. A saddle point evaluation[9] for $x^0 = x'^0$ and large $|\mathbf{x}-\mathbf{x}'|$ shows that it drops off

[8] The calculation can be done by expressing the fields in creation and annihilation operators, but it is simpler to use the Green function techniques to be developed in the next chapter.

[9] The calculation is similar to that in Problem 7.

exponentially fast on the scale of the Compton wavelength $1/m$,

$$\Delta^{(+)}(x - x') \propto \exp(-m\sqrt{(x - x')^2}), \tag{2.137}$$

where we have re-expressed the result in Lorentz invariant form. The amplitude $A(\mathbf{x}, t; \mathbf{x}', t')$ has the same behavior; it is even nonzero for $t < t'$!

So it seems that there is a problem with causality. The way out in field theory is again[10] the fact that coordinates $\mathbf{x}$ are not among the basic observables, which are to be constructed from the fields. There is no natural position operator representing a position measurement, of which $|\mathbf{x}\rangle$ is an eigenvector with eigenvalues $\mathbf{x}$. Field theory *is* causal in the sense that local measurements in spacetime commute if they are performed at spacelike separations. Such measurements cannot influence each other, which is an expression of causality.

To check this for free fields we evaluate the commutator of two scalar fields, $[\varphi(x), \varphi(y)]$. Using the representation of φ in terms of creation and annihilation operators (2.113) and the commutation relations (2.116) we get

$$[\varphi(x), \varphi(y)] = \int d\omega_p \left(e^{ip(x-y)} - e^{-ip(x-y)}\right) = -i\Delta(x - y). \tag{2.138}$$

The function $\Delta(x - y)$ was studied earlier in Problem 1.4: it vanishes for spacelike separations. It follows that local observables constructed out of $\varphi(x)$ and its derivatives (e.g. the energy-momentum tensor) also commute at spacelike separations.

In an interacting Lorentz invariant field theory the commutator of two fields is no longer a c-number as in (2.138), and in gauge theories quantized in noncovariant gaugees they may not even commute at spacelike separations. However, commutators of local gauge invariant observables constructed out of the fields still vanish for spacelike separations. This property is sometimes called 'local commutativity', or simply 'locality'. In mathematical approaches to field theory it is taken as an axiom in the formulation of the theory.

The function $\Delta^{(+)}(z)$ in (2.136) is the so-called positive frequency part of $\Delta(z)$, i.e. it is a superposition of exponentials $\exp(-ip^0 z^0)$ with only positive frequencies ($p^0 > 0$). Its nonvanishing at spacelike separations is canceled in $\Delta(z)$ by a negative frequency part:

$$\Delta = \Delta^{(+)} + \Delta^{(-)}, \tag{2.139}$$

$$\Delta^{(-)}(z) = -i\int d\omega_p\, e^{-ipx} = -i\int d\omega_p\, e^{+ip^0x^0 - i\mathbf{p}\mathbf{x}}. \tag{2.140}$$

This cancelation leading to causality is seen by important authors as a deep reason for the necessity of antiparticles, 'traveling backwards in time', with, for the present case of the real scalar field, 'particles being their own antiparticles'

[10] Cf. the discussion in sect. 2.8.

(cf. Weinberg I, Peskin; see also Feynman). To the author of these lecture notes such remarks are mystifying. Thus far we have not seen any need to introduce the concept of antiparticles (i.e. particles with opposite charge(s) and the same mass and spin). Antiparticles will appear naturally in the context of QED and more general gauge field theory. Instead, we see the locality of field theory (differential equations of motion, action and energy etc. written as integrals over space(time)), as the fundamental property leading to causality.

2.10 Classical field

Having familiarized ourselves with a quantum field we would like to understand its relation to the classical field. Classical behavior in a quantum theory may be expected only for states with special properties, involving large quantum numbers and coherence. Here we shall illustrate this with a simple example, the free scalar field coupled to an external source. The operator field equation is

$$(-\partial^2 + m^2)\varphi(x) = J(x), \tag{2.141}$$

where $J(x)$ is the external source which we assume to vanish outside a compact domain in spacetime, in particular

$$J(x) = 0, \quad x^0 > t_+ \quad \text{or} \quad x^0 < t_-. \tag{2.142}$$

The classical field will be an expectation value of the quantum field. Such expectation values vanish in any state with a definite number of particles because $\varphi(x)$, being linear in the creation and annihilation operators, changes the number of particles. An interesting state with classical properties is the ground state of the system before the source is acting, the so-called in-vacuum $|0\,\mathrm{in}\rangle$. We shall calculate the expectation value

$$\varphi_c(x) = \langle \mathrm{in}\,0|\varphi(x)|0\,\mathrm{in}\rangle, \tag{2.143}$$

and find that it has the properties of the classical field (hence the subscript c).

Let us solve the field equation with the help of the retarded Green function $G^R(x-y)$ (cf. Problem 1.3),

$$\varphi(x) = \varphi_{\mathrm{in}}(x) + \int d^4y\, G^R(x-y)J(y), \tag{2.144}$$

where $\varphi_{\mathrm{in}}(x)$ is the so-called incoming field, which is free, $(-\partial^2 + m^2)\varphi_{\mathrm{in}}(x) = 0$. The incoming field is the unique solution of the free Klein-Gordon equation (defined for all times) which is equal to the field $\varphi(x)$ at times before the source is acting, $\varphi(x) = \varphi_{\mathrm{in}}(x)$ for $x^0 < t_-$. Similarly, there is an outgoing field, which is free and defined for all times, with $\varphi_{\mathrm{out}}(x) = \varphi(x)$ for $x^0 > t_+$. Note that

$J(x)$ is a c-number while φ, $\varphi_{\rm in}$ and $\varphi_{\rm out}$ are q-numbers. The outgoing field can be expressed in terms of the incoming field by choosing $x^0 > t_+$ and using (cf. Problem 1.3)

$$G^R(x-y) = i\int d\omega_p\,(e^{ip(x-y)} - e^{-ip(x-y)}), \quad x^0 > y^0. \tag{2.145}$$

Then

$$\varphi_{\rm out}(x) = \varphi_{\rm in}(x) + \int d\omega_p\left[e^{ipx}\,i\tilde{J}(p) - e^{-ipx}\,i\tilde{J}(p)^*\right], \tag{2.146}$$

where

$$\tilde{J}(p) = \int d^4y\, e^{-ipy}\, J(y) \tag{2.147}$$

is the Fourier transform of the source.

We can now identify the incoming and outgoing vacua from the creation and annihilation operators of $\varphi(x)$:

$$\varphi(x) = \int d\omega_p\left[a(p,x^0)e^{ipx} + a^\dagger(p,x^0)e^{-ipx}\right]. \tag{2.148}$$

By choosing $x^0 > t_+$ or $x^0 < t_-$ we get just the a and $a^\dagger$ of the in and outgoing fields:

$$\begin{aligned} a(p,x^0) &= a_{\rm in}(p), \quad x^0 < t_-, \\ &= a_{\rm out}(p), \quad x^0 > t_+, \end{aligned} \tag{2.149}$$

$$\varphi_{\rm in,out}(x) = \int d\omega_p\left[a_{\rm in,out}(p)e^{ipx} + a^\dagger_{\rm in,out}(p)e^{-ipx}\right]. \tag{2.150}$$

From (2.146) – (2.150) we see that

$$a_{\rm out}(p) = a_{\rm in}(p) + i\tilde{J}(p). \tag{2.151}$$

The incoming vacuum satisfies

$$a_{\rm in}(p)|0\,{\rm in}\rangle = 0, \tag{2.152}$$

since it is the ground state at times $x^0 < t_-$. So, using (2.144), the expectation value (2.143) is simply given by

$$\varphi_c(x) = \int d^4y\, G^R(x-y)J(y). \tag{2.153}$$

It is indeed just the solution of the classical field equation with retarted boundary conditions $\varphi_c(x) = 0$, $x < t_-$.

On the other hand, if we analyze the in-vacuum in terms of particles at late times, it is a *coherent state*:

$$a_{\rm out}(p)|0\,{\rm in}\rangle = i\tilde{J}(p)|0\,{\rm in}\rangle, \tag{2.154}$$

as follows from (2.151). Such states are 'minimum uncertainty states'. They are superpositions of states with different particle numbers with well defined phase relations. This follows from the theory of coherent states, which tells us that the in-vacuum has the form

$$|0\,\mathrm{in}\rangle \propto \exp\left[\int d\omega_p\, i\tilde{J}(p)a^\dagger_{\mathrm{out}}(p)\right]\,|0\,\mathrm{out}\rangle. \quad (2.155)$$

Here $|0\,\mathrm{out}\rangle$ is the out-vacuum, the state with no particles at times $x^0 > t_+$, which satisfies

$$a_{\mathrm{out}}(p)|0\,\mathrm{out}\rangle = 0. \quad (2.156)$$

Information on coherent states can be found in Brown sects. (1.7), (1.8) and books on quantum mechanics.

Summarizing, the classical field can be viewed as the expectation value of the quantum field in a coherent state. The free field is equivalent to a collection of harmonic oscillators, which are known to possess states for which the classical approximation is exact. For interacting theories additional conditions need to be satisfied, such as effectively weak interactions and large quantum numbers.

2.11 Summary

In the quantum theory the canonical variables $\varphi(\mathbf{x})$ and $\pi(\mathbf{x})$ are operators in Hilbert space with canonical commutation relations. The conserved Noether 'charges' P^μ and $J^{\mu\nu}$ (which can be expressed in terms of φ and π) are operators with the commutation relations of the Poincaré algebra. They are the generators of a unitary representation $U(a,\ell)$ of Poincaré transformations in Hilbert space. In the Heisenberg picture the equations of motion are the Heisenberg equations for the operators φ and π. In principle these equations can differ from the classical field equations, but often they have the same form.

The free field is equivalent to a collection of harmonic oscillators and the energy-momentum operators P^μ can be simultanously diagonalized using creation and annihilation operators. The ground state is interpreted as the vacuum and its excitations are interpreted as particles. These particles have identical mass and zero spin, they are bosons. The (formally infinite) energy of the vacuum is normalized to zero by adjusting a parameter which plays the role of the cosmological constant in the theory with gravity.

The momentum basis gives the natural description of relativistic particles, while the position basis is not covariant. Poincaré invariance is however guaranteed by the existence of the $U(a,\ell)$.

Adding simple interaction terms to the free field action, the particles also interact and they can scattering amongst themselves. Particles with differing masses can be described by introducing a field for each particle type with corresponding mass parameter in the action. Decay processes are then a natural possibility.

The classical field can be understood as an expectation value of the quantum field in a quantum state with classical properties.

2.12 Problems

1. *Fourier modes*

   Given (2.9) verify (2.10) by doing the integrals; verify also the statement about the eigenfunctions of $\partial/\partial x$.

2. *Fock space*

   The Hilbert space for a system of arbitrarily many particles such as the free scalar field is called Fock space. A basis is given by $|0\rangle$, $|p\rangle$, $|p_1 p_2\rangle$, etc. The states are normalized as

$$\begin{aligned} \langle p|q\rangle &= 2p^0(2\pi)^3\delta(\mathbf{p}-\mathbf{q}), \quad p^0 = \sqrt{m^2+\mathbf{p}^2}, \\ \langle p_1 p_2|q_1 q_2\rangle &= \langle p_1|q_1\rangle\langle p_2|q_2\rangle + \langle p_1|q_2\rangle\langle p_2|q_1\rangle, \end{aligned}$$

   etc. In general we get a sum over all permutations π of $1, \cdots, n$ (the value of n will be clear from the context),

$$\langle p_1 \cdots p_m|q_1 \cdots q_n\rangle = \delta_{mn} \sum_\pi \langle p_1|q_{\pi 1}\rangle \cdots \langle p_n|q_{\pi n}\rangle. \tag{2.157}$$

   The completeness relation can be written as

$$\hat{1} = \hat{1}_0 + \hat{1}_1 + \hat{1}_2 + \cdots, \tag{2.158}$$

   where $\hat{1}_0 = |0\rangle\langle 0|$ and $\hat{1}_n$ is the unit operator in the n-particle subspace,

$$\hat{1}_n = \frac{1}{n!}\int d\omega_{p_1} \cdots d\omega_{p_n}\, |p_1 \cdots p_n\rangle\langle p_1 \cdots p_n|. \tag{2.159}$$

   Verify this by taking matrix elements with $|q_1 \cdots q_m\rangle$.

3. *Phase space integral and scattering*

   In this problem we evaluate the remaining integrals encountered in two-particle scattering and decay.

   The integral (called a phase space integral)

$$I(p) = \int d\omega_{q_1}\, d\omega_{q_2}\, (2\pi)^4\delta^4(q_1+q_2-p), \tag{2.160}$$

$$d\omega_{q_i} = \frac{d^3 q_i}{(2\pi)^3 2\sqrt{\mathbf{q}_i^2+m_i^2}}, \quad i = 1, 2, \tag{2.161}$$

is Lorentz invariant, $I(p) = I(\ell p)$. It is convenient to evaluate it in the center of mass frame defined by $\boldsymbol{p} = 0$, in the following steps:

a) integrate over $\mathbf{q}_2$ using the momentum conserving delta functions,
b) choose spherical coordinates

$$\mathbf{q}_1 \to (|\mathbf{q}|, \theta, \phi), \quad d^3q_1 = |\mathbf{q}|^2 d|\mathbf{q}|\, d\Omega, \quad d\Omega = d(\cos\theta)\, d\phi \tag{2.162}$$

c) for the $|\mathbf{q}|$ integral use the energy conserving delta function and the general formula

$$\int_a^b dx\, \delta(f(x))g(x) = \sum_j \frac{1}{|f'(x_j)|}\, g(x_j), \tag{2.163}$$

where the summation is over the zero(s) x_j of $f(x)$ in the interval (a, b). In the present case the argument of the delta function, $\sqrt{m_1^2 + |\mathbf{q}|^2} + \sqrt{m_2^2 + |\mathbf{q}|^2} - p^0$, has only one zero. We use $s \equiv -p^2$, which is p_0^2 in the center of mass.

Verify that the result is given by

$$I = \frac{|\mathbf{q}|}{16\pi^2\sqrt{s}} \int d\Omega = \frac{|\mathbf{q}|}{4\pi\sqrt{s}}, \quad s = -p^2, \tag{2.164}$$

with

$$|\mathbf{q}|^2 = \frac{s^2 + (m_1^2 - m_2^2)^2}{4s} - \frac{m_1^2 + m_2^2}{2}. \tag{2.165}$$

The application of $I(p)$ to two-particle decay is straightforward.

In the application to scattering $1 + 2 \to 3 + 4$, $q_1 \to p_3$, $q_2 \to p_4$, p is the total incoming momentum, $p = p_1 + p_2$, and θ may be the angle between $\mathbf{p}_1$ and $\mathbf{p}_3$. The invariant amplitude $|T|^2$ is a Lorentz invariant function of the momenta, so a function of the two independent invariants $s = -(p_1 + p_2)^2 = -(p_3 + p_4)^2$, and $t = -(p_1 - p_3)^2 = -(p_2 - p_4)^2$. The other invariant[11] $u = -(p_1 - p_4)^2 = -(p_2 - p_3)^2$ is not independent, because $s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2$ (verify).

The 'flux factor' is given by

$$4E_1E_2v_{12} = 4\sqrt{(p_1p_2)^2 - m_1^2m_2^2} = 4|\mathbf{p}_1|\sqrt{s}, \tag{2.166}$$

with $|\mathbf{p}_1|$ given by (2.165) with $|\mathbf{q}| \to |\mathbf{p}_1|$.

For the total cross section we need to multiply by 1/2 if the two particles in the final state are identical, to avoid double counting. This is the same factor 1/2! as in (2.159), $n = 2$.

[11] The invariants s, t and u are called Mandelstam variables.

For the differential cross section we do not integrate over θ and ϕ. In the center of mass frame

$$d\sigma = d\Omega \frac{1}{64\pi^2 s} \frac{|\mathbf{p}_3|}{|\mathbf{p}_1|} |T|^2. \tag{2.167}$$

Verify.

Alternatively, we can specify the invariant momentum transfer $t = -(p_3 - p_1)^2$. It is linearly related to $\cos\theta$, $dt = 2|\mathbf{p}_1|\,|\mathbf{p}_3|\,d\cos\theta$, so $d\sigma/dt$ can be simply be read off from $d\sigma/d\cos\theta = 2\pi d\sigma/d\Omega$. We can also insert a constraining delta function $\delta(t + (p_1 - p_3)^2)$ in the integral $I(p_1 + p_2)$ which gives the same result (see e.g. Brown section 3.4):

$$\frac{d\sigma}{dt} = \frac{1}{64\pi s\, \mathbf{p}_1^2} |T|^2, \tag{2.168}$$

which holds also in the unequal mass case.

4. *Commuting with generators*

   Verify the commutation relations (2.103), (2.104). Hint: Write out explicitly P_0, P_m, J_{mn}, J_{0n} in terms of the canonical φ and π. Since the generators are constant in time, choose the time in P_μ and $J_{\mu\nu}$ equal to the argument x^0 of $\varphi(x)$, and use the canonical equal time commutation relations (2.6), which are valid at any time t:

$$[\varphi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = i\delta^3(\mathbf{x} - \mathbf{y}), \quad [\varphi(\mathbf{x}, t), \varphi(\mathbf{y}, t)] = [\pi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = 0. \tag{2.169}$$

5. *Commutators*

   Consider operators q_a, p_a, with the canonical commutation relations $[q_a, p_b] = i\delta_{ab}$, $[q_a, q_b] = 0$, $[p_a, p_b] = 0$. In terms of these, and (c-number) matrices M_{ab}, which are zero on the diagonal, $M_{aa} = 0$ (no summation), define operators $A(M) = -iM_{ab}p_a q_b$, or in matrix notation $A(M) = -ip^T Mq$. Calculate the commutators $[q_a, A(M)]$, $[p_a, A(M)]$, and show that $[A(M_1), A(M_2)] = A([M_1, M_2])$.

   As an application, let a be a continuous index $a \to \mathbf{x}$, $\delta_{ab} \to \delta(\mathbf{x} - \mathbf{y})$. Obtain the commutator $[J_k, J_l]$, where $J_k = \frac{1}{2}\epsilon_{klm}J_{lm}$ is the generator for rotations (e.g. for J_k, $M_{ab} \to -i\epsilon_{klm}x_l\partial_m\delta(\mathbf{x} - \mathbf{y})$). Likewise, obtain the commutator of J_k with the translation generators, $[J_k, P_l] =$?

6. *Finite transformations*

   Let $U = \exp(iF)$ be unitary, $F = F^\dagger$. Use the identity

$$\begin{aligned} U^\dagger A U &= A + i[A, F] + \frac{i^2}{2}[[A, F], F] + \cdots \\ &= A + \sum_{n=1}^{\infty} \frac{i^n}{n!} [\cdots [A, F], \cdots, F] \end{aligned} \tag{2.170}$$

to prove (2.107).

7. *Relativistic one-particle hamiltonian*

Consider (2.127). Expanding the square root to finite order, $h(\mathbf{x}-\mathbf{y})$ would consist of derivatives of delta functions. For example in the non-relativistic approximation

$$h(\mathbf{x}-\mathbf{y}) = \left(m + \frac{-\nabla^2}{2m}\right)\delta(\mathbf{x}-\mathbf{y}). \tag{2.171}$$

Integration over $\mathbf{y}$ gives the usual free particle hamiltonian acting on $\psi(\mathbf{x},t)$ as a differential operator, in addition to the rest-energy contribution m. Using Fourier transformation we can express the full h in the form

$$h(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\, e^{i\mathbf{p}\mathbf{x}}\, \sqrt{m^2+\mathbf{p}^2}. \tag{2.172}$$

The integral does not converge but this merely reflects that $h(\mathbf{x})$ is a distribution, singular around the origin. For large $\mathbf{x}$ we can use the saddle point method to get the dominant behavior of $h(\mathbf{x})$, e.g. by writing $\mathbf{x} = r\hat{x}$,

$$h(\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3}\, \exp\left[i\mathbf{p}\hat{x}r + \frac{1}{2}\ln(m^2+\mathbf{p}^2)\right], \tag{2.173}$$

and expanding about the correct stationary point of the exponent for $r \to \infty$. Verify that the stationary point is given by $\mathbf{p} \approx im\hat{x}$ and that $h(\mathbf{x}) \propto \exp(-mr)$ for large mr.

8. *$O(2)$ model, complex scalar field*

Consider a model with two scalar fields ϕ_1 and ϕ_2 with action[12]

$$S = -\int d^4x\, \left[\frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi_k\partial_\nu\phi_k + \frac{1}{2}m^2\phi^2 + \frac{1}{4}\lambda(\phi^2)^2\right]. \tag{2.174}$$

We use the summation convention also for indices such as k: $\phi_k\phi_k \equiv \sum_{k=1}^2 \phi_k\phi_k \equiv \phi^2$. The above action is invariant under rotations in 'internal space'

$$\phi_1' = \cos\alpha\,\phi_1 - \sin\alpha\,\phi_2, \quad \phi_2' = \sin\alpha\,\phi_1 + \cos\alpha\,\phi_2, \quad S[\phi'] = S[\phi]. \tag{2.175}$$

Infinitesimal rotations can be written in the form

$$\delta\phi_k = -\epsilon_{kl}\phi_l\delta\alpha, \quad \epsilon_{12} = -\epsilon_{21} = 1, \quad \epsilon_{11} = \epsilon_{22} = 0, \tag{2.176}$$

with infinitesimal rotation angle $\delta\alpha$. To a continuous symmetry corresponds a conserved quantity, usually called 'charge' (Noether's theorem). This can

[12] All objects are 'bare', we ignore renormalization.

be seen as follows. Consider infinitesimal rotation angles depending on space-time: $\delta\alpha(x)$. The action is now in general not invariant anymore because α depends on x; for an infinitesimal rotation,

$$\delta S = \int d^4x\, j^\mu \partial_\mu \delta\alpha = -\int d^4x\, \partial_\mu j^\mu \delta\alpha. \tag{2.177}$$

However, if the ϕ_k satisfy the field equations (equations of motion), then $\delta S = 0$ and we have a local balance equation (a 'conserved current'),

$$\partial_\mu j^\mu = 0, \tag{2.178}$$

with a corresponding conserved charge

$$Q = \int d^3x\, j^0. \tag{2.179}$$

a. Show that the current is given by

$$j^\mu = \epsilon_{kl} \partial^\mu \phi_k\, \phi_l. \tag{2.180}$$

b. Derive the field equations from the stationary action principle.

c. Verify using the field equations that $\partial_\mu j^\mu = 0$.

d. In the quantum theory Q is an operator, which can be expressed in the creation and annihilation operators at time zero. Show that

$$Q = \sum_{\mathbf{p}} a^\dagger_{\mathbf{p}k} (-i\epsilon_{kl}) a_{\mathbf{p}l}. \tag{2.181}$$

e. Consider the free theory with $\lambda = 0$. Choosing the vacuum energy to be zero, the energy-momentum operator is given by

$$P^\mu = \sum_{\mathbf{p}} a^\dagger_{\mathbf{p}k} a_{\mathbf{p}k}\, p^\mu. \tag{2.182}$$

Since Q is time independent it should be possible to diagonalize Q and P^0 simultaneously. This can be done as follows. Define

$$a_{\mathbf{p}\pm} = \frac{1}{\sqrt{2}}(a_{\mathbf{p}1} \mp i a_{\mathbf{p}2}). \tag{2.183}$$

Show that in terms of the new creation and annihilation operators

$$Q = \sum_p (a^\dagger_{\mathbf{p}+} a_{\mathbf{p}+} - a^\dagger_{\mathbf{p}-} a_{\mathbf{p}-}), \tag{2.184}$$

$$P^\mu = \sum_p (a^\dagger_{\mathbf{p}+} a_{\mathbf{p}+} + a^\dagger_{\mathbf{p}-} a_{\mathbf{p}-})\, p^\mu. \tag{2.185}$$

f. Verify that $Q\, a^\dagger_{\mathbf{p}+}|0\rangle = +a^\dagger_{\mathbf{p}+}|0\rangle$

The usual words that go with this model is as follows: $a^\dagger_{\mathbf{p}+}$ is the creation operator for particles, $a^\dagger_{\mathbf{p}-}$ is the creation operator for antiparticles. The particles have charge $+1$, the antiparticles have charge -1 and Q counts the number of particles minus the number of antiparticles. Particles and antiparticles have the same mass.

In the generalization to n fields with $k = 1, 2, \ldots, n$ $(n > 2)$ the internal symmetry group is larger $(O(2) \to O(n))$, with $n(n-1)/2$ Noether charges, and no such natural division into 'particles' and 'antiparticles'.

Returning to $n = 2$, an often used notation is in terms of complex fields

$$\varphi \equiv \frac{1}{\sqrt{2}}\,(\phi_1 - i\phi_2). \tag{2.186}$$

In this case φ and φ^* are treated as independent variables.

g. Show that the action can be rewritten as

$$S = -\int d^4x \left[\partial_\mu\varphi^*\partial^\mu\varphi + m^2\varphi^*\varphi + \lambda(\varphi^*\varphi)^2\right]. \tag{2.187}$$

Rederive Noether's theorem and its consequences in this notation. Go through the canonical formalism, quantization, etc. Express φ and its canonical conjugate π_φ in terms of the creation and annihilation operators.

Calculate the matrix element $\langle\psi|j^\mu(x)|\psi\rangle$ for a wave packet state of a single particle as in (2.120) and compare with (2.129).

9. *Pauli-Villars regularization of* $\langle 0|T^{\mu\nu}|0\rangle$

The integral (cf. (2.40)) in

$$\begin{aligned}\langle 0|T^{00}|0\rangle &= \epsilon_0 + \frac{4\pi}{2(2\pi)^3}\int_0^\infty dp\, p^2 \sum_i c_i\sqrt{m_i^2 + p^2} \qquad (2.188)\\ &= \epsilon_0 + \frac{4\pi}{2(2\pi)^3}\int^\infty dp\, p^2 \sum_i c_i\left(p + \frac{1}{2}\frac{m_i^2}{p} - \frac{m_i^4}{8p^3} + \cdots\right),\end{aligned}$$

converges at the upper limit provided that

$$\sum_i c_i = 0, \quad \sum_i c_i m_i^2 = 0, \quad \sum_i c_i m_i^4 = 0. \tag{2.189}$$

where $c_1 = 1$ and $m_1 = m$. Verify that the above conditions can be satisfied with four c's and m's. Hint: make a suitable choice of the ratios $r_i = m_i/m$, $i = 2, 3, 4$, and solve for c_2, c_3 and c_4.

Verify that the regularized surface term in (2.43) $\propto \sum_i c_i\Lambda^3\sqrt{m_i^2 + \Lambda^2}$ vanishes in the limit $\Lambda \to \infty$.

Verify that $\langle 0|T^{kl}|0\rangle$ is also finite.

One can add the requirement that the integrand of the originally divergent integral should be recovered when the regulator masses are sent to infinity. To implement this an additional c_5 and m_5 would be needed such that in addition to the requirements (2.189) also

$$\sum_{i\geq 2} c_i m_i = 0. \tag{2.190}$$

Calculate ϵ in terms of the c_i and m_i. Hint 1: calculate $\langle 0|T^{\mu}{}_{\mu}|0\rangle = -4\epsilon$. Hint 2: check the identity

$$\frac{1}{2\omega} = \int_{-\infty}^{\infty} \frac{dp_4}{2\pi}\, \frac{1}{\omega^2 + p_4^2}. \tag{2.191}$$

Hint 3: write the integral for $\langle 0|T^{\mu}{}_{\mu}|0\rangle$ as a four-dimensional integral in euclidean momentum-space and evaluate this with the help of a spherical cutoff, using spherical coordinates,

$$\begin{aligned} 4(\epsilon - \epsilon_0) &= \int \frac{d^4p}{(2\pi)^4} \sum_i c_i m_i^2 \frac{1}{m_i^2 + \mathbf{p}^2 + p_4^2} \\ &= \frac{2\pi^2}{(2\pi)^4} \int_0^{\infty} dp\, p^3 \sum_i c_i m_i^2 \frac{1}{m_i^2 + p^2}, \\ &= \sum_i c_i m_i^2 \int_0^{\Lambda} dp\, p^3 \sum_i c_i m_i^2 \frac{1}{m_i^2 + p^2}, \qquad \Lambda \to \infty. \end{aligned} \tag{2.192}$$

In the second step we used that in four dimensions the integral over angles equals $2\pi^2$.

The remaining integral is elementary and the limit $\Lambda \to \infty$ can be taken without a problem because of the properties of the c_i and m_i. The result is (verify)

$$4\epsilon = 4\epsilon_0 + \frac{m^4}{16\pi^2} \ln \frac{m^2}{\mu^2} + \sum_{i\geq 2} c_i \frac{m_i^4}{16\pi^2} \ln \frac{m_i^2}{\mu^2}, \tag{2.193}$$

where μ^2 is some arbitrary mass scale introduced for convenient separation of the Λ dependence (the result does not depend on μ).

If we now let the regulator masses approach infinity, $m_i \to \infty$, $i = 2$, 3, 4, 5, the regulator part appears to diverge quartically, $\propto m_i^4$, which has to be compensated by ϵ_0 such that ϵ remains finite. In common parlor: ϵ_0 is renormalized to ϵ.

Example, using Mathematica with $m_2 = 2rm$, $m_3 = 3rm$, $m_4 = 4rm$, $m_5 = 5rm$, gives

$$4\epsilon = 4\epsilon_0 + m^4 \left(a_4 r^4 + a_2 r^2 + a_0 \ln r + a_0'\right) \tag{2.194}$$

plus terms that vanish as $r \to \infty$, with

$$a_4 = -\frac{30}{7}\,(32\ln 4 - 279\ln 9 + 640\ln 16 - 375\ln 25), \tag{2.195}$$

etc.

Chapter 3

Path integral methods

The path integral approach to quantization has resulted in a powerful language for quantum field theory. Here we shall try to be brief and concentrate on applications to perturbation theory. We have seen in simple cases that calculations need answers for the vacuum expectation value of products of field, $\langle 0|\hat{\varphi}(x_1)\cdots\hat{\varphi}(x_n)|0\rangle$, and we shall develop the calculational tools for this.

3.1 Path integral for quantum mechanics

Consider a simple system described by the Lagrange function $L = L(q,\dot{q},t)$ or the corresponding Hamilton function $H = H(p,q,t)$,

$$L = \frac{1}{2}m\dot{q}^2 - V(q,t), \quad H = \frac{p^2}{2m} + V(q,t), \tag{3.1}$$

where p and q are related by $p = \partial L/\partial\dot{q} = m\dot{q}$. In the quantum theory p and q become operators[1] $\hat{p}$, $\hat{q}$ with $[\hat{q},\hat{p}] = i\hbar$. The coordinate basis $|q\rangle$ is characterized by

$$\hat{q}|q\rangle = q|q\rangle, \tag{3.2}$$

$$\langle q'|q\rangle = \delta(q'-q), \quad \int dq\,|q\rangle\langle q| = 1, \tag{3.3}$$

The time evolution is described by the operator $\hat{U}(t',t'')$. In case the potential V does not depend explicitly on time, the hamiltonian $\hat{H}$ is time independent and the evolution operator is given by

$$\hat{U}(t',t'') = \exp[-i\hat{H}(t'-t'')/\hbar], \tag{3.4}$$

where we have temporarily made $\hbar$ explicit. Path integral quantization will give us a representation of the matrix element

[1]When helpful, we indicate operators in Hilbert space by a $\hat{}$.

$\langle q'|\hat{U}(t',t'')|q''\rangle$,

$$\langle q'|\hat{U}(t',t'')|q''\rangle = \int d[q]\, \exp\left\{\frac{i}{\hbar}S[q]\right\}. \tag{3.5}$$

Here S is the action functional of the system,

$$S[q] = \int_{t''}^{t'} dt\, L(q(t),\dot{q}(t)), \tag{3.6}$$

and $\int d[q]$ symbolizes an integration over all functions $q(t)$ such that

$$q(t') = q', \quad q(t'') = q''. \tag{3.7}$$

Note that $\hbar$ appears explicitly in (3.5), all other symbols are 'classical'. Instead of working with q-numbers (operators) $\hat{p}$ and $\hat{q}$, the path integral allows us to work with time dependent c-numbers (commuting numbers) $q(t)$.

The path integral is a summation over all 'paths' ('trajectories', 'histories') $q(t)$, with given end points q', q''. The classical path, which satisfies the equation of motion

$$\frac{\delta S}{\delta q(t)} = \frac{\partial L}{\partial q(t)} - \frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{q}(t)} = 0, \tag{3.8}$$

is only one out of infinitely many possible paths. Each path has a 'weight' $\exp(iS/\hbar)$. If $\hbar$ is relatively small such that the phase $\exp(iS/\hbar)$ varies rapidly over the paths, then a stationary phase approximation will be good in which the clasical path and its small neighborhood gives the dominant contribution. The other extreme is where the variation of $S/\hbar$ is of order 1 and the system is in the quantum regime. In the following we shall use units again in which $\hbar = 1$.

A formal definition of $\int d[q]$ is given by

$$\int d[q] = \prod_{t''<t<t'} \int dq(t), \tag{3.9}$$

i.e. for every $t \in (t'',t')$ we integrate over the domain of q, e.g. $-\infty < q < \infty$. The definition is formal because the continuous product $\prod_t$ still has to be defined. We shall give such a definition with the help of a discretization procedure.

3.2 Regularization by discretization

To define the path integral properly we discretize time in small units a, writing

$$t'-t'' = Na, \quad t_n = na, \quad n = 0,1,\ldots,N, \quad q(t_n) = q_n, \quad q_0 = q'', \quad q_N = q'. \tag{3.10}$$

For a smooth function $q(t)$ the time derivative $\dot{q}(t)$ can be approximated by $\dot{q}(t_n) \approx (q_{n+1}-q_n)/a$, such that the discretized Lagrange function may be written as

$$L_n^{\rm discr}(q_{n+1},q_n) = \frac{m}{2a^2}(q_{n+1}-q_n)^2 - \frac{1}{2}V_{n+1}(q_{n+1}) - \frac{1}{2}V_n(q_n). \tag{3.11}$$

We have devided the potential term equally between times n and $n+1$, and $V_n(q) \equiv V(q, t_n)$. Except at the end points, the two halves add up to one in the discretized action defined by

$$\begin{aligned} S_{\text{discr}}[q] &= a \sum_{n=0}^{N-1} L_n^{\text{discr}}(q_{n+1}, q_n) \\ &= -\frac{a}{2} V_N(q') + \frac{m}{2a}(q' - q_{N-1})^2 - aV_{N-1}(q_{N-1}) \\ &\quad + \frac{m}{2a}(q_{N-1} - q_{N-2})^2 - aV_{N-2}(q_{N-2}) \\ &\quad + \cdots + \frac{m}{2a}(q_1 - q'')^2 - \frac{a}{2} V_0(q''). \end{aligned} \tag{3.12}$$

A tentative definition of the path integral is now

$$\int d[q] \, \exp\{iS[q]\} = \lim_{N \to \infty} c^N \int (\prod_{n=1}^{N-1} dq) \exp\{iS_{\text{discr}}[q]\}, \tag{3.13}$$

where the constant c is still to be specified.

To show that this definition of the path integral gives us the evolution operator, we introduce a minimal step evolution operator $\hat{T}$ by giving its matrix elements

$$\langle q_1 | \hat{T}_n | q_2 \rangle = c \exp[iaL_n^{\text{discr}}(q_1, q_2)]. \tag{3.14}$$

The operator $\hat{T}$ is called the transfer operator, its matrix elements the transfer matrix. In terms of the transfer operator we have

$$\begin{aligned} c^N \int (\prod_{n=1}^{N-1} dq) \, e^{iS_{\text{discr}}[q]} &= \int dq_1 \cdots dq_{N-1} \langle q' | \hat{T}_{N-1} | q_{N-1} \rangle \cdots \langle q_2 | \hat{T}_1 | q_1 \rangle \langle q_1 | \hat{T}_0 | q'' \rangle \\ &= \langle q' | \hat{T}_{N-1} \hat{T}_{N-2} \cdots \hat{T}_0 | q'' \rangle. \end{aligned} \tag{3.15}$$

Now we are almost done. First, with a suitable choice of the constant c the transfer operator can be written in the form

$$\hat{T}_n = e^{-iaV_{n+1}(\hat{q})/2} \, e^{-ia\hat{p}^2/2m} \, e^{-iaV_n(\hat{q})/2}. \tag{3.16}$$

Taking matrix elements between $\langle q_1 |$ and $| q_2 \rangle$ and comparing with (3.11), we see that this formula is correct if

$$\langle q_1 | e^{-ia\hat{p}^2/2m} | q_2 \rangle = c \, e^{im(q_1 - q_2)^2/2a}. \tag{3.17}$$

Inserting eigenstates $|p\rangle$ of the momentum operator $\hat{p}$ using

$$\langle q | p \rangle = e^{ipq}, \quad \int \frac{dp}{2\pi} \, |p\rangle\langle p| = 1, \tag{3.18}$$

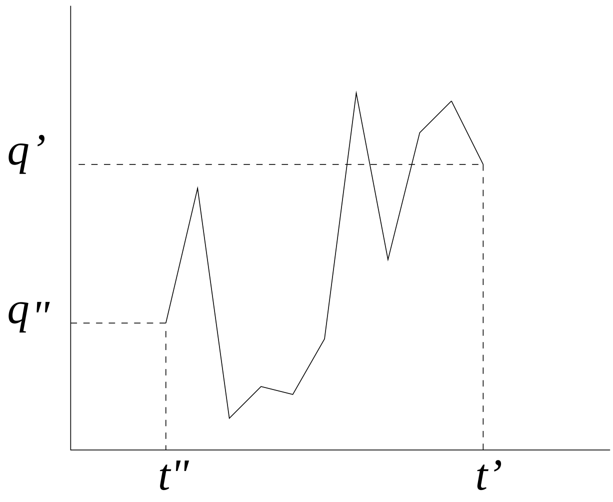

Figure 3.1: Typical path contributing to the discretized path integral.

and the gaussian integral[2]

$$\int_{-\infty}^{\infty} dp \, \exp\left(-\frac{1}{2}(ia/m)p^2 + irp\right) = \sqrt{\frac{2\pi}{ia/m}} \, \exp\left(\frac{1}{2}\frac{(ir)^2}{ia/m}\right) \tag{3.19}$$

we find that (3.17) is true provided that we choose

$$c = \sqrt{\frac{m}{2\pi ia}} = \sqrt{\frac{m}{2\pi a}} e^{-i\pi/4}. \tag{3.20}$$

Second, using the Baker-Cambell-Haussdorff formula it is easy to show that (cf. Problem 1)

$$\hat{T}_n = \exp[-ia\hat{H}_n + O(a^3)] = \hat{U}(t_{n+1}, t_n) + O(a^3). \tag{3.21}$$

Hence,

$$\hat{T}_{N-1}\hat{T}_{N-2}\cdots\hat{T}_0 = \hat{U}(t', t'') + O(a^2), \tag{3.22}$$

and together with (3.13), (3.15) we get the desired result

$$\int d[q] \, \exp\{iS[q]\} = \langle q'|\hat{U}(t', t'')|q''\rangle. \tag{3.23}$$

For finite N the errors are of order a^2. This is important in practical computations and is the reason why we have devided the potential term in (3.11) over two 'time slices'.

In case $V(q,t)$ does not depend explicitly on time, the subscript n in L_n, V_n, $\hat{T}_n$ and $\hat{H}_n$ can be dropped, and $\prod_n \hat{T}_n = \hat{T}^N = \exp[-i\hat{H}(t' - t'') + O(a^2)]$.

[2]This is a *convergent* integral. It is more appropriately called a Fresnel integral, but we shall generically call such integrals 'gaussian' as if the exponential were real.

In the limit $N \to \infty$, the discretized action $S_{\rm discr}[q]$ becomes equal to the continuum action $S[q]$ when we substitute smooth functions $q(t)$. However, because the q_n are integrated over on every time slice n, such smoothness is not present typically in the integrand of the path integral (typical paths q_n have highly discontinuous first derivatives cf. Fig. 3.1). Hence, a continuum limit $S_{discr}[q] \to S[q]$ cannot be taken in the integrand of the path integral. We regard the path integral as *defined* by the discretization procedure, with the limit $a \to 0$ ($N \to \infty$) taken only at suitable points in calculations.[3] In the following we shall mostly work formally, in a notation in which the limit has already been taken (even under the integral), which gives nice intuitive formulas. However, we may have to return sometimes to the original definition to avoid mathematical ambiguities. In the 'lattice field theory' method such ambiguities are avoided by working only with the well-defined discretized forms.

3.3 Imaginary time

It is interesting to make an analytic continuation to imaginary time according to the substitution $t \to -it$. This can be justified if the potential $V(q)$ is bounded from below. For example, this is the case for the anharmonic oscillator

$$V(q) = \frac{1}{2}m\omega^2 q^2 + \frac{1}{4}\lambda q^4, \quad \lambda \geq 0, \tag{3.24}$$

where $q \in (-\infty, \infty)$.

Consider the discretized path integral (3.15). The integrations over the variables q_n converge at large q_n, and continue to converge if we rotate a in the complex plane according to

$$a = |a|e^{-i\delta}, \ \ \delta : 0 \to \frac{\pi}{2}. \tag{3.25}$$

The reason is that for all $\delta \in (0, \pi)$ the real part of the exponent in (3.15) is negative:

$$\frac{i}{a} = \frac{i}{|a|e^{-i\delta}} = \frac{1}{|a|}(-\sin\delta + i\cos\delta), \quad -ia = -i|a|e^{-i\delta} = |a|(-\sin\delta - i\cos\delta). \tag{3.26}$$

The result of this analytic continuation in a is that the discretized path integral takes the form

$$\begin{aligned} \langle q'|\hat{T}^N|q''\rangle &= |c|^N \int (\prod_n dq_n) \exp[-S_I^{\rm discr}], \qquad (3.27)\\ S_I^{\rm discr} &= |a| \sum_{n=0}^{N-1} \left[\frac{m}{2|a|^2}(q_{n+1}-q_n)^2 + \frac{1}{2}V(q_{n+1}) + \frac{1}{2}V(q_n)\right]. \end{aligned}$$

[3] There are also other definitions possible, see e.g. Problem 2. The discretization method has the advantage that it can be applied also to gauge theories.

Here the subscript I denotes the imaginary time version of S. After transformation to imaginary time the transfer operator takes the hermitian form

$$\hat{T} = e^{-bV(\hat{q})/2}\, e^{-b\hat{p}^2/2m}\, e^{-bV(\hat{q})/2}, \quad b = |a|. \tag{3.28}$$

This is a positive operator, i.e. all its expectation values and hence also all its eigenvalues are positive. We may therefore define a discretized hermitian Hamilton operator $\hat{H}_{\rm discr}$ according to

$$\hat{T} = \exp[-b\hat{H}_{\rm discr}], \quad \hat{H}_{\rm discr} = \hat{H} + O(b^2), \tag{3.29}$$

and interpret this as the hamiltonian of the discrete system.

A natural object in the imaginary time formalism is the partition function

$$Z = {\rm Tr}\, e^{-\hat{H}(t_+ - t_-)} = \int dq \langle q|e^{-\hat{H}(t_+ - t_-)}|q\rangle = \lim_{N\to\infty} {\rm Tr}\, \hat{T}^N, \tag{3.30}$$

where we think of t_+ (t_-) as the largest (smallest) time under consideration, with $t_+ - t_- = Na$. From quantum statistical mechanics we recognize that Z is the canonical partition function corresponding to the temperature

$$T = (t_+ - t_-)^{-1}, \tag{3.31}$$

in units where Boltzmann's constant $k_B = 1$. The path integral representation of Z is obtained by setting in (3.27) $q_N = q_0 \equiv q$ ($q' = q'' \equiv q$) and integrating over q:

$$Z = \int_{pbc} d[q]\, e^{-S_I}. \tag{3.32}$$

Here '*pbc*' indicates the fact that the integration is now over all discretized functions $q(t)$, $t_- < t < t_+$, with periodic boundary conditions, $q(t_+) = q(t_-)$.

The integrand $\exp(-S_I)$ in the imaginary time path integral is real and bounded from above. This makes numerical calculations and theoretical analysis very much easier. Furthermore, in the generalization to field theory to be given later there is a direct connection to Statistical Physics, which has led to many fruitful developments. The imaginary time formulation can be sufficient to extract the relevant physical information, without the need to continue back to real time. A prime example is the mass spectrum of QCD, where the eigenstates of the hamiltonian are bound states of quarks and gluons. The analogy of $\exp(-S_I)$ with a Boltzmann factor makes it possible to address this complicated problem by Monte Carlo computations.

In the following we shall continue with ordinary but slightly modified real time: we assume the rotation angle $\delta > 0$ to be infinitesimal, i.e. very small and going to zero at suitable stages of the calculations. This has several advantages. For now we note that convergence is improved: integrals like

$$\int_{-\infty}^{\infty} dq\, \exp\left[\frac{i}{a}\frac{1}{2}m(q-q')^2 - ia\frac{1}{2}m\omega^2 q^2\right] q^n \tag{3.33}$$

are absolutely convergent for $\delta > 0$ for $n = 1, 2, \cdots$ and not only (conditionally convergent) for $n = 0$ (which is the case for $\delta = 0$). We also note at this point that it does not hurt to add another convergence device, which will be useful later: replace

$$\omega^2 \to \omega^2 - i\epsilon, \tag{3.34}$$

with $\epsilon > 0$ infinitesimal. This produces the convergence factor $\exp(-a\epsilon\frac{1}{2}mq^2)$ (we are assuming $m > 0$). We shall suppress ϵ and δ in our notation in the following, unless explicitely needed.

3.4 External force technique

An important technique in field theory is probing the response of the system to an external source, the analogue of which is here an external force. The potential is taken to be of the form

$$V(q,t) = V(q) - f(t)q, \tag{3.35}$$

where $f(t)$ is an arbitrary function of time. We redefine the notation and use S and L for the action and lagrangian of the unperturbed system,

$$S = \int dt\, L, \quad L = \frac{1}{2}m\dot{q}^2 - V(q). \tag{3.36}$$

From the equations of motion

$$m\ddot{q} = -\frac{\partial V}{\partial q} + f. \tag{3.37}$$

follows that $f(t)$ has the interpretation of an external force. The path integral for the evolution matrix is

$$\langle q'|\hat{U}(t',t'')|q''\rangle = \int d[q]\, e^{iS[q]+i\int dt\, f(t)q(t)}. \tag{3.38}$$

We are going to differentiate this expression with respect to f ('probing the system').

Since we shall be mostly using the Heisenberg picture in which the operators are time dependent,

$$\hat{q}(t) = \hat{U}(t,0)^\dagger \hat{q}\, \hat{U}(t,0), \tag{3.39}$$

it will be convenient to use corresponding coordinate basis vectors $|q,t\rangle$ in Hilbert space with the properties

$$\hat{q}(t)|q,t\rangle = q|q,t\rangle, \tag{3.40}$$

$$\langle q,t|q',t\rangle = \delta(q-q'), \quad \int dq\, |q,t\rangle\langle q,t| = \hat{1}. \tag{3.41}$$

This basis can be chosen as

$$|q,t\rangle = \hat{U}(t,0)^\dagger |q\rangle = \hat{U}(0,t)|q\rangle. \tag{3.42}$$

Note that $|q,t\rangle$ is not the Schrödinger state at time t which developed out of the initial $|q\rangle$ at time zero: this would be $\hat{U}(t,0)|q\rangle$.

For the differentiation with respect to f we use again the method of small time-steps a, $a = (t'-t'')/N$, $t_k = ka$, $k = 0, \cdots, N$, $N \to \infty$. We write

$$\begin{aligned}\langle q',t'|q'',t''\rangle &= \langle q'|\hat{U}(t',t'')|q''\rangle = \langle q'|\prod_{k=0}^{N-1}\hat{U}(t_{k+1},t_k)|q''\rangle \\ &= \int dq_1\cdots dq_{N-1}\prod_{k=0}^{N-1}\langle q_{k+1},t_{k+1}|q_k,t_k\rangle.\end{aligned} \tag{3.43}$$

The one-step transition amplitudes are approximated by the matrix elements of the transfer operator,

$$\begin{aligned}\langle q_{k+1},t_{k+1}|q_k,t_k\rangle &= \langle q_{k+1}|\hat{U}(t_k+a,t_k)|q_k\rangle \approx \langle q_1|\hat{T}_k|q_2\rangle \\ &= c\exp\left\{iaL(q_1,q_2) + ia\frac{1}{2}\left[f(t_{k+1})q_1 + f(t_k)q_2\right]\right\}.\end{aligned} \tag{3.44}$$

Assuming $t_j \approx t \in (t',t'')$ we then have

$$\begin{aligned}\frac{\delta}{i\delta f(t)}\langle q',t'|q'',t''\rangle &\approx \frac{\partial}{ia\partial f(t_j)}\int dq_1\cdots dq_{N-1}\prod_{k=0}^{N-1}\langle q_{k+1},t_{k+1}|q_k,t_k\rangle \\ &= \int dq_1\cdots dq_{N-1}\prod_{k=j+1}^{N-1}\langle q_{k+1},t_{k+1}|q_k,t_k\rangle \\ &\quad \frac{\partial}{ia\partial f(t_j)}\langle q_{j+1},t_{j+1}|q_j,t_j\rangle\langle q_j,t_j|q_{j-1},t_{j-1}\rangle \\ &\quad \prod_{k=0}^{j-2}\langle q_{k+1},t_{k+1}|q_k,t_k\rangle \\ &= \int dq_1\cdots dq_{N-1}\prod_{k=j+1}^{N-1}\langle q_{k+1},t_{k+1}|q_k,t_k\rangle \\ &\quad \langle q_{j+1},t_{j+1}|q_j,t_j\rangle\, q_j\, \langle q_j,t_j|q_{j-1},t_{j-1}\rangle \\ &\quad \prod_{k=0}^{j-2}\langle q_{k+1},t_{k+1}|q_k,t_k\rangle \\ &\to \int dq\, \langle q',t'|q,t\rangle\, q\, \langle q,t|q'',t''\rangle,\end{aligned} \tag{3.45}$$

as $N \to \infty$. Note that q above is just a dummy integration variable. The combination

$$\int dq\, |q,t\rangle\, q\, \langle q,t| = \hat{q}(t), \tag{3.46}$$

is the Heisenberg operator at time t. So we have the result

$$\frac{\delta}{i\delta f(t)}\langle q',t'|q'',t''\rangle = \langle q',t'|\hat{q}(t)|q'',t''\rangle. \tag{3.47}$$

Differentiating again, using the chain rule in (3.45) we get two contributions,

$$\begin{aligned}
\frac{\delta}{i\delta f(t_2)}\frac{\delta}{i\delta f(t_1)}\langle q',t'|q'',t''\rangle &= \frac{\delta}{i\delta f(t_2)}\int dq_1\,\langle q',t'|q_1,t_1\rangle\, q_1\,\langle q_1,t_1|q'',t''\rangle \\
&= \int dq_1\,\frac{\delta}{i\delta f(t_2)}\left(\langle q',t'|q_1,t_1\rangle\right)\, q_1\,\langle q_1,t_1|q'',t''\rangle \\
&\quad + \int dq_1\,\langle q',t'|q_1,t_1\rangle\, q_1\,\frac{\delta}{i\delta f(t_2)}\left(\langle q_1,t_1|q'',t''\rangle\right).
\end{aligned}$$

The first term only contributes for $t_2 > t_1$, the second term only contributes for $t_2 < t_1$ (recall that we only consider times $t'' < t_{1,2} < t'$). We can distinguish these two cases with the help of the Heavyside step function

$$\begin{aligned}
\theta(t) &= 1, \quad t > 0 \\
&= 0, \quad t < 0,
\end{aligned} \tag{3.48}$$

and the result of the differentiation can be expressed in the form (3.47),

$$\begin{aligned}
\frac{\delta}{i\delta f(t_2)}\frac{\delta}{i\delta f(t_1)}\langle q',t'|q'',t''\rangle &= \int dq_1\,[\theta(t_2-t_1)\langle q',t'|\hat{q}(t_2)|q_1,t_1\rangle\, q_1\,\langle q_1,t_1|q'',t''\rangle \\
&\quad + \theta(t_1-t_2)\langle q',t'|q_1,t_1\rangle\, q_1\,\langle q_1,t_1|\hat{q}(t_2)|q'',t''\rangle] \\
&= \theta(t_2-t_1)\langle q',t'|\hat{q}(t_2)\hat{q}(t_1)|q'',t''\rangle \\
&\quad + \theta(t_1-t_2)\langle q',t'|\hat{q}(t_1)\hat{q}(t_2)|q'',t''\rangle.
\end{aligned} \tag{3.49}$$

This can be compactly written with the help of the time ordering instruction symbol T:

$$\frac{\delta}{i\delta f(t_2)}\frac{\delta}{i\delta f(t_1)}\langle q',t'|q'',t''\rangle = \langle q',t'|T\hat{q}(t_1)\hat{q}(t_2)|q'',t''\rangle, \tag{3.50}$$

which is defined as

$$T\hat{q}(t_1)\hat{q}(t_2) = \theta(t_1-t_2)\hat{q}(t_1)\hat{q}(t_2) + \theta(t_2-t_1)\hat{q}(t_2)\hat{q}(t_1). \tag{3.51}$$

The above process evidently generalizes to

$$\frac{\delta}{i\delta f(t_1)}\cdots\frac{\delta}{i\delta f(t_n)}\langle q',t'|q'',t''\rangle = \langle q',t'|T\hat{q}(t_1)\cdots\hat{q}(t_n)|q'',t''\rangle, \tag{3.52}$$

with T the instruction to write the operators in order of decreasing times (writing this in terms of θ functions gets cumbersom for $n > 2$). Note that these so-called time ordered products are symmetric under interchanging labels $t_i \leftrightarrow t_j$.

Setting $f = 0$ after differentiation we get the expansion 'coefficients' of a 'Taylor series' in powers of f (called a Volterra series in case of functionals). So, summing the series we have

$$\begin{aligned}
\langle q', t'|q'', t''\rangle[f] &= \sum_{n=0}^{\infty} \frac{1}{n!} \int dt_1 \cdots dt_n \left(\frac{\delta}{\delta f(t_1)} \cdots \frac{\delta}{\delta f(t_n)} \langle q', t'|q'', t''\rangle \right) [0] \\
&\quad f(t_1) \cdots f(t_n) \\
&= \sum_{n=0}^{\infty} \frac{i^n}{n!} \int dt_1 \cdots dt_n \, \langle q', t'|T\hat{q}(t_1) \cdots \hat{q}(t_n)|q'', t''\rangle[0] \, f(t_1) \cdots f(t_n) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \langle q', t'|T \left[i \int dt \, f(t)\hat{q}(t) \right]^n |q'', t''\rangle[0] \\
&= \langle q', t'|T \, e^{i\int dt\, f(t)\hat{q}(t)}|q'', t''\rangle[0],
\end{aligned} \tag{3.53}$$

where we indicated the functional dependence on $f(t)$ in square brackets. Note that on the right hand side $\hat{q}(t)$ evolves in time according to the hamiltonian with the external force set to zero.

On the other hand, we have path integral representations for various expressions:

$$\langle q', t'|q'', t''\rangle[f] = \int d[q] \, e^{iS[q]+i\int dt\, f(t)q(t)}, \tag{3.54}$$

and

$$\langle q', t'|T\hat{q}(t_1) \cdots \hat{q}(t_n)|q'', t''\rangle[0] = \int d[q] \, e^{iS[q]} \, q(t_1) \cdots q(t_n), \tag{3.55}$$

because

$$\frac{\delta}{i\delta f(t_1)} \cdots \frac{\delta}{i\delta f(t_n)} \, e^{i\int dt\, f(t)q(t)} = q(t_1) \cdots q(t_n) \, e^{i\int dt\, f(t)q(t)}. \tag{3.56}$$

3.5 Ground state expectation values

In field theory we are especially interested in expectation values in the ground state. We now illustrate how these can be calculated in our quantum mechanical model. In the following $\hat{H}$ is the hamiltonian of the model with $f = 0$. The crucial assumption will be that the spectrum of $\hat{H}$ is bounded from below, i.e. there is a ground state with minimal energy:

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle, \quad \hat{H}|\psi_0\rangle = E_0|\psi_0\rangle \quad E_{n\neq 0} > E_0. \tag{3.57}$$

We have made the further simplifying assumption that there is an energy gap between the ground state energy E_0 and the first excited state energy E_1.

Suppose now that $f(t)$ is non-zero only in a bounded region $t_- < t < t_+$. Then the hamiltonian is the time-independent $\hat{H}$ for $t < t_-$ and $t > t_+$, and we may write

$$\begin{aligned}
\langle q'|\hat{U}(t', t'')|q''\rangle &= \langle q'|\hat{U}(t', t_+)\hat{U}(t_+, t_-)\hat{U}(t_-, t'')|q''\rangle \\
&= \langle q'|e^{-i\hat{H}(t'-t_+)}\hat{U}(t_+, t_-)e^{-i\hat{H}(t_- - t'')}|q''\rangle.
\end{aligned} \tag{3.58}$$

We now use completeness,

$$\sum_n |\psi_n\rangle\langle\psi_n| = \hat{1}, \tag{3.59}$$

and insert intermediate states to obtain for the right hand side

$$\sum_{mn} e^{-iE_m(t'-t_+)} \langle q'|\psi_m\rangle\langle\psi_m|\hat{U}(t_+,t_-)|\psi_n\rangle\langle\psi_n|q''\rangle e^{-iE_n(t_- -t'')}. \tag{3.60}$$

Consider now letting $t' \to \infty$, $t'' \to -\infty$, keeping in mind the fact that our times are rotated over a small negative angle δ away from the real axis, as motivated in Sect. 3.3. Let us make this temporarily explicit by writing

$$t = \tau(1 - i\delta), \quad \delta > 0. \tag{3.61}$$

With $\delta > 0$, $\tau' \to +\infty$ and $\tau'' \to -\infty$ the energy exponentials such as

$$e^{-iE_m(t'-t_+)} = e^{-(\tau'-\tau_+)E_m\delta}\, e^{-(\tau'-\tau_+)iE_m} \tag{3.62}$$

suppress the excited states relative to the ground state. So the ground state contribution dominates,

$$\begin{aligned} \langle q',t'|q'',t''\rangle &= \langle q'|\hat{U}(t',t'')|q''\rangle \qquad (3.63)\\ &\to e^{-iE_0(t'-t_+)} \langle q'|\psi_0\rangle\langle\psi_0|\hat{U}(t_+,t_-)|\psi_0\rangle\langle\psi_0|q''\rangle e^{-iE_0(t_- -t'')}. \end{aligned}$$

The neglected terms are suppressed by factors $\exp[-\tau'(E_m - E_0)\delta]$ and $\exp[\tau''(E_m - E_0)\delta]$, $m \neq 0$, relative to the ground state. Note the appearence of the ground-state wave function

$$\langle q|\psi_0\rangle = \psi_0(q), \tag{3.64}$$

which provides a way to calculate ψ_0 from the path integral.

In a similar fashion we can derive, using (3.53), (3.42) and $U(0,t) = e^{iHt}$,

$$\begin{aligned} \langle q',t'|q'',t''\rangle[f] &= \langle q',t'|T\, e^{i\int dt\, f(t)\hat{q}(t)}|q'',t''\rangle[0] \\ &\to e^{-iE_0t'} \langle q'|\psi_0\rangle\langle\psi_0|T\, e^{i\int dt\, f(t)\hat{q}(t)}|\psi_0\rangle[0]\, \langle\psi_0|q''\rangle e^{iE_0t''}. \end{aligned}$$

We can cancel the dependence on the boundaries in time by deviding by the same expression with $f = 0$, and arrive at the neat formula

$$\begin{aligned} \langle\psi_0|T\, e^{i\int dt\, f(t)\hat{q}(t)}|\psi_0\rangle &= \frac{\langle q',\infty|q'',-\infty\rangle[f]}{\langle q',\infty|q'',-\infty\rangle[0]} \\ &= \frac{\int d[q]\, e^{iS[q]+i\int dt\, f(t)q(t)}}{\int d[q]\, e^{iS[q]}}, \end{aligned} \tag{3.65}$$

where we used $\langle\psi_0|\psi_0\rangle = 1$. Since the boundary effects drop out in the ratio (3.65), we may as well choose boundary values $q' = 0$, $q'' = 0$, since this allows us to do partial integrations without having to worry about boundary terms. Thus we define the path integral for zero boundary values at infinite times,

$$Z[f] \equiv \langle 0, \infty | 0, -\infty \rangle = \int d[q]\, e^{iS[q]+i\int dt\, f(t)q(t)}, \tag{3.66}$$

in terms of which

$$\langle \psi_0 | T\, e^{i\int dt\, f(t)\hat{q}(t)} | \psi_0 \rangle = \frac{Z[f]}{Z[0]}. \tag{3.67}$$

In field theory the analogous expressions will turn out to be very useful.

3.6 Harmonic oscillator

The results so far will now be illustrated by explicit calculations for the important case of the harmonic oscillator. We shall be brief and only give a formal evaluation of the ratio of path integrals with and without external force. Our first task is to evaluate $Z[f]$ in (3.66), or rather the ratio $Z[f]/Z[0]$. The action is given by

$$S[q] = \int dt \left(\frac{1}{2} m\dot{q}^2 - \frac{1}{2} m\omega^2 q^2 \right) = -\frac{1}{2} \int dt\, q \left(m\frac{\partial^2}{\partial t^2} + m\omega^2 \right) q. \tag{3.68}$$

On the r.h.s. we made a partial integration, for which the boundary terms vanish in the case (3.66). We now write this in suggestive matrix form,

$$S = -\frac{1}{2} q^T G^{-1} q, \tag{3.69}$$

with continuous matrix index t and

$$G^{-1}(t, t') = \left(m\frac{\partial^2}{\partial t^2} + m\omega^2 \right) \delta(t - t'). \tag{3.70}$$

For example,

$$\left(G^{-1} q \right)(t) = \int dt'\, G^{-1}(t, t') q(t') = \left(m\frac{\partial^2}{\partial t^2} + m\omega^2 \right) q(t). \tag{3.71}$$

Note that G^{-1} is symmetric:

$$G^{-1}(t, t') = G^{-1}(t', t). \tag{3.72}$$

The path integral now looks like a multiple gaussian integral[4]

$$Z[f] = \int d[q]\, \exp\left(-\frac{1}{2} i q^T G^{-1} q + i f^T q \right), \tag{3.73}$$

[4] In discretized form it *is* an ordinary multiple gaussian integral, and $G^{-1}(t_n, t_{n'})$ is just an ordinary matrix.

and we shall evaluate the dependence on f by a shift of integration variables,

$$q = q' + Gf, \tag{3.74}$$

where G is the inverse of G^{-1} under the given boundary conditions,

$$G^{-1}G = 1. \tag{3.75}$$

This inverse is also symmetric, $G^T = G$, which will be verified below. The shift makes the exponent purely quadratic in q',

$$-\frac{1}{2}q^T G^{-1} q + f^T q = -\frac{1}{2}q'^T G^{-1} q' + \frac{1}{2} f^T G f. \tag{3.76}$$

The integration measure is invariant under shifts,

$$d[q] = d[q'], \quad \text{or} \quad \int_{-\infty}^{\infty} \prod_t dq(t)\, F[q] = \int_{-\infty}^{\infty} \prod_t dq'(t)\, F[q' + Gf]. \tag{3.77}$$

So we get

$$Z[f] = Z[0] \exp\left(i\frac{1}{2} f^T G f\right), \tag{3.78}$$

or more explicitly

$$Z[f] = Z[0] \exp\left[i\frac{1}{2}\int dt\, dt'\, f(t) G(t,t') f(t')\right]. \tag{3.79}$$

From the explicit form of (3.75),

$$\int dt''\, G^{-1}(t,t'')G(t'',t') = \left(m\frac{\partial^2}{\partial t^2} + m\omega^2\right) G(t,t') = \delta(t-t'), \tag{3.80}$$

we see that G is the Green function corresponding to the differential operator G^{-1}. It can be found by Fourier transformation. Assuming $G(t,t')$ to depend on the difference $t - t'$ we try

$$G(t,t') = \int \frac{dp}{2\pi}\, G(p) e^{-ip(t-t')}. \tag{3.81}$$

Since the Fourier transform of the Dirac delta function is 1 we have

$$m(-p^2 + \omega^2)G(p) = 1, \quad \text{or} \quad G(p) = \frac{1}{m(\omega^2 - p^2)}. \tag{3.82}$$

A prescription is needed to deal with the pole in the integrand of (3.81). The prescription for the present situation follows from the discussion about convergence given at the end of Sect. 3.3: we give ω^2 a small negative imaginary part:

$$\omega^2 \to \omega^2 - i\epsilon. \tag{3.83}$$

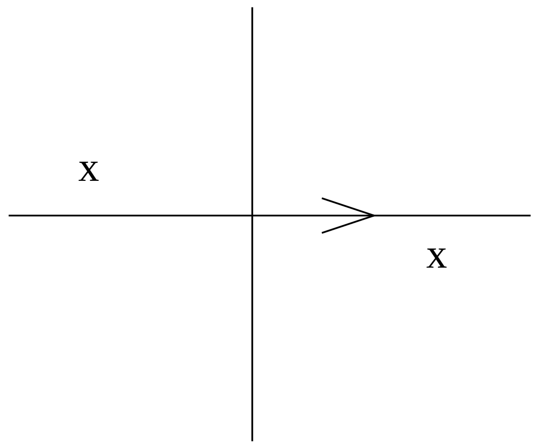

Figure 3.2: Position of the poles of $(\omega^2 - p^2 - i\epsilon)^{-1}$ in the complex p plane.

Evaluating the integral by contour integration leads to (cf. figure 3.2)

$$-i \int \frac{dp}{2\pi} \frac{1}{m} \frac{e^{-ip(t-t')}}{\omega^2 - p^2 - i\epsilon} = \theta(t-t') \frac{e^{-i\omega(t-t')}}{2m\omega} + \theta(t'-t) \frac{e^{i\omega(t-t')}}{2m\omega}. \qquad (3.84)$$

Hence,

$$-iG(t,t') = \frac{e^{-i\omega|t-t'|}}{2m\omega}. \qquad (3.85)$$

It is symmetric in $t \leftrightarrow t'$. It is complex. It is the analogue of the Feynman propagator for a scalar field.

The remaining path integral $Z[0]$ is formally just a multiple gaussian integral. It can be evaluated by expanding $q(t)$ in terms of a complete set of orthonormal eigenfunctions of G^{-1}:

$$\begin{aligned} G^{-1} u_\alpha &= \lambda_\alpha u_\alpha, \quad u_\alpha^T u_\beta = \delta_{\alpha\beta}, \\ q(t) &= \sum_\alpha q_\alpha u_\alpha(t), \quad q^T G^{-1} q = \sum_\alpha \lambda_\alpha q_\alpha^2. \end{aligned}$$

The orthonormal eigenfunctions form a unitary matrix $u_\alpha(t)$, and

$$\begin{aligned} d[q] &\propto \prod_\alpha dq_\alpha, \\ Z[0] &\propto \int_{-\infty}^{\infty} \prod_\alpha dq_\alpha \, e^{-i\frac{1}{2}\sum_\alpha \lambda_\alpha q_\alpha^2} \propto \left[\prod_\alpha \lambda_\alpha\right]^{-1/2} = [\det G]^{1/2}. \qquad (3.86) \end{aligned}$$

The various factors of proportionality can be replaced by explicit factors in a given regularization. We shall not need $Z[0]$ in our introductory course. However, note that in the imaginary time formulation it contains all the physics of the quantum canonical partition function (cf. Problem 2).

A few remarks are here in order.

The above derivation was elegant but formal. The results must and have been checked in various ways. For instance, one can evaluate the path integral carefully in discretized form and study the continuum limit in various quantities. It is instructive to evaluate the path integral for generic boundary conditions at finite time. Then the results (also G) do depend on the boundary values (which can be used to calculate $\psi_0[q]$), and one can show that this dependence disappears in the limit of infinite extent in time, as derived in the previous section. See e.g. Brown sect. 1.5.

From (3.67) and (3.79) we now have an explicit formula for the 'generating functional of ground state expectation values of time ordered products of the coordinates':

$$\langle\psi_0|T\, e^{i\int dt\, f(t)\hat{q}(t)}|\psi_0\rangle = e^{i\frac{1}{2}\int dt\, dt'\, f(t)G(t,t')f(t')}. \tag{3.87}$$

Differentiating twice with respect to f and setting $f = 0$ afterwards gives

$$\langle\psi_0|T\hat{q}(t)\hat{q}(t')|\psi_0\rangle = -iG(t,t'). \tag{3.88}$$

There are some instructive checks. First, apply the differential operator G^{-1} to the left hand side using the identity

$$\frac{\partial}{\partial t}\,\theta(t-t') = \delta(t-t'), \tag{3.89}$$

the canonical commutation relations and the Heisenberg equations of motion: the result is $-i\delta(t-t')$, as it should according to (3.80) (cf. Problem 3). This shows that the left hand side of (3.88) is a Green function, but it does not check the boundary conditions. Second, a complete check can be made by evaluating the left hand side of (3.88) explicitly using creation and annihilation operators:

$$\begin{aligned} \hat{q}(t) &= e^{i\hat{H}t}\,\hat{q}\,e^{-i\hat{H}t} = \frac{1}{\sqrt{2m\omega}}\left(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t}\right), \\ [\hat{a},\hat{a}^\dagger] &= 1, \quad [\hat{a},\hat{a}] = 0, \quad \hat{a}|\psi_0\rangle = 0, \quad \text{etc.} \end{aligned} \tag{3.90}$$

The result is (3.88) with (3.85).

3.7 Scalar field and Ising model

We now generalize to the case of infinitely many 'coordinates' $\varphi(\mathbf{x})$ labeled by the continuous index $\mathbf{x}$, i.e. the scalar field. We write $|\varphi\rangle$ for the basis vectors in the 'coordinate representation' (2.7). Then

$$\hat{\varphi}(\mathbf{x})|\varphi\rangle = \varphi(\mathbf{x})|\varphi\rangle \tag{3.91}$$

and formally

$$\begin{aligned} |\varphi\rangle &= \prod_{\mathbf{x}} \otimes |\varphi(\mathbf{x})\rangle, \\ \langle\varphi|\varphi'\rangle &= \prod_{\mathbf{x}} \delta[\varphi(\mathbf{x}) - \varphi'(\mathbf{x})], \\ \int_{-\infty}^{\infty} \prod_{\mathbf{x}} d\varphi(\mathbf{x})\, |\varphi\rangle\langle\varphi| &= \hat{1}. \end{aligned} \tag{3.92}$$

The path integral representation for the evolution kernel has the form

$$\langle\varphi', t'|\varphi'', t''\rangle = \langle\varphi'|\hat{U}(t', t'')|\varphi''\rangle = \int d[\varphi]\, e^{iS[\varphi]}, \tag{3.93}$$

with formally

$$d[\varphi] = \prod_x d\varphi(x). \tag{3.94}$$

These formal equations become concrete upon discretization. Since we not only have continuous time but also space this naturally leads to a *lattice* in spacetime. We illustrate this here briefly for the case of imaginary time, $x^0 \to -ix_4$. Then time is equivalent to space, spacetime is replaced by four dimensional space and it is natural to use a simple hypercubic lattice for discretization, with lattice spacing b equal in all directions. The finite volume system at inverse temperature L_4 and spatial volume L^3 is then approximated by a system on an $L^3 \times L_4$ lattice. Furthermore it becomes natural to use *lattice units*, i.e. units in which the lattice spacing is $b = 1$. With these conventions the partition function

$$\begin{aligned} Z &= \int d[\varphi]\, e^{-S_I^{\text{discr}}[\varphi]}, \\ S_I^{\text{discr}} &= \sum_x \sum_{\mu=1}^{4} \frac{1}{2}(\varphi_{x+\hat{\mu}} - \varphi_x)^2 + \sum_x V(\varphi_x), \end{aligned}$$

is well defined.

Models defined in this way have been studied by various analytical and numerical (Monte Carlo) methods. This has led to a fairly complete (and *nonperturbative*) understanding of four dimensional scalar field theory. For more information on this *Lattice Field Theory* approach see Creutz, Münster and Montvay, Rothe, Smit, Le Bellac and Barton, and the proceedings of the yearly Lattice 'XX meetings.

There is a close connection with the spin models studied in statistical physics, in particular the Ising model. Consider the φ^4 model

$$V = \frac{1}{2}\kappa_0\varphi^2 + \frac{1}{4}\lambda_0\varphi^4. \tag{3.95}$$

Using the change of variables, generalizing to d euclidean dimensions,

$$\varphi = \sqrt{2\alpha}\,\phi, \quad \kappa_0 = \frac{1-2\gamma}{\alpha} - 2d, \quad \lambda_0 = \frac{\gamma}{\alpha^2}, \tag{3.96}$$

the partition function can be written in the form (dropping an overal constant factor)

$$Z = \int \prod_x d\mu(\phi_x)\, \exp(2\alpha \sum_{x\mu} \phi_x \phi_{x+\hat{\mu}}), \tag{3.97}$$

$$d\mu(\phi) = d\phi \exp[-\phi^2 - \gamma(\phi^2-1)^2]. \tag{3.98}$$

We see that α parametrizes the coupling between neighboring 'spins' ϕ. The spins have a single site probability distribution given by (3.98). For $\gamma \to \infty$ the spins are restricted to values ± 1:

$$\frac{\int d\mu(\phi) f(\phi)}{\int d\mu(\phi)} \to \frac{1}{2}[f(1)+f(-1)], \tag{3.99}$$

and the model goes over into the Ising model in d dimensions. It turns out that the Ising model is indeed a very good formulation of the relativistic φ^4 theory in four dimensions! (For more information see the books by Le Bellac and Barton, Münster and Montvay, Smit, and more generally on the field theoretic description of critical phenomena: Zinn-Justin, Parisi, Drouffe and Itzykson.)

3.8 Free scalar field

We continue with real time and set $\varphi' = \varphi'' = 0$, so that partial integrations do not involve boundary terms at t', t''. For space the same effect is obtained by choosing periodic boundary conditions. To obtain ground state expectation values we take the limit $t' \to \infty(1-i\delta)$ and $t'' \to -\infty(1-i\delta)$, as explained in sect. 3.5. Introducing the external source $J(x)$, the analogue of the external force $f(t)$, we can immediately write down the analogue of (3.66) and (3.67),

$$\langle 0,\infty|0,-\infty\rangle[J] \equiv Z[J] = \int d[\varphi]\, e^{iS[\varphi]+i\int d^4x\, J(x)\varphi(x)}, \tag{3.100}$$

$$\langle\psi_0|Te^{i\int d^4x\, J(x)\hat{\varphi}(x)}|\psi_0\rangle = \frac{Z[J]}{Z[0]} = \frac{\int d[\varphi]\, e^{iS[\varphi]+i\int d^4x\, J(x)\varphi(x)}}{\int d[\varphi]\, e^{iS[\varphi]}}. \tag{3.101}$$

For the free field these formulas can be worked out elegantly, assuming for simplicity infinite space. The action is

$$\begin{aligned} S &= -\frac{1}{2}\int d^4x\, (\partial_\mu\varphi\partial^\mu\varphi + m^2\varphi^2) = -\frac{1}{2}\int d^4x\, \varphi(-\partial^2+m^2)\varphi \\ &= -\frac{1}{2}\int d^4x\, d^4x'\, \varphi(x) G^{-1}(x,x')\varphi(x'), \\ &\equiv -\frac{1}{2}\varphi^T G^{-1}\varphi, \end{aligned} \tag{3.102}$$

in matrix notation, with

$$G^{-1}(x,x') = (-\partial^2+m^2)\delta^4(x-x'). \tag{3.103}$$

Then (3.79) generalizes to

$$Z[J] = Z[0] \exp\left[i\frac{1}{2}\int d^4x\, d^4x'\, J(x)G(x,x')J(x')\right].$$

The inverse of G^{-1} is given by

$$G(x,x') = \int \frac{d^4p}{(2\pi)^4}\, e^{ip(x-x')}\, G(p) \tag{3.104}$$

with

$$G(p) = \frac{1}{p^2+m^2} \to \frac{1}{-p_0^2+\mathbf{p}^2+m^2-i\epsilon}, \tag{3.105}$$

and

$$\begin{aligned} -iG(x,x') &= \int \frac{d^3p}{(2\pi)^3}\, \frac{e^{i\mathbf{p}(\mathbf{x}-\mathbf{x}')-iE(\mathbf{p})|x^0-x'^0|}}{2E(\mathbf{p})}, \quad E(\mathbf{p}) = \sqrt{m^2+\mathbf{p}^2}, \\ &= \int d\omega_p\, e^{ip(x-x')}, \quad x^0 > x'^0, \\ &= \int d\omega_p\, e^{-ip(x-x')}, \quad x^0 < x'^0. \end{aligned}$$

This is analogous to (3.82 – 3.85) for the harmonic oscillator, with $\omega_{\rm HO} \to E(\mathbf{p})$ and $m_{\rm HO} \to 1$. In the last line we replaced the dummy $\mathbf{p}$ by $-\mathbf{p}$. The Green function $G(x,x')$ is called the *Feynman propagator*.

Differentiating twice with respect to the external source and setting it to zero afterwards gives the basic vacuum expectation value of the time ordered product of two fields, the *two-point function*,

$$\langle\psi_0|T\hat\varphi(x_1)\hat\varphi(x_2)|\psi_0\rangle = -iG(x_1,x_2). \tag{3.106}$$

For the generalization to more fields in the time ordered product we just differentiate more times. It is convenient to introduce a suggestive notation for such *n-point functions*:

$$\begin{aligned} \langle\psi_0|T\hat\varphi(x_1)\cdots\hat\varphi(x_n)|\psi_0\rangle &= \frac{\int d[\varphi]\, e^{iS[\varphi]}\, \varphi(x_1)\cdots\varphi(x_n)}{\int d[\varphi]\, e^{iS[\varphi]}} \\ &\equiv \langle\varphi(x_1)\cdots\varphi(x_n)\rangle \end{aligned} \tag{3.107}$$

Note that the brackets in this notation do not mean an ordinary statistical average because $\exp(iS)$ is complex, and neither an ordinary quantum mechanical expectation value because of the time ordering involved. Using repeated differentiation with respect to the external source, setting it to zero afterwars, one finds (Prob. 3**d**)

$$\begin{aligned} \langle\varphi(x_1)\rangle &= 0, \\ \langle\varphi(x_1)\varphi(x_2)\rangle &= -iG(x_1,x_2), \\ \langle\varphi(x_1)\varphi(x_2)\varphi(x_3)\rangle &= 0, \\ \langle\varphi(x_1)\varphi(x_2)\varphi(x_3)\varphi(x_4)\rangle &= (-i)^2[G(x_1,x_2)G(x_3,x_4) + G(x_1,x_3)G(x_2,x_4) \\ &\quad + G(x_1,x_4)G(x_2,x_3)], \end{aligned} \tag{3.108}$$

etc. The general formula goes under the name 'Wick's theorem',

$$\begin{aligned}\langle\psi_0|T\varphi(x_1)\cdots\varphi(x_n)|\psi_0\rangle &\equiv \langle\varphi(x_1)\cdots\varphi(x_n)\rangle \\ &= \langle\varphi(x_1)\varphi(x_2)\rangle\cdots\langle\varphi(x_{n-1})\varphi(x_n)\rangle + \text{permutations},\end{aligned}$$

where every distinct permutation (partition into pairs (x_i, x_j)) occurs once. In practise we often abbreviate the notation further for clarity, e.g.

$$\langle\varphi_1\varphi_2\varphi_3\varphi_4\rangle = (-i)^2[G_{12}G_{34} + G_{13}G_{24} + G_{14}G_{23}]. \tag{3.109}$$

3.9 Summary

In the path integral formalism the action takes a prominent role by specifying the 'weight factor' $\exp(iS)$. In imaginary time this has the form of a Boltzmann factor $\exp(-S_I)$, and for a field theory in d spatial dimensions, the imaginary time path integral (with periodic boundary conditions in time) has the form of a *classical* partition function in $d+1$ dimensions, while it represents the *quantum* partition function in d dimensions.

For free fields the path integral can be explicitly evaluated, because it is 'just' a multiple gaussian integral. The external source technique is useful for dealing with correlation functions, which are vacuum expectation values of time ordered products of field operators. The two-point correlation function, also called Feynman propagator, is a Green function, similar to the retarded and advanced Green functions. In contrast to the latter two, the Feynman propagator $G(x, x')$ is complex, and it does *not* vanish outside the light cone (i.e. for $(x-x')^2 > 0$).

The path integral has great intuitive appeal and properly defined its formalism is equivalent to the canonical operator formulation (provided the latter is also properly defined!). This can be done with the lattice discretization, which is very useful for non-perturbative computations. However, a common practise in perturbation theory is to go ahead formally and ignore possible infinities temporarily, expecting to deal with them later 'on the fly'.

3.10 Problems

1. *Relation* $\hat{T}$ – $\hat{H}$

   Consider the Cambell-Baker-Haussdorff formula

   $$e^A\, e^B = e^{A+B+\frac{1}{2}[A,B]+\cdots}, \tag{3.110}$$

   where the $\cdots$ consist of multiple commutators of higher order in A and B ($[A,[A,B]]$, $[B,[A,B]]$, etc.). Let

   $$A = -ia\frac{1}{2}\,V(\hat{q}), \quad B = -ia\frac{\hat{p}^2}{2m}. \tag{3.111}$$

Verify that $\cdots = O(a^3)$, and

$$e^{-ia\frac{1}{2}V(\hat{q})}\, e^{-ia\frac{\hat{p}^2}{2m}}\, e^{-ia\frac{1}{2}V(\hat{q})} = e^{-ia\hat{H}+O(a^3)}, \quad \hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{q}). \tag{3.112}$$

2. *Partition function of the harmonic oscillator*

Consider the harmonic oscillator with unit mass given by the hamiltonian

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2. \tag{3.113}$$

We calculate the partition function

$$Z = \mathrm{Tr}\, e^{-\beta\hat{H}} \tag{3.114}$$

in several ways.

a. Using the energy representation to evaluate the trace and the well known energy spectrum verify that

$$\ln Z = -\frac{1}{2}\beta\omega - \ln(1 - e^{-\beta\omega}). \tag{3.115}$$

We can also calculate Z from the path integral by continuation to imaginary time, $t \to -i\tau$,

$$\begin{aligned} Z &= \int_{-\infty}^{\infty} dq_+\, \langle q_+|e^{-iHt_+}|q_+\rangle && (3.116)\\ &= \int dq_+ \int [dq] \exp\left\{ i\int_0^{t_+} dt\, \left[\frac{1}{2}\left(\frac{dq}{dt}\right)^2 - \frac{1}{2}\omega^2 q^2\right]\right\} && (3.117)\\ &\to \int dq_+\, \langle q_+|e^{-\beta H}|q_+\rangle && (3.118)\\ &= \int_{\mathrm{pbc}} [dq]\, \exp\left\{ -\int_0^{\beta} d\tau\, \left[\frac{1}{2}\left(\frac{dq}{d\tau}\right)^2 + \frac{1}{2}\omega^2 q^2\right]\right\}, && (3.119) \end{aligned}$$

where $\beta = \tau_+$. The path integral is now over functions $q(\tau)$ with $q(0) = q(\beta) = q_+$ and we also integrate over q_+; this is indicated by the subscript 'pbc', which means 'periodic boundary conditions'. Introducing

$$G^{-1}(\tau,\tau') = \left[-\frac{d^2}{d\tau^2} + \omega^2\right]\delta(\tau - \tau'), \tag{3.120}$$

be can write the partition function as

$$Z = [\det G^{-1}]^{-1/2}, \tag{3.121}$$

up to a constant.

To evaluate this determinant we introduce a suitable complete set of periodic functions in the interval $[0, \beta]$,

$$g_n(\tau) = \frac{1}{\sqrt{\beta}} \exp(i\omega_n \tau), \quad \omega_n = \frac{2\pi n}{\beta}, \quad n = 0, \pm 1, \pm 2, \cdots. \tag{3.122}$$

b. Obtain the matrix elements

$$G^{-1}_{mn} = \int_0^\beta d\tau d\tau' \, g^*_m(\tau) G^{-1}(\tau, \tau') g_n(\tau'). \tag{3.123}$$

c. The determinant $\det G^{-1}$ as given by the product over its eigenvalues diverges badly. We shall regulate the determinant by deviding it by a similar determinant with a different frequency ω'. Using the infinite product formula

$$\prod_{n=1}^{\infty} \left(1 + \frac{x^2}{n^2\pi^2}\right) = \frac{\sinh x}{x}, \tag{3.124}$$

verify that

$$\frac{Z(\omega)}{Z(\omega')} = \frac{\sinh(\beta\omega'/2)}{\sinh(\beta\omega/2)}. \tag{3.125}$$

If we now let ω' approach infinity, then in the trace representation (3.114) for $Z(\omega')$ the ground state contribution dominates: $Z(\omega') \to \exp(-\beta\omega'/2)$. Hence, verify that this leads again to the partition function (3.115) obtained earlier.

Using the discrete time slicing definition of the path integral the divergence of $\det G^{-1}$ is avoided and it leads of course to the right answer, but we shall not go into details. It is instructive anyway to play with divergent continuum expressions.

3. *Feynman propagator*

Eq. (3.85) is the Green function introduced by Feynman (usually called the propagator) for the case of the harmonic oscillator. We summarize:

$$Z[f] \equiv \int [dq] \exp\left[i \int_{-\infty}^{\infty} dt \, (\frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega^2 q^2 + fq)\right], \tag{3.126}$$

$$= Z[0] \exp\left[i\frac{1}{2} \int dt \, dt' \, f(t) G(t-t') f(t')\right], \tag{3.127}$$

$$\frac{Z[f]}{Z[0]} = \langle 0 | T \exp\left[i \int dt \, f(t) \hat{q}(t)\right] |0\rangle, \tag{3.128}$$

$$-iG(t-t') = \theta(t-t') \frac{e^{-i\omega(t-t')}}{2\omega} + \theta(t'-t) \frac{e^{+i\omega(t-t')}}{2\omega}. \tag{3.129}$$

We have set the mass $m = 1$. Furthermore $t \in (-\infty, \infty)$ with the understanding that we supply convergence for large times in a suitable way when necessary.

a. Using $d\theta(t)/dt = \delta(t)$ verify that

$$\left(\frac{d^2}{dt^2} + \omega^2\right) G(t - t') = \delta(t - t'). \tag{3.130}$$

Hint: after differentiating once with respect to t, simplify by using the identity $F(t)\delta(t) = 0$, which is valid for functions $F(t)$ which vanish at $t = 0$.

b. Calculate

$$\left(\frac{d^2}{dt^2} + \omega^2\right) [\theta(t - t')\langle 0|\hat{q}(t)\hat{q}(t')|0\rangle + \theta(t' - t)\langle 0|\hat{q}(t')\hat{q}(t)|0\rangle] \tag{3.131}$$

by using the equation of motion for $\hat{q}(t)$ and the canonical commutation relations $[\dot{\hat{q}}(t), \hat{q}(t)] = -i$. Compare with

$$\langle 0|T\hat{q}(t)\hat{q}(t')|0\rangle = -iG(t - t'). \tag{3.132}$$

c. By supplying convergence factors $\exp(\mp\epsilon t)$ for large times $t \to \pm\infty$, with infinitesimal $\epsilon > 0$, verify that the Fourier transform of G is given by

$$\int_{-\infty}^{\infty} dt\, e^{ikt}\, G(t) = \frac{1}{\omega^2 - k^2 - i\epsilon}. \tag{3.133}$$

The idea is to let $\epsilon \to +0$ in a later stage of calculations involving G, so ϵ and e.g. 5ϵ have the same meaning.

d. Express

$$\langle 0|T\hat{q}(t_1)\hat{q}(t_2)\hat{q}(t_3)\hat{q}(t_4)|0\rangle \tag{3.134}$$

in terms of products of Feynman propagators by functionally differentiating $Z[f]$ and setting $f = 0$ afterwards.

Chapter 4

Perturbation theory, diagrams and renormalization

In the previous chapter we have seen, in the example of the free scalar field, how the path integral can be used to calculate expectation values of time-ordered products of field operators in the ground (vacuum) state:

$$\langle\psi_0|T\hat{\varphi}(x_1)\cdots\hat{\varphi}(x_n)|\psi_0\rangle \equiv \langle\varphi(x_1)\cdots\varphi(x_n)\rangle.$$

From these so-called n-point functions we can obtain many quantities of physical interest. Important for us are scattering and decay amplitudes, and a simple recipe for obtaining these will be given in section 4.6, deferring its derivation to chapter 9.

We are now ready for the perturbative calculation of $\langle\varphi(x_1)\cdots\varphi(x_n)\rangle$ in interacting theories. Using the φ^4 theory as the simplest example, we introduce Feynman diagrams – which are an invaluable tool for representing the occurring mathematical expressions – and vertex functions, and evaluate the first few of these in order to clarify renormalization and counterterms.

4.1 Preparation

We recall that the generating functional for the n-point functions has the path-integral representation (3.101), i.e.

$$\langle\psi_0|Te^{i\int d^4x\, J(x)\hat{\varphi}(x)}|\psi_0\rangle = \frac{Z[J]}{Z[0]} = \frac{\int d[\varphi]\, e^{iS[\varphi]+i\int d^4x\, J(x)\varphi(x)}}{\int d[\varphi]\, e^{iS[\varphi]}} \qquad (4.1)$$

The action is written as the sum of a free part S_0 and an interacting part S_1,

$$S = S_0 + S_1, \qquad (4.2)$$

where S_0 is quadratic in the fields and S_1 is of higher order. Linear terms can be absorbed in the external source J, if needed, so we will not include such terms in

S. Note that a cosmological constant drops out in the ratio $Z[J]/Z[0]$, and we will often ignore such a term. In φ^4 theory

$$S_0 = -\int d^4x\, \frac{1}{2}\varphi(x)(-\partial^2 + m^2)\varphi(x), \tag{4.3}$$

$$S_1 = -\frac{\lambda}{4}\int d^4x\, \varphi(x)^4 + \Delta S. \tag{4.4}$$

Here m and λ are renormalized parameters and ΔS contains so-called counterterms: it is the difference between the terms shown explicitly and the original action formulated in terms of bare parameters κ_0, λ_0 and field φ_0:

$$S = -\int d^4x \left[\frac{1}{2}(\partial\varphi_0)^2 + \frac{1}{2}\kappa_0\varphi_0^2 + \frac{1}{4}\lambda_0\varphi_0^4 + \epsilon_0\right] \tag{4.5}$$

$$= -\int d^4x \left[\frac{1}{2}Z(\partial\varphi)^2 + \frac{1}{2}Z\kappa_0\varphi^2 + \frac{1}{4}Z^2\lambda_0\varphi^4 + \epsilon_0\right], \quad \varphi_0 = \sqrt{Z}\,\varphi, \tag{4.6}$$

$$= -\int d^4x \left[\frac{1}{2}(\partial\varphi)^2 + \frac{1}{2}m^2\varphi^2 + \frac{1}{4}\lambda\varphi^4\right] + \Delta S, \tag{4.7}$$

$$\Delta S = \int d^4x \left[\frac{1}{2}(1-Z)(\partial\varphi)^2 + \frac{1}{2}(m^2 - Z\kappa_0)\varphi^2 + \frac{1}{4}(\lambda - Z^2\lambda_0)\varphi^4 - \epsilon_0\right] \tag{4.8}$$

$$\equiv \int d^4x \left[\frac{1}{2}\delta Z(\partial\varphi)^2 + \frac{1}{2}\delta m^2\varphi^2 + \frac{1}{4}\delta\lambda\varphi^4 + \delta\epsilon\right]. \tag{4.9}$$

The so-called renormalized field φ is related to the original variable φ_0, by the multiplicative renormalization $\varphi_0 = \sqrt{Z}\,\varphi$. In fact, we will calculate in renormalized perturbation theory the n-point functions of φ, not of φ_0. The counterterms will be specified later on, after we have calculated we few vertex functions. As usual in minkowkian field theory, we have set renormalized vacuum energy to zero, $\epsilon = 0$.

The renormalized perturbation expansion is an expansion in powers of λ, not λ_0. The basic idea of counterterms is to choose the starting point of the perturbation expansion close to the final answer. For example, we may choose m to be the exact mass of the particles and λ the exact coupling constant (possible definitions of λ will be given later). So the effect of the interactions on m and λ is exactly cancelled by the counterterms. We may then hope that more general interaction effects will be small when λ is small. In particular, infinities, which which pop up because of our cavalier dealing with infinitely many degrees of freedom, are to be cancelled by the counterterms. It is hard to explain this properly at this point so we shall ignore the counterterms for the moment ($\Delta S \to 0$) and return to them later.

To further prepare for the perturbative expansion we write, in condensed notation,

$$\langle\varphi_1\cdots\varphi_n\rangle = \frac{\int d[\varphi]\, e^{iS_0+iS_1}\,\varphi_1\cdots\varphi_n}{\int d[\varphi]\, e^{iS_0+iS_1}} = \frac{\langle\varphi_1\cdots\varphi_n\, e^{iS_1}\rangle_0}{\langle e^{iS_1}\rangle_0}, \tag{4.10}$$

where the subscript 0 denotes the free field 'average'

$$\langle F[\varphi]\rangle_0 = \frac{\int d[\varphi]\, e^{iS_0}\, F[\varphi]}{\int d[\varphi]\, e^{iS_0}} \tag{4.11}$$

(we have devided numerator and denominator by $\int d[q]\, \exp(iS_0)$).

4.2 Diagrams for φ^4 theory.

Ignoring the counterterms, expanding in λ means expanding the exponential $\exp(iS_1)$. Consider first the denominator $\langle e^{iS_1}\rangle_0$ in (4.10):

$$\begin{aligned}
\langle e^{iS_1}\rangle_0 &= 1 - i\frac{\lambda}{4}\int d^4x\, \langle\varphi(x)^4\rangle_0 + O(\lambda^2) \\
&= 1 - i\frac{\lambda}{4}3\int d^4x\, (-i)^2 G(x,x)^2 + O(\lambda^2), &(4.12)\\
&= 1 + \frac{1}{8}(-i6\lambda)\int d^4x\, (-i)^2 G(x,x)^2 + O(\lambda^2), &(4.13)
\end{aligned}$$

where we used (3.108). The three terms in (3.108) give identical contributions. Consider next the numerator of the two point function ($n = 2$):

$$\begin{aligned}
\langle\varphi_1\varphi_2\, e^{iS_1}\rangle_0 &= \langle\varphi_1\varphi_2\rangle_0 - i\frac{\lambda}{4}\int d^4x\, \langle\varphi_1\varphi_2\,\varphi(x)^4\rangle_0 + \cdots \\
&= -iG(x_1,x_2) - i\frac{\lambda}{4}3\int d^4x\, (-i)^3 G(x_1,x_2)\, G(x,x)^2 &(4.14)\\
&\quad - i\frac{\lambda}{4}12\int d^4x\, (-i)^3 G(x_1,x)G(x_2,x)G(x,x) + \cdots &(4.15)
\end{aligned}$$

A pairing of fields is called a contraction. The factor 3 in (4.14) has the same origin as in (4.12), namely three possible contractions of four fields giving identical contributions, and indeed, we recognize the denominator as a factor in (4.14),

$$-iG(x_1,x_2)\left[1 + \frac{1}{8}(-i6\lambda)\int d^4x\, (-i)^2 G(x,x)^2 + O(\lambda^2)\right]. \tag{4.16}$$

The factor 12 in (4.15) comes from first contracting, say, φ_1 with one of the $\varphi(x)$, which gives 4 identical contributions, and then φ_2 with one of the remaining three $\varphi(x)$, which gives 3 identical contributions; then there remain two $\varphi(x)$'s which simply give $G(x,x)$.

The numerator of the four point function,

$$\langle\varphi_1\varphi_2\varphi_3\varphi_4 e^{iS_1}\rangle_0 = \langle\varphi_1\varphi_2\varphi_3\varphi_4\rangle_0 - i\frac{\lambda}{4}\int d^4x\, \langle\varphi_1\varphi_2\varphi_3\varphi_4\varphi(x)^4\rangle_0 + O(\lambda^2), \tag{4.17}$$

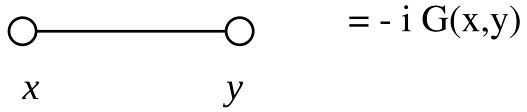

Figure 4.1: The propagator.

gives rise to more contributions. The first term is just (3.108). Applying Wick's theorem to the second, term there are contractions of only the $\varphi_1, \ldots, \varphi_4$ among themselves. This gives

$$\langle\varphi_1\varphi_2\varphi_3\varphi_4\rangle_0 \left[1 + \frac{1}{8}(-i6\lambda)\int d^4x\, G(x,x)^2 + O(\lambda^2)\right], \tag{4.18}$$

which has again the zero point contribution (4.13) found in the evaluation of the denominator. Then there are contributions in which only two of the $\varphi_1, \ldots, \varphi_4$ contract with $\varphi(x)^4$,

$$(-i)G(x_1,x_2)\left(-i\frac{\lambda}{4}\right) 12 \int d^4x\, (-i)^3 G(x_3,x)G(x,x)G(x,x_4) + 5 \text{ permutations.} \tag{4.19}$$

The term shown, with $\{x_1,x_2,x_3,x_4\} \to \{x_3,x_4,x_1,x_2\}$, contains (4.15) as a factor. There is one new contribution in which all of the $\varphi_1, \ldots, \varphi_4$ contract with $\varphi(x)^4$:

$$\langle\varphi_1\varphi_2\varphi_3\varphi_4\, iS_1\rangle_0^{\text{conn}} = -i\frac{\lambda}{4}\, 24 \int d^4x\, (-i)^4 G(x_1,x)G(x_2,x)G(x_3,x)G(x_4,x), \tag{4.20}$$

where the meaning of the label 'conn' (=connected) will get clear in the following.

At this point the expressions call out loudly for a representation in terms of diagrams. The propagator $G(x,y)$ times the factor $(-i)$ is represented by a line joining points labeled x and y. See Fig. 4.1. The little circles at the end of the line indicate that the line represents a propagator and not a two-point vertex function (to be defined below). Where arguments coincide, such as in $G(x,x)$, lines come together in the same point we have a vertex. This is most clear in the expression (4.20), see Fig. 4.2. The vertex represents the factor $-24i\lambda/4 = -i6\lambda$, as well as an integration over $\int d^4x$. The vertex is represented in Fig. 4.3, where the four lines indicate that it is a four point vertex, i.e. four propagators can be attached. Note that the lines of the vertex do not have the little circles at the end, whereas the circles in Fig. 4.2 indicate the presence of the propagators.[1] The diagrams for the denominator, the numerator of the two-point function and the numerator of the four-point functions are shown in Figs. 4.4, 4.5, 4.6, respectively.

[1]This convention is taken over from the books by Bjorken and Drell.

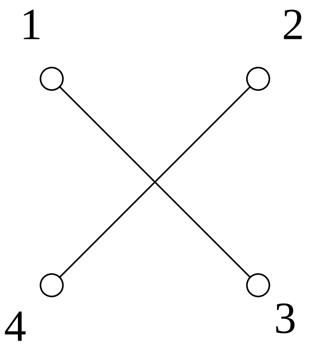

Figure 4.2: Lowest order diagram for the connected four point function $\langle\varphi_1 \cdots \varphi_4\rangle^{\text{conn}}$.

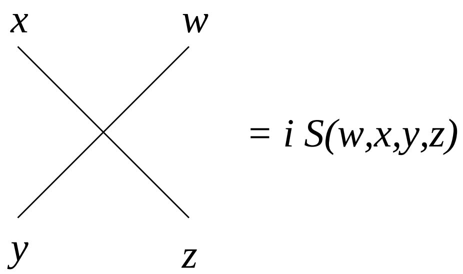

Figure 4.3: The four point vertex.

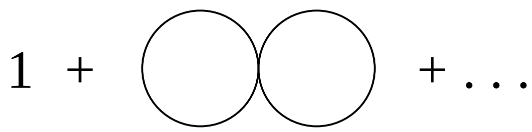

Figure 4.4: Expansion of $\langle e^{iS_1}\rangle_0$ to order λ.

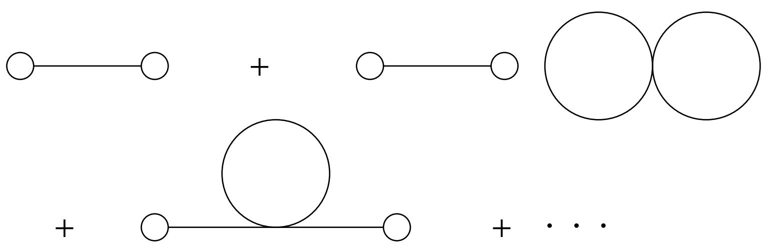

Figure 4.5: Expansion of $\langle\varphi_1\varphi_2\, e^{iS_1}\rangle_0$ to order λ.

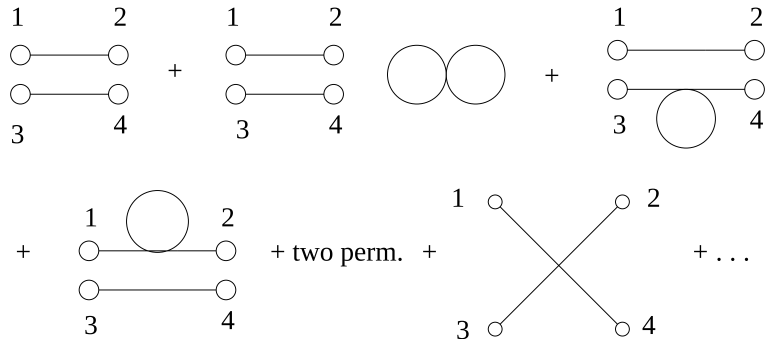

Figure 4.6: Expansion of $\langle\varphi_1 \cdots \varphi_4\, e^{iS_1}\rangle_0$ to order λ.

More precisely, we associate with the four-point vertex $i\times$ a *vertex function* $S(x_1, x_2, x_3, x_4)$, defined by the 4-th functional derivative of the action:

$$S(x_1,\ldots,x_4) \equiv \frac{\delta^4 S}{\delta\varphi(x_1)\cdots\delta\varphi(x_4)} = -6\lambda\int d^4x\, \delta^4(x_1-x)\cdots\delta^4(x_4-x) \tag{4.21}$$
$$= -6\lambda\delta^4(x_1-x_2)\delta^4(x_1-x_3)\delta^4(x_1-x_4).$$

The first form maintains the symmetry under permutations of the x's. There is also a two-point vertex function (cf. (1.148)),

$$S(x,y) = \left[\frac{\delta^2 S}{\delta\varphi(x)\delta\varphi(y)}\right]_{\varphi=0} = (\partial^2 - m^2)\delta^4(x-y). \tag{4.22}$$

A propagator-line emerging from a vertex now simply means integrating over the common argument of the vertex function and the propagator. For instance, the contribution (4.15) to the two-point function can be written as

$$\frac{1}{2}\int d^4y_1 \cdots d^4y_4\,(-i)G(x_1,y_1)\, iS(y_1,y_2,y_3,y_4)(-i)G(y_2,x_2)(-i)G(y_3,y_4). \tag{4.23}$$

The diagrams greatly clarify the formulas. There is a one-to-one correspondence between diagrams and the mathematical expressions, provided we keep in mind the accompanying combinatorical factors such as 1/8 in the zero-point function, and 1/2 in the two-point function. Such factors are related to the symmetry of a diagram and there are rules for determining them (see e.g. Peskin and Schroeder sect. 4.4), but it is often easy enough to get them by explicit counting of identical contributions, as we have done above.

One furthermore observes:

- A diagram may be disconnected, i.e. consist of several *connected* parts that are not connected by any line. The connected parts represent *factors* in the mathematical expression. We usually concentrate on the connected parts.

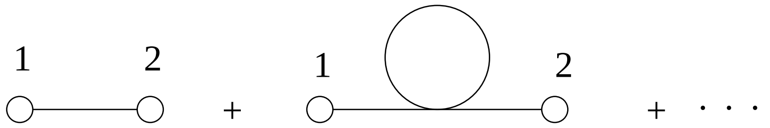

Figure 4.7: Expansion of $\langle\varphi_1\varphi_2\rangle$ to order λ.

- The diagrams for the denominator in (4.10) (often called vacuum diagrams), which do not have any external lines, cancel against similar disconnected parts in the numerator. This is true to all orders. Hence, for the evaluation of n-point functions with $n > 0$, we can forget about disconnected contributions containing diagrams without external lines. For example, the diagrams for the two-point function are given in Fig. 4.7.

- The diagrams can be decomposed into one particle irreducible (1PI) components and one particle reducible (1PR) components. A 1PR diagram becomes disconnected when cutting a single line, whereas a 1PI diagram remains connected. The 1PI components are called dressed vertex functions, and denoted by $\Gamma(x_1, x_2, \cdots)$. To lowest order they are equal to the bare vertex functions. There is a special notation for the two-point vertex function:
$$\Gamma(x_1, x_2) = S(x_1, x_2) - \Sigma(x_1, x_2). \tag{4.24}$$
and Σ is called the 'selfenergy'.

  The perturbative expansion of the 2- and 4-point dressed vertex functions is given in Fig. 4.8 and 4.9.

  The typical 1PR component is the exact two-point function, which is called the full, or fully dressed, propagator. It's expansion in terms of 1PI parts is given in Fig. 4.10. Denoting the full propagator by G', Fig. 4.10 represents the expansion
$$\begin{aligned}-iG'(x_1, x_2) &= -iG(x_1, x_2) + \int d^4y_1\, d^4y_2\, (-i)G(x_1, y_1)(-i)\Sigma(y_1, y_2)\\ &\quad (-i)G(y_2, x_2) + \int d^4y_1 \cdots d^4y_4\, (-i)G(x_1, y_1)\\ &\quad (-i)\Sigma(y_1, y_2)(-i)G(y_2, y_3)(-i)\Sigma(y_3, y_4)(-i)G(y_4, x_2)\\ &\quad + \cdots,\end{aligned} \tag{4.25}$$
or in matrix notation
$$\begin{aligned}G' &= G - G\Sigma G + G\Sigma G\Sigma G + \cdots = G\left[1 + \sum_{n=1}^{\infty}(-\Sigma G)^n\right]\\ &= G(1 + \Sigma G)^{-1} = (G^{-1} + \Sigma)^{-1}.\end{aligned} \tag{4.26}$$
In the last line we used the identity $A^{-1}B^{-1} = (BA)^{-1}$. We assumed that the geometric series converged, which may be true for (some, not all) matrix

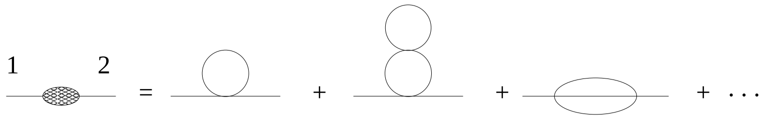

Figure 4.8: Expansion of the selfenergy $-i\Sigma(x_1, x_2)$ to $O(\lambda^2)$.

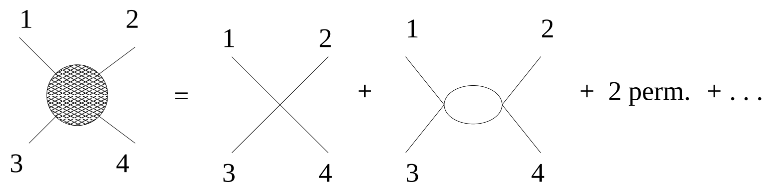

Figure 4.9: Expansion of the four point vertex function $i\Gamma(x_1, \cdots, x_4)$ to lowest non-trivial order.

elements, e.g. obtained by Fourier transformation, as will be done in the next section.

- The four-point function, dressed to all orders in λ, can now be summarized in terms of the full propagators and 4-point vertex function as in Fig. 4.11.

- More generally, an arbitrary diagram may be expressed in terms of full propagators and vertex functions. Expanding the full propagators and vertex functions in powers of λ, and then re-expanding the total, gives the correct expansion of the diagram.

- The perturbative expansion of vertex functions is a 'loop expansion': increasing the order in λ corresponds to increasing the number of loops in the diagrams. This is true to all orders.

  Keeping track of $\hbar$ it can be shown that with each additional loop the accompanying power of $\hbar$ increases. For this reason the loop expansion is also called a semiclassical expansion. The lowest order vertex functions are classical in the sense that the bare vertex functions are equal to the functional derivatives of the classical action.

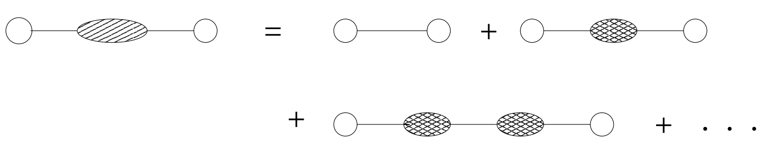

Figure 4.10: Structure of the two-point function $\langle\varphi_1\varphi_2\rangle$ in terms of 1PI two-point vertex functions.

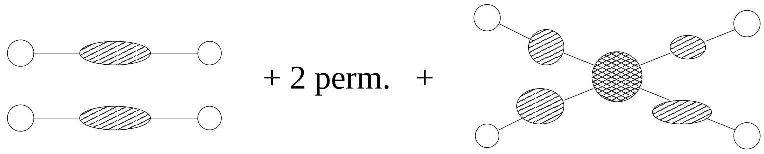

Figure 4.11: Structure of the four point function $\langle\varphi_1\cdots\varphi_4\rangle$.

4.3 Diagrams in momentum space

Because of translation invariance the n-point functions $\langle\varphi(x_1)\cdots\varphi(x_n)\rangle$ depend only on differences of coordinates. We can then simplify expressions by taking their Fourier transform, or in the jargon: going into 'momentum space'. The Fourier transform is defined as in the following example of the full n-point vertex function,[2]

$$\int d^4x_1\cdots d^4x_n\, e^{-i(p_1x_1+\cdots+p_nx_n)}\,\Gamma(x_1,\cdots,x_n) = (2\pi)^4\delta(p_1+\cdots+p_n)\,\Gamma(p_1,\cdots,p_n). \tag{4.27}$$

We have anticipated the presence of a momentum-conserving delta function.[3] For example, for the bare vertex function we have, using (4.21),

$$\begin{aligned}\int d^4x_1\cdots d^4x_4\, e^{-i(p_1x_1+\cdots+p_4x_4)}\,S(x_1,\cdots,x_4) &= -6\lambda\int d^4x\, e^{-i(p_1+p_2+p_3+p_4)x}\\ S(p_1,p_2,p_3,p_4) &= -6\lambda.\end{aligned} \tag{4.28}$$

So the bare vertex in momentum space is just the constant -6λ. There is a special notation for two point functions since these depend only on one independent momentum. For the bare two-point vertex function we have

$$S(p,-p)\equiv -G(p)^{-1} = -(m^2+p^2-i\epsilon), \tag{4.29}$$

where we have added the infinitesimal regulator term $i\epsilon$ (cf. (3.34)), while the propagator is given by

$$G(p,-p)\equiv G(p) = \frac{1}{m^2+p^2-i\epsilon}. \tag{4.30}$$

For the full two-point vertex function we have according to (4.24),

$$\Gamma(p,-p) = -[m^2+p^2+\Sigma(p)] \tag{4.31}$$

(including the $-i\epsilon$ in $\Sigma(p)$). See Figs. 4.12 – 4.14 for the corresponding diagrams.

[2]For simplicity the same symbol (here Γ) is used for the Fourier transform.

[3]In line with ubiquitous jargon, we call the Fourier variables: 'momenta'. The reason is that in applications to scattering these variables become the four-momenta of ('real' or 'virtual') particles.

An easy way to obtain the mathematical expression for the diagrams in momentum space is to represent everywhere the propagator by its Fourier transform

$$G(x,y) = \int \frac{d^4p}{(2\pi)^4}\, e^{ip(x-y)}\, G(p), \tag{4.32}$$

carry out the integrations over the coordinates of the vertex functions, and take the Fourier transform with respect to the external line variables. This leaves the momentum integrations associated with the propagators as in (4.32), with a four-momentum conservation delta function at every vertex. The result of integrating out as much as possible the propagator momenta using the delta functions, is one $(2\pi)^4\delta^4(\cdots)$ expressing overall conservation of the momenta at the external lines (which is to be extracted according to (4.27)), plus a number of remaining integrations associated with closed loops in the diagrams.

So the rules for obtaining the formula from a diagram in momentum space are:

- lines carry momenta p and a propagator $-iG(p)$;
- there is momentum conservation at each vertex, with a factor $iS(p_1,\cdots,p_4) = -i6\lambda$;
- integrate over the free momenta.

The selfenergy diagram in Fig. 4.14 is equal to

$$-i\Sigma(p) = \frac{1}{2}\,(-i6\lambda) \int \frac{d^4q}{(2\pi)^4}\, \frac{-i}{m^2+q^2-i\epsilon}. \tag{4.33}$$

The full propagator structure in Fig. 4.10 is just a geometric series in momentum space:

$$\begin{aligned} -iG'(p) &\equiv -iG(p) + (-i)G(p)(-i)\Sigma(p)(-i)G(p) \\ &\quad + (-i)G(p)(-i)\Sigma(p)(-i)G(p)(-i)\Sigma(p)(-i)G(p) + \cdots \\ &= -iG(p)\,\frac{1}{1+\Sigma(p)G(p)} = \frac{-i}{m^2+p^2+\Sigma(p)}. \end{aligned} \tag{4.34}$$

Comparing with (4.31) we see that the full propagator $G'(p)$ is the inverse of the two-point vertex function, as for free fields.

Fig. 4.15 shows the four point function to order λ^2. The corresponding formula is

$$\begin{aligned} i\Gamma(p_1,p_2,p_3,p_4) &= -i6\lambda \\ &\quad + \frac{1}{2}\,(-i6\lambda)^2 \int \frac{d^4q}{(2\pi)^4}\, \frac{-i}{m^2+q^2-i\epsilon}\, \frac{-i}{m^2+(q+p_1+p_2)^2-i\epsilon} \\ &\quad + 2 \text{ permutations} \end{aligned} \tag{4.35}$$

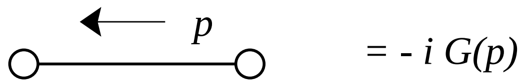

Figure 4.12: Diagram representing $-iG(p)$.

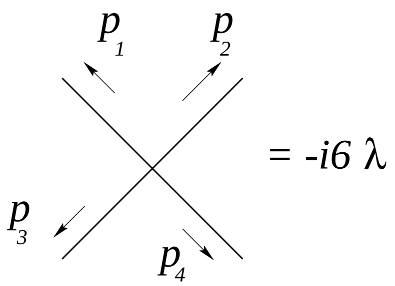

Figure 4.13: Vertex in momentum space.

Finally Fig. 4.16 shows a general $n > 4$ point vertex function in the one-loop approximation. Note that these are zero classically: there are no bare vertex functions with $n > 4$ in the φ^4 model (they would be represented by a tree-like vertex with more than four external lines).

4.4 One-loop vertex functions

We shall now evaluate the one loop diagrams for the two- and four-point vertex functions. The integrals in the expression for the two-point function (selfenergy) (4.33), and four-point function (4.35), diverge, so we first need to regularize them. This can be done in a variety of ways, leading to a variety of answers. This arbitrariness is to be absorbed in the bare parameters, similar to the case of the cosmological constant in sect. 2.3, such that the final physical answer is unambiguous. To be able to discuss this we first need to evaluate the diagrams. Here we shall use a simple method, which consists of making a 'Wick rotation', such that the integrals take a euclidean four-dimensional form, and then imposing

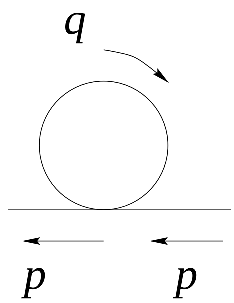

Figure 4.14: Order λ contribution to the selfenergy $-i\Sigma(p)$.

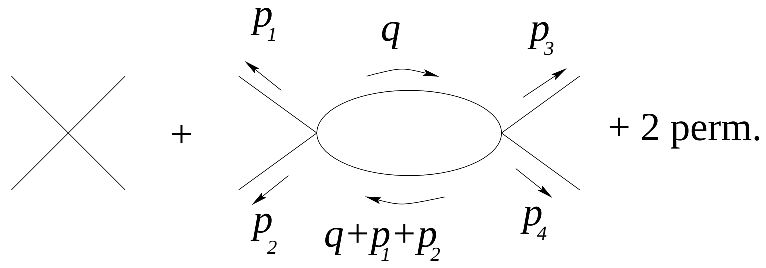

Figure 4.15: The 4-point vertex function $i\Gamma(p_1, p_2, p_3, p_4)$ to order λ^2.

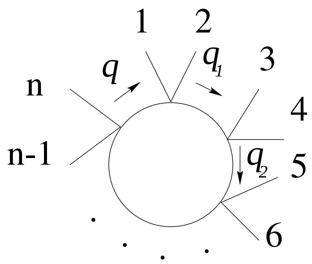

Figure 4.16: Contribution to the n-point vertex function $\Gamma(p_1, \cdots, p_n)$ in one loop order, $q_1 = q - p_1 - p_2$, $q_2 = q - p_1 - p_2 - p_3 - p_4$, etc.

a spherical cutoff $q^2 < \Lambda^2$.

Consider the integral in the self energy (4.33),

$$I = -i \int_{-\infty}^{\infty} dq^0 \, \frac{1}{m^2 + \mathbf{q}^2 - (q^0)^2 - i\epsilon} = -i \int_{-\infty}^{\infty} dq_0 \, \frac{1}{m^2 + \mathbf{q}^2 - q_0^2 - i\epsilon}, \qquad (4.36)$$

where we changed variables to $q_0 = -q^0$. The position of the poles of the integrand are illustrated in Fig. 4.17. Rotating the integration contour counter-clock wise over an angle $\pi/2$ (the 'Wick rotation') it is easy to verify that the above integral is unchanged. 'During the rotation' we do not cross any singularity. (Adding arcs at infinity, which do not contribute, the rotated contour is just a deformation of the original one along the real axis, see the figure on the right hand side.) Hence, setting $q_0 = iq_4$, with real q_4, we get

$$I = \int_{-\infty}^{\infty} dq_4 \, \frac{1}{m^2 + \mathbf{q}^2 + q_4^2}, \qquad q_4 = -iq_0, \qquad (4.37)$$

where we have set ϵ to zero because it is not needed anymore as the denominator does not vanish. It follows that the selfenergy can be written in euclidean form (i.e. a form where the metric is euclidean and q_4 enters equivalently to $q_1, \ldots q_3$):

$$\Sigma = 3\lambda \int \frac{d^4q}{(2\pi)^4} \, \frac{1}{m^2 + q^2}, \qquad (4.38)$$

where now $q^2 = q_1^2 + \cdots + q_4^2$, $d^4q = dq_1 \cdots dq_4$. We regularize this integral by introduce a cutoff $\sqrt{q^2} < \Lambda$. Introducing four dimensional spherical coordinates,

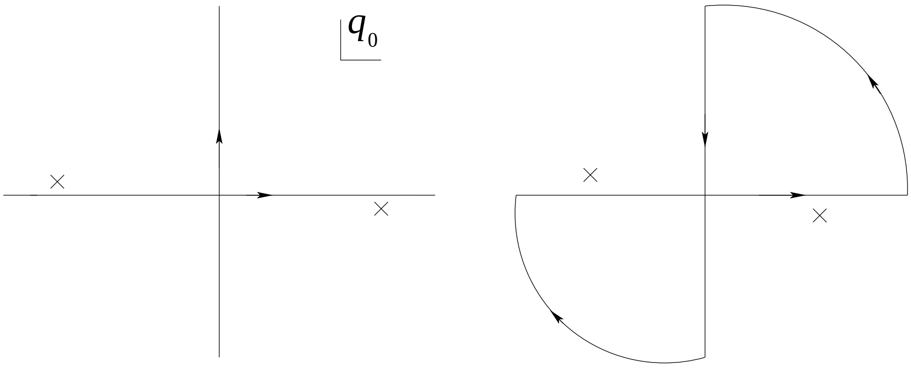

Figure 4.17: Left: Integration contours in the complex q_0 plane for the integral (4.36). The crosses indicate the position of the poles at $\pm(\sqrt{m^2+\mathbf{q}^2}-i\epsilon)$. The original contour along the real axis is rotated to the contour along the imaginary axis. Right: Integration contour giving a zero result, which can be used to prove that the two contours on the left give the same answer.

the integral over angles is[4] $2\pi^2$ and we are left with

$$\begin{aligned}\Sigma &= \frac{3\lambda}{(2\pi)^4}\, 2\pi^2 \int_0^\Lambda dq\, q^3\, \frac{1}{m^2+q^2} \qquad (4.39)\\ &= \frac{3\lambda}{16\pi^2}\left[\Lambda^2 - m^2 \ln\left(1+\frac{\Lambda^2}{m^2}\right)\right] = \frac{3\lambda}{16\pi^2}\left[\Lambda^2 - m^2\ln\frac{\Lambda^2}{m^2} + O(\Lambda^{-2})\right].\end{aligned}$$

The above evaluation was particularly easy because the one loop $\Sigma(p)$ is independent of p in φ^4 theory. The four-point vertex function is more interesting. The typical integral here is

$$J(p) = -i\int \frac{d^4q}{(2\pi)^4}\,\frac{1}{(m^2+q^2-i\epsilon)(m^2+(q+p)^2-i\epsilon)}, \qquad (4.40)$$

in terms of which

$$\Gamma(p_1,p_2,p_3,p_4) = -6\lambda + \frac{1}{2}(6\lambda)^2\left[J(p_1+p_2)+J(p_1+p_3)+J(p_1+p_4)\right], \quad (4.41)$$

to order λ^2. First we combine the denominators by using the identity (cf. Appendix)

$$\frac{1}{AB} = \int_0^1 dx\, \frac{1}{[xA+(1-x)B]^2}. \qquad (4.42)$$

Applying this identity for $A = m^2+q^2-i\epsilon$, $B = m^2+(q+p)^2-i\epsilon$ gives

$$J(p) \equiv -i\int\frac{d^4q}{(2\pi)^4}\int_0^1 dx\,\frac{1}{[x(m^2+q^2-i\epsilon)+(1-x)(m^2+(q+p)^2-i\epsilon)]^2}. \qquad (4.43)$$

[4]The ‘surface’ of the unit ball in n dimensions is $2\pi^{n/2}/\Gamma(n/2)$.

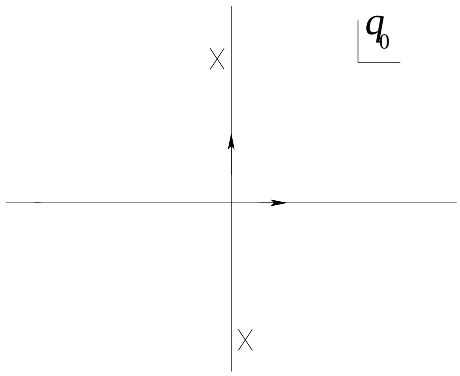

Figure 4.18: Integration contours in the complex q_0 plane for $J(p)$ for the case $p^2 < -4m^2$. The crosses indicate again the position of the poles and the original contour along the real axis can also in this case be rotated into the contour along the imaginary axis.

We now complete the square in the denominator by shifting the integrations over q,

$$q \to q - (1-x)p, \tag{4.44}$$

while ignoring for the moment the cutoff we are imposing on the integrations (we come back to this later). Then,

$$J(p) = -i \int \frac{d^4q}{(2\pi)^4} \int_0^1 dx \, \frac{1}{[m^2 + x(1-x)p^2 + q^2 - i\epsilon]^2}. \tag{4.45}$$

Doing the integration over q^0 first, we see similar pole positions as in the case of the selfenergy (here these are double poles), provided that $m^2 + x(1-x)p^2 > 0$, which is true for all x if $p^2 > -4m^2$. In case $p^2 < -4m^2$ the poles are situated as in Fig. 4.18 and the Wick rotation can still be done. Performing the Wick rotation and the q integrations with the euclidean cutoff $q^2 < \Lambda^2$ leads to

$$J(p) = \frac{1}{16\pi^2} \int_0^1 dx \left[\ln \frac{\Lambda^2}{m^2 + x(1-x)p^2 - i\epsilon} - 1 + O(\Lambda^{-2}) \right] \tag{4.46}$$

The remaining integration over the *Feynman parameter* x is elementary,

$$J(p) = \frac{1}{16\pi^2} \left(1 - \ln \frac{\sqrt{4m^2 + p^2 - i\epsilon} + \sqrt{p^2 - i\epsilon}}{\sqrt{4m^2 + p^2 - i\epsilon} - \sqrt{p^2 - i\epsilon}} + \ln \frac{\Lambda^2}{m^2} + O(\Lambda^{-2}) \right), \tag{4.47}$$

but this explicit answer does improve the clarity of the analytic structure of $J(p)$. The integral representation (4.46) is sufficient for our purpose.

We end this section with more comments:

- The maximum power of Λ in the expression for the vertex functions is the same as their dimension in mass units, i.e. 2 for the two-point function and

0 ($\leftrightarrow \ln\Lambda$ + const.) for the four-point function. The one-loop six- and higher-point functions (cf. Fig. 4.16) are finite.

- The regularization procedure used above has some arbitrariness, because the answer may depend on the 'moment' we introduce the cutoff Λ. It was introduced after the Wick rotation because then space and time are completely equivalent and we get Lorentz invariant answers. But even then Λ can be introduced either before or after shifting integration variables as in (4.44), whereas we do not shift the integration region accordingly. In addition there is more than one way to assign momenta to the propagators in the loop integrals (such that momentum conservation at the vertices is satisfied), which are related by transformations of the variables of the loop integrals. It can be shown[5] (see also Prob. 1) that in the limit $\Lambda \to \infty$ the arbitrariness is confined to a *polynomial in the external momenta, of degree equal to the mass dimension of the vertex function.* For dimension zero (i.e. logarithmically divergent integrals), shifts of the integration variable (as in (4.44)) do not to influence the momentum dependence of the regularized answer as $\Lambda \to \infty$, because each power of external momentum would imply a factor Λ^{-1}.

- It is evidently very advantageous to employ a regularization in which one can freely make shifts and other transformations of variables without having to worry about boundary effects. Important examples are dimensional regularization, Pauli-Villars regularization and the lattice. Dimensional regularization is very convenient and described in many books, e.g. De Wit & Smith, Veltman, Peskin & Schroeder.

- It is instructive to compare results with the lattice, where one finds (see e.g. Smit)

$$\Sigma = 3\lambda\left[C_0 a^{-2} - C_2 m^2 + \frac{1}{16\pi^2} m^2 \ln(a^2 m^2) + O(a^2)\right], \quad (4.48)$$

$$J(p) = C_2 - \frac{1}{16\pi^2}\int_0^1 dx\, \left[\ln\{a^2[m^2 + x(1-x)p^2 - i\epsilon]\} + 1\right] + O(a^2), \quad (4.49)$$

$$C_0 = 0.154933..., \quad C_2 = 0.0303457..., \quad (4.50)$$

where a is the lattice spacing. Note the correspondence $\Lambda \leftrightarrow a^{-1}$. Apart from a, all the lattice details reside in the constants C_0 and C_2 (neglecting terms vanishing in the continuum limit).

[5]The possible differences in the regularized expressions come from the integration region near the cutoff. In this region we can expand the propagator denominators in powers of the external momenta, with accompanying inverse powers of Λ.

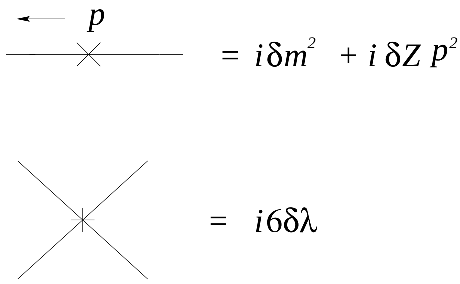

Figure 4.19: Counterterms in momentum space.

The arbitrariness in regularization method reflects itself in polynomials in p, m and Λ or a^{-1}, of degree equal to the dimension of the diagram, except for degree-zero terms where we have in addition logarithms (e.g. $\ln\Lambda$ or $\ln a^{-1}$) accompanying Λ^0 or a^0.

4.5 Renormalization and counterterms

In this section we will incorporate the effect of the counterterms and make a convenient choice for the constants δZ, δm^2 and $\delta\lambda$ in (4.8,4.9). The counterterms lead to additional vertices indicated by a cross in Fig. 4.19. To lowest order they contribute only to the two and four point vertex functions. For the selfenergy this means adding $-\delta m^2 - \delta Z\, p^2$ to $\Sigma(p)$,

$$\Sigma(p) = \frac{3\lambda}{16\pi^2}\left(\Lambda^2 - m^2 \ln\frac{\Lambda^2}{m^2}\right) - \delta m^2 - \delta Z\, p^2 + O(\lambda^2), \tag{4.51}$$

or for the complete two point function:

$$-\Gamma(p,-p) = m^2 - \delta m^2 + (1-\delta Z)p^2 + \frac{3\lambda}{16\pi^2}\left(\Lambda^2 - m^2 \ln\frac{\Lambda^2}{m^2}\right) + O(\lambda^2). \tag{4.52}$$

The counterterms can now be chosen such that $\Gamma(p,-p)$ is finite as $\Lambda\to\infty$. This determines the Λ dependence of δm^2 and we see that δZ is not needed to cancel infinities at this stage, since there is no divergent p^2 term (it is, however, needed at two-loop order). The possibility of finite parts in the counterterms induces a polynomial arbitrariness in $\Gamma(p,-p)$, of the form $\alpha + \beta p^2$. We now choose the counterterms such that $\Gamma(p,-p)$ satisfies so-called renormalization conditions. We need two conditions to fix the polynomial. A convenient choice is determined by an expansion about $p^2 = -m^2$:

$$\Gamma(p,-p) = -(m^2+p^2) + O((m^2+p^2)^2), \tag{4.53}$$

in which the constant term is chosen zero and the coefficient of $m^2 + p^2$ is chosen one. It follows that

$$\begin{aligned} \delta m^2 &= \frac{6\lambda}{32\pi^2}\left(\Lambda^2 - m^2 \ln\frac{\Lambda^2}{m^2}\right) + O(\lambda^2), && (4.54) \\ \delta Z &= O(\lambda^2). && (4.55) \end{aligned}$$

Similarly, the four point vertex function gets an additional contribution $6\delta\lambda$,

$$\begin{aligned} \Gamma(p_1,p_2,p_3,p_4) &= -6(\lambda - \delta\lambda) + \frac{1}{2}(6\lambda)^2\left[J(p_1+p_2) + J(p_1+p_3) + J(p_1+p_4)\right] \\ &\quad + O(\lambda^3), && (4.56) \end{aligned}$$

with the function $J(p)$ given by (4.46). Requiring this to be finite determines the Λ dependence of $\delta\lambda$ and leaves an ambiguity of the form of a constant, which is eliminated by imposing a renormalization condition on the four-point vertex function. A convenient condition is

$$\Gamma(0,0,0,0) = -6\lambda, \quad (4.57)$$

for which

$$6\delta\lambda = -\frac{3}{2}\frac{(6\lambda)^2}{16\pi^2}\left(\ln\frac{\Lambda^2}{m^2} - 1\right) + O(\lambda^3). \quad (4.58)$$

We now have the following results for the vertex functions:

$$\begin{aligned} \Gamma(p,-p) &= -m^2 - p^2 + O(\lambda^2), && (4.59) \\ \Gamma(p_1,p_2,p_3,p_4) &= -6\lambda - \frac{(6\lambda)^2}{32\pi^2}\int_0^1 dx\, \ln\frac{m^2 + x(1-x)(p_1+p_2)^2 - i\epsilon}{m^2} \\ &\quad + 2 \text{ permutations}. && (4.60) \end{aligned}$$

Keeping in mind that the $n > 4$ point vertex functions are finite, the approximation at one-loop order is now determined and unambigous. All regularization dependence has disappeared (it is easy to check for example that the lattice method leads to the same final result under the same renormalization conditions).

We end this section with some more remarks:

- As can be seen from (4.60), for $\lambda < 1$ the perturbative $O(\lambda^2)$ correction to the leading order four-point vertex function (i.e. -6λ) is small for momenta of order m.

- We can now understand better the meaning of the renormalized parameter λ: it is simply the value of the four point vertex function at zero momentum. We could make another choice for λ, say λ_1 is the value of $\Gamma(p_1,p_2,p_3,p_4)$ at some other standard set of momenta $p_1, \ldots, p_4$. Then $\lambda_1 = \lambda + O(\lambda^2)$. So other renormalization conventions imply a re-ordering of the perturbation series. This freedom in choice of expansion parameter is more fully exploited in the method of the Renormalization Group.

- With the above choice of renormalization condition for $\Gamma(p,-p)$, the parameter m is the mass of the particles. This follows from the fact that the pole in the full propagator $G'(p) = [m^2 + p^2 + \Sigma(p)]^{-1}$ as a function of p^0 is at $p^0 = \pm(\sqrt{m^2 + \mathbf{p}^2} - i\epsilon)$. The relation,

$$\text{position of pole in } G'(p) \leftrightarrow \text{particle mass}, \tag{4.61}$$

will be explained more fully in chapter 9. Other renormalization conditions are possible in which m is not the particle mass, but just a parameter with dimension of mass, related to the particle mass by a relation of the form: $m_{\text{part}} = m + O(\lambda)$. To see this, consider the equation

$$m^2 + p^2 + \Sigma(p) = 0. \tag{4.62}$$

Let $p^2 = -m_{\text{part}}^2$ be the zero of this equation (using p^2 as variable instead of p^0, which is simpler). Expanding the selfenergy about the zero $p^2 = -m_{\text{part}}^2$,

$$\Sigma(p) = \sigma_0 + \sigma_1(m_{\text{part}}^2 + p^2) + O((m_{\text{part}}^2 + p^2)^2), \tag{4.63}$$

gives

$$m_{\text{part}}^2 = m^2 + \sigma_0 = m^2 + O(\lambda). \tag{4.64}$$

Then the pole behavior is

$$G'(p) = \frac{1}{m^2 + p^2 + \Sigma(p)} \to \frac{Z_{\text{pole}}}{m_{\text{part}}^2 + p^2}, \quad Z_{\text{pole}} = \frac{1}{1+\sigma_1}. \tag{4.65}$$

With our previous renormalization conditions (4.53),

$$m = m_{\text{part}}, \quad \sigma_0 = \sigma_1 = 0, \quad Z_{\text{pole}} = 1. \tag{4.66}$$

- The rescaling of the field $\varphi_0 = \sqrt{Z}\,\varphi$ is called a *wave function renormalization.*

  The counterterm δZ is chosen to cancel divergencies in $\Gamma(p,-p)$ of the form βp^2, which occur in higher loop order. Above we assumed its finite part to be fixed such that the residue of the pole in $G'(p)$ is 1. So in general $Z \neq 1$.

  As we have seen above, other renormalization conditions may lead to a different residue Z_{pole}. The value of Z_{pole} is not a physical parameter, it just specifies the scale of φ. After all, also the scale of φ_0, the starting field, does not matter. It is a coordinate in configuration space and physical results should not depend on the choice (scale) of coordinates. This will be evident in chapter 9, where we make the connection between $\langle\varphi(x_1)\cdots\varphi(x_n)\rangle$ and the scattering cross-section.

- It has been shown that the infinities can be cancelled by polynomial counterterms to arbitrary order in perturbation theory. For the φ^4 theory these have the form used in this chapter, with the interpretation of an adjustment of the bare parameters of the theory. Such theories are called renormalizable. For non-renormalizable theories the number of different counterterms that are needed increases with the order of the loop expansion, and the number of free parameters ($\leftrightarrow$ renormalization conditions) increases accordingly, thus limiting the predictive power of the theory. The Standard Model is a renormalizable theory. Einstein quantum gravity is non-renormalizable. The proofs of these statements are complicated.

- The nonperturbative regularization provided by the lattice gives a different view on renormalization. The Compton wavelength in lattice units, $1/am$, is similar to the correlation length in spin models. The continuum region $am \ll 1$ corresponds to the critical region with large correlation lengths. The arbitrariness in the regularization procedure (different formulations giving the same physics) corresponds to the universality of critical phenomena. Universality is a nonperturbative version of renormalizability.

- Last, but not least:

  It is instructive to express the bare parameters in terms of the renormalized ones (cf. eqs. (4.8,4.9):

$$\kappa_0 = Z^{-1}(m^2 - \delta m^2) = m^2 - \frac{6\lambda}{32\pi^2}\left(\Lambda^2 - m^2 \ln\frac{\Lambda^2}{m^2}\right) + O(\lambda^2) \quad (4.67)$$

$$\lambda_0 = Z^{-2}(\lambda - \delta\lambda) = \lambda + \frac{1}{4}\frac{(6\lambda)^2}{16\pi^2}\left(\ln\frac{\Lambda^2}{m^2} - 1\right) + O(\lambda^3). \quad (4.68)$$

  where we used $Z = 1 + O(\lambda^2)$. We see that κ_0 has the tendency to get negative, and λ_0 has the tendenct to grow (more positive), if we let Λ grow. This is indeed what happens in a nonperturbative analysis on the lattice. One finds that λ_0 goes to infinity as the lattice spacing diminishes. Recalling sect. 3.7, this means that φ^4 theory is the 4D Ising model in disguise. With no bare coupling to adjust anymore (λ_0 being infinite) the renormalized coupling λ turns out still to depend on the lattice spacing, but very weakly, while continuum behavior, Lorentz invariance, etc. can be excellent. One finds that in the strict continuum *limit* the renormalized coupling vanishes $\propto 1/\ln(am)$. So we should *not* take this limit.

  This at first sight shocking phenomenon is called 'triviality' (the theory being trivial at zero renormalized coupling). It is not a property of the regularization but of the model itself. The interpretation is that the model breaks down at the regularization scale (cutoff). As we know now, this is also a property of our best phenomenological theory, the 'Standard Model'.

However, in practise there is no problem, since the cutoff can be quite large, e.g. larger than the Planck scale $m_{\text{Planck}} = G^{-1/2} = 1.2\ 10^{19}$ GeV, depending on the renormalized couplings, some of which are not yet known.

Note that, for $\Lambda/m = m_{\text{Planck}}/1$ GeV, the perturbative correction between λ_0 and λ is actually smaller than 1 if $\lambda < 1$.

4.6 Recipe for scattering and decay amplitudes

The question, how to obtain scattering and decay amplitudes from n-point functions will be answered later in chapter 9. The discussion there is somewhat lengthy and a simplified derivation suitable for the tree graph approximation will be given also later in sect. 6.6. Here we just give some rules which allow us to go ahead and apply the diagrammatic techniques to calculate the physically measurable cross sections and decay rates:

1. Let Z_{pole} be the residue of the pole in the two-point function,
$$G'(p) \to \frac{Z_{\text{pole}}}{m_{\text{part}}^2 + p^2}. \tag{4.69}$$
where m_{part} is the particle mass. For simplicity we shall assume to have arranged things such that $m_{\text{part}} = m$ and $Z_{\text{pole}} = 1$. In the tree graph approximation this is the standard convention, in higher orders it implies a suitable choice of counterterms with a suitable rescaling of the fields.

2. Let $G'(p_1, \cdots, p_n)$ be the Fourier transform (as in (4.27)) of the correlation function $G'(x_1, \cdots, x_n)$ defined by
$$\langle \varphi_1 \cdots \varphi_n \rangle_{\text{conn}} \equiv (-i)^{n-1} G'(x_1, \cdots, x_n) \tag{4.70}$$
This function has the structure (cf. Fig. 4.11)
$$G'(p_1, \cdots, p_n) = G'(p_1) \cdots G'(p_n)\, H(p_1, \cdots, p_n), \tag{4.71}$$
i.e. one can always identify full propagators at the external lines. Removing these leaves H, which is sometimes called the 'amputated' n-point correlation function. Note that H is a symmetric function of its arguments.

3. The scattering amplitude for the process $1 + 2 \to 3 + \cdots n$ is simply given by
$$T(p_3, \cdots, p_n; p_1, p_2) = -H(p_3, \cdots, p_n, -p_1, -p_2), \tag{4.72}$$
where now $p_j^0 = \sqrt{m_j^2 + \mathbf{p}_j^2}$, $j = 1, \cdots, n$ (for the φ^4 model the m_j's are all equal). The overall minus sign is a convention.

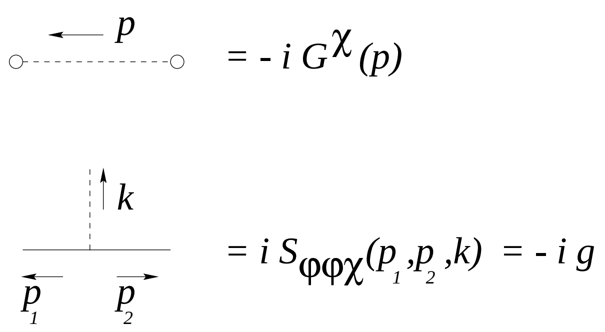

Figure 4.20: Propagator and vertex of χ.

4. For the decay amplitude $1 \to 2 + \cdots + n$ we have a similar formula, in the approximation which treats the decaying particle initially as a stable particle with mass m_1:

$$T(p_2, \cdots, p_n; p_1) = -H(p_2, \cdots, p_n, -p_1). \tag{4.73}$$

We now give some applications. The scattering amplitude of the φ^4 model is given by

$$T(p_3, p_4; p_1, p_2) = -\Gamma(p_3, p_4, -p_1, -p_2), \tag{4.74}$$

and from (4.60) we see that $T = 6\lambda + O(\lambda^2)$, in accordance with the lowest order result (2.89) obtained earlier. Note that the $O(\lambda^2)$ contribution is in general complex, because $(p_1 + p_2)^2 \leq -4m^2$, with an imaginary part determined by the $i\epsilon$ in the logarithm in (4.60). (N.B. In the broken phase of the φ^4 model (corresponding to the magnetized phase of the Ising model) there is also a three-point vertex and T is not just equal to $-\Gamma$.)

As a different application, consider the two-field model described by the classical action

$$S[\chi, \varphi] = -\int d^4x \left(\frac{1}{2}(\partial\chi)^2 + \frac{1}{2}M^2\chi^2 + \frac{1}{2}(\partial\varphi)^2 + \frac{1}{2}m^2\varphi^2 + \frac{\lambda}{4}\varphi^4 + \frac{g}{2}\varphi^2\chi\right), \tag{4.75}$$

which was used in sect. 2.6 to illustrate the possibility of particle decay. Here we assume the χ particle to be stable, $M < 2m$. The free part of the action leads to an additional propagator for the χ field,

$$G^\chi(p) \equiv G^{\chi\chi}(p, -p) = \frac{1}{M^2 + p^2 - i\epsilon}, \tag{4.76}$$

To distinguish it from the φ propagator $G^\varphi(p)$ we use a dashed line in diagrams. The interaction S_1 leads to an additional bare three point vertex as in Fig. 4.20.

The following will be the subject of Problem 2:

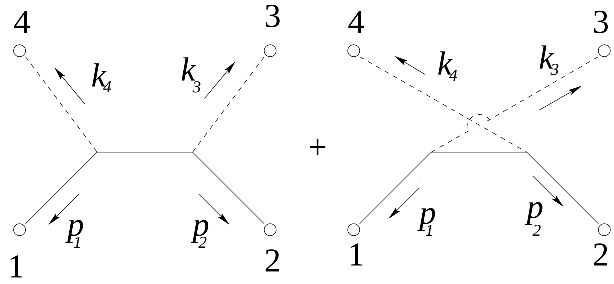

Figure 4.21: Tree graph contribution to $\langle\varphi_1\varphi_2\chi_3\chi_4\rangle^{\text{conn}}$.

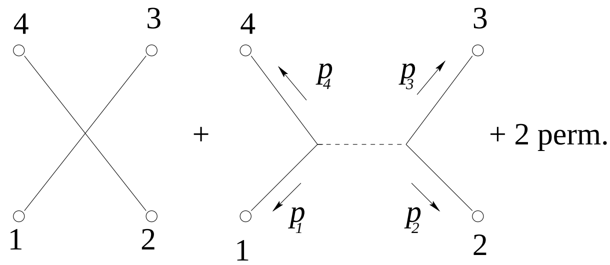

Figure 4.22: Tree graph contribution to $\langle\varphi_1\varphi_2\varphi_3\varphi_4\rangle^{\text{conn}}$.

- The tree graphs in Fig. 4.21 represent the zero loop contribution to $\langle\varphi_1\varphi_2\chi_3\chi_4\rangle^{\text{conn}}$.
- The tree graph contribution to $\langle\varphi_1\cdots\varphi_4\rangle^{\text{conn}}$ is shown in Fig. 4.22. The correspondig scattering amplitude is given by

$$T(p_3,p_4;p_1,p_2) = 6\lambda - g^2\left(\frac{1}{M^2-s}+\frac{1}{M^2-t}+\frac{1}{M^2-u}\right), \tag{4.77}$$

with $s=-(p_1+p_2)^2$, $t=-(p_1-p_3)^2$ and $u=-(p_1-p_4)^2$ in standard notation. The differential cross section in the center of mass frame is given by

$$\begin{aligned}\left[\frac{d\sigma}{d\Omega}\right]_{\text{cm}} &= \frac{1}{64\pi^2 s}\left[6\lambda - g^2\left(\frac{1}{M^2-s}+\frac{1}{M^2+2|\mathbf{p}|^2(1+\cos\theta)}\right.\right.\\ &\qquad \left.\left.+\frac{1}{M^2+2|\mathbf{p}|^2(1-\cos\theta)}\right)\right]^2. \end{aligned}\tag{4.78}$$

Note the symmetry under $\theta\to\pi-\theta$, $\phi\to\phi+\pi$ corresponding to the identical particles.

4.7 Summary

Expanding $\exp(iS_1)$ in the path integral we can express the full correlation functions in terms of free field propagators and vertex functions. Diagrams are a great help in clarifying these expressions and we can establish a 1-1 correspondence between the diagrams and the mathematical expressions. The correlation functions can be decomposed in terms of 1PI vertex functions and full propagators (i.e. two-point correlation functions).

The cavalier treatment of the infinite number of degrees of freedom in field theory leads to infinities in loop diagrams. These infinities can be cancelled by counterterms and in a renormalizable theory the counterterms can be absorbed in the parameters appearing in the action. The parameters in the action are called bare parameters (in the φ^4 model these are ϵ_0, Z, m_0^2 and λ_0), while the physical parameters are defined in terms of the renormalized vertex functions ('dressed' by loop diagrams). The arbitrariness in the finite parts of diagrams and counterterms is fixed by imposing some standard normalization condition on the renormalized vertex functions.

A simple recipe can be followed to obtain scattering or decay amplitudes from the correlation functions.

4.8 Appendix

The identity (4.42) can be derived as follows. First use the exponential representation

$$\frac{1}{A^\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty dt\, t^{\alpha-1} e^{-tA}, \tag{4.79}$$

to get

$$\frac{1}{A^\alpha B^\beta} = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty dt\, du\, t^{\alpha-1} u^{\beta-1}\, e^{-tA-uB}. \tag{4.80}$$

Next, make the transformation of variables

$$t = xr, \quad u = (1-x)r, \quad r \in (0,\infty), \quad x \in (0,1), \tag{4.81}$$

and use

$$\int_0^\infty dr\, r^{\alpha+\beta-1}\, e^{-r[xA+(1-x)B]} = \frac{\Gamma(\alpha+\beta)}{[xA+(1-x)B]^{\alpha+\beta}}, \tag{4.82}$$

which gives

$$\frac{1}{A^\alpha B^\beta} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 dx\, \frac{x^{\alpha-1}(1-x)^{\beta-1}}{[xA+(1-x)B]^{\alpha+\beta}}. \tag{4.83}$$

The derivation is valid for postive A, B and α, β, and furthermore anywhere else that can be reached by analytic continuation such that the integral over x remains convergent. The generalization to

$$\begin{aligned} \frac{1}{A_1^{\alpha_1} \cdots A_n^{\alpha_n}} &= \frac{\Gamma(\alpha_1 + \cdots + \alpha_n)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_n)} \int_0^1 dx_1 \cdots dx_n\, \delta(1 - x_1 - \cdots - x_n) \\ &\quad \frac{x_1^{\alpha_1-1}\cdots x_n^{\alpha_n-1}}{[x_1 A_1 + \cdots + x_n A_n]^{\alpha_1+\cdots+\alpha_n}} \end{aligned} \tag{4.84}$$

is straightforward.

4.9 Problems

1. *Shifting integration variables in divergent integrals*

Consider the quadratically divergent integral

$$I(p) = \int \frac{d^4q}{(2\pi)^4} \frac{1}{m^2 + (q+p)^2}, \tag{4.85}$$

in euclidean momentum-space. Shifting the integration variable $q \rightarrow q - p$ one would conclude that $I(p) = I(0)$. This is indeed true for dimensional, Pauli-Villars, and lattice regularization. However, suppose we regulate the integral by imposing a spherical cutoff:

$$I(p) \equiv \int_{q^2<\Lambda^2} \frac{d^4q}{(2\pi)^4} \frac{1}{m^2 + (q+p)^2}. \tag{4.86}$$

The dependence on p can be studied by makin a Taylor expansion about $p = 0$:

$$I(p) = I(0) + p_\mu \left[\partial^p_\mu I(p)\right]_{p=0} + \frac{1}{2} p_\mu p_\nu \left[\partial^p_\mu \partial^p_\nu I(p)\right]_{p=0} + \cdots \tag{4.87}$$

Show that for large Λ,

$$I(p) = I(0) + \frac{1}{8\pi^2} p^2 + O(\Lambda^{-2}). \tag{4.88}$$

Let $K(p)$ be given by

$$K(p) = \int_{q^2<\Lambda^2} \frac{d^4q}{(2\pi)^4} \frac{1}{[m^2 + (q+p)^2]^2}. \tag{4.89}$$

Neglecting terms that vanish as $\Lambda \rightarrow \int d^4y$, what is the relation between $K(p)$ and $K(0)$?

Hint: use symmetry arguments to conclude that,

$$\int_{q^2<\Lambda^2} \frac{d^4q}{(2\pi)^4} \frac{q_\mu}{[m^2+q^2]^n} = 0, \tag{4.90}$$

$$\int_{q^2<\Lambda^2} \frac{d^4q}{(2\pi)^4} \frac{q_\mu q_\nu}{[m^2+q^2]^n} = \frac{\delta_{\mu\nu}}{4} \int_{q^2<\Lambda^2} \frac{d^4q}{(2\pi)^4} \frac{q^2}{[m^2+q^2]^n}. \tag{4.91}$$

The above examples I and K may seem a little academic, but in gauge theories the self-energy diagrams typically lead to quadratically divergent integrals of the type

$$\int \frac{d^4q}{(2\pi)^4} \frac{(2q+p)_\mu (2q+p)_\nu}{[m^2+q^2][m^2+(q+p)^2]}, \tag{4.92}$$

which one would like to evaluate with the method (4.42), with subsequent shifting of integration variable as in (4.44). For such integrals the ambiguities of sharp momentum cutoffs have to be well understood and absorbed into the renormalization prcedure. It is however, much simpler to use a regularization that is invariant under shifts of integration variables. A convenient method is dimensional regularization, which is described in many text books on relativistic quantum fields.

2. *Revisiting the two-scalar-field model*

   a. Verify Fig. 4.20, i.e. derive the bare vertex function in position space and Fourier transform this to momentum space.

   b. Verify that the tree graphs in Fig. 4.21 represent the zero loop contribution to $\langle\varphi_1\varphi_2\chi_3\chi_4\rangle^{\rm conn}$ and determine the corresponding mathematical expression (in momentum space) using the path integral.

   c. Similarly, verify the tree graph contribution to $\langle\varphi_1\cdots\varphi_4\rangle^{\rm conn}$ shown in Fig. 4.22 and write down the corresponding mathematical expression.

   d. Verify the expression (4.78) for the differential cross section (cf. Prob. 2.3).

   e. Assume that in the quantum theory the action of the two scalar field model (4.75) is the total action without counterterms. Introduce the counterterms corresponding to mass, coupling and wave function renormalization and write down S_0, S_1 and ΔS for this model.

3. *Nonzero vacuum expectation value*

   The φ^4 model and also the model (4.75) has a discrete symmetry $\varphi \to -\varphi$. From this one may conclude that $\langle\varphi\rangle = 0$. However, the symmetry may be broken spontaneously, as in the Ising model. This does not happen for free fields, but with interactions it is possible, depending on the value of the parameters of the model. Here we assume no spontaneous symmetry breaking (we have implicitly made this assumption throughout this chapter).

   On the other hand, there is no $\chi \to -\chi$ symmetry in the model (4.75). Therefore we can expect that $\langle\chi(x)\rangle \neq 0$. Under homogeneous circumstances it will be independent of x. The full propagator is defined as a *correlation function*,

$$G'^{\chi}(x,y) = \langle\chi(x)\chi(y)\rangle - \langle\chi(x)\rangle\langle\chi(y)\rangle. \tag{4.93}$$

   a. Verify that in one-loop order $\langle\chi\rangle$ is given by the diagram in Fig. 4.23 and determine the corresponding mathematical expression.

   b. Verify that the selfenergy vertex function for the χ field, $\Sigma_\chi(p)$, is given in one loop order by the diagram in Fig. 4.24

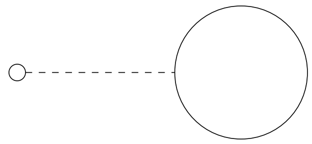

Figure 4.23: One loop contribution to $\langle\chi\rangle$.

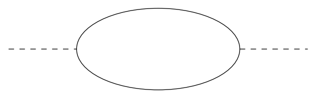

Figure 4.24: One loop contribution to $-i\Sigma_\chi(p)$.

c. Evaluate $\Sigma_\chi(p)$ with the cutoff method.

d. Determine the counterterms to make $\Sigma_\chi(p)$ finite, such that

$$G'^{\chi}(p) \to \frac{1}{M^2+p^2}, \tag{4.94}$$

near its pole.

Chapter 5

Spinor fields

Up to now we have only considered in detail fields that have integer spin under rotations. There are also representations of the rotation group with half integer spin which are embedded into representations of the Lorentz group.[1] The simplest have spin 1/2, these are called spinor representations. Majorana fields and Dirac fields are real and complex spinor fields, respectively. This chapter provides an introduction to the classical fields, the next chapter deals with the quantum case.

5.1 Spinors

To construct spinor representations we introduce Dirac matrices. These are four 4×4 matrices γ^μ which satisfy the anticommutation relations[2]

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}\, 1. \tag{5.1}$$

Hence,

$$\begin{aligned} \gamma_\mu \gamma_\nu &= -\gamma_\nu \gamma_\mu, &(5.2)\\ \gamma_0^2 &= -1, \quad \gamma_k^2 = 1, \quad k = 1, 2, 3. &(5.3) \end{aligned}$$

Below we shall give explicit examples of matrices γ_μ satisfying the above relations. Let us now first show that generators $S_{\alpha\beta}$ of Lorentz transformations can be constructed in terms of them. Consider

$$\begin{aligned} S_{\alpha\beta} = -S_{\beta\alpha} &= -i\frac{1}{4}[\gamma_\alpha, \gamma_\beta], \quad \text{spinor representation} &(5.4)\\ &= 0, \quad \alpha = \beta &(5.5)\\ &= -i\frac{1}{2}\gamma_\alpha\gamma_\beta, \quad \alpha \neq \beta. &(5.6) \end{aligned}$$

[1] Recall the discussion below eq. (1.43).

[2] This relation reads more explicitly $(\gamma_\mu)_{ab}(\gamma_\nu)_{bc} + (\gamma_\nu)_{ab}(\gamma_\mu)_{bc} = 2\eta_{\mu\nu}\delta_{ac}$, where $a, b, c = 1, \ldots, 4$ are the matrix indices. We usually denote the identity matrix $(\mathbb{1})_{ac} = \delta_{ac}$ simply by 1, and sometimes even omit this symbol as well in compact notation.

Using the anticommutation relations[3] (5.1) we have for $\alpha \neq \beta$,

$$
\begin{aligned}
[S_{\alpha\beta}, \gamma_\mu] &= -i\frac{1}{2}[\gamma_\alpha\gamma_\beta, \gamma_\mu] && (5.7)\\
&= -i\frac{1}{2}\gamma_\alpha\{\gamma_\beta, \gamma_\mu\} + i\frac{1}{2}\{\gamma_\alpha, \gamma_\mu\}\gamma_\beta && (5.8)\\
&= -i(\eta_{\beta\mu}\gamma_\alpha - \eta_{\alpha\mu}\gamma_\beta), && (5.9)
\end{aligned}
$$

and in addition for $\mu \neq \nu$,

$$
\begin{aligned}
[S_{\alpha\beta}, S_{\mu\nu}] &= -i\frac{1}{2}[S_{\alpha\beta}, \gamma_\mu\gamma_\nu] && (5.10)\\
&= -i\frac{1}{2}\left([S_{\alpha\beta}, \gamma_\mu]\gamma_\nu + \gamma_\mu[S_{\alpha\beta}, \gamma_\nu]\right) && (5.11)\\
&= -\frac{1}{2}(\eta_{\beta\mu}\gamma_\alpha\gamma_\nu - \eta_{\alpha\mu}\gamma_\beta\gamma_\nu + \eta_{\beta\nu}\gamma_\mu\gamma_\alpha - \eta_{\alpha\nu}\gamma_\mu\gamma_\beta) && (5.12)\\
&= i(-\eta_{\beta\mu}S_{\alpha\nu} + \eta_{\alpha\mu}S_{\beta\nu} + \eta_{\beta\nu}S_{\alpha\mu} - \eta_{\alpha\nu}S_{\beta\mu}). && (5.13)
\end{aligned}
$$

In the last line we used

$$\gamma_\alpha\gamma_\nu = 2iS_{\alpha\nu} + \eta_{\alpha\nu}\,1, \tag{5.14}$$

etc. Hence, the matrices $S_{\alpha\beta}$ satisfy the commutation relations (1.53) and therefore they provide a representation of the Lorentz group.

We now go into more detail on the properties of the Dirac matrices. The gamma matrices are unitary,

$$\gamma_\mu^\dagger\gamma_\mu = 1, \quad \text{no sum over } \mu, \tag{5.15}$$

which implies with (5.3),

$$\gamma_0^\dagger = -\gamma_0, \quad \gamma_k^\dagger = \gamma_k, \quad k = 1, 2, 3. \tag{5.16}$$

As a special notation for $i\gamma^0$ we introduce a matrix β

$$\beta = i\gamma^0 = -i\gamma_0 = \beta^\dagger, \tag{5.17}$$

in terms of which we can write

$$\gamma_\mu^\dagger = -\beta\gamma_\mu\beta. \tag{5.18}$$

The notation

$$\alpha^\mu = i\beta\gamma^\mu, \quad (\alpha^0 = -\alpha_0 = 1) \tag{5.19}$$

also occurs. An important combination is furthermore the matrix γ_5,

$$\gamma_5 \equiv \gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma_5^\dagger, \tag{5.20}$$

[3] Basic identities such as $[ab, c] = a[b, c] + [a, c]b$ and $[ab, c] = a\{b, c\} - \{a, c\}b$ are useful here.

which anticommutes with all the γ_μ,

$$\gamma_5\gamma_\mu = -\gamma_\mu\gamma_5, \quad \mu = 0, \cdots, 3. \tag{5.21}$$

There is an infinite number of realizations of the Dirac matrices, which differ by unitary transformations. We shall give two: a so-called chiral representation in which γ_5 is diagonal, and a Majorana representation in which the γ^μ are real. It is convenient to use a tensor product notation in terms of Pauli matrices σ_k, $k = 1, 2, 3$. We also introduce 4×4 matrices ρ_k and σ_k which are essentially just the Pauli matrices. Introducing a two-index notation for the Dirac index a, $a \leftrightarrow \alpha\beta$, with $\alpha, \beta = 1, 2$ (so $a = 1, \ldots, 4 \leftrightarrow \alpha\beta = 11, 21, 12, 22$), they are given by

$$(\sigma_k)_{\alpha\beta,\alpha'\beta'} = (\sigma_k)_{\alpha\alpha'}\,\delta_{\beta\beta'}, \quad (\rho_k)_{\alpha\beta,\alpha'\beta'} = (\sigma_k)_{\beta\beta'}\,\delta_{\alpha\alpha'}, \tag{5.22}$$

where the σ_k on the right hand side are the 2×2 Pauli matrices. Their action on a spinor ψ is given by

$$(\sigma_k\rho_l\psi)_{\alpha\beta} = (\sigma_k)_{\alpha\alpha'}\,(\sigma_l)_{\beta\beta'}\,\psi_{\alpha'\beta'}, \tag{5.23}$$

$$(\rho_k\psi)_{\alpha\beta} = (\sigma_k)_{\beta\beta'}\,\psi_{\alpha\beta'}, \quad (\sigma_k\psi)_{\alpha\beta} = (\sigma_k)_{\alpha\alpha'}\,\psi_{\alpha'\beta}, \tag{5.24}$$

with the usual summation over repeated indices. For simplicity we use the same symbol for the 4×4 σ_k as for the 2×2 Pauli matrices and the unit matrix is often omitted. A chiral representation is now given by

$$\beta = i\gamma^0 = \begin{pmatrix} 0 & 1\!\!1 \\ 1\!\!1 & 0 \end{pmatrix} = \rho_1, \tag{5.25}$$

$$\gamma^k = \begin{pmatrix} 0 & i\sigma_k \\ -i\sigma_k & 0 \end{pmatrix} = -\rho_2\,\sigma_k, \tag{5.26}$$

$$\gamma_5 = \begin{pmatrix} 1\!\!1 & 0 \\ 0 & -1\!\!1 \end{pmatrix} = \rho_3, \tag{5.27}$$

$$\alpha^k = \begin{pmatrix} \sigma_k & 0 \\ 0 & -\sigma_k \end{pmatrix} = \rho_3\,\sigma_k, \quad \alpha^0 = 1. \tag{5.28}$$

where we have also indicated the matrices in 2×2 block form. A Majorana representation is given by

$$\gamma^1 = -\sigma_3, \ \ \gamma^2 = -\rho_2\sigma_2, \ \ \gamma^3 = \sigma_1, \ \ \gamma^0 = i\rho_3\sigma_2, \tag{5.29}$$

which are real, $\gamma_\mu = \gamma_\mu^*$, and imply

$$\gamma_5 = \rho_1\sigma_2 = -\gamma_5^*. \tag{5.30}$$

The anticommutation relations (5.1) can be easily checked from these representations, since $[\rho_k, \sigma_l] = 0$. For example, in the chiral representation,

$$\{\gamma_k, \gamma_l\} = \rho_2^2\{\sigma_k, \sigma_l\} = 2\delta_{kl}, \ \ \{\gamma_0, \gamma_k\} = -i\{\rho_1, \rho_2\}\sigma_k = 0, \ \ \gamma_0^2 = -\rho_1^2 = -1. \tag{5.31}$$

A spinor field $\psi(x)$ is now a field transforming under Lorentz transformations $x' = \ell x$ as

$$\psi'(x') = D\psi(x), \quad (\psi'_a(x') = D_{ab}\,\psi_b(x), \;\; a, b = 1, \ldots, 4), \tag{5.32}$$

with

$$\ell = \exp(i\frac{1}{2}\omega^{\alpha\beta}S^{\rm def}_{\alpha\beta}), \quad D(\ell) = \exp(i\frac{1}{2}\omega^{\alpha\beta}S_{\alpha\beta}) = \exp(\frac{1}{4}\omega^{\alpha\beta}\gamma_\alpha\gamma_\beta), \tag{5.33}$$

where $S^{\rm def}_{\alpha\beta}$ are the generators of the defining representation (1.52) and $S_{\alpha\beta}$ the generators in a spinor representation.[4] We see that D is real in a Majorana representation of the gamma matrices. Accordingly, it can be convenient use such representations for Majorana fields:

$$\psi(x)^* = \psi(x), \;\; D = D^*, \quad \text{Majorana}. \tag{5.34}$$

From two Majorana fields $\psi_{1,2}$ we can construct a complex field, a Dirac field ψ,

$$\psi = \frac{1}{\sqrt{2}}\,(\psi_1 - i\psi_2). \tag{5.35}$$

It is then generally convenient to allow the γ_μ to be complex. When we make a unitary transformation from a Majorana representation to a representation where the γ_μ are complex, the ψ in (5.35) stays of course complex but the new $\psi_{1,2}$ become also complex. In general then, Majorana fields are complex but satisfy a reality constraint, cf. (5.120) in Appendix 5.5. We shall continue with the Dirac case, from which it is straightforward to specialize to the Majorana case.

Let us now see how the Dirac field transforms under rotations and confirm that the spin of the representation is 1/2. Rotations only involve ω^{kl}, $k, l = 1, 2, 3$. Now in the chiral representation (5.28) we have explicitly

$$S_{kl} = \frac{1}{2}\sigma_m = \frac{1}{2}\begin{pmatrix} \sigma_m & 0 \\ 0 & \sigma_m \end{pmatrix}, \quad (k, l, m = 1, 2, 3 \text{ cycl.}) \tag{5.36}$$

and

$$D = \begin{pmatrix} e^{i\frac{1}{2}\boldsymbol{\varphi}\cdot\boldsymbol{\sigma}} & 0 \\ 0 & e^{i\frac{1}{2}\boldsymbol{\varphi}\cdot\boldsymbol{\sigma}} \end{pmatrix}, \quad \varphi_k \equiv \frac{1}{2}\epsilon_{klm}\omega_{lm}, \quad \text{rotation}, \tag{5.37}$$

which shows that the representation has spin 1/2. For a special Lorentz transformation we have similarly

$$S_{0k} = -i\frac{1}{2}\rho_3\sigma_k = \frac{1}{2}\begin{pmatrix} -i\sigma_k & 0 \\ 0 & i\sigma_k \end{pmatrix}, \tag{5.38}$$

[4]For notational convenience we suppress the label 'spin' on $S^{\rm spin}_{\alpha\beta}$ and $D^{\rm spin}$.

and

$$D = \begin{pmatrix} e^{\frac{1}{2}\boldsymbol{\chi}\cdot\boldsymbol{\sigma}} & 0 \\ 0 & e^{-\frac{1}{2}\boldsymbol{\chi}\cdot\boldsymbol{\sigma}} \end{pmatrix}, \quad \chi_k \equiv \omega^{0k}, \quad \text{special Lorentz transformation.} \tag{5.39}$$

The 2×2 block diagonal form shows that D is a reducible representation of the proper Lorentz group. The representation in the upper 2×2 block is inequivalent to that in the lower block. In standard notation these representations are denoted by (1/2,0) and (0,1/2). Another notation is R and L, $D^{(1/2,0)} = D_R$, $D^{(0,1/2)} = D_L$. The fields ψ_R and ψ_L transforming under the irreducible D_R and D_L can be obtained from ψ by applyting the *chiral projectors* $P_{R,L}$ defined by

$$P_R = \frac{1}{2}(1+\gamma_5), \quad P_L = \frac{1}{2}(1-\gamma_5, \tag{5.40}$$

with the properties

$$P_R^2 = P_R, \quad P_L^2 = P_L, \quad P_R + P_L = 1, \quad P_R P_L = 0. \tag{5.41}$$

In the chiral representation

$$\psi_R = P_R \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ 0 \\ 0 \end{pmatrix}, \quad \psi_L = P_L \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \psi_3 \\ \psi_4 \end{pmatrix}. \tag{5.42}$$

In the Majorana representation (5.29)

$$\psi_R = P_R \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \psi_1 - i\psi_4 \\ \psi_2 + i\psi_3 \\ -i(\psi_2 + i\psi_3) \\ i(\psi_1 - i\psi_4) \end{pmatrix}, \quad \psi_L = \psi_R^*, \tag{5.43}$$

where the latter relation is true for real ψ. We see that a single real Majorana field is equivalent to a field ψ_R with two independent complex components. We also see that ψ_R^* is equivalent to a ψ_L.

The representation $D = D_R + D_L$ is irreducible if parity is included: we shall see below that the parity transformation is represented by $D_P = \gamma^0$, which interchanges the two blocks of the chiral representation.

In Nature, parity is violated, so why does not one just use one of these irreducible representations? The reason is that we need both representations for making Lorentz tensors, in particular scalars such as the action and, in the quantum theory, Lorentz invariant amplitudes. It is a curious fact that the 4×4 D in the Majorana representation is irreducible within the real numbers. So in Majorana language the question posed above does seem to not come up.

We now turn to the construction of Lorentz invariant or covariant combinations of ψ and ψ^*. Instead of ψ^* we shall use $\psi^\dagger = \psi^{*T}$, which is more convenient. As in (5.32) we may consider ψ to be column vector on which the Dirac matrices can act from the left, and $\psi^\dagger$ a row vector on which the matrices can act from the right. Now $\psi^\dagger$ transforms as $\psi'(x')^\dagger = \psi(x)^\dagger D^\dagger$. However, $\psi'^\dagger\psi' \neq \psi^\dagger\psi$ since D is not unitary in general: $D^\dagger \neq D^{-1}$ (it is known that finite dimensional representations of the Lorentz group cannot be unitary). Using (5.18) we have instead

$$\begin{aligned} D^\dagger &= \exp(\frac{1}{4}\omega^{\mu\nu}\gamma_\nu^\dagger\gamma_\mu^\dagger) = \beta\exp(\frac{1}{4}\omega^{\mu\nu}\gamma_\nu\gamma_\mu)\beta && (5.44)\\ &= \beta\exp(-\frac{1}{4}\omega^{\mu\nu}\gamma_\mu\gamma_\nu)\beta && (5.45)\\ &= \beta D^{-1}\beta, && (5.46)\end{aligned}$$

or using $\beta^2 = 1$,

$$D^\dagger\beta = \beta D^{-1}. \tag{5.47}$$

This shows that $\psi^\dagger\beta\psi$ is a scalar:

$$\psi'(x')^\dagger\beta\psi'(x') = \psi^\dagger(x)D^\dagger\beta D\psi(x) = \psi^\dagger(x)\beta\psi(x). \tag{5.48}$$

We see that the matrix β plays the role of a metric in spinor space. It is customary to use a special notation for the combination $\psi^\dagger\beta$,

$$\bar\psi \equiv \psi^\dagger\beta, \tag{5.49}$$

which transforms as

$$\bar\psi'(x') = \bar\psi(x)D^{-1}. \tag{5.50}$$

As the notation anticipated, γ^μ is a vector matrix, in the sense that

$$D^{-1}\gamma^\mu D = \ell^\mu{}_\nu\gamma^\nu. \tag{5.51}$$

This follows by writing the left hand side in infinitesimal form,

$$(1 - i\frac{1}{2}\omega^{\alpha\beta}S_{\alpha\beta})\gamma^\mu(1 + i\frac{1}{2}\omega^{\alpha\beta}S_{\alpha\beta}) = \gamma^\mu + i\frac{1}{2}\omega^{\alpha\beta}[\gamma^\mu, S_{\alpha\beta}] + \ldots \tag{5.52}$$

and using (5.9), in which we recognize on the right hand side the generators in the vector representation (1.52),

$$[\gamma^\mu, S_{\alpha\beta}] = -i(\eta^\mu_\alpha\eta_{\beta\nu} - \eta^\mu_\beta\eta_{\alpha\nu})\gamma^\nu. \tag{5.53}$$

It follows that $\bar\psi(x)\gamma^\mu\psi(x)$ is a vector field,

$$\bar\psi'(x')\gamma^\mu\psi'(x') = \ell^\mu{}_\nu\bar\psi(x)\gamma^\nu\psi(x). \tag{5.54}$$

Similarly, $\gamma^\mu\gamma^\nu$ is a tensor matrix; the antisymmetric part is proportional to $S^{\mu\nu}$, while the symmetric part is proportional to $\eta^{\mu\nu}$.

Let us introduce at this point suitable representatives of parity and time reversal,

$$D_P = \gamma^0, \tag{5.55}$$
$$D_T = i\gamma^0\gamma_5 \tag{5.56}$$

(which are both real in a Majorana representation). These are reasonable definitions as can be seen from

$$D_P^{-1}\gamma^\mu D_P = (\ell_P)^\mu{}_\nu\gamma^\nu, \quad \bar\psi'(x')\gamma^\mu\psi'(x') = (\ell_P)^\mu{}_\nu\bar\psi(x)\gamma^\nu\psi(x), \tag{5.57}$$
$$D_T^{-1}\gamma^\mu D_T = (\ell_T)^\mu{}_\nu\gamma^\nu, \quad \bar\psi'(x')\gamma^\mu\psi'(x') = (\ell_T)^\mu{}_\nu\bar\psi(x)\gamma^\nu\psi(x). \tag{5.58}$$

The matrix γ_5 is a Lorentz scalar,

$$D^{-1}\gamma_5 D = \gamma_5, \tag{5.59}$$

since it commutes with the generators $S_{\alpha\beta}$. It is in fact a *pseudoscalar* under parity and time reversal:

$$D_P^{-1}\gamma_5 D_P = -\gamma_5, \quad D_T^{-1}\gamma_5 D_T = -\gamma_5. \tag{5.60}$$

The antisymmetrized product of three Dirac matrices

$$\frac{1}{3!}\epsilon_{\kappa\lambda\mu\nu}\gamma^\lambda\gamma^\mu\gamma^\nu = -i\gamma_\kappa\gamma_5 \tag{5.61}$$

is a pseudovector.

The 16 matrices 1, γ_μ, $2iS_{\mu\nu}$, $i\gamma_\mu\gamma_5$ and γ_5 are linearly independent and form a basis for the 4×4 matrices: an arbitrary 4×4 matrix can be written as a linear superposition of them. We shall not go into details here.

5.2 Action, Dirac equation and Noether charges

Consider the problem of writing down an action S for spinor fields. To get linear field equations from the stationary action principle, S has to be quadratic or bilinear in the fields. We also want it to be local, Lorentz and parity invariant, i.e. of the form

$$S = \int d^4x\,\mathcal{L}(x), \tag{5.62}$$

with a real Lagrange density $\mathcal{L}$ which is a scalar under Lorentz transformations as well as under the parity transformation P. The reality condition will be relevant for Dirac fields, which are complex, so consider first this case. Candidate terms in $\mathcal{L}$ are $\bar\psi\psi$, $\bar\psi\gamma^\mu\partial_\mu\psi + \text{c.c}$, $\bar\psi\gamma^\mu\gamma_5\partial_\mu\psi + \text{c.c.}$, $\partial_\mu\bar\psi\partial^\mu\psi$ and so on. A convenient rule is

$$[\psi_1^\dagger M\psi_2]^* = \psi_2^\dagger M^\dagger\psi_1, \tag{5.63}$$

where M is any matrix. So $\bar\psi\psi = \psi^\dagger\beta\psi$ is real, while

$$[\bar\psi\gamma^\mu\partial_\mu\psi]^* = [\psi^\dagger\beta\gamma^\mu\partial_\mu\psi]^* = \partial_\mu\psi^\dagger\gamma^{\mu\dagger}\beta\psi = -\partial_\mu\bar\psi\gamma^\mu\psi, \tag{5.64}$$

where we used (5.18). Keeping a minimal (> 0) number of derivatives and avoiding parity violation (γ_5 leads to a sign change under P), the minimal form for $\mathcal{L}$ is given by

$$-\mathcal{L} = m\bar\psi\psi + \frac{1}{2}\bar\psi\gamma^\mu\partial_\mu\psi - \frac{1}{2}\partial_\mu\bar\psi\gamma^\mu\psi \equiv \bar\psi(m + \gamma^\mu \overleftrightarrow{\partial_\mu})\psi, \quad \text{Dirac.} \tag{5.65}$$

We introduced the convenient shorthand

$$f_1 \overleftrightarrow{\partial_\mu} f_2 = \frac{1}{2}(f_1\partial_\mu f_2 - \partial_\mu f_1\, f_2). \tag{5.66}$$

The parameter m is real and the coefficient of the derivative term has been chosen unity, which is a normalization convention for the fields. The need for the overall minus sign will be verified later.

Requiring S to be stationary leads to the Dirac equation. In the Dirac case we treat ψ and $\bar\psi$ as independent, since the real and imaginary parts $\psi_{1,2}$ of ψ are independent variables. The relation of $\psi_{1,2}$ to ψ and ψ^* can be written as a linear transformation of variables

$$\begin{pmatrix} \psi_a \\ \psi_a^* \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}\begin{pmatrix} \psi_{1a} \\ \psi_{2a} \end{pmatrix} \tag{5.67}$$

($a = 1, \cdots, 4$ is the spinor index). The variation with respect to $\bar\psi$ gives

$$\delta_{\bar\psi} S = -\int d^4x\, \delta\bar\psi(m + \gamma^\mu \overleftrightarrow{\partial_\mu})\psi = -\int d^4x\, \delta\bar\psi(m + \gamma^\mu\partial_\mu)\psi, \tag{5.68}$$

where we made a partial integration (with $\delta\bar\psi = 0$ at the boundary of the integration volume, as usual). Requiring the action to be stationary gives the *Dirac equation*

$$(m + \gamma^\mu\partial_\mu)\psi = 0. \tag{5.69}$$

The conjugate equation

$$0 = \delta_\psi S = -\int d^4x\, \bar\psi(m + \gamma^\mu \overleftrightarrow{\partial_\mu})\delta\psi \Rightarrow m\bar\psi - \partial_\mu\bar\psi\gamma^\mu = 0, \tag{5.70}$$

can also be obtained by hermition conjugation of (5.69).

Symmetries of the action lead to Noether currents and conserved quantities. These can be found by the methods used in sect. 1.7 for the scalar field. Infinitesimal spacetime dependent translations

$$\delta\psi = -\epsilon^\mu \partial_\mu \psi, \quad \delta\bar{\psi} = -\epsilon^\mu \partial_\mu \bar{\psi}, \tag{5.71}$$

lead to the energy-momentum tensor

$$T^{\mu\nu} = \bar{\psi}\gamma^\mu \overleftrightarrow{\partial^\nu} \psi + \eta^{\mu\nu}\mathcal{L}, \quad \partial_\mu T^{\mu\nu} = 0. \tag{5.72}$$

Similarly, the infinitesimal spacetime dependent Lorentz transformations

$$\delta\psi = -\omega^{\alpha\beta}(x_\beta\partial_\alpha - i\frac{1}{2}S_{\alpha\beta})\psi, \quad \delta\bar{\psi} = -\omega^{\alpha\beta}(x_\beta\partial_\alpha\bar{\psi} + i\frac{1}{2}\bar{\psi}S_{\alpha\beta}), \tag{5.73}$$

lead to the generalized angular momentum currents

$$J^{\mu\alpha\beta} = x^\alpha T^{\mu\beta} - x^\beta T^{\mu\alpha} + \frac{i}{2}\bar{\psi}\{\gamma^\mu, S^{\alpha\beta}\}\psi, \quad \partial_\mu J^{\mu\alpha\beta} = 0. \tag{5.74}$$

The conserved energy-momentum and generalized angular momenta are as usual given by

$$P^\nu = \int d^3x\, T^{0\nu}, \quad J^{\alpha\beta} = \int d^3x\, J^{0\alpha\beta}. \tag{5.75}$$

Notice that the above $T^{\mu\nu}$ is not symmetric: it cannot be used in the Einstein equations. The route of finding a better energy-momentum tensor via a Dirac action in general relativity is too involved to be exposed here. There are other ways to improve the above $T^{\mu\nu}$ into a symmetric one, see for instance Weinberg I sect. 7.4. Such improvement does not affect the total P^μ and $J^{\mu\nu}$.

The action has more symmetries. It is invariant under phase transformations

$$\psi'(x) = e^{i\omega}\,\psi(x), \quad \bar{\psi}'(x) = e^{-i\omega}\,\bar{\psi}(x). \tag{5.76}$$

Using spacetime dependent infinitesimal transformations

$$\delta\psi = i\epsilon\psi, \quad \delta\bar{\psi} = -i\epsilon\bar{\psi}, \tag{5.77}$$

we find a conserved current from

$$\delta S = -\int d^4x\, j^\mu \partial_\mu \epsilon = 0 \Rightarrow \partial_\mu j^\mu = 0. \tag{5.78}$$

The current and corresponding charge are given by

$$j^\mu = \bar{\psi} i\gamma^\mu \psi, \quad Q = \int d^3x\, j^0 = \int d^3x\, \psi^\dagger \psi. \tag{5.79}$$

If the parameter m vanishes we have one more continuous symmetry:

$$\psi'(x) = e^{i\omega\gamma_5}\,\psi(x), \quad \bar{\psi}'(x) = \bar{\psi}(x)e^{i\omega\gamma_5}, \tag{5.80}$$

with corresponding conserved current and charge

$$j_5^\mu = \bar\psi i\gamma^\mu\gamma_5\psi, \quad Q_5 = \int d^3x\, j_5^0 = \int d^3x\, \psi^\dagger\gamma_5\psi. \tag{5.81}$$

We end this section with an at first sight puzzling phenomenon. For the Majorana case we assume ψ to be real and choose a Majorana representation of the γ^μ. Then with the conventional factor 1/2 for real fields, the Lagrange density becomes

$$\mathcal{L} = -\frac{1}{2}\psi^T\beta(m + \gamma^\mu \overleftrightarrow{\partial}_\mu)\psi, \quad \text{Majorana.} \tag{5.82}$$

Now, since β is hermitian and imaginary (in a Majorana representation) it is antisymmetric, $\beta = -\beta^T$. Therefore the term $\psi^T\beta\psi$ is zero if ψ is an ordinary real number. The derivative term has a similar property, as follows from $\alpha^\mu \equiv i\beta\gamma^\mu = \alpha^{\mu T}$ (in a Majorana representation) and the fact that $\overleftrightarrow{\partial}_\mu$ is antisymmetric. The Majorana action is identically zero for a real field! This raises the question how to use the action in a path integral. Or is it not true that a Dirac field is equivalent to two Majorana fields? The surprising resolution of this puzzle is, that spinor fields in a path integral have to be *anticommuting*, i.e. $\psi_a(x)\psi_b(y) = -\psi_b(y)\psi_a(x)$. For such objects the Majorana action is *not* identically zero.

5.3 Solutions of the Dirac equation

To get familiar with spinors and also for later use, we shall now study plane wave solutions of the Dirac equation

$$(m + \gamma^\mu\partial_\mu)\psi = 0. \tag{5.83}$$

Multiplying from the left by $m - \gamma^\mu\partial_\mu$ we encounter

$$m^2 - (\gamma\partial)^2 = m^2 - \frac{1}{2}\{\gamma^\mu, \gamma^\nu\}\partial_\mu\partial_\nu = m^2 - \eta^{\mu\nu}\partial_\mu\partial_\nu, \tag{5.84}$$

so a solution of the Dirac equation is also a solution of the Klein-Gordon equation. The Dirac equation can be put into hamiltonian form[5] by multiplication from the left by β:

$$i\partial_0\psi = (m\beta - i\boldsymbol{\alpha}\nabla)\psi = H_D\psi, \quad H_D \equiv m\beta - i\boldsymbol{\alpha}\nabla, \tag{5.85}$$

where $\alpha_k = \alpha_k^\dagger = i\beta\gamma_k$. The *Dirac hamiltonian* H_D is a hermitian differential operator in the usual inner product,

$$\int d^3x\, \psi_1^\dagger H_D\psi_2 = \left[\int d^3x\, \psi_2^\dagger H_D\psi_1\right]^*, \tag{5.86}$$

[5] The similarity with the Schrödinger wave equation will be commented upon at the end of sect. 6.2

assuming infinite volume or appropriate boundary conditions allowing partial intgration without surface terms. For simplicity we shall assume for a while periodic boundary conditions in an L^3 box. Because of its hermiticity we may expect H_D to possess a complete set of eigenfunctions, which can be found among the eigenfunctions $\exp(i\mathbf{px})$ of ∇:

$$H_D\, w_{\mathbf{p}} e^{i\mathbf{px}} = e^{i\mathbf{px}}\, H_D(\mathbf{p})\, w_{\mathbf{p}}, \quad H_D(\mathbf{p}) \equiv m\beta + \boldsymbol{\alpha}\mathbf{p}. \tag{5.87}$$

Here $H_D(\mathbf{p})$ is just a hermitian 4×4 matrix which we expect to have four independent eigenvectors. Since α_k and β satisfy the anticommutation relations

$$\{\alpha_k, \alpha_l\} = 2\delta_{kl}, \quad \{\alpha_k, \beta\} = 0, \tag{5.88}$$

we have the identity

$$H_D(\mathbf{p})^2 = m^2 + mp_k\{\beta, \alpha_k\} + p_k p_l \frac{1}{2}\{\alpha_k, \alpha_l\} = m^2 + \mathbf{p}^2. \tag{5.89}$$

Hence, the eigenvalues of $H_D(\mathbf{p})$ are equal to $\pm\sqrt{m^2 + \mathbf{p}^2}$, which are presumably two-fold degenerate:

$$H_D(\mathbf{p})\, w_{\mathbf{p}\lambda\epsilon} = \epsilon E_{\mathbf{p}} w_{\mathbf{p}\lambda\epsilon}, \quad \epsilon = \pm, \quad \lambda = \pm, \quad E_{\mathbf{p}} = \sqrt{m^2 + \mathbf{p}^2}. \tag{5.90}$$

Here λ distinguishes the two eigenvectors with the same eigenvalue of H_D. To specify this further we can look for another matrix which commutes with H_D and use its eigenvalues to pin down the w's with the same ϵ. A suitable candidate is the hermitian matrix $\hat{p}\cdot\boldsymbol{S}$ ($S_k = \epsilon_{klm} S_{lm}/2$ is the spin matrix and $\hat{p} = \mathbf{p}/|\mathbf{p}|$ is the unit vector in the direction $\mathbf{p}$). This commutes with $H_D(\mathbf{p})$ and has eigenvalues $\pm 1/2$:

$$[m\beta + \alpha_k p_k, \hat{p}_l S_l] = 0, \quad (\hat{p}_k S_k)^2 = 1/4 \tag{5.91}$$

(which follow from the anticommutation relations (5.88), or by using the chiral representation of the gamma matrices). So we may choose the $w_{\mathbf{p}\lambda\epsilon}$ to satisfy

$$\hat{p}\cdot\boldsymbol{S}\, w_{\mathbf{p}\lambda\epsilon} = \frac{1}{2}\lambda\, w_{\mathbf{p}\lambda\epsilon}. \tag{5.92}$$

Such spinors are called *helicity spinors*. In the zero mass case $[H_D, \gamma_5] = 0$, so then the helicity spinors can be chosen to be also eigenvectors of γ_5:

$$\gamma_5\, w_{\mathbf{p}\lambda\epsilon} = \lambda\epsilon\, w_{\mathbf{p}\lambda\epsilon}, \quad m = 0 \text{ only} \tag{5.93}$$

(it can be seen e.g. in the chiral representation that $\boldsymbol{\alpha} = 2\gamma_5\mathbf{S}$).

Another possibility for specifying the meaning of λ is to start with $\mathbf{p} = 0$ and take the w's to be basis vectors with the usual properties under the action of $\mathbf{S}$:

$$\mathbf{S}\, w_{\mathbf{0}\lambda\epsilon} = \frac{1}{2}\boldsymbol{\sigma}_{\lambda'\lambda}\, w_{\mathbf{0}\lambda'\epsilon}, \tag{5.94}$$

$$S_3\, w_{\mathbf{0}\lambda\epsilon} = \frac{1}{2}\lambda\, w_{\mathbf{0}\lambda\epsilon}. \tag{5.95}$$

Then the $\mathbf{p} \neq 0$ w's can be obtained by boosting $w_{\mathbf{0}\lambda\epsilon}$ with the special Lorentz transformation taking $(m, \mathbf{0})$ into $(E_{\mathbf{p}}, \mathbf{p})$ (cf. Appendix 5.6).

In a Majorana representation where γ^μ is real,

$$H_D(\mathbf{p})^* = -H_D(-\mathbf{p}), \quad \mathbf{S}^* = -\mathbf{S}, \tag{5.96}$$

so we can choose to relate the positive and negative energy spinors by

$$w_{\mathbf{p},\lambda,-} = w^*_{-\mathbf{p},-\lambda,+}. \tag{5.97}$$

In a general representation there is a similar relation involving the so-called charge conjugation matrix, which is explained in the appendix 5.5.

The orthonormality and completeness relations can be chosen to have the (non-covariant) form

$$w^\dagger_{\mathbf{p}\lambda\epsilon} w_{\mathbf{p}\lambda'\epsilon'} = \delta_{\lambda\lambda'}\,\delta_{\epsilon\epsilon'}, \tag{5.98}$$

$$\sum_{\lambda\epsilon} w_{\mathbf{p}\lambda\epsilon} w^\dagger_{\mathbf{p}\lambda\epsilon} = \mathbb{1}. \tag{5.99}$$

Putting things together we have for the time-dependent equation

$$\begin{aligned} i\partial_0\, w_{\mathbf{p}\lambda\epsilon} e^{ipx} &= p^0\, w_{\mathbf{p}\lambda\epsilon} e^{ipx} \\ &= H_D(\mathbf{p})\, w_{\mathbf{p}\lambda\epsilon}\, e^{ipx} = \epsilon E_{\mathbf{p}}\, w_{\mathbf{p}\lambda\epsilon}\, e^{ipx}, \end{aligned} \tag{5.100}$$

so the frequencies are given by $p^0 = \epsilon E_{\mathbf{p}}$.

For infinite volume we shall use covariantly normalized spinors defined by

$$u(p,\lambda) = \sqrt{2E_{\mathbf{p}}}\, w_{\mathbf{p},\lambda,+}, \tag{5.101}$$

$$v(p,\lambda) = \sqrt{2E_{\mathbf{p}}}\, w_{-\mathbf{p},-\lambda,-}, \tag{5.102}$$

Note the minus signs as in (5.97). In a general representation u and v are related by the charge conjugation transformation (5.119). Their properties are the subject of Appendix 5.6 and Problems 1, 2.

5.4 Summary

The 4×4 spinor representation of the Lorentz group is conveniently constructed with the help of Dirac matrices. This representation is reducible into 2×2 blocks, from which it can be seen that rotations are represented in the same way as the spin 1/2 representation in non-relativistic quantum mechanics. Including parity, the 4×4 representation is irreducible.

Spinor fields are fields transforming in the spinor representation. It is possible to choose a real representation of the gamma matrices, called the Majorana representation, with corresponding real spinor fields. Two real Majorana fields make one complex Dirac field.

The action for a free spinor field can be constructed such that it is Poincaré- and parity-invariant. It leads to a field equation called the Dirac equation. The continuous symmetries of the action give rise to Noether 'charges' P^μ and $J^{\mu\nu}$ in the usual way, which are conserved once the field equations are satisfied. The Dirac action has an additional U(1) symmetry, with conserved charge Q and, in case of zero mass parameter, a second U(1) symmetry with charge Q_5. The latter (but not the former) is also present in case of a single Majorana field.

The Dirac equation can be rewritten in a form resembling the Schrödinger equation, which was the motivation for Dirac for inventing this equation. However, in the present context it is just a classical field equation which has nothing to do with quantum mechanics yet.

A plane wave ansatz for solutions to the Dirac equation leads to a (4×4) matrix-eigenvalue equation. The solutions are called polarization spinors (by analogy to polarization vectors in the Maxwell case), which are labelled by the sign of the 'energy' (positive or negative frequency), the 'momentum' (wave vector), and the sign of the helicity (the eigenvalue of the z-component of the spin matrix in the rest frame).

5.5 Appendix: Charge conjugation matrix C

In general, hermiticity properties are preserved under a change of representation. In a real representation these become symmetry properties under transposition. It is possible to express symmetry and reality properties of Dirac matrices in a representation-independent form. For this we need the so-called charge conjugation matrix C.

In any representation there is a unitary antisymmetric matrix C,

$$C^\dagger C = 1, \quad C^T = -C, \tag{5.103}$$

relating γ^μ and $(\gamma^\mu)^T$ according to

$$\gamma^{\mu T} = -C^\dagger \gamma^\mu C. \tag{5.104}$$

In the Majorana representation $\gamma^{\mu T} = \gamma^{\mu \dagger}$ and (cf. (5.18))

$$C = \beta = i\gamma^0 \tag{5.105}$$

$(= -\rho_3\sigma_2)$. In any other representation obtained by a unitary transformation (indicating the Majorana representation by the ˆ for the moment):

$$\begin{aligned} \gamma^\mu &= U\,\hat\gamma^\mu\, U^\dagger, \quad U^\dagger U = 1, && (5.106)\\ \gamma^{\mu T} &= (U\,\hat\gamma^\mu\, U^\dagger)^T = U^*\,\hat\gamma^{\mu T}\, U^T = -U^*\,\hat\beta\hat\gamma^\mu\hat\beta\, U^T && (5.107)\\ &= -U^* U^\dagger\,\beta\gamma^\mu\beta\, U U^T, && (5.108) \end{aligned}$$

and we obtain C in the form

$$C = \beta U U^T \equiv \beta \tilde{C}, \quad \tilde{C} = U U^T. \tag{5.109}$$

We then also have in any representation

$$\begin{aligned} \gamma^{\mu *} &= (\gamma^\mu)^{\dagger T} = -(\beta \gamma^\mu \beta)^T = C^\dagger \beta \gamma^\mu \beta C & (5.110) \\ &= \tilde{C}^\dagger \gamma^\mu \tilde{C}. & (5.111) \end{aligned}$$

In the Majorana representation $\tilde{C} = 1$. In our chiral representation a suitable U is given by

$$U = e^{-i\frac{\pi}{4}\rho_2\sigma_2} = \frac{1}{\sqrt{2}}(1 - i\rho_2\sigma_2) \tag{5.112}$$

(e.g. $U\hat{\gamma}_1 U^\dagger = -U\sigma_3 U^\dagger = -U^2\sigma_2 = i\rho_2\sigma_2\sigma_3 = -\rho_2\sigma_1$), and

$$\tilde{C} = U U^T = e^{-i\frac{\pi}{2}\rho_2\sigma_2} = -i\rho_2\sigma_2, \quad C = \rho_3\sigma_2 = -\gamma^0\gamma^2. \tag{5.113}$$

The charge conjugation matrix is useful for relating D^T to D^{-1},

$$D^T = C^\dagger D^{-1} C, \tag{5.114}$$

and D^* to D,

$$D^* = D^{\dagger T} = C^\dagger \beta \, D \, \beta C. \tag{5.115}$$

For example, $C^\dagger_{ab}\psi_a\psi_b = \psi^T C^\dagger \psi$ is a scalar.

For general complex spinors ψ and $\bar{\psi} = \psi^\dagger\beta$ the so-called charge conjugate spinors $\psi^{(c)}$ and $\bar{\psi}^{(c)}$ are defined as

$$\psi^{(c)} = (\bar{\psi} C)^T = \beta C \psi^* = \tilde{C}\,\psi^*, \quad \bar{\psi}^{(c)} = -(C^\dagger\psi)^T \tag{5.116}$$

(where the formula for $\bar{\psi}^{(c)}$ follows from $\psi^{(c)}$). Under Lorentz transformations $\psi^{(c)}$ transforms like ψ. In particular the antiparticle spinors $v(p,\lambda)$ are defined as the charge conjugates of $u(p,\lambda)$. Denoting again the spinors in the Majorana representation by a ˆ, we have

$$\hat{v}(p,\lambda) = \hat{u}^*(p,\lambda), \quad \text{(Majorana representation)} \tag{5.117}$$

and in a general representation

$$\begin{aligned} u(p,\lambda) &= U\hat{u}(p,\lambda), \quad v(p,\lambda) = U\hat{v}(p,\lambda), & (5.118) \\ v(p,\lambda) &= U\hat{v}(p,\lambda) = U\hat{u}^*(p,\lambda) = UU^T\,u^*(p,\lambda) = \tilde{C}\,u^*(p,\lambda) & \\ &= u^{(c)}(p,\lambda), \quad \text{arbitrary representation.} & (5.119) \end{aligned}$$

The Majorana property of a spinor field (i.e. ψ is real in a Majorana representation) can be expressed in representation independent form: the Majorana field is self (charge) conjugate,

$$\psi^{(c)} = \psi, \quad \bar{\psi}^{(c)} = \bar{\psi}. \tag{5.120}$$

5.6 Appendix: Polarization spinors

The main text introduced polarization spinors as eigenvectors of the Dirac hamiltonian. We make a fresh start here and present a construction based solely on the spinor representation of the Lorentz group. (See also Problem 2.)

A particle at rest transforms under rotations like a two component spinor χ_λ,

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{5.121}$$

From these two-component spinors we make a four-component spinor for a particle at rest in the chiral representation,

$$u(\bar{p}, \lambda) = \sqrt{2m}\,\xi_\lambda, \quad \xi_\lambda = \frac{1}{\sqrt{2}} \begin{pmatrix} \chi_\lambda \\ \chi_\lambda \end{pmatrix}, \tag{5.122}$$

where m is the particle mass and we use the notation

$$\bar{\mathbf{p}} = \mathbf{0}, \;\; \bar{p}^0 = m. \tag{5.123}$$

The curious normalization factor $\sqrt{2m}$ is put in for later convenience. We can characterize $u(\bar{p}, \lambda)$ by the eigenvalues of the two commuting matrices σ_3 and β,

$$\sigma_3\, u(\bar{p}, \lambda) \;=\; \lambda u(\bar{p}, \lambda), \quad \sigma_3\, \xi_\lambda = \lambda \xi_\lambda \tag{5.124}$$

$$\beta u(\bar{p}, \lambda) \;=\; u(\bar{p}, \lambda), \quad \beta \xi_\lambda = \xi_\lambda. \tag{5.125}$$

These relations together with $u(\bar{p}, -) = \frac{1}{2}(\sigma_1 - i\sigma_2)\, u(\bar{p}, +)$ serve to characterize $u(\bar{p}, \lambda)$ also in a general representation.

Polarization spinors $u(p, \lambda)$ for arbitrary momentum p now follow by applying a standard boost ℓ_p which takes $\bar{p}$ into p (cf. (5.39) and Problem 1.10, see also Prob. 2):

$$\ell_p\, \bar{p} \;=\; p, \quad p^0 = \sqrt{m^2 + \mathbf{p}^2} \tag{5.126}$$

$$D_p \;\equiv\; D(\ell_p) = e^{\boldsymbol{\chi}\cdot\boldsymbol{\sigma}\gamma_5/2}, \quad \boldsymbol{\chi} = \chi\, \hat{\mathbf{p}}, \;\; \hat{\mathbf{p}} = \frac{\mathbf{p}}{|\mathbf{p}|}, \quad \tanh\chi = \frac{|\mathbf{p}|}{p^0}. \tag{5.127}$$

Applying this standard boost to $u(\bar{p}, \lambda)$ we get

$$u(p, \lambda) \;=\; D_p\, u(\bar{p}, \lambda) \tag{5.128}$$

$$= \; (\cosh\frac{\chi}{2} + \sinh\frac{\chi}{2}\, \hat{\mathbf{p}} \cdot \boldsymbol{\sigma}\gamma_5) u(\bar{p}, \lambda) \tag{5.129}$$

$$= \; (\sqrt{p^0 + m} + \sqrt{p^0 - m}\, \hat{\mathbf{p}} \cdot \boldsymbol{\sigma}\gamma_5)\xi_\lambda. \tag{5.130}$$

We shall also need conjugate spinors related to $u(p, \lambda)$ by charge conjugation (cf. (5.116)),

$$v(p, \lambda) \;\equiv\; u^{(c)}(p, \lambda) \tag{5.131}$$

$$= \; \beta C\, u(p, \lambda)^* = [\bar{u}(p, \lambda) C]^T, \tag{5.132}$$

$$\bar{v}(p, \lambda) \;=\; \bar{u}^{(c)}(p, \lambda) = -[C^\dagger u(p, \lambda)]^T. \tag{5.133}$$

In the Majorana representation $C = \beta$, giving simply

$$v(p,\lambda) = u(p,\lambda)^*, \quad \text{Majorana rep.} \tag{5.134}$$

Since charge conjugate spinors transform under Lorentz transformations like ordinary spinors we have

$$\begin{aligned} v(p,\lambda) &= D_p\, v(\bar{p},\lambda) & (5.135)\\ &= (\sqrt{p^0+m} + \sqrt{p^0-m}\,\hat{\mathbf{p}}\cdot\boldsymbol{\sigma}\gamma_5)\xi_\lambda^{(c)}, & (5.136)\\ \xi_\lambda^{(c)} &= \beta C\,\xi_\lambda^*. & (5.137) \end{aligned}$$

Furthermore, at rest

$$\begin{aligned} \sigma_3 v(\bar{p},\lambda) &= -\lambda\, v(\bar{p},\lambda), \quad \sigma_3\,\xi_\lambda^{(c)} = -\lambda\,\xi_\lambda^{(c)}, & (5.138)\\ \beta\, v(\bar{p},\lambda) &= -v(\bar{p},\lambda), \quad \beta\,\xi_\lambda^{(c)} = -\xi_\lambda^{(c)}, & (5.139) \end{aligned}$$

and

$$\begin{aligned} \bar{u}(\bar{p},\lambda)\, i\gamma^\mu\, u(\bar{p},\lambda') &= 2\bar{p}^\mu\,\delta_{\lambda\lambda'}, & (5.140)\\ \bar{v}(\bar{p},\lambda)\, i\gamma^\mu\, v(\bar{p},\lambda') &= 2\bar{p}^\mu\,\delta_{\lambda\lambda'}, & (5.141)\\ \bar{u}(\bar{p},\lambda)\, u(\bar{p},\lambda') &= 2m\,\delta_{\lambda\lambda'}, & (5.142)\\ \bar{v}(\bar{p},\lambda)\, v(\bar{p},\lambda') &= -2m\,\delta_{\lambda\lambda'}, & (5.143)\\ \bar{u}(\bar{p},\lambda)\, v(\bar{p},\lambda') &= \bar{v}(\bar{p},\lambda)\, u(\bar{p},\lambda') = 0. & (5.144) \end{aligned}$$

The orthogonality of a $u(\bar{p},\lambda)$ and a $v(\bar{p},\lambda')$ follow from the fact that they are eigenvectors of β with different eigenvalues. From the above follow the relations for general p:

$$\begin{aligned} \bar{u}(p,\lambda)\, i\gamma^\mu\, u(p,\lambda') &= 2p^\mu\,\delta_{\lambda\lambda'}, & (5.145)\\ \bar{v}(p,\lambda)\, i\gamma^\mu\, v(p,\lambda') &= 2p^\mu\,\delta_{\lambda\lambda'}, & (5.146)\\ \bar{u}(p,\lambda)\, u(p,\lambda') &= 2m\,\delta_{\lambda\lambda'}, & (5.147)\\ \bar{v}(p,\lambda)\, v(p,\lambda') &= -2m\,\delta_{\lambda\lambda'}, & (5.148)\\ \bar{u}(p,\lambda)\, v(p,\lambda') &= \bar{v}(p,\lambda)\, u(p,\lambda') = 0. & (5.149) \end{aligned}$$

For example,

$$\begin{aligned} \bar{u}(p,\lambda)\, i\gamma^\mu\, u(p,\lambda') &= \bar{u}(\bar{p},\lambda) D_p^{-1} i\gamma^\mu D_p u(\bar{p},\lambda) = \ell^\mu_{p\ \nu}\, 2\bar{p}^\nu\,\delta_{\lambda\lambda'}\\ &= 2p^\mu\,\delta_{\lambda\lambda'}. & (5.150) \end{aligned}$$

Since $\bar{u}(p,\lambda) i\gamma^0 = u(p,\lambda)^\dagger$ we can interprete (5.145) and (5.146) for $\mu = 0$ as orthogonality relations. The u's are orthogonal to the v's in the sense

$$\begin{aligned} u(p,\lambda)^\dagger\, v(\tilde{p},\lambda') &= u(\bar{p},\lambda)^\dagger D_p^\dagger\, D_{\tilde{p}}\, v(\bar{p},\lambda') & (5.151)\\ &= 0, \quad \tilde{p} \equiv (-\mathbf{p},\, p^0), & (5.152) \end{aligned}$$

where we used $D_p^\dagger = D_p = D_{\tilde{p}}^{-1}$.

Similarly, we have completeness type relations at rest,

$$\begin{aligned}\sum_\lambda u(\bar{p},\lambda)\,\bar{u}(\bar{p},\lambda) &= 2m\sum_\lambda \xi_\lambda\,\xi_\lambda^\dagger = m(1+\beta)\\ &= m - i\gamma^\mu \bar{p}_\mu,\end{aligned} \tag{5.153}$$

and for general momentum

$$\sum_\lambda u(p,\lambda)\,\bar{u}(p,\lambda) = m - i\gamma^\mu p_\mu, \tag{5.154}$$

$$\sum_\lambda v(p,\lambda)\,\bar{v}(p,\lambda) = -m - i\gamma^\mu p_\mu. \tag{5.155}$$

The second relation follows from the first and the definition of $v(p,\lambda)$,

$$\begin{aligned}\sum_\lambda v_a(p,\lambda)\,\bar{v}_b(p,\lambda) &= \sum_\lambda [\bar{u}(p,\lambda)C]_a\,[-C^\dagger u(p,\lambda)]_b = -[C^\dagger\,(m - i\gamma^\mu p_\mu)\,C]_{ba}\\ &= -[C^\dagger\,(m - i\gamma^\mu p_\mu)\,C]^T_{ab} = -[C\,(m - i\gamma^\mu p_\mu)^T\,C^\dagger]_{ab}\\ &= -(m + i\gamma^\mu p_\mu)_{ab}.\end{aligned} \tag{5.156}$$

In the Majorana representation these relations follow more easily from the reality of the γ^μ and $v = u^*$.

Because of the orthogonality relations (5.145), (5.146) and (5.152) the completeness relation in four dimensional spinor space reads

$$\sum_\lambda [u(p,\lambda)u(p,\lambda)^\dagger + v(\tilde{p},\lambda)v(\tilde{p},\lambda)^\dagger] = 2p^0. \tag{5.157}$$

Eqs. (5.125) and (5.139) generalize to arbitrary p as,

$$i\gamma^\mu p_\mu\, u(p,\lambda) = -mu(p,\lambda), \quad i\gamma^\mu p_\mu\, v(p,\lambda) = mv(p,\lambda), \tag{5.158}$$

which turn out to be the free Dirac equation in momentum space.

We conclude this appendix with the zero-mass limit of the polarization spinors (which is also their approximate form for high energies). From eqs. (5.130) and (5.136) we see that for $m \to 0$,

$$u(p,\lambda) \;\to\; \sqrt{|\mathbf{p}|}\,(1 + \hat{\mathbf{p}}\cdot\boldsymbol{\sigma}\gamma_5)\,\xi_\lambda, \tag{5.159}$$

$$v(p,\lambda) \;\to\; \sqrt{|\mathbf{p}|}\,(1 + \hat{\mathbf{p}}\cdot\boldsymbol{\sigma}\gamma_5)\,\xi_\lambda^{(c)}. \tag{5.160}$$

Note that the quantity within parenthesis is essentially a projector: $[(1 + \hat{\mathbf{p}}\cdot\boldsymbol{\sigma}\gamma_5)/2]^2 = (1 + \hat{\mathbf{p}}\cdot\boldsymbol{\sigma}\gamma_5)/2$. Let us change the specification of the ξ_λ such that they become eigenvectors of the helicity matrix

$$\frac{1}{2}\hat{\mathbf{p}}\cdot\boldsymbol{\sigma}, \tag{5.161}$$

with eigenvectors $\lambda/2$. This can be done by a standard rotation which brings the three axis along $\hat{\mathbf{p}}$,

$$\begin{aligned} \xi_\lambda(\theta,\phi) &= e^{-i\phi\frac{1}{2}\sigma_3}\, e^{-i\theta\frac{1}{2}\sigma_2}\, \xi_\lambda, && (5.162)\\ \hat{\mathbf{p}} &= (\sin\theta\cos\phi,\, \sin\theta\sin\phi,\, \cos\theta). && (5.163) \end{aligned}$$

Then λ is the sign of the helicity,

$$\hat{\mathbf{p}}\cdot\boldsymbol{\sigma}\,\xi_\lambda(\theta,\phi) = \lambda\xi_\lambda(\theta,\phi), \quad \hat{\mathbf{p}}\cdot\boldsymbol{\sigma}\,\xi^{(c)}_\lambda(\theta,\phi) = -\lambda\xi^{(c)}_\lambda(\theta,\phi), \tag{5.164}$$

and the helicity is tied to γ_5,

$$\begin{aligned} u(|\mathbf{p}|,\theta,\phi,\lambda) &= \sqrt{|\mathbf{p}|}\,(1+\lambda\gamma_5)\,\xi_\lambda(\theta,\phi), && (5.165)\\ v(|\mathbf{p}|,\theta,\phi,\lambda) &= \sqrt{|\mathbf{p}|}\,(1-\lambda\gamma_5)\,\xi^{(c)}_\lambda(\theta,\phi). && (5.166) \end{aligned}$$

We recognize here chiral projectors $P_{L,R} = (1\mp\gamma_5)/2$. Since γ_5 commutes with $\boldsymbol{\sigma}$ we can choose the helicity ξ's to be eigenvectors of γ_5. The eigenvalue χ of γ_5, which takes values ±1, is called the chirality ('handedness'). We see that for the u-spinors $\chi=\lambda$, whereas for the v-spinors $\chi=-\lambda$. Then a right handed spinor $u_R = P_R\,u$, which in the chiral representation has only the upper two components nonzero, has positive helicity, while a left handed spinor $u_L = P_L\,u$, which in the chiral representation has only the lower two components nonzero, has negative helicity, and vice versa for $v_{R,L}$.

5.7 Appendix: Traces of gamma matrices and other identities

Calculations of unpolarized cross sections or decay rates of processes involving fermions can be carried out in terms of traces over products of gamma matrices. The following identities can be derived (see for example Bjorken & Drell I sect. 7.2 and De Wit & Smith sect. E.4):

$$\begin{aligned} \mathrm{Tr}\,\gamma^{\mu_1}\cdots\gamma^{\mu_n} &= 0, \quad n=\text{odd}, && (5.167)\\ \mathrm{Tr}\,1 &= 4, && (5.168)\\ \mathrm{Tr}\,\gamma^\mu\gamma^\nu &= 4\eta^{\mu\nu}, && (5.169)\\ \mathrm{Tr}\,\gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu &= 4(\eta^{\kappa\lambda}\eta^{\mu\nu} - \eta^{\kappa\mu}\eta^{\lambda\nu} + \eta^{\kappa\nu}\eta^{\lambda\mu}), && (5.170)\\ \mathrm{Tr}\,\gamma_5\gamma^{\mu_1}\cdots\gamma^{\mu_n} &= 0, \qquad n=0,1,2,3, && (5.171)\\ &= 4i\epsilon^{\mu_1\cdots\mu_4}, \quad n=4, && (5.172)\\ \gamma_\mu\gamma^\mu &= 4, && (5.173)\\ \gamma_\mu\gamma^\kappa\gamma^\mu &= -2\gamma^\kappa, && (5.174)\\ \gamma_\mu\gamma^\kappa\gamma^\lambda\gamma^\mu &= 4\eta^{\kappa\lambda}, && (5.175)\\ \gamma_\mu\gamma^\alpha\gamma^\kappa\gamma^\beta\gamma^\mu &= -2\gamma^\beta\gamma^\kappa\gamma^\alpha. && (5.176) \end{aligned}$$

5.8 Problems

1. *Projectors*

   A projector P satisfies the equation $P^2 = P$; its eigenvalues are 0 and 1. Polarization sums over outer products of Dirac spinors produce projectors onto the positive and negative energy subspace of the Dirac Hamiltonian $H_D(\mathbf{p})$. Verify

$$\sum_\lambda w_{\mathbf{p}\lambda\pm} w^\dagger_{\mathbf{p}\lambda\pm} = \frac{E_\mathbf{p} \pm H_D(\mathbf{p})}{2E_\mathbf{p}} \equiv P_\pm, \tag{5.177}$$

$$P_\pm^2 = P_\pm, \quad P_+P_- = 0, \quad P_+ + P_- = 1 \tag{5.178}$$

   A covariant version of these relations is given by

$$\sum_\lambda u(p,\lambda)\bar{u}(p,\lambda) = m - i\gamma p = (m - i\gamma p)^2/2m, \tag{5.179}$$

$$\sum_\lambda v(p,\lambda)\bar{v}(p,\lambda) = -(m + i\gamma p) = -(m + i\gamma p)^2/2m. \tag{5.180}$$

2. *Polarization spinors*

   (See also Appendix 5.6.)

   In the following the derivations can be carried out using only algebraic properties; explicit representations are not required (but e.g. the chiral representation can be useful if one gets stuck).

   Let ℓ_p be the special Lorentz transformation that transforms $\bar{p} \equiv (m, \mathbf{0})$ into p: $\ell^\mu{}_\nu \bar{p}^\nu = p^\mu$, $p^0 = \sqrt{m^2 + \mathbf{p}^2}$. In matrix notation we have

$$\ell_p = e^{i\chi^k S_{0k}}, \quad \chi^k = \chi\,\frac{p^k}{|\mathbf{p}|}, \quad \tanh\chi = \frac{|\mathbf{p}|}{p^0}, \quad \text{defining representation.} \tag{5.181}$$

   This can be seen as follows. First choose $\mathbf{p}$ in the 3-direction. Then the boost $B \equiv \exp(i\chi S_{03})$ reproduces (1.35) (cf. Problem 1.10). To get $\mathbf{p}$ to point in an arbitrary direction, apply a suitable rotation R: $p^k = R^k{}_3|\mathbf{p}|$. Then $p = RB\,\bar{p} = RBR^{-1}\,\bar{p}$, since $\bar{p}$ is invariant under rotations. Hence, $\ell_p = RBR^{-1}$ is a candidate for ℓ_p. That this is correct indeed follows from the fact that the boost generator is a 3-vector under rotations, $R\,S_{0k}\,R^{-1} = R^l{}_k\,S_{0l}$ (cf. Problem 1.11).

   From now on let $S_{\alpha\beta}$ be the generators in the *spinor representation*. Then

$$D_p \equiv D(\ell_p) = e^{i\chi^k S_{0k}}, \quad \text{spinor representation.} \tag{5.182}$$

   We can use the special boost D_p to define 'rest-frame specified' polarization spinors, als follows.

Let $u(\bar{p},\lambda)$ be a spinor 'at rest', characterized by (recall $S_k \equiv \frac{1}{2}\epsilon_{klm}S_{lm}$)

$$\begin{aligned} \beta\, u(\bar{p},\lambda) &= u(\bar{p},\lambda), & (5.183)\\ S_3\, u(\bar{p},\lambda) &= \frac{1}{2}\lambda u(\bar{p},\lambda), & (5.184)\\ u^\dagger(\bar{p},\lambda)u(\bar{p},\lambda') &= 2\bar{p}^0\,\delta_{\lambda\lambda'} \quad (\bar{p}^0 = m). & (5.185)\end{aligned}$$

Note that $[\beta, S_3] = 0$. Define

$$u(p,\lambda) = D_p u(\bar{p},\lambda). \tag{5.186}$$

Verify that

$$\begin{aligned} \bar{u}(p,\lambda)\, u(p,\lambda') &= 2m\,\delta_{\lambda\lambda'}, & (5.187)\\ \bar{u}(p,\lambda)\, i\gamma^\mu\, u(p,\lambda') &= 2p^\mu\,\delta_{\lambda\lambda'}, & (5.188)\\ \sum_\lambda u(\bar{p},\lambda)\bar{u}(\bar{p},\lambda) &= m(1+\beta) = m - i\gamma^\mu\bar{p}_\mu, & (5.189)\\ \sum_\lambda u(p,\lambda)\bar{u}(p,\lambda) &= m - i\gamma^\mu p_\mu, & (5.190)\\ i\gamma^\mu p_\mu u(p,\lambda) &= -m\, u(p,\lambda) & (5.191)\end{aligned}$$

This last equation shows that $u(p,\lambda)\exp(ipx)$ satisfies the Dirac equation and the u are positive-energy polarization spinors.

In a Majorana representation $\gamma_\mu^* = \gamma_\mu$, and the negative-energy spinors v may simply be defined by $v(p,\lambda) = u(p,\lambda)^*$ (cf. (5.97)). We then have (verify)

$$\begin{aligned} \beta u(\bar{p},\lambda)^* &= -u(\bar{p},\lambda)^*, \quad S_3 u(\bar{p},\lambda)^* = -\frac{1}{2}\,\lambda\, u(\bar{p},\lambda)^*, & (5.192)\\ v(p,\lambda) &\equiv u(p,\lambda)^* = D_p^* u(\bar{p},\lambda)^* = D_p u(\bar{p},\lambda)^*, & (5.193)\\ \sum_\lambda v(p,\lambda)\bar{v}(p,\lambda) &= -(m + i\gamma^\mu p_\mu). & (5.194)\end{aligned}$$

Verify also that

$$u^\dagger(p,\lambda)u(\tilde{p},\lambda')^* = 0, \quad \tilde{p} = (p^0, -\mathbf{p}). \tag{5.195}$$

In an arbitrary representation of the γ^μ this becomes (cf. (5.119))

$$u^\dagger(p,\lambda)v(\tilde{p},\lambda') = 0, \quad \tilde{p} = (p^0, -\mathbf{p}). \tag{5.196}$$

This is in accordance with the fact that the positive-energy eigenfunctions of the Dirac hamiltonian H_D are orthogonal to the negative-energy eigenfunctions, since they correspond to different eigenvalues of a hermitian operator:

$$\int d^3x\, [u(p,\lambda)e^{i\mathbf{px}}]^\dagger v(p',\lambda')e^{-i\mathbf{p}'\mathbf{x}} = 0. \tag{5.197}$$

Chapter 6

Quantized spinor fields: fermions

In this chapter we quantize the spinor fields. We shall find that their quanta have the (to be expected) interpretation as spin 1/2 particles. We shall go through the essentials of the arguments leading to the spin-statistics theorem, which states that the spin 1/2 particles described by a relativistic spinor field have to be *fermions*. Functional techniques and the path integral lead to the introduction of *anticommuting 'numbers'*, resulting in elegant analogues of many formulas of the bosonic (commuting) case.

6.1 Canonical quantization: wrong

The real (Majorana) field has presumably half the number of degrees of freedom of the complex (Dirac) field and would therefore be a good starting point for quantization. But below (5.82) we deduced that the action for the real field vanishes identically, so how to proceed? It seems best to face this problem head-on: make the logical jump by assuming that the spinor field is anticommuting. However, let us first try to evade the problem, by noting that it seems to be absent in the case of complex fields. Proceeding this way we are led by the canonical method into a *cul-de-sac*, from which we will recover by going back and changing the quantization rules in accordance with the statistics expected for spin 1/2 particles.

Since the action is linear in the time derivative it is natural to assume it has the $p\dot{q}$ form (1.90). Let us rewrite the Dirac action in terms of the real and imaginary parts of ψ,

$$\psi = \frac{1}{\sqrt{2}}\,(\psi_1 - i\psi_2), \tag{6.1}$$

and use a Majorana representation for γ^μ, such that the symmetries of various objects are explicit. This gives[1]

$$S = \int dt\, L, \quad L = \int d^3x\, \psi_1^T \dot{\psi}_2 - H, \tag{6.2}$$

$$H = \int d^3x\, \psi^\dagger \overset{\leftrightarrow}{H}_D \psi = \int d^3x\, i\psi_2^T \overset{\leftrightarrow}{H}_D \psi_1, \tag{6.3}$$

$$\overset{\leftrightarrow}{H}_D = m\beta - i\boldsymbol{\alpha}\cdot \overset{\leftrightarrow}{\nabla}, \tag{6.4}$$

Recall that $\beta = -\beta^* = -\beta^T$ and $\alpha_k = \alpha_k^T = \alpha_k^*$ in a Majorana representation, so L and H are real. The functional H is evidently the hamiltonian (cf. (1.90), which is in accordance with (5.75),

$$H = \int d^3x\, T^{00}. \tag{6.5}$$

The canonical variables are ψ_{2a} and $\pi_{2a} = \delta L/\delta\dot{\psi}_{2a} = \psi_{1a}$.

We now quantize the theory by requiring the fields to be operators in Hilbert space with the canonical commutation relations ($\pi_{2a} = \psi_{1a}$)

$$[\psi_{2a}(\mathbf{x}), \psi_{1b}(\mathbf{y})] = i\delta_{ab}\delta(\mathbf{x}-\mathbf{y}), \quad [\psi_{2a}(\mathbf{x}), \psi_{2b}(\mathbf{y})] = [\psi_{1a}(\mathbf{x}), \psi_{1b}(\mathbf{y})] = 0. \tag{6.6}$$

Going back to the original Dirac field we get the commutation relations

$$[\psi_a(\mathbf{x}), \psi_b^\dagger(\mathbf{y})] = \delta_{ab}\delta(\mathbf{x}-\mathbf{y}), \quad [\psi_a(\mathbf{x}), \psi_b(\mathbf{y})] = [\psi_a^\dagger(\mathbf{x}), \psi_b^\dagger(\mathbf{y})] = 0. \tag{6.7}$$

Having derived these relations in a Majorana representation for γ^μ, they remain valid under unitary transformations and from now on we shall allow arbitrary representations, unless otherwise stated.

Of immediate interest are the eigenstates and eigenvalues of the hamiltonian and of the momentum operator

$$\boldsymbol{P} = \int d^3x\, \psi^\dagger(-i\nabla)\psi. \tag{6.8}$$

Note that, although ψ and $\psi^\dagger$ do not commute, there is no ambiguity in the ordering of these operators since $\nabla\delta(\mathbf{x}-\mathbf{y})$ is an antisymmetric 'matrix'. For the hamiltonian the order also does not matter because α_k and β are traceless matrices. We shall start out again with a finite L^3 box with periodic boundary conditions, such that we can freely do partial integration and do not encounter Dirac delta functions with vanishing argument. We expand the fields in a complete set of eigenfunctions of H_D and $-i\nabla$, say at time zero,

$$\psi(\mathbf{x}) = \sum_{\mathbf{p}\lambda\epsilon} a_{\mathbf{p}\lambda\epsilon}\, w_{\mathbf{p}\lambda\epsilon}\, \frac{e^{i\mathbf{px}}}{\sqrt{L^3}}, \tag{6.9}$$

$$a_{\mathbf{p}\lambda\epsilon} = \int d^3x\, \frac{e^{-i\mathbf{px}}}{\sqrt{L^3}}\, w_{\mathbf{p}\lambda\epsilon}^\dagger \psi(\mathbf{x}), \tag{6.10}$$

[1] We made a partial integration to convert all time derivatives to ψ_2. Otherwise we would get $L = \int d^3x\, \frac{1}{2}(\psi_1\dot{\psi}_2 - \psi_2\dot{\psi}_1) - H$, which would lead to canonical momenta $\pi_1 = \delta L/\delta\dot{\psi}_1 = -\psi_2/2$, $\pi_2 = \psi_1/2$, giving conflicting canonical commutation relations like $0 = [\psi_1(\mathbf{x}), \psi_2(\mathbf{y})] = [\psi_1(\mathbf{x}), -2\pi_1(\mathbf{y})] = -2i\delta(\mathbf{x}-\mathbf{y})$.

with the conjugate expressions obtained by hermitian conjugation. The 'expansion coefficients' have the commutation relations of the creation and annihilation operators of a harmonic oscillator for each mode,

$$[a_{\mathbf{p}\lambda\epsilon}, a^\dagger_{\mathbf{p}'\lambda'\epsilon'}] = \delta_{\mathbf{p}\mathbf{p}'}\,\delta_{\lambda\lambda'}\,\delta_{\epsilon\epsilon'}, \tag{6.11}$$

$$[a_{\mathbf{p}\lambda\epsilon}, a_{\mathbf{p}'\lambda'\epsilon'}] = [a^\dagger_{\mathbf{p}\lambda\epsilon}, a^\dagger_{\mathbf{p}'\lambda'\epsilon'}] = 0. \tag{6.12}$$

They are represented in Hilbert space as creation and annihilation operators indeed, by assuming an empty state $|\emptyset\rangle$ with the property

$$a_{\mathbf{p}\lambda\epsilon}|\emptyset\rangle = 0, \tag{6.13}$$

and letting the $a^\dagger_{\mathbf{p}\lambda\epsilon}$ act on $|\emptyset\rangle$ any number of times. The hamiltonian and momentum operator take the following form upon inserting the expansions of $\psi(\mathbf{x})$ and $\psi^\dagger(\mathbf{x})$,

$$H = \sum_{\mathbf{p}\lambda}\left[a^\dagger_{\mathbf{p}\lambda+}a_{\mathbf{p}\lambda+} - a^\dagger_{\mathbf{p}\lambda-}a_{\mathbf{p}\lambda-}\right] E_{\mathbf{p}}, \tag{6.14}$$

$$\mathbf{P} = \sum_{\mathbf{p}\lambda\epsilon} a^\dagger_{\mathbf{p}\lambda\epsilon}a_{\mathbf{p}\lambda\epsilon}\,\mathbf{p}, \tag{6.15}$$

where $E_{\mathbf{p}} = \sqrt{m^2 + \mathbf{p}^2}$ is the eigenvalue of the Dirac hamiltonian found in (5.90).

We now see a disastrous problem. By application of arbitrarily many negative-energy creation operators $a^\dagger_{\mathbf{p}\lambda-}$ on the empty state $|\emptyset\rangle$, we get a state with arbitrarily negative energy: the hamiltonian has no state with lowest energy, it has no ground state, no vacuum state! This also excludes Statistical Physics in terms of a canonical partition function $Z = \mathrm{Tr}\, e^{-H/T}$, with positive temperature T. Lacking a ground state, the theory developed so far cannot have a physical interpretation and has to be abandoned.

The reason is not the infinite number of degrees of freedom in field theory. Imposing a cutoff $|\mathbf{p}| < \Lambda$ leads to the same problem (even keeping only a single momentum mode). The basic reason is the fact that an operator like H_D has eigenvalues coming in pairs $\pm E$. Even the classical hamiltonian is unbounded from below if the ψ and ψ^* are ordinary complex numbers.

6.2 Canonical quantization: right

The problem was found and resolved by Dirac, not in present quantum field-theoretic setting, but in a formulation starting from a single particle. We shall re-phrase his resolution in the present field theory context.

First we note that the states

$$|\mathbf{p}\lambda+\rangle = a^\dagger_{\mathbf{p}\lambda+}|\emptyset\rangle \tag{6.16}$$

are like particle states with energy $E_{\mathbf{p}} = \sqrt{m^2 + \mathbf{p}^2}$ and momentum $\mathbf{p}$, hence their mass is m.[2] We expect these particles to have spin 1/2. To show this, consider the angular momentum operator $J_k = \epsilon_{klm} J_{lm}/2$:

$$\mathbf{J} = \int d^3x\, \psi^\dagger(-i\mathbf{x} \times \nabla + \mathbf{S})\psi \tag{6.17}$$

(there is no operator ordering ambiguity because $\mathrm{Tr}\,\mathbf{S} = 0$). It's commutator with $\psi^\dagger$ is given by

$$[\mathbf{J}, \psi^\dagger(\mathbf{x})] = +i\mathbf{x} \times \nabla\psi^\dagger(\mathbf{x}) + \psi^\dagger(\mathbf{x})\mathbf{S}, \tag{6.18}$$

from which we infer, using the conjugate of (6.10),

$$[\mathbf{J}, a^\dagger_{\mathbf{p}\lambda\epsilon}] = \int d^3x\, \psi^\dagger(\mathbf{x})\, (-i\mathbf{x} \times \nabla + \mathbf{S})\; w_{\mathbf{p}\lambda\epsilon}\, \frac{e^{i\mathbf{p}\mathbf{x}}}{\sqrt{L^3}}. \tag{6.19}$$

Since ∇ acting on the exponential gives $i\mathbf{p}$ is it is simplest here to consider $\mathbf{p} = 0$ and the spinor specification (5.94), which gives

$$[\mathbf{J}, a^\dagger_{\mathbf{0}\lambda\epsilon}] = \frac{1}{2}\boldsymbol{\sigma}_{\lambda'\lambda}\, a^\dagger_{\mathbf{0}\lambda'\epsilon}. \tag{6.20}$$

Using this result we calculate

$$\mathbf{J}|\mathbf{0}\lambda+\rangle = [\mathbf{J}, a^\dagger_{\mathbf{0}\lambda+}]|\emptyset\rangle = \frac{1}{2}\boldsymbol{\sigma}_{\lambda'\lambda}|\mathbf{0}\lambda'+\rangle, \tag{6.21}$$

and we conclude that the particles have spin 1/2. Aternatively, we can concentrate on the *helicity operator* $\hat{p} \cdot \mathbf{J}$, use the helicity spinors (5.92) and reach the same conclusion.

Because the particles have spin 1/2 we expect them to be *fermions* and the states to be antisymmetric under interchange of labels,

$$|\mathbf{p}_1\lambda_1+, \mathbf{p}_2\lambda_2+\rangle = -|\mathbf{p}_2\lambda_2+, \mathbf{p}_1\lambda_1+\rangle. \tag{6.22}$$

But this requires $a^\dagger_{\mathbf{p}_1\lambda_1+} a^\dagger_{\mathbf{p}_2\lambda_2+} = -a^\dagger_{\mathbf{p}_2\lambda_2+} a^\dagger_{\mathbf{p}_1\lambda_1+}$ i.e. the creation operators have to anticommute instead of commute. We therefore discard the canonical commutation relations and assume instead the so-called canonical *anticommutation relations*

$$\{\psi_a(\mathbf{x}), \psi_b^\dagger(\mathbf{y})\} = \delta_{ab}\delta(\mathbf{x} - \mathbf{y}), \quad \{\psi_a(\mathbf{x}), \psi_b(\mathbf{y})\} = \{\psi_a^\dagger(\mathbf{x}), \psi_b^\dagger(\mathbf{y})\} = 0, \tag{6.23}$$

which imply

$$\{a_{\mathbf{p}\lambda\epsilon}, a^\dagger_{\mathbf{p}'\lambda'\epsilon'}\} = \delta_{\mathbf{p}\mathbf{p}'}\,\delta_{\lambda\lambda'}\,\delta_{\epsilon\epsilon'}, \tag{6.24}$$

$$\{a_{\mathbf{p}\lambda\epsilon}, a_{\mathbf{p}'\lambda'\epsilon'}\} = \{a^\dagger_{\mathbf{p}\lambda\epsilon}, a^\dagger_{\mathbf{p}'\lambda'\epsilon'}\} = 0. \tag{6.25}$$

[2]Strickly speaking, the mass is $|m|$. We shall assume $m > 0$, the case $m < 0$ can be shown to be equivalent.

Let us pause for a moment and consider a single mode $\mathbf{p}, \lambda, \epsilon$. Suppressing the labels we have

$$\{a, a^\dagger\} = 1, \quad a^2 = a^{\dagger 2} = 0. \tag{6.26}$$

This can be represented as

$$a|\emptyset\rangle = 0, \quad a^\dagger|\emptyset\rangle = |1\rangle, \quad a|1\rangle = |\emptyset\rangle, \quad a^\dagger|1\rangle = 0, \tag{6.27}$$

and we see that we have a two-dimensional Hilbert space. The operators are nilpotent (in accordance with the Pauli principle) and their matrix elements are given by $a_{00} = a_{10} = a_{11} = 0$, $a_{01} = 1$, and similar for $a^\dagger$, or

$$a \to \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad a^\dagger \to \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{6.28}$$

Generalizing to N different a and $a^\dagger$, we get the tensor product of N two dimensional Hilbert spaces, which has dimension 2^N. However, we can re-interpret the state $|1\rangle$ as the empty state. Then $|\emptyset\rangle = a|1\rangle$ is re-interpreted as the filled state and the role of the creation and annihilation operators are interchanged. This can be made explicit by writing

$$d = a^\dagger, \quad d^\dagger = a, \quad \{d, d^\dagger\} = 1, \quad d|1\rangle = 0, \quad d^\dagger|1\rangle = |\emptyset\rangle. \tag{6.29}$$

Consider again the hamiltonian (6.14), this time with the anticommutation rules for the a's and $a^\dagger$'s. Assuming for the moment a cutoff $|\mathbf{p}| < \Lambda$ on the number of modes we have a well defined ground state $|0\rangle$ given by

$$\begin{aligned} |0\rangle &= \prod_{\mathbf{p}\lambda} a^\dagger_{\mathbf{p}\lambda-}\,|\emptyset\rangle, && (6.30)\\ a_{\mathbf{p}\lambda+}|0\rangle &= 0, \quad a^\dagger_{\mathbf{p}\lambda-}|0\rangle = 0, && (6.31)\\ H|0\rangle &= E_0|0\rangle, \quad E_0 = -\sum_{\mathbf{p}\lambda} E_{\mathbf{p}}. && (6.32) \end{aligned}$$

All the negative energy states are filled: this state is called the *Dirac sea*. A 'hole in the sea', $a_{\mathbf{p}\lambda-}|0\rangle$, means an excitation of the ground state with positive energy compared to the ground state. This is interpreted as a different type of particle.

We can represent the situation clearer by a change of notation, writing

$$\begin{aligned} b_{\mathbf{p}\lambda} &= a_{\mathbf{p}\lambda+}, \quad b^\dagger_{\mathbf{p}\lambda} = a^\dagger_{\mathbf{p}\lambda+}, \quad \{b_{\mathbf{p}\lambda}, b^\dagger_{\mathbf{p}'\lambda'}\} = \delta_{\mathbf{p}\mathbf{p}'}\,\delta_{\lambda\lambda'}, && (6.33)\\ d_{\mathbf{p}\lambda} &= a^\dagger_{-\mathbf{p},-\lambda,-}, \quad d^\dagger_{\mathbf{p}\lambda} = a_{-\mathbf{p},-\lambda,-}, \quad \{d_{\mathbf{p}\lambda}, d^\dagger_{\mathbf{p}'\lambda'}\} = \delta_{\mathbf{p}\mathbf{p}'}\,\delta_{\lambda\lambda'}, && (6.34)\\ \{b, b'\} &= \{d, d'\} = \{b, d'\} = \{b, d'^\dagger\} = \ldots = 0. && (6.35) \end{aligned}$$

Then the energy and momentum operators are given by

$$\begin{aligned} H &= \sum_{\mathbf{p}\lambda} \left(b^\dagger_{\mathbf{p}\lambda} b_{\mathbf{p}\lambda} + d^\dagger_{\mathbf{p}\lambda} d_{\mathbf{p}\lambda}\right) E_{\mathbf{p}} + E_0, && (6.36)\\ \mathbf{P} &= \sum_{\mathbf{p}\lambda} \left(b^\dagger_{\mathbf{p}\lambda} b_{\mathbf{p}\lambda} + d^\dagger_{\mathbf{p}\lambda} d_{\mathbf{p}\lambda}\right) \mathbf{p}. && (6.37) \end{aligned}$$

Note the $d^\dagger d$ order with positive coefficients, for H as well as for $\mathbf{P}$. We see that $b^\dagger_{\mathbf{p}\lambda}$ and $d^\dagger_{\mathbf{p}\lambda}$ each create (different) particles with the same energy $E_\mathbf{p}$ and momentum $\mathbf{p}$ out of the vacuum $|0\rangle$, which satisfies

$$b_{\mathbf{p}\lambda}|0\rangle = d_{\mathbf{p}\lambda}|0\rangle = 0. \tag{6.38}$$

It is also straightforward to check along the lines leading to (6.21) that the b and d particles have identical spin properties (thanks to the $-\lambda$ in (6.34).

As the above expressions for H and $\mathbf{P}$ show, the particle states are now obtained by application of $b^\dagger$ and $d^\dagger$ to the ground state. e.g.

$$|\mathbf{p}\lambda\rangle = b^\dagger_{\mathbf{p}\lambda}\,|0\rangle, ; \quad |\overline{\mathbf{p}\lambda}\rangle = d^\dagger_{\mathbf{p}\lambda}\,|0\rangle, \tag{6.39}$$

$$|\mathbf{p}_1\lambda_1, \mathbf{p}_2\lambda_2\rangle = -|\mathbf{p}_2\lambda_1, \mathbf{p}_1\lambda_1\rangle = b^\dagger_{\mathbf{p}_1\lambda_1}\, b^\dagger_{\mathbf{p}_2\lambda_2}\,|0\rangle, \tag{6.40}$$

etc. We distinguish the d states with a 'bar' over its labels.

Consider next the Noether charge Q found in (5.79). It's operator form has an ordering ambiguity which we resolve by (anti)symmetrization:

$$Q \equiv \int d^3x\, \frac{1}{2}\left(\psi^\dagger_a\psi_a - \psi_a\psi^\dagger_a\right). \tag{6.41}$$

In terms of the creation and annihilation operators Q is given by

$$Q = \sum_{\mathbf{p}\lambda}\left(b^\dagger_{\mathbf{p}\lambda}b_{\mathbf{p}\lambda} - d^\dagger_{\mathbf{p}\lambda}d_{\mathbf{p}\lambda}\right). \tag{6.42}$$

It just counts the number of b-particles minus the number of d-particles, with the same momentum and spin.

In the context of QED, the Noether current j^μ turns out to be the electromagnetic current and Q is the charge in units of the elementary charge e. The b- and d-particles have opposite charge. For this reason the d-particles are called *antiparticles* (and the b's just 'particles').

In the zero mass case the operator Q_5 (cf. (5.81)) is also conserved. Using helicity spinors we have

$$Q_5 = \int d^3x\, \psi^\dagger\gamma_5\psi = \sum_{\mathbf{p}\lambda}\left(b^\dagger_{\mathbf{p}\lambda}b_{\mathbf{p}\lambda} - d^\dagger_{\mathbf{p}\lambda}d_{\mathbf{p}\lambda}\right)\,\lambda. \tag{6.43}$$

So Q_5 counts the number right handed ($\lambda = +$) particles minus the number of left handed ($\lambda = -$) particles, plus the number of left handed antiparticles minus the number of right handed antiparticles.

In terms of the b's and d's the expansion of the Dirac field takes the form

$$\psi(\mathbf{x}) = \sum_{\mathbf{p}\lambda}\left(b_{\mathbf{p}\lambda}\, w_{\mathbf{p}\lambda+}\,\frac{e^{i\mathbf{p}\mathbf{x}}}{\sqrt{L^3}} + d^\dagger_{\mathbf{p}\lambda}\, w_{-\mathbf{p},-\lambda-}\,\frac{e^{-i\mathbf{p}\mathbf{x}}}{\sqrt{L^3}}\right). \tag{6.44}$$

For infinite volume we use the covariantly normalized spinors defined by

$$u(p,\lambda) = \sqrt{2p^0}\, w_{\mathbf{p},\lambda,+}, \quad p^0 = \sqrt{m^2+\mathbf{p}^2}, \tag{6.45}$$

$$v(p,\lambda) = \sqrt{2p^0}\, w_{-\mathbf{p},-\lambda,-}, \tag{6.46}$$

and corresponding creation and annihilation operators

$$b(p,\lambda) = \sqrt{2p^0L^3}\, b_{\mathbf{p}\lambda}, \quad d(p,\lambda) = \sqrt{2p^0L^3}\, d_{\mathbf{p}\lambda}, \tag{6.47}$$

etc., such that,

$$\{b(p,\lambda), b^\dagger(p',\lambda')\} = 2p^0(2\pi)^3\delta(\mathbf{p}-\mathbf{p}')\,\delta_{\lambda\lambda'}, \tag{6.48}$$

etc. Including time dependence we have for the free field[3]

$$\psi(x) = \sum_\lambda \int d\omega_p \left[b(p,\lambda)\,u(p,\lambda)\,e^{ipx} + d^\dagger(p,\lambda)\,v(p,\lambda)\,e^{-ipx}\right], \tag{6.49}$$

$$\bar\psi(x) = \sum_\lambda \int d\omega_p \left[b^\dagger(p,\lambda)\,\bar u(p,\lambda)\,e^{-ipx} + d(p,\lambda)\,\bar v(p,\lambda)\,e^{ipx}\right]. \tag{6.50}$$

Subtracting the vacuum energy, we get for the energy-momentum operators

$$P^\mu = \sum_\lambda \int d\omega_p \left[b^\dagger(p,\lambda)b(p,\lambda) + d^\dagger(p,\lambda)d(p,\lambda)\right] p^\mu, \tag{6.51}$$

and the charge

$$Q = \sum_\lambda \int d\omega_p \left[b^\dagger(p,\lambda)b(p,\lambda) - d^\dagger(p,\lambda)d(p,\lambda)\right]. \tag{6.52}$$

The anticommutation relations have led to a sound physical interpretation of the Dirac field. Only if the particles described by the field are fermions did we get a satisfactory situation. Another approach to the relation between spin and statistics, that stresses the importance of locality, is described in the appendix 6.8, whereas in the next section we also get the same results from the path integral.

It is instructive to rewrite these results in terms of hermitian Majorana fields $\psi_{1,2}$ related to ψ by $\psi = (\psi_1 - i\psi_2)/\sqrt{2}$ (using real γ^μ). We shall be very brief. A Dirac field is simply equivalent to two Majorana fields. For a single Majorana field $\psi = \psi^\dagger$, and we note:

$$v(p,\lambda) = u^*(p,\lambda), \quad \text{(cf. (5.97,5.102))}, \tag{6.53}$$

$$\psi(x) = \sum_\lambda \int d\omega_p \left[b(p,\lambda)\,u(p,\lambda)\,e^{ipx} + b^\dagger(p,\lambda)\,u^*(p,\lambda)\,e^{-ipx}\right], \tag{6.54}$$

$$j^\mu = 0, \tag{6.55}$$

$$j_5^\mu = \frac{1}{2}\psi^T\beta i\gamma^\mu\gamma_5\psi \neq 0. \tag{6.56}$$

[3]Because of the analogy with the plane wave expansion of the free electromagnetic field the u's and v's are called polarization spinors.

We end this section with some comments. Dirac formulated his equation for a Schrödinger quantum-mechanical wave function of a single electron: his ψ was *not* a classical field. Interpreting ψ in eq. (5.85) as a Schrödinger wave function means conceptually something very different from a field equation, and it led to several difficulties, one of which was how to interpret the negative energy solutions. Dirac's solution in postulating a filled sea of fermionic states led to the prediction of antiparticles as holes in the sea. However, this gave additional difficulties since it brought the theory outside the single-particle framework. The quantum field theory version does not have these difficulties. Since it looked to the field theory pioneers like a quantization of a single particle Schrödinger wave function ψ, which was already supposed to describe a quantum theory, the name "second quantization" got into use as a name for field theory.

We like to mention again that antiparticles are *not* predicted by mere relativistic field theory. There is no reason for the doubling of degrees of freedom embodied in the Dirac field, a single Majorana field is the simplest possibility and (as for a single scalar field) it implies no antiparticles. However, we shall see later in chapter 8 that the coupling to the electromagnetic field *does* require the Dirac field (or two Majorana fields), and hence predicts the existence of antiparticles. But in non-relativistic quantum electrodynamics antiparticles are not needed, as we review in section 6.4.

6.3 Path integral quantization

We shall now follow an independent approach to the quantization of spinor fields, using the notion of a path integral. As mentioned below (5.82), the action vanishes identically for a single real-valued Majorana field. So if a path integral with integrand $\exp(iS)$ is going to make sense, the Majorana field $\psi_a(x)$ cannot consist of ordinary real numbers. For the action not to vanish, the variable ψ appearing in the path integral has to be *anticommuting*:

$$\psi_a(x)\psi_b(y) = -\psi_b(y)\psi_a(x), \quad [\psi_a(x)]^2 = 0. \tag{6.57}$$

This is what we going to assume from now on. We enlarge the system of complex numbers with such anticommuting objects. They are called 'Grassmann variables', somewhat inappropriately, because they do not vary at all like ordinary variables. Instead they can be given meaning as generators of a Grassmann algebra. They are also called 'anticommuting numbers', also somewhat inappropriately, because they are not numbers. We shall gradually develop a reasonable mathematical framework for them. To start with, we assume the rule

$$[\psi_a(x)\psi_b(y)]^* = \psi_b^*(y)\psi_a^*(x) \tag{6.58}$$

which we used in (5.63) to get a real Lagrange density. For ordinary complex numbers the above relation is of course trivial, but for noncommuting objects we have to define it. The rule above is similar to the operator relation $(AB)^\dagger = B^\dagger A^\dagger$. We shall also assume that the anticommuting numbers commute with ordirary numbers, and hence also with bosonic field operators, while they anticommute with fermionic operators:[4]

$$\psi_a(x)\hat{\psi}_b(y) = -\hat{\psi}_b(y)\psi_a(x), \tag{6.59}$$

As a first consequence we see that the Dirac field is equivalent to two Majorana fields indeed, since, writing $\psi = (\psi_1 - i\psi_2)/\sqrt{2}$, $\bar{\psi} = (\bar{\psi}_1 + i\bar{\psi}_2)/\sqrt{2}$, and using (6.58), the Dirac action is just the sum of two identical Majorana actions:

$$\begin{aligned} S &= -\int d^4x\, \bar{\psi}(m + \gamma^\mu \overleftrightarrow{\partial}_\mu)\psi \\ &= \int d^4x \left[\frac{1}{2}\psi_1^T \beta(m + \gamma^\mu \overleftrightarrow{\partial}_\mu)\psi_1 + \frac{1}{2}\psi_2^T \beta(m + \gamma^\mu \overleftrightarrow{\partial}_\mu)\psi_2\right] \end{aligned} \tag{6.60}$$

and similar for the hamiltonian. Note that this is very different from (6.2,6.3). It would now be logically clearest to continue with the case of a single Majorana field, but for conciseness we shall instead go directly to the doubled case of the Dirac field.

We proceed heuristically and assume the following relations by analogy to the scalar field,

$$Z[\eta, \bar{\eta}] = \int [d\bar{\psi} d\psi]\, e^{iS[\psi,\bar{\psi}] + i\int d^4x\, (\bar{\eta}\psi + \bar{\psi}\eta)}, \tag{6.61}$$

$$\frac{Z[\eta, \bar{\eta}]}{Z[0,0]} = \langle 0|T e^{i\int d^4x\, (\bar{\eta}\hat{\psi} + \hat{\bar{\psi}}\eta)}|0\rangle, \tag{6.62}$$

in which S is the Dirac action, η and $\bar{\eta}$ are anticommuting external sources, and $|0\rangle$ is the ground state (vacuum). The meaning of the symbols $\int[d\bar{\psi}d\psi]$ has to be specified further. For now we make the crucial assumption that the 'measure' $[d\bar{\psi}d\psi]$ is translation invariant, such that

$$\int [d\bar{\psi}d\psi] F[\psi, \bar{\psi}] = \int [d\bar{\psi}d\psi] F[\psi + \xi, \bar{\psi} + \bar{\xi}], \tag{6.63}$$

for any anticommuting $\xi, \bar{\xi}$ and functional F. Then, using the shorthand notation[5]

$$\int d^4x\, \bar{\psi}(m + \gamma^\mu \overleftrightarrow{\partial}_\mu)\psi \equiv \bar{\psi} G^{-1} \psi, \tag{6.64}$$

$$G^{-1}(x,y) = (m + \gamma^\mu \partial_\mu)\delta^4(x-y), \tag{6.65}$$

[4]In this section (and whenever needed for clarity) we denote operators in Hilbert space by a caret. So $\hat{\psi}_b(x)$ denotes the quantized fermion-operator field and $\psi_a(x)$ denotes an anticommuting number.

[5]Notice that we are freely making partial integrations without picking up boundary terms, implying appropriate boundary conditions in space and time (as in the bosonic case).

and making the transformation of variables

$$\psi = \psi' + G\eta, \quad \bar{\psi} = \bar{\psi}' + \bar{\eta}G, \quad G^{-1}G = 1, \tag{6.66}$$

such that

$$-\bar{\psi}G^{-1}\psi + \bar{\eta}\psi + \bar{\psi}\eta = -\bar{\psi}'G^{-1}\psi' + \bar{\eta}G\eta, \tag{6.67}$$

we have the result

$$Z[\eta, \bar{\eta}] = Z[0,0]\, e^{i\bar{\eta}G\eta} = Z[0,0]\, e^{i\int d^4x\, d^4y\, \bar{\eta}(x)G(x,y)\eta(y)}. \tag{6.68}$$

Comparing coefficients of multinomials in η and $\bar{\eta}$ we will get the Wick formula for fermions, relating vacuum expectation values of time ordered products of field operators to sums of products of Green functions G. From the properties of the latter we can then infer once more the interpretation in terms of fermions.

Consider the first few terms in η and $\bar{\eta}$,

$$\begin{aligned} \frac{Z[\eta,\bar{\eta}]}{Z[0,0]} &= 1 + i\int d^4x\, d^4y\, \bar{\eta}(x)G(x,y)\eta(y) + \cdots \\ &= 1 + i\langle 0|\int d^4x\, \bar{\eta}(x)\hat{\psi}(x)|0\rangle + i\langle 0|\int d^4x\, \hat{\bar{\psi}}(x)\eta(x)|0\rangle \\ &\quad + \frac{i^2}{2}\langle 0|\left[\int d^4x\, \bar{\eta}(x)\hat{\psi}(x)\right]^2|0\rangle + \frac{i^2}{2}\langle 0|\left[\int d^4x\, \hat{\bar{\psi}}(x)\eta(x)\right]^2|0\rangle \\ &\quad + i^2\int d^4x \int d^4y\, \langle 0|T\bar{\eta}(x)\hat{\psi}(x)\,\hat{\bar{\psi}}(y)\eta(y)|0\rangle + \cdots, \end{aligned} \tag{6.69}$$

where we kept the combinations $\bar{\eta}\psi$ and $\bar{\psi}\eta$ in first instance together since, as a product of two anticommuting objects, they are of the commuting type. Note that this implies that even products such as $\bar{\eta}\hat{\psi}$ and $\hat{\bar{\psi}}\eta$ can be freely interchanged under the time-ordering operator, i.e. $T\bar{\eta}(x)\hat{\psi}(x)\,\hat{\bar{\psi}}(y)\eta(y) = T\hat{\bar{\psi}}(y)\eta(y)\,\bar{\eta}(x)\hat{\psi}(x)$, as for the scalar-field case. Comparing coefficients of $\eta, \bar{\eta}$ we conclude that

$$\langle 0|\hat{\psi}|0\rangle = \langle 0|\hat{\bar{\psi}}|0\rangle = 0. \tag{6.70}$$

Comparing the bilinear $\bar{\eta}(x)\eta(y)$ terms we have to be careful about signs: for $x^0 > y^0$,

$$T\bar{\eta}_a(x)\hat{\psi}_a(x)\hat{\bar{\psi}}_b(y)\eta_b(y) = \bar{\eta}_a(x)\hat{\psi}_a(x)\hat{\bar{\psi}}_b(y)\eta_b(y) = \bar{\eta}_a(x)\eta_b(y)\hat{\psi}_a(x)\hat{\bar{\psi}}_b(y) \tag{6.71}$$

whereas for $x^0 < y^0$,

$$T\bar{\eta}_a(x)\hat{\psi}_a(x)\hat{\bar{\psi}}_b(y)\eta_b(y) = \hat{\bar{\psi}}_b(y)\eta_b(y)\,\bar{\eta}_a(x)\hat{\psi}_a(x) = -\bar{\eta}_a(x)\eta_b(y)\hat{\bar{\psi}}_b(y)\hat{\psi}_a(x). \tag{6.72}$$

Since the product $\bar{\eta}_a(x)\eta_b(y)$ is commuting and can be taken in front of the T symbol we conclude that for fermion fields

$$\begin{aligned} T\hat{\psi}_a(x)\hat{\bar{\psi}}_b(y) &= \hat{\psi}_a(x)\hat{\bar{\psi}}_b(y), \quad x^0 > y^0 \\ &= -\hat{\bar{\psi}}_b(y)\hat{\psi}_a(x), \quad x^0 < y^0. \end{aligned} \tag{6.73}$$

Comparison of the bilinear terms in $\eta, \bar{\eta}$ gives

$$\langle 0|T\hat{\psi}_a(x)\hat{\bar{\psi}}_b(y)|0\rangle = -iG_{ab}(x,y). \tag{6.74}$$

The generalization to more fields is given by, in shorthand notation,

$$\begin{aligned}
\langle 0|T\hat{\psi}_1\hat{\bar{\psi}}_1\hat{\psi}_2\hat{\bar{\psi}}_2\cdots\hat{\psi}_n\hat{\bar{\psi}}_n|0\rangle &\equiv \langle\psi_1\bar{\psi}_1\psi_2\bar{\psi}_2\cdots\psi_n\bar{\psi}_n\rangle &(6.75)\\
&= \frac{\int[d\bar{\psi}d\psi]\,e^{iS}\,\psi_1\bar{\psi}_1\cdots\psi_n\bar{\psi}_n}{\int[d\bar{\psi}d\psi]\,e^{iS}} &(6.76)\\
&= (-i)^n\sum_\pi(-1)^\pi\,G_{1\,\pi1}\cdots G_{n\,\pi n}, &(6.77)
\end{aligned}$$

where

$$G_{1\,\pi1} = G_{a_1\,b_{\pi1}}(x_1, y_{\pi1}), \quad \text{etc.}, \tag{6.78}$$

and the summation is over all permutations π of $1,\ldots,n$, with signature $(-1)^\pi$. Instead of comparing coefficients of $\eta, \bar{\eta}$ we can also use repeated differentation with respect η and $\bar{\eta}$; see appendix 6.9 for fermionic differentiation rules. So in general the fermion field operators behave as anticommuting inside a time ordered product, and in a given time ordering we get a sign that is equal to the signature of the permutation needed to order the operators according to decreasing times from left to right.

To get a nonvanishing result there have to be an equal number of ψ's and $\bar{\psi}$'s. This is a consequence of the U(1) phase-symmetry (5.76). Making the transformation of variables

$$\psi = e^{i\omega}\,\psi', \quad \bar{\psi} = e^{-i\omega}\,\bar{\psi}' \tag{6.79}$$

and assuming the fermionic integration measure $[d\bar{\psi}d\psi]$ to be invariant it follows that

$$\langle\psi_1\psi_2\cdots\psi_m\bar{\psi}_1\cdots\bar{\psi}_n\rangle = e^{i(m-n)\omega}\,\langle\psi_1\psi_2\cdots\psi_m\bar{\psi}_1\cdots\bar{\psi}_n\rangle, \tag{6.80}$$

so the result vanishes if $m \neq n$. For a single Majorana field this U(1) symmetry is absent and the analogue of (6.77) would be similar to the scalar-field case, except for signs.

Next we study the fermion Green function G, which has to satisfy $G^{-1}G = 1$, or

$$(m + \gamma\partial)\,G(x,y) = \delta^4(x-y). \tag{6.81}$$

Writing G as a Fourier transform

$$G(x,y) = \int \frac{d^4p}{(2\pi)^4}\,e^{ip(x-y)}\,G(p), \tag{6.82}$$

we get the equation

$$(m + i\gamma p)\,G(p) = 1. \tag{6.83}$$

Using (cf. (5.84))

$$(m + i\gamma p)(m - i\gamma p) = m^2 + (\gamma p)^2 = m^2 + p^2, \tag{6.84}$$

we see that

$$G(p) = \frac{m - i\gamma p}{m^2 + p^2 - i\epsilon}, \tag{6.85}$$

is a possible solution, where we have used the same $i\epsilon$ prescription for dealing with the poles at $p^2 = -m^2$ as for the scalar field. We shall see shortly that this leads to consistent expressions. Thus the spinor field Green function is a linear combination of the scalar field propagator and its derivatives,

$$G(x,y) = \int \frac{d^4p}{(2\pi)^4}\, e^{ip(x-y)}\, \frac{m - i\gamma p}{m^2 + p^2 - i\epsilon} \tag{6.86}$$

$$= (m - \gamma^\mu \partial_\mu) \int \frac{d^4p}{(2\pi)^4}\, e^{ip(x-y)}\, \frac{1}{m^2 + p^2 - i\epsilon}. \tag{6.87}$$

We shall now show that the results above that followed from the path-integral approach to quantization lead to the same particle interpretation as the 'canonical quantization' in the previous section. Choosing $x^0 > y^0$ we have

$$\langle 0|\hat{\psi}(x)\hat{\bar{\psi}}(y)|0\rangle = (m - \gamma\partial) \int d\omega_p\, e^{ip(x-y)} \tag{6.88}$$

$$= \int d\omega_p\, e^{ip(x-y)}\, (m - i\gamma p), \tag{6.89}$$

$$= \sum_\lambda \int d\omega_p\, e^{ip(x-y)}\, u(p,\lambda)\bar{u}(p,\lambda), \tag{6.90}$$

where we used the $i\epsilon$ prescription in reducing the four-dimensional Fourier integral to a three-dimensional one, and also (5.179) in Problem 5.1. We insert a complete set of states between the field operators on the left-hand side, assuming this can be written in terms of a zero-particle (vacuum) contribution, one-particle contributions, two-particle contributions, etc. (cf. (2.158):

- The vacuum does not contribute because $\langle 0|\hat{\psi}|0\rangle = \langle 0|\hat{\bar{\psi}}|0\rangle = 0$, as we have learned from the coefficients linear in $\eta, \bar{\eta}$ in (6.69);
- The one particle contribution is

$$\sum_\lambda \int d\omega_p \langle 0|\hat{\psi}(x)|p\lambda\rangle\langle p\lambda|\hat{\bar{\psi}}(y)|0\rangle. \tag{6.91}$$

Using translation invariance

$$\hat{\psi}(x) = e^{-i\hat{P}x}\, \hat{\psi}(0)\, e^{i\hat{P}x}, \quad \hat{\bar{\psi}}(y) = e^{-i\hat{P}y}\, \hat{\bar{\psi}}(0)\, e^{i\hat{P}y}, \tag{6.92}$$

this can be written as

$$\sum_\lambda \int d\omega_p\, e^{ip(x-y)}\, \langle 0|\hat{\psi}(0)|p\lambda\rangle\langle p\lambda|\hat{\bar{\psi}}(0)|0\rangle, \tag{6.93}$$

which has the same exponentials as (6.90).
-The multiple particle contributions would give exponentials $\exp[ip(x-y)]$ (by the same translation invariance argument), with, since p^μ is in this case the total four-momentum of a multiparticle state, $-p^2 \geq 4m^2$. Hence, these contributions have to zero because the time-dependence cannot fit that of (6.90).
We conclude that only one particle states contribute, with

$$\langle 0|\hat{\psi}(x)|p\lambda\rangle = u(p,\lambda)\, e^{ipx}, \tag{6.94}$$

$$\langle p,\lambda|\hat{\bar{\psi}}(y)|0\rangle = \bar{u}(p,\lambda)\, e^{-ipy}. \tag{6.95}$$

Repeating the reasoning for $x^0 < y^0$ gives

$$-iG_{ab}(x,y) = (m-\gamma\partial)_{ab} \int d\omega_p\, e^{-ip(x-y)} \tag{6.96}$$

$$= \int d\omega_p\, e^{-ip(x-y)}\, (m+i\gamma p)_{ab} \tag{6.97}$$

$$= -\sum_\lambda \int d\omega_p\, e^{-ip(x-y)}\, v_a(p,\lambda)\bar{v}_b(p,\lambda) \tag{6.98}$$

$$= -\langle 0|\hat{\bar{\psi}}_b(y)\hat{\psi}_a(x)|0\rangle = -\sum_\lambda \int d\omega_p\, \langle 0|\bar{\psi}_b(y)|\overline{p,\lambda}\rangle\langle\overline{p,\lambda}|\hat{\psi}_a(x)|0\rangle + \ldots \tag{6.99}$$

where we assumed that there are different particles (indicated by the 'bar') with identical momentum and spin properties, because of the Dirac field being complex. Comparing the first and last lines we conclude that again the multiple particle contributions vanish and that

$$\langle 0|\hat{\bar{\psi}}(y)|\overline{p\lambda}\rangle = \bar{v}(p,\lambda)\, e^{ipy}, \tag{6.100}$$

$$\langle\overline{p,\lambda}|\hat{\psi}(x)|0\rangle = v(p,\lambda)\, e^{-ipx}. \tag{6.101}$$

The $i\epsilon$ specification of the fermion Green function is evidently the correct one for being able to match the left- and right-hand side of (6.74).

Taking limits of equal time, we easily derive the vacuum expectation value of canonical *anti*commutation relations between the field operators (see also appendix 6.8),

$$\langle 0|\{\hat{\psi}_a(\mathbf{x},t), \hat{\psi}_b^\dagger(\mathbf{y},t)\}|0\rangle = \delta_{ab}\,\delta(\mathbf{x}-\mathbf{y}). \tag{6.102}$$

These can presumably be extended to operator relations (i.e. holding true for arbitrary matrix elements) by the Wick formula (6.77), thus reconstructing fermionic Fock space. Note that we could easily have got a minus sign on the right hand side above, if G^{-1} had the opposite sign. This would be unacceptable as it would imply a negative metric in state vector space (it would not be a Hilbert space), leading to negative 'probabilities'. Obtaining the correct sign is the reason for the sign choice $\mathcal{L} = +i\psi^* \overleftrightarrow{\partial_0}\, \psi + \cdots$ in the lagrange density (5.65).

So the path integral formula for the vacuum expectation value of time-ordered products (6.62), together with some assumed basic properties of fermionic integration, provides an alternative formulation of the theory. The exposition so far is sufficient for us to proceed with perturbation theory for interacting theories with fermions. Appendix 6.9 gives further results on anticommuting numbers and fermionic integration.

Let us sketch the evaluation of the remaining $Z[0,0]$. Expanding the fields in terms of eigenfunction of the hermitian operator βG^{-1}, where G^{-1} is the Dirac operator $m + \gamma^\mu \partial_\mu$,

$$\beta G^{-1} f_\alpha = \lambda_\alpha f_\alpha, \quad \int d^4x\, f_\alpha^\dagger(x) f_\beta(x) = \delta_{\alpha\beta}, \tag{6.103}$$

$$\psi(x) = \sum_\alpha \psi_\alpha f_\alpha(x), \quad \psi^\dagger(x) = \sum_\alpha \psi_\alpha^* f_\alpha^\dagger(x), \tag{6.104}$$

brings the action into the form

$$S = -\sum_\alpha \lambda_\alpha \psi_\alpha^* \psi_\alpha. \tag{6.105}$$

The transformation of variables $\psi(x) \to \psi_\alpha$ is formally unitary, so its determinant is a phase factor which is expected to cancel in the combination $d\bar{\psi} d\psi$, and using also $\det \beta = 1$ we have formally

$$[d\bar{\psi}\, d\psi] = \prod_{xa} d\bar{\psi}_a(x)\, d\psi_a(x) = \prod_{xa} d\psi_a^*(x)\, d\psi_a(x) = \prod_\alpha d\psi_\alpha^*\, d\psi_\alpha, \tag{6.106}$$

Then, with the rules derived in appendix 6.9,

$$\int [d\bar{\psi}\, d\psi]\, e^{iS} = \int \left[\prod_\alpha d\psi_\alpha^*\, d\psi_\alpha\right] e^{-i\sum_\alpha \lambda_\alpha \psi_\alpha^* \psi_\alpha} \tag{6.107}$$

$$\propto \prod_\alpha \lambda_\alpha \equiv \det[\beta G^{-1}] = \det G^{-1}. \tag{6.108}$$

Note that $\det G^{-1}$ is in the numerator, and not in the denominator as for Bose fields. There is no square root because the Dirac field is complex. This formula has been checked (e.g. for the partition function) also by other means.

We warn the reader that in case of fermions the formal character of the above derivations is known to easily give rise to problems[6], especially with gauge-field interactions involving γ_5, and careful regularizations of the otherwise ambigouous infinities are all the more important.

We have not yet given a thorough definition of the time-evolution operator $\hat{U}(t_1, t_2)$ in terms of the fermionic path integral, using spacetime discretization. This turns out to be remarkably involved, because of the so-called fermion doubling phenomenon. Problems arise with the implementation of chiral symmetries

[6]One then speaks of 'anomalies', which turn out to have important physical implications.

like (5.80). Because of the importance for a proper mathematical definition of quantum field theory as well as for applications to nonperturbative computations, much effort has been spent on solving these problems, which involve deep properties of gauge theories and fermions, and research into this area of 'lattice fermions' is still in progress. For an introduction, see the books on lattice field theory.

6.4 Non-relativistic fields and antiparticles

In the non-relativistic regime it is better to use a reduced description that measures kinetic energies of the particles relative to their energy at rest: $E = \mathbf{p}^2/2m$. For example, this is customary in the field-theoretic formulation of condensed-matter physics.

For free fields the reduction can be conveniently done in terms of the explicit form for the generating functional for the n-point functions. Consider the case of a single Majorana field,

$$\begin{aligned} Z[\eta] &= \int [d\psi]\, \exp\left\{ i \int d^4x\, \left[-\frac{1}{2}\psi^T\beta(m+\gamma^\mu\partial_\mu)\psi + \eta^\dagger\beta\psi \right] \right\}, && (6.109)\\ &\equiv Z[0]\exp\{iW[\eta]\}, && (6.110)\\ W[\eta] &= \frac{1}{2}\int d^4x\, d^4y\, \eta(x)^T\beta G(x,y)\eta(y) \\ &= \frac{1}{2}\int \frac{d^4p}{(2\pi)^4}\,\eta(-p)^T\,\beta\frac{m-i\gamma^\mu p_\mu}{m^2+p^2-i\epsilon}\,\eta(p), && (6.111) \end{aligned}$$

where

$$\eta(p) = \eta(-p)^* = \int d^4x\, e^{-ipx}\,\eta(x). \qquad (6.112)$$

The sources η can be thought of as mimicking the effects of other particles in an interacting theory. In the non-relativistic regime $\eta(p)$ is non-zero only for frequencies near $\pm$ the rest-energy, $p^0 \approx \pm m$, and momenta such that the non-relativistic kinetic energy $\mathbf{p}^2/2m \ll m$. This means that in the integral in (6.111), only values $p^0 \approx \pm m$ and momenta $|\mathbf{p}| \ll m$ contribute. To implement this, let us first rewrite (6.111) using

$$\begin{aligned} &\int_{-\infty}^0 dp^0 \int d^3p\,\eta(-p)^T\frac{\beta(m-i\gamma_\mu p^\mu)}{m^2+p^2-i\epsilon}\,\eta(p) \\ = &\int_0^\infty dp^0 \int d^3p\,\eta(-p)^T\frac{\beta(m-i\gamma_\mu p^\mu)}{m^2+p^2-i\epsilon}\,\eta(p), \qquad (6.113) \end{aligned}$$

where we changed variables $p^\mu \to -p^\mu$, interchanged the sources and used the symmetry properties $\beta^T = -\beta$, $(\beta\gamma_\mu)^T = \beta\gamma_\mu$ that hold in a Majorana representation. Using the above identity (6.111) can be written as

$$W[\eta] = \int_0^\infty \frac{dp^0}{2\pi}\int \frac{d^3p}{(2\pi)^3}\,\eta(p)^\dagger\,\beta\frac{m-i\gamma^\mu p_\mu}{m^2+p^2-i\epsilon}\,\eta(p). \qquad (6.114)$$

Measuring frequencies relative to the rest energy is expressed by writing

$$p^0 = m + p'^0. \tag{6.115}$$

The non-relativistic reduction in the denominator in (6.114) is then implemented by

$$m^2 + p^2 - i\epsilon = m^2 + \mathbf{p}^2 - (m + p'^0)^2 - i\epsilon \approx \mathbf{p}^2 - 2mp'^0 - i\epsilon, \tag{6.116}$$

where we neglected $(p'^0)^2$ relative to m^2. Note that we also assumed p'^0 to be of the same order as the kinetic energy $\mathbf{p}^2/2m$. For the numerator we get

$$m - i\gamma^\mu p_\mu = m + i\gamma^0 p^0 - i\gamma^k p^k \approx m(1 + i\gamma^0) = 2m(1 + \beta)/2, \tag{6.117}$$

where we neglected $p^{'0}$ and $|\mathbf{p}|$ compared to m. The result of this approximation is

$$W[\eta] \approx \int_{-m}^{\infty} \frac{dp'^0}{2\pi} \int \frac{d^3p}{(2\pi)^3}\, \eta(m+p'^0, \mathbf{p})^\dagger \frac{(1+\beta)/2}{\mathbf{p}^2/2m - (p'^0 + i\epsilon)}\, \eta(m+p'^0, \mathbf{p}). \tag{6.118}$$

Now

$$\frac{1+\beta}{2} \equiv P_+ = P_+^2 \tag{6.119}$$

is a projector, it projects onto the subspace of eigenvalue +1 of β. This means that only two of the four Dirac-components of η contribute in the non-relativistic form. To make this explicit, we note that $\eta(p)$ and $\eta(p)^*$ are for $p^0 > 0$ not constraint by the reality condition $\eta(p)^* = \eta(-p)$. So we can consider the real and imaginary parts of $\eta(p)$, or equivalently, $\eta(p)$ and $\eta(p)^*$, to be independent Grassmann variables in (6.118), and we can change the representation of the Dirac matrices by a unitary transformation,

$$\eta(p) \to U^\dagger \eta(p), \quad \eta(p)^\dagger \to \eta(p)^\dagger U, \tag{6.120}$$

to one in which β is diagonal, the so-called Dirac representation:

$$U\beta U^\dagger = \rho_3, \quad U\alpha_k U^\dagger = \rho_1 \sigma_k, \quad \text{Dirac representation} \tag{6.121}$$

In this representation

$$P_+ = \frac{1+\rho_3}{2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.122}$$

and the two-component nature of $P_+\eta(p)$ and $\eta(p)^\dagger P_+$ is manifest. We can clarify the situation further by introducing two-component sources ζ and ζ^*,

$$\zeta_a(p'^0, \mathbf{p}) = \eta_a(m + p'^0, \mathbf{p}), \quad \zeta_a(p'^0, \mathbf{p})^* = \eta_a(m + p'^0, \mathbf{p})^*, \quad a = 1, 2. \tag{6.123}$$

We can also extend the lower limit of integral over p'^0 in (6.118) from $-m$ to $-\infty$, since $-m$ is already out of the regime were the sources are practically non-vanishing. Dropping now the prime on the dummy-integration variable p'^0, we can rewrite (6.118) in the form

$$\begin{aligned} W[\zeta, \zeta^*] &= \int \frac{d^4p}{(2\pi)^4} \zeta_a(p)^* \frac{1}{\mathbf{p}^2/2m - (p^0 + i\epsilon)} \zeta_a(p) \qquad (6.124) \\ &= \int d^4x \int d^4y \, \zeta_a^*(x) G_{\rm nr}^R(x, y) \zeta_a(y), \qquad (6.125) \end{aligned}$$

where

$$\zeta(x) = \int \frac{d^4p}{(2\pi)^4} e^{ipx} \zeta(p), \quad \zeta^\dagger(x) = \int \frac{d^4p}{(2\pi)^4} e^{-ipx} \zeta^\dagger(p), \qquad (6.126)$$

and

$$G_{\rm nr}^R(x, y) = \int \frac{d^4p}{(2\pi)^4} \frac{e^{ip(x-y)}}{\mathbf{p}^2/2m - (p^0 + i\epsilon)} \qquad (6.127)$$

is a the *retarded* Green function of the differential operator $i\partial_0 - \nabla^2/2m$. It is not difficult to find the path integral that reproduces the non-relativistic n-point functions specified by (6.125),

$$\begin{aligned} Z[\zeta, \bar{\zeta}] &= \int [d\psi \, d\psi^*] \, e^{iS_{\rm nr} + i \int d^4x \, (\bar{\zeta}_a^* \psi_a + \psi_a^* \zeta_a)}, \qquad (6.128) \\ S_{\rm nr} &= \int d^4x \left(\psi_a^* i\partial_0 \psi_a - \frac{1}{2m} \partial_k \psi_a^* \partial_k \psi_a \right), \qquad (6.129) \end{aligned}$$

where ψ_a, $a = 1, 2$, is a two-component non-relativistic spinor field. The index a is the spin index on which the Pauli matrices σ_k can act.

The above action is the same as that of the scalar-field version (1.63), apart form the doubling by the spin index a and the fact that here we are dealing with Grassmann variables. We end this section with the following observations:

- It is straightforward to derive, from the n-point functions obtained from the generating functional Z, the (anti)commutation relations of the field operators in the operator version of the theory ($x^0 = y^0$):

$$\begin{aligned} [\hat\psi(x), \hat\psi^\dagger(y)] &= \delta(\mathbf{x} - \mathbf{y}), \\ [\hat\psi(x), \hat\psi(y)] &= [\hat\psi^\dagger(x), \hat\psi^\dagger(y)] = 0, \quad \text{Bose}, \qquad (6.130) \\ \{\hat\psi_a(x), \hat\psi_b^\dagger(y)\} &= \delta_{ab}\delta(\mathbf{x} - \mathbf{y}), \\ \{\hat\psi_a(x), \hat\psi_b(y)\} &= \{\hat\psi_a^\dagger(x), \hat\psi_b^\dagger(y)\} = 0, \quad \text{Fermi}. \qquad (6.131) \end{aligned}$$

- The fact that a retarded Green function appears in the expression for two-point function implies that $\hat\psi$ annihilates the ground state (unlike the relativistic field):

$$\langle 0|T\hat\psi_a(x)\hat\psi_b^\dagger(y)|0\rangle = -iG_{\rm nr}^R(x, y) \Rightarrow \langle 0|\hat\psi_b^\dagger(y)\hat\psi_a(x)|0\rangle = 0. \qquad (6.132)$$

The stronger statement $\hat{\psi}_a(x)|0\rangle = 0$ follows from taking all n-point functions into consideration. It follows from the commutation relations that $\hat{\psi}^\dagger(x)$ acts like a creation operator; it creates a particle at position x.

- The non-relativistic field is complex, in contrast to the real Majorana field. The action (6.129) has an extra $U(1)$ symmetry that the Majorana action in (6.109) does not possess:

$$\psi(x) \to e^{i\omega}\, \psi(x), \quad \psi^\dagger(x) \to e^{-i\omega}\, \psi^\dagger(x). \tag{6.133}$$

The corresponding Noether-current is given by

$$j^0 = \psi^\dagger\psi, \quad j_k = \frac{1}{2m}(-\psi^\dagger i\partial_k\psi + i\partial_k\psi^\dagger\psi), \tag{6.134}$$

and the conserved charge simply counts the number of particles. This complex-number property of non-relativistic fields is just right for representations of the *Gallilei group*, the group of translations, rotations and non-relativistic boosts; see e.g. Weinberg I.

The action (6.129) even has an $SU(2)$ symmetry corresponding to transformations on the spin index, independent of spatial rotations: $\psi_a(x) \to \Omega_{ab}\psi_b$, $\psi^*_a(x) \to \psi^*_b(x)\Omega^\dagger_{ba}$, $\Omega \in SU(2)$. This symmetry is an artifact of the free field and it is usually broken by interactions. For example, a magnetic moment interaction $\propto \psi^\dagger(x)\sigma_k\psi(x)B_k(x)$ to an external magnetic field $\mathbf{B}(x)$ is invariant under spatial rotations that affect both the spin index a and the spatial index $\mathbf{x}$ in the usual way.

- Suppose the non-relativistic field describes electrons. As a consequence of the field being complex, we can couple it gauge-invariantly to the electromagnetic field by the minimal substitution rule $\partial_\mu \to \partial_\mu + ieA_\mu$, adding also a magnetic-moment coupling,

$$\begin{aligned} S[\psi, \psi^*, A] &= \int d^4x \left[\psi^\dagger(i\partial_0 - eA_0)\psi - \frac{1}{2m}(\partial_k - ieA_k)\psi^\dagger(\partial_k + ieA_k)\psi \right. \\ &\qquad \left. -\frac{ge}{4m}\,\psi^\dagger\sigma_k\psi\, B_k\right], \end{aligned} \tag{6.135}$$

where g is the Landé g factor (gyromagnetic ratio).

Such a procedure is not possible for the real Majorana field. Instead (cf. chapter 8), the complex Dirac field is needed to be able to formulate a gauge invariant coupling to the Maxwell field. The relativistic theory also gives a precise value for the coupling constant g, which is a free parameter in the non-relativistic theory.

As we have seen, the Dirac field describes two kinds of particles with identical masses and opposite charges — it turns out that the eigenvalue ± 1 of the Noether charge is indeed the charge of a particle.

Thus, whereas antiparticles are not needed in non-relativistic electrodynamics, the relativistic quantum field theory of charged particles *predicts the existence of antiparticles.*

For more information on non-relativistic quantum fields see, e.g. the book by Brown, and specialized books on application to condensed matter physics.

6.5 Yukawa models

As an example for illustration of perturbative methods we shall use a so-called Yukawa model described by the action[7]

$$S = -\int d^4x \left(\frac{1}{2}\,\partial_\mu\phi\partial^\mu\phi + \frac{1}{2}\,m^2\phi^2 + \bar{\psi}\gamma^\mu\partial_\mu\psi + M\bar{\psi}\psi + g\bar{\psi}\Gamma\psi\phi\right) + \Delta S, \quad (6.136)$$

where either $\Gamma = 1$ or $\Gamma = i\gamma_5$, and ΔS denotes the counterterms . For $\Gamma = i\gamma_5$ this action is still parity invariant, but with a minus sign in the transformation of the scalar field:

$$\begin{aligned} \psi'(x) &= D_P\psi(\ell_P^{-1}x), \quad \bar{\psi}'(x) = \bar{\psi}(\ell_P^{-1}x)\,D_P^{-1}, \quad \phi'(x) = -\phi(\ell_P^{-1}x), \\ \bar{\psi}'\gamma_5\psi'\phi' &= -\bar{\psi}\gamma_0^{-1}\gamma_5\gamma_0\psi\phi = \bar{\psi}\gamma_5\psi\phi \quad (D_P = \gamma^0). \end{aligned} \quad (6.137)$$

This phase factor $\pm$ in the transformation rule for ϕ (+ for $\Gamma = 1$, − for $\Gamma = i\gamma_5$) is called the intrinsic parity of the scalar field. In case of the plus sign the field ϕ and the corresponding particles are called scalars, in case of the minus sign pseudoscalars. In the pseudoscalar case we can think of ψ, $\bar{\psi}$, and ϕ as giving a crude description of the interactions of protons (p), antiprotons ($\bar{p}$) and neutral pions (π^0).

Writing

$$S = S_0 + S_1, \quad (6.138)$$

$$S_1 = -\int d^4x\, g\bar{\psi}\Gamma\psi\phi + \Delta S, \quad (6.139)$$

the perturbative expansion of $\langle\psi_1\bar{\psi}_2\phi_3\cdots\rangle$ leads to expressions which are again well represented by diagrams. To distinguish the bosonic and fermionic propagators we shall change the notation of the latter to S:

$$S_{ab}(x,y) = \int \frac{d^4p}{(2\pi)^4}\, e^{ip(x-y)}\, S_{ab}(p), \quad S_{ab}(p) = \frac{M\delta_{ab} - ip_\mu\gamma^\mu_{ab}}{M^2 + p^2 - i\epsilon}. \quad (6.140)$$

The fermion propagator $S_{ab}(x,y)$ is not a symmetric function of its arguments, $S_{ab}(x,y) \neq S_{ba}(y,x)$ or $S_{ab}(p) \neq S_{ba}(-p)$, and to be able to distinguish its arguments we draw an arrow in its line in the diagrams, as in Fig. 6.1: the arrow

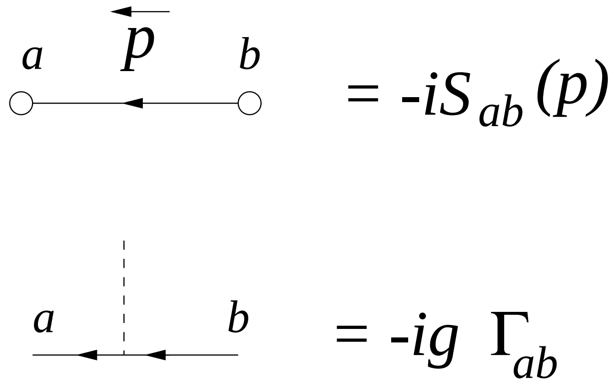

Figure 6.1: Fermion propagator $S_{ab}(p) = (M - i\gamma p)_{ab}/(M^2 + p^2 - i\epsilon)$ and bare vertex function $-ig\Gamma_{ab}$ ($\Gamma = 1$ or $i\gamma_5$) of the Yukawa models. The fermion line is drawn, the scalar line is dashed. The arrow in the fermion line points to the row-index a.

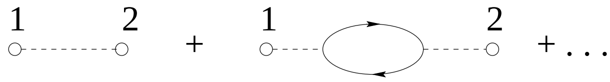

Figure 6.2: First two nonzero terms in the perturbative expansion of $\langle\phi(x_1)\phi(x_2)\rangle$ in a Yukawa model (without bare ϕ^4 interaction).

points in the direction of the 'row index' of S. The interaction S_1 leads to a bare vertex function in position space

$$S_{\bar\psi_a\psi_b\phi}(u,v,w) = -g\Gamma_{ab}\int d^4x\,\delta(x-u)\,\delta(x-v)\,\delta(x-w), \tag{6.141}$$

or in momentum space

$$S_{\bar\psi_a\psi_b\phi}(p,q,k) = -g\Gamma_{ab}, \tag{6.142}$$

as illustrated in Fig. 6.1. The calculation of n-point functions goes along the same lines as in scalar field theory. We write

$$\langle\psi_1\bar\psi_2\phi_3\cdots\rangle = \frac{\langle\psi_1\bar\psi_2\phi_3\cdots e^{iS_1}\rangle_0}{\langle e^{iS_1}\rangle_0}, \tag{6.143}$$

with $\langle\cdots\rangle_0$ the free field 'average', and expand the exponential. Disconnected vacuum diagrams again cancel between numerator and denominator. For example the first few diagrams for the scalar two point function are given in Fig. 6.2. In position space these represent

$$\begin{aligned}\langle\phi_1\phi_2\rangle &= \langle\phi_1\phi_2\rangle_0 + \frac{(-ig)^2}{2!}\int d^4x\,d^4y\,\langle\phi_1\phi_2\bar\psi_a(x)\Gamma_{ab}\psi_b(x)\phi(x)\\ &\quad \bar\psi_c(y)\Gamma_{cd}\psi_d(y)\phi(y)\rangle_0^{\rm conn} + \cdots\end{aligned} \tag{6.144}$$

[7]There is no special reason why we switched using the notation ϕ in stead of φ.

$$\begin{aligned} &= -iG(x_1,x_2) - 2\frac{(-ig)^2}{2!}\int d^4x\,d^4y\,(-i)^4 G(x_1,x)G(x_2,y) \\ &\quad \Gamma_{ab}S_{bc}(x,y)\Gamma_{cd}S_{da}(y,x) + \cdots, \end{aligned} \tag{6.145}$$

where the factor 2 comes from similar contractions with $x \leftrightarrow y$ which give an identical contribution, and the minus sign in front of it comes from putting the fermion fields in standard $\psi\bar{\psi}$ order as for the propagator ($\bar{\psi}_a\psi_b\bar{\psi}_c\psi_d = -\psi_b\bar{\psi}_c\psi_d\bar{\psi}_a$). Note that such minus signs are typical for closed fermion loops. In momentum space we get

$$G'(p) = G(p) - G(p)\Sigma(p)G(p) + \cdots \tag{6.146}$$

$$-i\Sigma(p) = -\int \frac{d^4q}{(2\pi)^4}\,\mathrm{Tr}\,(-ig\Gamma)(-i)S(q)(-ig\Gamma)(-i)S(p+q). \tag{6.147}$$

6.6 Scattering in the tree graph approximation

The calculation of scattering amplitudes and cross sections will here be illustrated in the Yukawa model. We concentrate on the leading contributions which are represented by tree graphs. First we derive some convenient formulas involving creation and annihilation operators.

We define for the scalar field

$$\phi^{(\pm)}(p,x^0) = \pm\int d^3x\,e^{\mp ipx}(i\overrightarrow{\partial_0} - i\overleftarrow{\partial_0})\phi(x), \quad p^0 = \sqrt{m^2+\mathbf{p}^2}, \tag{6.148}$$

Recalling the expansion

$$\phi_{\mathrm{free}}(x) = \int d\omega_p\,\left[a(p)\,e^{ipx} + a^\dagger(p)\,e^{-ipx}\right] \tag{6.149}$$

of the fee scalar field, it is easy to check that if $\phi(x)$ satisfies the free Klein-Gordon equation with mass m, then $\phi^{(\pm)}(p,x^0)$ are actually time independent and equal to the usual covariantly normalized annihilation and creation operators:

$$\phi^{(+)}(p,x^0) \equiv a(p,x^0) = a(p) \quad \text{if } (m^2-\partial^2)\,\phi = 0, \tag{6.150}$$

$$\phi^{(-)}(p,x^0) \equiv a^\dagger(p,x^0) = a^\dagger(p) \quad \text{if } (m^2-\partial^2)\,\phi = 0. \tag{6.151}$$

The notation $\phi^{(\pm)}(p,x^0)$ stresses that we are just taking linear combinations of the otherwise arbitray scalar field.[8] Because they are linear combinations we can make contractions with them just as for the fields themselves. For instance,

$$\langle\phi^{(+)}(p,x^0)\phi(y)\rangle_0 = \int d^3x\,e^{-ipx}(i\overrightarrow{\partial_0} - i\overleftarrow{\partial_0})(-i)G(x,y), \tag{6.152}$$

[8]For the free case the $\pm$ may be thought of indicating the positive and negative frequency components of the field.

where the subscript 0 on the left-hand side denotes the free case as usual. For clarity in applications we shall use the $a(p, x^0)$ notation in the following. Using the above expression in terms of $G(x, y)$, or the interpretation as creation or annihilation operators, we now have the convenient formulas[9]

$$\begin{aligned}
\langle a(p,\infty)\,\phi(x)\rangle_0 &= \langle 0|a(p)\,\phi(x)|0\rangle_0 = \langle p|\phi(x)|0\rangle_0 = e^{-ipx}, &(6.153)\\
\langle a(p,-\infty)\,\phi(x)\rangle_0 &= \langle 0|\phi(x)a(p)|0\rangle_0 = 0, &(6.154)\\
\langle a^*(p,\infty)\,\phi(x)\rangle_0 &= \langle 0|a^\dagger(p)\phi(x)|0\rangle_0 = 0, &(6.155)\\
\langle a^*(p,-\infty)\,\phi(x)\rangle_0 &= \langle 0|\phi(x)\,a^\dagger(p)|0\rangle_0 = \langle 0|\phi(x)|p\rangle_0 = e^{ipx}. &(6.156)
\end{aligned}$$

Here $\pm\infty$ means that the time has been sent to $\pm\infty$.

Similarly, we define for the fermion field

$$\begin{aligned}
b(p,\lambda,x^0) &= \psi^{(+)}(p,\lambda,x^0) = \int d^3x\, e^{-ipx}\,\bar{u}(p,\lambda)\,i\gamma^0\,\psi(x), &(6.157)\\
b^\dagger(p,\lambda,x^0) &= \bar{\psi}^{(-)}(p,\lambda,x^0) = \int d^3x\,\bar{\psi}(x)\,i\gamma^0\,u(p,\lambda)\,e^{ipx}, &(6.158)\\
d^\dagger(p,\lambda,x^0) &= \psi^{(-)}(p,\lambda,x^0) = \int d^3x\, e^{ipx}\,\bar{v}(p,\lambda)\,i\gamma^0\,\psi(x), &(6.159)\\
d(p,\lambda,x^0) &= \bar{\psi}^{(+)}(p,\lambda,x^0) = \int d^3x\,\bar{\psi}(x)\,i\gamma^0\,v(p,\lambda)\,e^{-ipx}, &(6.160)
\end{aligned}$$

where

$$p^0 = \sqrt{M^2 + \mathbf{p}^2}. \qquad (6.161)$$

Recalling the expansions

$$\begin{aligned}
\psi_{\text{free}}(x) &= \sum_\lambda \int d\omega_p \left[b(p,\lambda)\,u(p,\lambda)\,e^{ipx} + d^\dagger(p,\lambda)\,v(p,\lambda)\,e^{-ipx}\right], &(6.162)\\
\bar{\psi}_{\text{free}}(x) &= \sum_\lambda \int d\omega_p \left[b^\dagger(p,\lambda)\,\bar{u}(p,\lambda)\,e^{-ipx} + d(p,\lambda)\,\bar{v}(p,\lambda)\,e^{ipx}\right], &(6.163)
\end{aligned}$$

we see that if the fermion field satisfies the free Dirac equation with mass M, then $b(p, \lambda, x^0)$, ..., $d^\dagger(p, \lambda, x^0)$ are actually time-independent and equal to the annihilation, ..., creation operators of particles, ..., antiparticles. We have in the free case

$$\langle b(p,\lambda;x^0)\bar{\psi}(y)\rangle_0 = \int d^3x\, e^{-ipx}\,\bar{u}(p,\lambda)\,i\gamma^0\,(-i)S(x,y), \qquad (6.164)$$

etc., and using such expressions, or the operator versions, the following formulas can be derived which are convenient for later use:

$$\begin{aligned}
\langle b(p,\lambda,\infty)\,\bar{\psi}(x)\rangle_0 &= \langle 0|b(p,\lambda)\,\bar{\psi}(x)|0\rangle_0 = \bar{u}(p,\lambda)\,e^{-ipx}, &(6.165)\\
\langle b(p,\lambda,\infty)\,\psi(x)\rangle_0 &= 0, &(6.166)
\end{aligned}$$

[9]Inside $\langle\cdots\rangle$ we write a^* in stead of $a^\dagger$ because in the path integral these are not operators.

$$
\begin{aligned}
\langle \psi(x)\, b^*(p,\lambda,-\infty)\rangle_0 &= \langle 0|\psi(x)\, b^\dagger(p,\lambda)|0\rangle_0 = u(p,\lambda)\, e^{ipx}, &(6.167)\\
\langle \bar\psi(x)\, b^*(p,\lambda,-\infty)\rangle_0 &= 0, &(6.168)\\
\langle d(p,\lambda,\infty)\, \psi(x)\rangle_0 &= \langle 0|d(p,\lambda)\, \psi(x)|0\rangle_0 = v(p,\lambda)\, e^{-ipx}, &(6.169)\\
\langle d(p,\lambda,\infty)\, \bar\psi(x)\rangle_0 &= 0, &(6.170)\\
\langle \bar\psi(x)\, d^*(p,\lambda,-\infty)\rangle_0 &= \langle 0|\bar\psi(x)\, d^\dagger(p,\lambda)|0\rangle_0 = \bar v(p,\lambda)\, e^{ipx}, &(6.171)\\
\langle \psi(x)\, d^*(p,\lambda,-\infty)\rangle_0 &= 0, &(6.172)
\end{aligned}
$$

with the remaining time-specifications vanishing.

These formulas will now be used for the calculation of scattering amplitudes. The formulation and calculation of scattering processes in field theory will be discussed in more detail in chapter 9, but in the tree graph approximation the following reasoning is plausibly correct. Consider the proton-pion scattering process $(p,\lambda)+k \to (p',\lambda')+k'$. The probability amplitude $\langle p'\lambda', k'\,\mathrm{out}|p\lambda, k\,\mathrm{in}\rangle$ for finding the outgoing state in the incoming state can be calculated as

$$
\begin{aligned}
\langle p'\lambda', k'\,\mathrm{out}|p\lambda, k\,\mathrm{in}\rangle &= \langle a(k',t')b(p'\lambda',t')b^*(p,\lambda,t)a^*(k,t)\rangle\\
&= \frac{\langle a(k',t')b(p'\lambda',t')b^*(p,\lambda,t)a^*(k,t)e^{iS_1}\rangle_0}{\langle e^{iS_1}\rangle_0}, \qquad (6.173)
\end{aligned}
$$

where $t\to-\infty$, $t'\to+\infty$ and $\langle\cdots\rangle_0$ denotes the free field 'average'

$$
\langle\cdots\rangle_0 = \frac{\int [d\bar\psi d\psi][d\phi]\, e^{iS_0}\,\cdots}{\int [d\bar\psi d\psi][d\phi]\, e^{iS_0}}, \qquad (6.174)
$$

as in (4.10). Expanding

$$
e^{iS_1} = 1+\cdots+\frac{(-ig)^2}{2!}\int d^4x\, d^4x'\, \bar\psi(x)\Gamma\psi(x)\phi(x)\, \bar\psi(x')\Gamma\psi(x')\phi(x')+\cdots \qquad (6.175)
$$

we encounter (assuming the final state to be different from the initial state, such that the disconnected contributions leading to $\langle p'\lambda'|p\lambda\rangle\langle k'|k\rangle$ vanish):

$$
\begin{aligned}
&\langle a(k',t')b(p'\lambda',t')\bar\psi(x)\Gamma\psi(x)\phi(x)\, \bar\psi(x')\Gamma\psi(x')\phi(x')\, b^*(p\lambda,t)a^*(k,t)\rangle_0\\
=\;& \Big[\langle b(p'\lambda',\infty)\bar\psi(x)\rangle_0\,\Gamma\, \langle\psi(x)\bar\psi(x')\rangle_0\,\Gamma\, \langle\psi(x')b^*(p\lambda,-\infty)\rangle_0 + x\leftrightarrow x'\Big]\\
&[\langle a(k',\infty)\phi(x)\rangle_0\, \langle\phi(x')a^*(k,-\infty)\rangle_0 + x\leftrightarrow x']\\
&+\text{ contractions leading to loop diagrams}\\
=\;& e^{-ip'x}\,\bar u(p',\lambda')\Gamma(-i)S(x,x')\Gamma u(p,\lambda)e^{ipx'}[e^{-ik'x}\, e^{ikx'} + e^{-ik'x'}\, e^{ikx}]\\
&+x\leftrightarrow x' + \text{contractions leading to loop diagrams.} \qquad (6.176)
\end{aligned}
$$

Note the condensed notation in which the summation over Dirac indices is suppressed. For example, we have more explicitly

$$
\begin{aligned}
&\langle b(p'\lambda',\infty)\bar\psi(x)\rangle_0\,\Gamma\, \langle\psi(x)\bar\psi(x')\rangle_0\,\Gamma\, \langle\psi(x')b^*(p\lambda,-\infty)\rangle_0 \qquad (6.177)\\
=\;& \langle b(p'\lambda',\infty)\bar\psi_a(x)\rangle_0\,\Gamma_{ab}\, \langle\psi_b(x)\bar\psi_c(x')\rangle_0\,\Gamma_{cd}\, \langle\psi_d(x')b^*(p\lambda,-\infty)\rangle_0,
\end{aligned}
$$

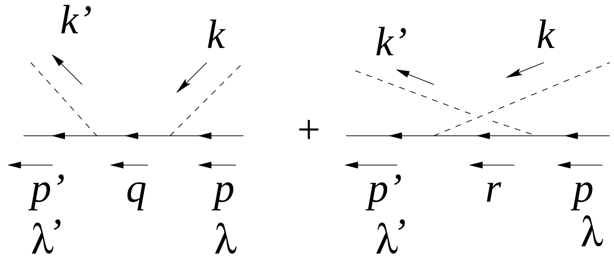

Figure 6.3: Tree diagrams for fermion-(pseudo)scalar scattering in the Yukawa model; $q = p + k$, $r = p - k'$.

with the usual summation convention for a, b, c, d. Substituting

$$S(x, x') = \int \frac{d^4 s}{(2\pi)^4} \, e^{is(x-x')} \, \frac{M - i\gamma s}{M^2 + s^2} \tag{6.178}$$

and integrating over x and x' we encounter

$$\delta^4(-p' - k' + s) \, \delta^4(p + k - s) = \delta^4(p + k - p' - k') \, \delta^4(p + k - s) \tag{6.179}$$

for the first contribution and

$$\delta^4(-p' + k + s) \, \delta^4(p - k' - s) = \delta^4(p + k - p' - k') \, \delta^4(p - k' - s), \tag{6.180}$$

for the second. The $x \leftrightarrow x'$ terms give the same contribution and we find the form

$$\langle p'\lambda', k' \, \text{out} | p\lambda, k \, \text{in} \rangle = -i(2\pi)^4 \delta^4(p' + k' - p - k) \, T(p', \lambda', k'; p, \lambda, k), \tag{6.181}$$

with the result for the scattering amplitude (invariant amplitude) T:

$$T = \bar{u}(p', \lambda')(-ig\Gamma) \left[\frac{M - i\gamma(p + k)}{M^2 + (p + k)^2} + \frac{M - i\gamma(p - k')}{M^2 + (p - k')^2} \right] (-ig\Gamma) u(p, \lambda), \tag{6.182}$$

The two diagrams corresponding to this process are given in Fig. 6.3.

Defining

$$\overline{|T|^2} = \frac{1}{4} \sum_{\lambda\lambda'} |T|^2 \tag{6.183}$$

as the average of $|T|^2$ over initial and final spins, the differential cross section takes the same form as for spinless particles

$$\left[\frac{d\sigma}{d\Omega} \right]_{\text{cm}} = \frac{1}{64\pi^2 s} \, 2\overline{|T|^2}, \tag{6.184}$$

where the factor 2 comes from the summation over final spin states. The further calculation of $\overline{|T|^2}$ (which is a Lorentz-invariant function of the momenta) will not be done here (similar calculations will be carried out later in QED).

However, we like to draw the attention to the angular dependence coming from the denominator of the second fermion propagator: since $M^2 + (p - k')^2 = -m^2 + 2\sqrt{M^2 + \mathbf{p}^2}\,\sqrt{m^2 + \mathbf{p}^2}\, + 2|\mathbf{p}|^2 \cos\theta$ in the center of mass frame (θ being the angle between $\mathbf{p}$ and $\mathbf{p}'$), we get a *backward peak* at high energies, where the masses may be neglected:

$$\frac{1}{M^2 + (p - k')^2} \to \frac{1}{|\mathbf{p}|^2(\pi - \theta)^2}, \quad \theta \approx \pi. \tag{6.185}$$

Consider next antifermion-(pseudo)scalar scattering. Here (6.173) gets replaced by

$$\langle \overline{p'\lambda'}, k' \,\text{out}|\overline{p\lambda}, k\,\text{in}\rangle = \frac{\langle a(k',\infty)d(p'\lambda',\infty)d^*(p,\lambda,-\infty)a^*(k,-\infty)e^{iS_1}\rangle_0}{\langle e^{iS_1}\rangle_0}. \tag{6.186}$$

To make the contractions slightly more transparant we now first put the 'expectation value' in standard $\psi\bar{\psi}\cdots\psi\bar{\psi}$ order (recall $d \leftrightarrow \bar{\psi}$ and $d^\dagger \leftrightarrow \psi$):

$$\begin{aligned} &\langle a(k',\infty)d(p'\lambda',\infty)d^*(p,\lambda,-\infty)a^*(k,-\infty)e^{iS_1}\rangle_0 \\ = \;& -\langle a(k',\infty)d^*(p,\lambda,-\infty)d(p',\lambda',\infty)a^*(k,-\infty)e^{iS_1}\rangle_0. \end{aligned} \tag{6.187}$$

We now encounter

$$\langle a(k',\infty)d^*(p\lambda,-\infty)\bar{\psi}(x)\Gamma\psi(x)\phi(x)\,\bar{\psi}(x')\Gamma\psi(x')\phi(x')\,d(p'\lambda',\infty)a^*(k,-\infty)\rangle_0$$

$$\begin{aligned} = \;& \langle d^*(p\lambda,-\infty)\bar{\psi}(x)\rangle_0\,\Gamma\,\langle\psi(x)\bar{\psi}(x')\rangle_0\,\Gamma\,\langle\psi(x')d(p'\lambda',\infty)\rangle_0 \\ & [\langle a(k',\infty)\phi(x)\rangle_0\,\langle\phi(x')a^*(k,-\infty)\rangle_0 \\ & + \langle a(k',\infty)\phi(x')\rangle_0\,\langle\phi(x)a^*(k,-\infty)\rangle_0]] \\ & + x \leftrightarrow x' \\ & + \text{contractions leading to loop diagrams} \\ = \;& e^{ipx}\,\bar{v}(p,\lambda)\Gamma(-i)S(x,x')\Gamma v(p',\lambda')e^{-ip'x'}[e^{-ik'x}\,e^{ikx'} + e^{-ik'x'}\,e^{ikx}] \\ & + x \leftrightarrow x' + \text{contractions leading to loop diagrams.} \end{aligned} \tag{6.188}$$

The integrations over x and x' now produce

$$\delta^4(p - k' + s)\,\delta^4(-p' + k - s) = \delta^4(p + k - p' - k')\,\delta^4(p - k' + s) \tag{6.189}$$

for the first contribution and

$$\delta^4(p + k + s)\,\delta^4(-p' - k' - s) = \delta^4(p + k - p' - k')\,\delta^4(p + k + s), \tag{6.190}$$

for the second, which leads to

$$T = -\bar{v}(p,\lambda)(-ig\Gamma)\left[\frac{M + i\gamma(p - k')}{M^2 + (p - k')^2} + \frac{M + i\gamma(p + k)}{M^2 + (p + k)^2}\right](-ig\Gamma)v(p',\lambda'), \tag{6.191}$$

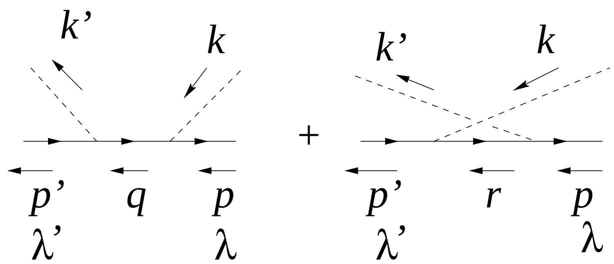

Figure 6.4: Tree diagrams for antifermion-(pseudo)scalar scattering; $q = p + k$, $r = p - k'$.

with the identification $s = k' - p$ for the first contribution and $s = -p - k$ for the second. We see an overall minus sign, and $\bar{v}$ and v describe the initial state and final state, respectively. The diagrams for this process are shown in Fig. 6.4. Notice that the momenta q and r flow in this case against the direction of the arrow in the fermion line.

6.7 Summary

Canonical quantization of the Dirac field is not straightforward. Physical requirements (the existence of a ground state, locality, positivity of Hilbert space metric) led to *anticommutation relations* for the field operators. These can be realized in Hilbert space with creation and annihilation operators of the anticommuting type, such that occupation numbers obey the Pauli principle and basis vectors are antisymmetric under simultaneous exchange of labels. The excitations of the ground state are particles with spin 1/2 obeying Fermi-Dirac statistics.

The Dirac field describes two distinct types of particles with the same mass and spin. The particle states can be eigenstates of Q (which commutes with the operator P^μ of total four-momentum), and those with $Q = +1$ (-1) are conventionally called particles (antiparticles). Q itself equals the total number of particles minus antiparticles.

A single Majorana field has only one type of particle of a given spin (and correspondingly, no conserved Q). The concept of antiparticles does not arise in this case, except that for zero mass Q_5 is conserved and it can take the role of Q in the Dirac case. In the non-relativistic limit of a single Majorana field a new U(1) symmetry emerges with a Noether charge corresponding to conservation of particles. Relativistically only the number of particles minus the number of antiparticles is conserved. Non-relativistic quantum electrodynamics does not require the existence of antiparticles, but, as we shall see in chapter 8, the relativistic formulation does predict antiparticles indeed.

The path-integral description of fermions requires the introduction of anticommuting 'numbers', which have nice properties and allow for an easy evaluation of the correlation functions in the free case. Perturbation theory leads again to diagrams, here with Dirac propagators as well. A heuristic derivation of scattering amplitudes shows how spin dependence is represented by the polarization spinors.

6.8 Appendix: More on spin and statistics

We give here another reasoning that a description of spin 1/2 particles in terms of spinor fields requires the particles to obey Fermi-Dirac statistics.

Assume a theory of free spin 1/2 particles, in which there is a vacuum state $|0\rangle$ with zero energy-momentum, and one particle states $|p\lambda\rangle$ with energy-momentum p^μ,

$$P^\mu|0\rangle = 0, \quad P^\mu|p\lambda\rangle = p^\mu|p\lambda\rangle \tag{6.192}$$

(as before P^μ is the energy-momentum operator and $\lambda = \pm$ is a spin index). The conventions are such that these states are obtained by the action of standard boosts ℓ_p (as in Appendix 5.5.6) on a particle state at rest,

$$|p\lambda\rangle = U(\ell_p)|\bar{p}\lambda\rangle, \quad \bar{\mathbf{p}} = \mathbf{0}, \ \ \bar{p}^0 = m, \tag{6.193}$$

where $U(\ell_p)$ is the unitary operator representing ℓ_p in Hilbert space. The index $\lambda = \pm$ labels the eigenvalue of the third component of angular momentum J_3 in the rest frame of the particle,

$$J_3|\bar{p}\lambda\rangle = \frac{1}{2}\lambda\,|\bar{p}\lambda\rangle. \tag{6.194}$$

Let $\psi(x)$ now be a Majorana spinor-field operator. In the Majorana representation for the gamma matrices this is a hermitian spinor-field

$$\psi_a^\dagger(x) = \psi_a(x). \tag{6.195}$$

By analogy to the scalar field we assume $\psi(x)$ to annihilate spin 1/2 particles to the vacuum according to

$$\langle 0|\psi_a(x)|p,\lambda\rangle = u_a(p,\lambda)\,e^{ipx}. \tag{6.196}$$

The form of this equation is dictated by translation invariance (the factor $\exp(ipx)$) and Lorentz invariance (the factor $u_a(p,\lambda)$, because $|p,\lambda\rangle$ and $u(p,\lambda)$ are constructed in exactly the same way with the boost ℓ_p). The remaining factor ($= 1$) is a normalization condition for $\psi(x)$. In general we may have an additional factor $\sqrt{Z_\psi}$, which we choose to be 1 in case of no interactions. Taking the complex conjugate of (6.196) and multiplying by by β, gives

$$\langle 0|\bar{\psi}(x)|p,\lambda\rangle = \bar{v}(p,\lambda)\,e^{ipx}. \tag{6.197}$$

We have seen that free fields create only single particle states out of the vacuum. If we assume this to be the case of our free spinor field as well, we can derive the vacuum expectation value of equal time commutator or anticommutator relations as follows. Using completeness we have

$$\begin{aligned}
\langle 0|\psi(x)\,\bar{\psi}(y)|0\rangle &= \sum_\lambda \int d\omega_p\,\langle 0|\psi(x)|p,\lambda\rangle\langle p,\lambda|\bar{\psi}(y)|0\rangle && (6.198)\\
&= \sum_\lambda \int d\omega_p\, e^{ip(x-y)}\,u(p,\lambda)\,\bar{u}(p,\lambda) && (6.199)\\
&= \int d\omega_p\, e^{ip(x-y)}\,(m - ip^\mu\gamma_\mu). && (6.200)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\langle 0|\bar{\psi}_b(y)\,\psi_a(x)|0\rangle &= \sum_\lambda \int d\omega_p\,\langle 0|\bar{\psi}_b(y)|p,\lambda\rangle\langle p,\lambda|\psi_a(x)|0\rangle && (6.201)\\
&= \sum_\lambda \int d\omega_p\, e^{ip(y-x)}\,v_a(p,\lambda)\,\bar{v}_b(p,\lambda) && (6.202)\\
&= -\int d\omega_p\, e^{ip(y-x)}\,(m + ip^\mu\gamma_\mu)_{ab}. && (6.203)
\end{aligned}$$

From these relations now follow the vacuum expectation values of equal time commutators or anticommutators:

$$\begin{aligned}
&\langle 0|[\psi_a(x)\,\bar{\psi}_b(y) \pm \bar{\psi}_b(y)\,\psi_a(x)]|0\rangle_{x^0=y^0} && (6.204)\\
&= \int d\omega_p\,[e^{i\mathbf{p}(\mathbf{x}-\mathbf{y})}\,(m - ip^\mu\gamma_\mu) \mp [e^{-i\mathbf{p}(\mathbf{x}-\mathbf{y})}\,(m + ip^\mu\gamma_\mu)]_{ab} && (6.205)\\
&= \int d\omega_p\, e^{i\mathbf{p}(\mathbf{x}-\mathbf{y})}\,[(m - ip^k\gamma_k)(1\mp 1) + ip^0\gamma^0(1\pm 1)]_{ab}. && (6.206)
\end{aligned}$$

It follows that the vacuum expectation value of the commutator $[\psi_a(x), \bar{\psi}_b(y)]$ is given by

$$\langle 0|[\psi_a(x), \bar{\psi}_b(y)]|0\rangle_{x^0=y^0} = 2(m - \gamma^k\partial_k)_{ab}\int d\omega_p\, e^{i\mathbf{p}(\mathbf{x}-\mathbf{y})}. \tag{6.207}$$

This does not vanish for $\mathbf{x} \neq \mathbf{y}$: the commutator is not 'local' in the sense that it vanishes for spacelike distances. On the other hand, the vacuum expectation value of the anticommutator is simple and local:

$$\begin{aligned}
\langle 0|\{\psi_a(x), \bar{\psi}_b(y)\}|0\rangle_{x^0=y^0} &= i\gamma^0_{ab}\,\delta(\mathbf{x}-\mathbf{y}), && (6.208)\\
\langle 0|\{\psi_a(x), \psi_b^\dagger(y)\}|0\rangle_{x^0=y^0} &= \delta_{ab}\,\delta(\mathbf{x}-\mathbf{y}). && (6.209)
\end{aligned}$$

The spinor operators at different points in space evidently do not commute at equal times: they anticommute! This then leads to the conclusion that multiparticle states created by repeated application of the spinor field on the vacuum state are antisymmetric in interchange of labels: the particles have to be fermions.

Let us list the important ingredients which went into this celebrated spin-statistics connection:

- Hilbert space (of course with positive metric);
- a vacuum state $|0\rangle$ and one particle states $|p, \lambda\rangle$ with the expected energy momentum eigenvalues (6.192);
- translation invariance and Lorentz invariance;
- locality.

We stress here the relevance of the locality principle. Imagine constructing local observables $O(x)$ out of the spinor field. We want these to be local, i.e. they should commute for spacelike separations,

$$[O(x), O(y)] = 0, \quad (x-y)^2 > 0. \tag{6.210}$$

The spinor fields are *not* local in this sense, because anticommutators are not commutators, and apparently spinor fields are not observables. However, 'bilinears' of the type (Γ is some combination of Dirac matrices)

$$O(x, \Gamma) = \bar{\psi}(x)\Gamma\psi(x), \tag{6.211}$$

(and generalizations thereoff, e.g. involving derivatives), do satisfy locality. This follows from the fact that we can express commutators of field bilinears in terms of anticommutators of fields. The anticommutators satisfy locality, and therefore also the commutators of the bilinears,

$$[O(x, \Gamma_1), O(y, \Gamma_2)] = 0, \quad (x-y)^2 > 0. \tag{6.212}$$

Familiar observables like currents and the energy momentum tensor can indeed be expressed as 'bilinears'. Had we insisted on commutation relations for $\psi(x)$, we would have had to assume a nonlocal commutator $[\psi_a(x), \bar{\psi}_b(y)]$, as follows from its vacuum expectation value (6.207), and we could not have satisfied the locality principle.

6.9 Appendix: Anticommuting variables

Because fermion variables anticommute, the variation of the action can be written in two equivalent but different ways

$$\delta S = \int d^4x\, \delta\psi_a(x)\, \frac{\delta S}{\delta\psi_a(x)} = \int d^4x\, S\, \frac{\overleftarrow{\delta}}{\delta\psi_a(x)}\, \delta\psi_a(x), \tag{6.213}$$

and correspondingly we have to distinguish between left and right derivatives. To see this in more detail let us write the Majorana action in a condensed notation,

$$S = -\int d^4x\, \frac{1}{2}\, \psi^T\beta(m + \gamma\partial)\psi \equiv \frac{1}{2}\, S_{kl}\, \psi^k\psi^l, \tag{6.214}$$

where $S_{kl} = -S_{lk}$. Then

$$\begin{aligned}
\delta S &= S(\psi + \delta\psi) - S(\psi) \\
&= \frac{1}{2} S_{kl}\left[(\psi^k + \delta\psi^k)(\psi^l + \delta\psi^l) - \psi^k\psi^l\right] \\
&= \frac{1}{2} S_{kl}\,(\delta\psi^k\,\psi^l + \psi^k\,\delta\psi^l) \\
&= S_{kl}\,\delta\psi^k\,\psi^l = -S_{kl}\psi^l\,\delta\psi^k. \qquad (6.215)
\end{aligned}$$

Hence

$$\frac{\delta}{\delta\psi^k}\,S = S_{kl}\,\psi^l, \quad S\,\frac{\overleftarrow{\delta}}{\delta\psi^k} = -S_{kl}\,\psi^l. \qquad (6.216)$$

The differentiations also behave like anticommuting variables, e.g.

$$\frac{\delta}{\delta\psi^k}\,\psi^l = \delta^l_k, \qquad (6.217)$$

$$\frac{\delta}{\delta\psi^k}\,(\psi^l\psi^m) = \delta^l_k\,\psi^m - \psi^l\,\delta^m_k. \qquad (6.218)$$

We shall always use left derivatives. Notice that $\delta^{(2)}S/\delta\psi^l\delta\psi^k$ (i.e. first $\delta/\delta\psi^k$ and then $\delta/\delta\psi^l$) equals S_{kl}.

To define integration, consider two anticommuting variable ζ and η. Since $\zeta^2 = \eta^2 = 0$, the most general function of ζ and η has the form

$$f(\zeta, \eta) = f_0 + f_1\zeta + f_2\eta + f_{12}\zeta\eta, \qquad (6.219)$$

with the f's ordinary real or complex numbers. In the path integral is was crucial to assume translation invariance of the 'measure', so we require the integral over ζ to satisfy

$$\int d\zeta\, f(\zeta, \eta) = \int d\zeta\, f(\zeta + \eta, \eta). \qquad (6.220)$$

Suppose $f_{12} = 0$. Comparing the coefficients of η we get

$$\int d\zeta\, f_2\eta = \int d\zeta\,(f_1 + f_2)\eta, \qquad (6.221)$$

from which we conclude

$$\int d\zeta = 0. \qquad (6.222)$$

The remaining term

$$\int d\zeta\, f_1\zeta \qquad (6.223)$$

needs to be given some value. The simplest choice is

$$\int d\zeta\,\zeta = 1, \qquad (6.224)$$

such that $\int d\zeta\, f_1\zeta = f_1$. Then, if f_{12} is also non-zero,

$$\int d\zeta\, f(\zeta,\eta) = f_1 + f_{12}\eta. \tag{6.225}$$

Symbols like $d\zeta$ are also treated as anticommuting. The odd looking integration rules (6.222,6.224) are known as Berezin integration. Note that fermionic integration is just differentiation!

Consider now a finite set of anticommuting numbers ζ_k, $k = 1, \cdots, n$. These generate a so-called Grassmann algebra: indeed, anticommuting 'variables' are nothing but generators of such an algebra. We enlarge the set by new generators ζ_k^* and stipulate that under 'complex conjugation'

$$(c\zeta_k)^* = c^*\zeta_k^*, \quad (c\zeta_k\zeta_l)^* = c^*\zeta_l^*\zeta_k^*, \tag{6.226}$$

where c is a complex number. The above rule has already been used in the discussion of the reality of the Dirac action.

We now have the interesting formula

$$\int d\zeta^* d\zeta\, e^{-\zeta^\dagger M\zeta} = \det M, \quad d\zeta^* d\zeta \equiv \prod_{k=1}^{n} d\zeta_k^* d\zeta_k, \tag{6.227}$$

where M is any square matrix. The derivation is straightforward. We have

$$\begin{aligned}
\int d\zeta^* d\zeta\, \zeta_{k_1}\zeta_{l_1}^* \cdots \zeta_{k_n}\zeta_{l_n}^* &= \int d\zeta_1^* \cdots d\zeta_n^* d\zeta_n \cdots d\zeta_1\, \zeta_{k_1} \cdots \zeta_{k_n}\zeta_{l_n}^* \cdots \zeta_{l_1}^* \\
&= \int d\zeta_1^* \cdots d\zeta_n^* d\zeta_n \cdots d\zeta_1\, \epsilon_{k_1\cdots k_n}\epsilon_{l_1\cdots l_n}\zeta_1 \cdots \zeta_n\zeta_n^* \cdots \zeta_1^* \\
&= \epsilon_{k_1\cdots k_n}\epsilon_{l_1\cdots l_n} = \sum_\pi (-1)^\pi\, \delta_{k_1,l_{\pi 1}} \cdots \delta_{k_n,l_{\pi n}}. \qquad (6.228)
\end{aligned}$$

Then

$$\begin{aligned}
\int d\zeta^* d\zeta\, e^{-\zeta_l^* M_{lk}\zeta_k} &= \int d\zeta^* d\zeta\, \frac{(+1)^n}{n!}\, \zeta_{k_1}\zeta_{l_1}^* \cdots \zeta_{k_n}\zeta_{l_n}^* M_{l_1k_1} \cdots M_{l_nk_n} \\
&= \frac{1}{n!}\, \epsilon_{k_1\cdots k_n}\epsilon_{l_1\cdots l_n} M_{l_1k_1} \cdots M_{l_nk_n} = \epsilon_{k_1\cdots k_n} M_{1k_1} \cdots M_{nk_n} \\
&= \det M. \qquad (6.229)
\end{aligned}$$

Using translation invariance[10] we can also derive

$$\int d\zeta^* d\zeta\, e^{-\zeta^\dagger M\zeta + \eta^\dagger\zeta + \zeta^\dagger\eta} = \det M\, e^{\eta^\dagger M^{-1}\eta}. \tag{6.230}$$

These formulas need of course modification if M does not have an inverse. However, note that the expression on the left-hand side is defined for *arbitrary M*

[10] We shall always assume a sufficiently large set of anticommuting generators to be able to introduce 'sources' η, η^*, etc.

Writing an arbitray function $f(\zeta,\zeta^*)$ as

$$f(\zeta,\zeta^*) = \sum_{p,q=0}^{n} \frac{1}{p!q!}\, \zeta_{k_1}\cdots\zeta_{k_p}\zeta^*_{l_q}\cdots\zeta^*_{l_1} f_{k_1\cdots k_p,l_1\cdots l_q}, \tag{6.231}$$

with coefficients completely antisymmetric in permutations on the k's or l's, it is easy to derive along similar lines that

$$\begin{aligned}\int d\zeta^* d\zeta\, f(\zeta,\zeta^*) &= f_{1\cdots n,1\cdots n} &(6.232)\\ &= \int d\zeta^* d\zeta\, f(\zeta+\eta,\zeta^*+\eta^*), &(6.233)\\ \int d\zeta^* d\zeta\, f(A\zeta, B\zeta^*) &= \det A \det B \int d\zeta^* d\zeta\, f(\zeta,\zeta^*), &(6.234)\end{aligned}$$

which we may interprete as a property of the fermionic 'measure' under linear transformations of variables:

$$d(A\zeta+\eta) = (\det A)^{-1} d\zeta, \quad d(B\zeta^*+\eta^*) = (\det B)^{-1} d\zeta^*. \tag{6.235}$$

Note that the effect of the matrices A and B is opposite to the bosonic case, for which $d^n(A\phi) = \det A\; d^n\phi$.

Corresponding formulas can be derived for the real case and antisymmetric M, $M = -M^T$. Since the complex case can be viewed as a doubling of the real case we get (up to a possible sign depending on the ordering in the definition of $d\zeta$)

$$\int d\zeta\, e^{-\frac{1}{2}\zeta^T M \zeta} = \sqrt{\det M}, \tag{6.236}$$

which is a monomial in the matrix elements $M_{kl} = -M_{lk}$. It is called a pfaffian.

Last, but not least, we can formulate quantum mechanics of fermions in terms of wave functions $\phi(\zeta^*) = \langle\zeta|\phi\rangle$ representing a state $|\phi\rangle$ and matrix elements $O(\zeta_1^*,\zeta_2) = \langle\zeta_1|\hat{O}|\zeta_2\rangle$ representing operators $\hat{O}$ in Hilbert space. This is usually done in terms of fermionic coherent states. See for example the book by this author, appendix C.

6.10 Problems

1. *Exercize*

   Verify eq. (6.37).

2. *Consistency check*

   In the correct quantization with anticommutators, verify the analogue of (6.21) for antiparticle states, for J_3:

$$J_3\, |\overline{p\lambda}\rangle = \frac{1}{2}\,\lambda\, |\overline{p\lambda}\rangle, \quad \bar{p} \equiv (m, \mathbf{0}) \tag{6.237}$$

(where $|\overline{p\lambda}\rangle = d^\dagger(p,\lambda)\,|0\rangle$). Assume the vacuum to be rotationally invariant, $\mathbf{J}|0\rangle = 0$.

3. *Matrix elements of currents*

a. For a free Dirac field, verify the following matrix elements of the current $j^\mu = \bar\psi i\gamma^\mu\psi$:

$$\begin{aligned}
\langle 0|j^\mu(x)|0\rangle &= 0, &(6.238)\\
\langle p'\lambda'|j^\mu(x)|p\lambda\rangle &= \bar u' i\gamma^\mu u\, e^{i(p-p')x}, &(6.239)\\
\langle\overline{p'\lambda'}|j^\mu(x)|\overline{p\lambda}\rangle &= -\bar v i\gamma^\mu v'\, e^{i(p-p')x}, &(6.240)\\
\langle p'\lambda', \overline{p\lambda}|j^\mu(x)|0\rangle &= \bar u' i\gamma^\mu v\, e^{-i(p+p')x}, &(6.241)\\
\langle 0|j^\mu(x)|p\lambda, \overline{p'\lambda'}\rangle &= \bar v' i\gamma^\mu u\, e^{i(p+p')x}, &(6.242)
\end{aligned}$$

where $u = u(p,\lambda)$, $\bar u' = \bar u(p',\lambda')$, etc.

Hint:[11] convert the matrix elements to 'path integral averages', e.g.

$$\langle p'\lambda'|j^\mu(x)|p\lambda\rangle = \langle b(p',\lambda',\infty)\,\bar\psi(x)i\gamma^\mu\psi(x)\,b^*(p,\lambda,-\infty)\rangle. \qquad (6.243)$$

b. Using the charge conjugation matrix C introduced in Appendix 5.5 to relate the u's with the v's, verify from the explicit answers obtained above that

$$\langle p'\lambda'|j^\mu(x)|p\lambda\rangle = -\langle\overline{p'\lambda'}|j^\mu(x)|\overline{p\lambda}\rangle. \qquad (6.244)$$

c. Verify $\partial_\mu j^\mu = 0$ in the above matrix elements of the current j^μ.

d. For the explicit expressions obtained above for the matrix elements of j^μ verify that

$$\langle p'\lambda'|Q|p\lambda\rangle = \langle p'\lambda'|p\lambda\rangle, \qquad (6.245)$$

etc., where $Q = \int d^3x\, j^0(x)$.

e. *Gordon decomposition*

Recall that for general μ and ν,

$$S_{\mu\nu} = \frac{-i}{4}[\gamma_\mu, \gamma_\nu]. \qquad (6.246)$$

Let $u = u(p,\lambda)$, $\bar u' = \bar u(p',\lambda')$. Verify that

$$2(p-p')^\nu\, \bar u'\, S_{\mu\nu}\, u = 2m\, \bar u'\gamma_\mu u + i(p+p')_\mu\, \bar u' u, \qquad (6.247)$$

and the Gordon decomposition

$$\bar u' i\gamma^\mu u = \frac{1}{2m}\,[(p+p')^\mu \bar u'\, u + 2i(p-p')_\nu\, \bar u'\, S^{\mu\nu}\, u]. \qquad (6.248)$$

[11] Note that this technique can also be used for other matrix elements, for example, the matrix element (2.133) of the energy-momentum tensor of the scalar field.

4. *Commutators revisited*

   This problem is the analog of Problem 2.5.

   Consider fermion operators ψ_A and $\psi_A^\dagger$ with the anticommutation relations

$$\{\psi_A, \psi_B\} = 0, \quad \{\psi_A^\dagger, \psi_B^\dagger\} = 0, \quad \{\psi_A, \psi_B^\dagger\} = \delta_{AB}. \tag{6.249}$$

   Let M be a matrix with c-number elements M_{AB}, with vanishing trace, $\mathrm{tr}\, M = 0$, and define $O(M) = M_{AB}\psi_A^\dagger\psi_B$ (summation over repeated indices). Verify that

$$[O(M_1), O(M_2)] = O([M_1, M_2]). \tag{6.250}$$

   Hint, for arbitray operators X, Y and Z we have (check) $[XY, Z] = X\{Y, Z\} - \{X, Z\}Y$, and also $[XY, Z] = X[Y, Z] + [X, Z]Y$.

   As an application, let A be the pair of indices $(a, \mathbf{x})$ and $\psi_A = \psi_a(\mathbf{x})$ a Dirac field. Obtain the commutation relations for the angular-momentum operators $J_k = \epsilon_{klm}J_{lm}/2$, for which $M_{AB} \to -i\epsilon_{klm}x_l\partial_m\delta(\mathbf{x}-\mathbf{y})\delta_{ab} + \delta(\mathbf{x}-\mathbf{y})(S_k)_{ab}$.

5. *Fermion-(anti)fermion scattering*

   Derive, along the lines in sect. 6.5, the diagrams and scattering amplitudes for the processes $(p_1\lambda_1) + (p_2\lambda_2) \to (p_1'\lambda_1') + (p_2'\lambda_2')$ and $(p_1\lambda_1) + \overline{(p_2\lambda_2)} \to (p_1'\lambda_1') + \overline{(p_2'\lambda_2')}$. Put all relevant labels in the diagrams, as in Figs. 6.3 and 6.4.

   Answers for these (anti)fermion scattering amplitudes:

$$\begin{aligned} -iT(p_1'\lambda_1', p_2'\lambda_2'; p_1\lambda_1, p_2\lambda_2) &= \bar{u}_1'(-ig\Gamma)u_1\,\bar{u}_2'(-ig\Gamma)u_2\,\frac{-i}{m^2 + (p_1 - p_1')^2} \\ &\quad - \bar{u}_2'(-ig\Gamma)u_1\,\bar{u}_1'(-ig\Gamma)u_2\,\frac{-i}{m^2 + (p_1 - p_2')^2} \end{aligned} \tag{6.251}$$

   and

$$\begin{aligned} -iT(p_1'\lambda_1', \overline{p_2'\lambda_2'}; p_1\lambda_1, \overline{p_2\lambda_2}) &= -\bar{u}_1'(-ig\Gamma)u_1\,\bar{v}_2(-ig\Gamma)v_2'\,\frac{-i}{m^2 + (p_1 - p_1')^2} \\ &\quad + \bar{u}_1'(-ig\Gamma)v_2'\,\bar{v}_2(-ig\Gamma)u_1\,\frac{-i}{m^2 + (p_1 + p_2)^2}, \end{aligned} \tag{6.252}$$

   where $\bar{u}_2' = \bar{u}(p_2', \lambda_2')$ etc.

6. *Fermion-antifermion annihilation and creation*

Give the diagrams (with labels) and scattering amplitudes for the processes $(p\lambda) + \overline{(p'\lambda')} \to k + k'$ and $k + k' \to (p\lambda) + \overline{(p'\lambda')}$. In these two problems do not give the derivation but make sure you are able to do so.

Answers for these fermion-antifermion annihilation or creation amplitudes:

$$\begin{aligned} -iT(k, k'; p\lambda, \overline{p'\lambda'}) &= \bar{v}(p', \lambda')(-ig\Gamma)(-i)\frac{M - i\gamma(p-k)}{M^2 + (p-k)^2}(-ig\Gamma)u(p, \lambda) \\ &\quad + k \leftrightarrow k', \end{aligned} \tag{6.253}$$

and

$$\begin{aligned} -iT(p\lambda, \overline{p'\lambda'}; k, k') &= \bar{u}(p, \lambda)(-ig\Gamma)(-i)\frac{M - i\gamma(p-k)}{M^2 + (p-k)^2}(-ig\Gamma)v(p', \lambda') \\ &\quad + k \leftrightarrow k'. \end{aligned} \tag{6.254}$$

Chapter 7

Quantized electromagnetic field

The quantization of the electromagnetic field poses special problems due to gauge invariance. We first quantize canonically in the Coulomb gauge. The resulting formulation is logically simple, but its lack of manifest covariance can be cumbersome in practical calculations. This problem is remedied by using the path integral in transforming to a generalized covariant gauge. In the resulting formulation feature unphysical 'particles', named ghosts, and Hilbert space has to be extended into a vector space with indefinite metric. However, the connection with the Coulomb gauge should guarantee correct physics, whereas manifest covariance will make the calculation of perturbative corrections in interacting gauge-field theory a lot simpler.

7.1 Quantization in the Coulomb gauge

We rewrite the Maxwell action with external current in the form

$$S = \int dt\, L, \tag{7.1}$$

$$L = \int d^3x\, (-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + J^\mu A_\mu) \tag{7.2}$$

$$= \int d^3x\, \Big(\frac{1}{2}\dot{A}_m\dot{A}_m - \dot{A}_m\partial_m A_0 + \frac{1}{2}\partial_n A_m\partial_m A_n - \frac{1}{2}\partial_m A_n\partial_m A_n + \frac{1}{2}\partial_m A_0\partial_m A_0 + J^0 A_0 + J^m A_m\Big) \tag{7.3}$$

There is now a complication in the canonical formalism which is typical for gauge theories: $\dot{A}_0$ is lacking in L, so the canonical conjugate to A_0 vanishes, $\Pi_0 \equiv \delta L/\delta\dot{A}_0 = 0$. One way to deal with this is to eliminate A^0 itself as a dynamical variable by choice of gauge condition. Consider the equation of motion that follows from varying the action with respect to A_0,

$$0 = \delta S = \int d^4x\, (\partial_\mu F^{\mu 0} + J^0)\delta A_0. \tag{7.4}$$

This is Gauss's law, or Coulomb's law,

$$\begin{aligned} 0 = \frac{\delta S}{\delta A_0} &= \partial_m F^{m0} + J^0 = -\nabla \cdot \mathbf{E} + J^0 \\ &= -\partial_m(-\partial_m A^0 - \partial_0 A^m) + J^0. \end{aligned} \tag{7.5}$$

Now impose the Coulomb gauge condition

$$\partial_m A_m = 0. \tag{7.6}$$

This has the effect that the time derivative drops out of (7.5), such that (7.5) takes the form

$$-\Delta A^0 = J^0, \quad \Delta \equiv \nabla^2. \tag{7.7}$$

Since this equation does not contain time derivatives, it is not a dynamical equation anymore, but an equation of constraint at every instant in time. With suitable spatial boundary conditions the potential A^0 is completely determined in terms of J^0. For infinite space

$$A^0(\mathbf{x}, t) = \int d^3y\, \frac{1}{4\pi|\mathbf{x} - \mathbf{y}|} J^0(\mathbf{y}, t), \tag{7.8}$$

where we used the fact that the Coulomb potential is a Green function for the laplacian Δ:

$$-\Delta\, \frac{1}{4\pi|\mathbf{x} - \mathbf{y}|} = \delta(\mathbf{x} - \mathbf{y}). \tag{7.9}$$

Hence, A^0 is not a dynamical variable in the Coulomb gauge.

We shall use the Coulomb gauge for the canonical formalism and continue to write A^0, for simplicity, keeping in mind that it is a given function of J^0. In this gauge we can rewrite the lagrangian in the form

$$L = \int d^3x \left[\frac{1}{2}\dot{A}_m \dot{A}_m - \frac{1}{2}A_m(-\Delta)A_m + J_m A_m\right] - E_C, \tag{7.10}$$

$$E_C = \int d^3x \left(-\frac{1}{2}\partial_m A^0 \partial_m A^0 + J^0 A^0\right) = \int d^3x\, \frac{1}{2} J^0 A^0. \tag{7.11}$$

We used $\partial_m A_m = 0$, $\Delta A^0 = -J^0$ and made partial integrations of ∂_m assuming boundary conditions such that surface terms vanish. The quantity E_C is the Coulomb energy. Using (7.8) it can be written as

$$E_C = \frac{1}{2} \int d^3x\, d^3y\, J^0(\mathbf{x}, t) \frac{1}{4\pi|\mathbf{x} - \mathbf{y}|} J^0(\mathbf{y}, t). \tag{7.12}$$

The lagrangian is now in the form $L(q, \dot{q})$ with $q_k(t) \to A_m(\mathbf{x}, t)$.

Next we have to deal with the continuous nature of the index $\mathbf{x}$ and the constraint $\partial_m A_m(\mathbf{x}, t) = 0$. This can be done by expanding the potentials into a discrete set of basis functions $f_\alpha^m(\mathbf{x})$ satisfying $\partial_m f_\alpha^m(\mathbf{x}) = 0$. Let us enclose the system in a cubic box $-L/2 \leq x_m \leq L/2$ with periodic boundary conditions. For a large enough box its finiteness and the type of boundary conditions should not matter. Periodic boundary conditions are convenient because with it all boundary terms in partial integrations vanish (the box has no boundary) and they do not spoil translation invariance. We can use the discrete set of eigenfunctions of the laplacian Δ to construct the $f_\alpha^m(\mathbf{x})$. The real eigenfunctions of the laplacian correspond to products of the standing waves $\cos(k_1 x_1)\cos(k_2 x_2)\cos(k_3 x_3)$, $\sin(k_1 x_1)\cos(k_2 x_2)\cos(k_3 x_3), \ldots, \sin(k_1 x_1)\sin(k_2 x_2)\sin(k_3 x_3)$, with $k_m = 2\pi n_m/L$, $n_m = 0, 1, 2, \ldots$, and the eigenvalues are given by $-\Delta \to \omega_\alpha^2 = \mathbf{k}_\alpha^2$. Out of these eigenfunctions the real $f_\alpha^m(\mathbf{x})$ can be constructed satisfying $\partial_m f_\alpha^m(\mathbf{x}) = 0$. The details of this are tedious and not needed in the following and we shall just record their properties:

$$-\Delta f_\alpha^m(\mathbf{x}) = \omega_\alpha^2 f_\alpha^m(\mathbf{x}), \quad \partial_m f_\alpha^m(\mathbf{x}) = 0, \tag{7.13}$$

$$\int d^3x\, f_\alpha^m(\mathbf{x})^* f_\beta^m(\mathbf{x}) = \delta_{\alpha\beta}, \tag{7.14}$$

$$\sum_\alpha f_\alpha^m(\mathbf{x}) f_\alpha^n(\mathbf{y})^* = P_{mn}^T(\mathbf{x}, \mathbf{y}). \tag{7.15}$$

We have written these equations in general complex form, because we usually use the plane waves that are eigenfunctions of the gradient operator ∂_n, which are complex. In the next section we shall give an explicit construction of these basis functions. For the moment have to keep in mind that the $f_\alpha^m(\mathbf{x})$ are real. The object $P_{mn}^T(\mathbf{x}, \mathbf{y})$ is a projector on the space of 'transverse' vector functions, i.e. a projector, $P^2 = P$, or

$$\int d^3y\, P_{kl}^T(\mathbf{x}, \mathbf{y})\, P_{lm}^T(\mathbf{y}, \mathbf{z}) = P_{km}^T(\mathbf{x}, \mathbf{z}), \tag{7.16}$$

that is transverse, $\partial_m P_{mn}^T(\mathbf{x}, \mathbf{y}) = 0$. It is the identity operator for transverse vector functions (satisfying $\partial_m A_m(\mathbf{x}) = 0$),

$$\int d^3y\, P_{mn}^T(\mathbf{x}, \mathbf{y}) A_n(\mathbf{y}) = A_m(\mathbf{x}), \quad \text{if } \partial_m A_m = 0. \tag{7.17}$$

An explicit expression for P^T will be given in the next section (cf. (7.42)). In the summation $\sum_\alpha$ we exclude the 'zero mode' $\mathbf{k} = (0, 0, 0)$ (this would be automatic with Dirichlet boundary conditions). This means that we exclude here potentials A_m which are constant in space. Such potentials complicate the (otherwise interesting) mathematics and we usually do not need them in physical applications.

In terms of these basis functions we can now expand the potentials in normal modes,

$$A_m(\mathbf{x},t) = \sum_\alpha q_\alpha(t) f_\alpha^m(\mathbf{x}), \tag{7.18}$$

$$q_\alpha(t) = \int d^3x\, f_\alpha^m(\mathbf{x}) A_m(\mathbf{x},t), \tag{7.19}$$

and in terms of the new coordinates q_α the lagrangian takes the form, for $J^\mu = 0$,

$$L = \sum_\alpha \left(\frac{1}{2}\dot{q}_\alpha \dot{q}_\alpha - \frac{1}{2}\omega_\alpha^2 q_\alpha q_\alpha \right). \tag{7.20}$$

This shows that the electromagnetic field is equivalent to an infinite set of harmonic oscillators, with unit mass and frequencies ω_α. The canonical description is now an obvious generalization of the case of one harmonic oscillator,

$$p_\alpha = \partial L/\partial \dot{q}_\alpha = \dot{q}_\alpha, \tag{7.21}$$

$$H = \sum_\alpha \left(\frac{1}{2} p_\alpha p_\alpha + \frac{1}{2}\omega_\alpha^2 q_\alpha q_\alpha \right), \tag{7.22}$$

$$(q_\alpha, p_\beta) = \delta_{\alpha\beta}, \quad (q_\alpha, q_\beta) = (p_\alpha, p_\beta) = 0, \tag{7.23}$$

$$\dot{p}_\alpha = (p_\alpha, H), \quad \dot{q}_\alpha = (q_\alpha, H) = p_\alpha, \tag{7.24}$$

where (A, B) is the Poisson bracket. Evidently, the canonical conjugate to the field $A_m(\mathbf{x})$ is

$$\begin{aligned} \Pi_m(\mathbf{x}) &= (A_m(\mathbf{x}), H) = \dot{A}_m(\mathbf{x}) \\ &= \sum_\alpha p_\alpha f_\alpha^m(\mathbf{x}). \end{aligned} \tag{7.25}$$

The system is quantized by imposing canonical commutation relations between the p's and q's,

$$[q_\alpha, p_\beta] = i\delta_{\alpha\beta}, \quad [q_\alpha, q_\beta] = [p_\alpha, p_\beta] = 0. \tag{7.26}$$

The quantized electromagnetic field is now an operator in Hilbert space. The commutation relations between the p_α and q_α imply the following relations between A_m and Π_m,

$$[A_m(\mathbf{x}), \Pi_n(\mathbf{y})] = iP^T_{mn}(\mathbf{x},\mathbf{y}), \quad [A_m(\mathbf{x}), A_n(\mathbf{y})] = [\Pi_m(\mathbf{x}), \Pi_n(\mathbf{y})] = 0. \tag{7.27}$$

For example,

$$\begin{aligned} [A_m(\mathbf{x}), \Pi_n(\mathbf{y})] &= \sum_{\alpha\beta} [q_\alpha, p_\beta]\, f_\alpha^m(\mathbf{x}) f_\beta^n(\mathbf{y})^* = i \sum_\alpha f_\alpha^m(\mathbf{x}) f_\alpha^n(\mathbf{y})^* \\ &= iP^T_{mn}(\mathbf{x}.\mathbf{y}). \end{aligned} \tag{7.28}$$

7.2 Energy-momentum eigenstates

To guide our physical interpretation we shall use the energy momentum P^μ of the field, which is now also an operator, and determine its eigenstates and eigenvalues. In the Coulomb gauge A_0 vanishes when $J^\mu = 0$, cf. (7.8). Then (cf. (1.78))

$$\begin{aligned} T^{00} &= \frac{1}{2}(E_m E_m + B_m B_m) \\ &= \frac{1}{2}(\dot{A}_m \dot{A}_m + \partial_n A_m \partial_n A_m - \partial_n A_m \partial_m A_n), \qquad (7.29) \\ T^{0n} &= \epsilon_{nmp} E_m B_p = -\dot{A}_m \partial_n A_m + \dot{A}_m \partial_m A_n, \qquad (7.30) \end{aligned}$$

giving

$$\begin{aligned} P^0 &= \int d^3x\, T^{00} = \int d^3x\, \left[\frac{1}{2}\Pi_m \Pi_m + \frac{1}{2} A_m(-\Delta) A_m\right], \qquad (7.31) \\ P^n &= \int d^3x\, T^{0n} = \int d^3x\, (-\Pi_m \partial_n A_m). \qquad (7.32) \end{aligned}$$

We used the Coulomb gauge condition $\partial_m A_m = 0$ and $\dot{A}_m = \Pi_m$. Notice that there is no operator-ordering ambiguity in P_m: we can also write Π_m to the right of A_m, the difference involves the derivative of the commutator, $\partial_m \delta(\mathbf{x}-\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = 0$. Using the normal mode expansion we find

$$\begin{aligned} P^0 &= \sum_\alpha \left(\frac{1}{2} p_\alpha p_\alpha + \frac{1}{2}\omega_\alpha^2 q_\alpha q_\alpha\right) \\ &= H. \qquad (7.33) \end{aligned}$$

The momentum operator is less easy to express in terms of the normal modes because the real mode functions $f_\alpha^m(\mathbf{x})$ are not eigenfunctions of ∂_n. Therefore we now introduce a different set $f_\alpha^m(\mathbf{x})$ which are eigenfunctions of the hermitian differential operators $i\partial_n$ and Δ, and satisfy $\partial_m f_\alpha^m(\mathbf{x}) = 0$. They are complex and have the form[1] ($\alpha \to (\mathbf{k}, \lambda)$)

$$f_{\mathbf{k},\lambda}^m(\mathbf{x}) = e^m(\mathbf{k}, \lambda) e^{i\mathbf{k}\mathbf{x}}, \quad k_m = n_m 2\pi/L, \quad n_m = 0, \pm 1, \pm 2, \ldots, \qquad (7.34)$$

These are clearly eigenfunctions of ∂_n and Δ. Recall that the n_m have to be integers to satisfy periodic boundary conditions in a box of size L^3. To satisfy $\partial_m f^m = 0$, the $e^m(\mathbf{k}, \lambda)$ have to be orthogonal to $\mathbf{k}$ (hence the terminology 'transverse'), as illustrated in fig. 7.1,

$$\mathbf{k} \cdot \mathbf{e}(\mathbf{k}, \lambda) = 0. \qquad (7.35)$$

For example for $\mathbf{k} = (0, 0, |\mathbf{k}|) = |\mathbf{k}|\hat{3}$, $\mathbf{e}(\mathbf{k}, 1) = (1, 0, 0) = \hat{1}$, $\mathbf{e}(\mathbf{k}, 2) = (0, 1, 0) =$

[1] We use here a normalization of the Fourier modes that treats the volume-factors L^3 differently from the scalar and fermion cases (for no good reason).

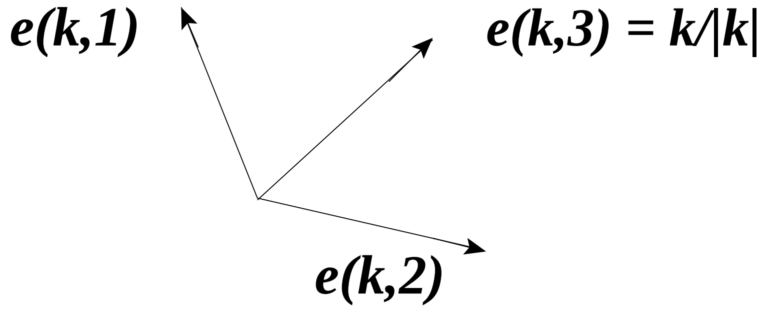

Figure 7.1: Real transverse polarization vectors $\mathbf{e}(\mathbf{k},1)$ and $\mathbf{e}(\mathbf{k},2)$, and the longitudinal unit vector $\mathbf{e}(\mathbf{k},3)$.

$\hat{2}$. In general the $\mathbf{e}(\mathbf{k},\lambda)$ may be obtained from this by a rotation, a standard rotation that takes $(0,0,|\mathbf{k}|)$ into $\mathbf{k}$. Another set, which is also well-known from classical electrodynamics, consists of the right- and left-handed polarization vectors

$$\mathbf{e}(\mathbf{k},\pm) = \mp\frac{1}{\sqrt{2}}[\mathbf{e}(\mathbf{k},1) \pm i\mathbf{e}(\mathbf{k},2)]. \tag{7.36}$$

The polarization vectors satisfy

$$e_m(\mathbf{k},\lambda)^* e_m(\mathbf{k},\lambda') = \delta_{\lambda\lambda'}, \tag{7.37}$$

$$\sum_\lambda e_m(\mathbf{k},\lambda)e_n(\mathbf{k},\lambda)^* = \left(\delta_{mn} - \frac{k_m k_n}{\mathbf{k}^2}\right) \equiv P^T_{mn}(\mathbf{k}). \tag{7.38}$$

The basis functions are orthogonal and complete in the sense (7.15), with

$$\alpha \rightarrow (\mathbf{k},\lambda), \tag{7.39}$$

$$\delta_{\alpha\alpha'} \rightarrow \delta_{\lambda\lambda'}\, L^3 \delta_{\mathbf{k},\mathbf{k}'}, \tag{7.40}$$

$$\sum_\alpha \rightarrow \frac{1}{L^3}\sum_{\mathbf{k},\lambda}, \tag{7.41}$$

$$P^T_{mn}(\mathbf{x},\mathbf{y}) = \frac{1}{L^3}\sum_{\mathbf{k}} e^{-i\mathbf{k}\mathbf{x}+i\mathbf{k}\mathbf{y}}\left(\delta_{mn} - \frac{k_m k_n}{\mathbf{k}^2}\right),. \tag{7.42}$$

where the zero mode $\mathbf{k} = \mathbf{0}$ is absent again.

We now expand the A_m and Π_m in terms of these basis functions as follows,

$$A_m(\mathbf{x}) = \frac{1}{L^3}\sum_{\mathbf{k},\lambda}\frac{1}{2k^0}[e^{i\mathbf{k}\mathbf{x}}e^m(\mathbf{k},\lambda)a(\mathbf{k},\lambda) + e^{-i\mathbf{k}\mathbf{x}}e^m(\mathbf{k},\lambda)^* a^\dagger(\mathbf{k},\lambda)], \tag{7.43}$$

$$\Pi_m(\mathbf{x}) = \frac{1}{L^3}\sum_{\mathbf{k},\lambda}\frac{1}{2k^0}[-ik^0 e^{i\mathbf{k}\mathbf{x}}e^m(\mathbf{k},\lambda)a(\mathbf{k},\lambda) + ik^0 e^{-i\mathbf{k}\mathbf{x}}e^m(\mathbf{k},\lambda)^* a^\dagger(\mathbf{k},\lambda)],$$

where

$$k^0 = |\mathbf{k}| \tag{7.44}$$

and the factor $1/2k^0$ anticipates covariant normalization. The form of (7.43) is guided by the harmonic oscillator forms,

$$q_\alpha = \frac{1}{2\omega_\alpha}\sqrt{2\omega_\alpha}\,(a_\alpha + a_\alpha^\dagger), \tag{7.45}$$

$$p_\alpha = \frac{1}{2\omega_\alpha}\sqrt{2\omega_\alpha}\,(-i\omega_\alpha a_\alpha + i\omega_\alpha a_\alpha^\dagger). \tag{7.46}$$

The relations (7.43) define $a(\mathbf{k},\lambda)$ and $a^\dagger(\mathbf{k},\lambda)$, and may be inverted as follows. We write

$$a_m(\mathbf{k}) = \sum_\lambda e_m(\mathbf{k},\lambda)a(\mathbf{k},\lambda), \tag{7.47}$$

$$a(\mathbf{k},\lambda) = e_m(\mathbf{k},\lambda)^* a_m(\mathbf{k}). \tag{7.48}$$

Then

$$\int d^3x\, e^{-i\mathbf{kx}} A_m(\mathbf{x}) = \frac{1}{2k^0}[a_m(\mathbf{k}) + a_m^\dagger(-\mathbf{k})], \tag{7.49}$$

$$\int d^3x\, e^{-i\mathbf{kx}} \Pi_m(\mathbf{x}) = \frac{1}{2}[-ia_m(\mathbf{k}) + ia_m^\dagger(-\mathbf{k})], \tag{7.50}$$

giving

$$a_m(\mathbf{k}) = \int d^3x\, e^{-i\mathbf{kx}}[k^0 A_m(\mathbf{x}) + i\Pi_m(\mathbf{x})], \tag{7.51}$$

$$a_m^\dagger(-\mathbf{k}) = \int d^3x\, e^{-i\mathbf{kx}}[k^0 A_m(\mathbf{x}) - i\Pi_m(\mathbf{x})]. \tag{7.52}$$

The commutation relations between $a_m(\mathbf{k})$ and $a_m^\dagger(\mathbf{k})$ can now be calculated from (7.27) to be

$$\begin{aligned} [a_m(\mathbf{k}), a_n^\dagger(\mathbf{l})] &= P_{mn}^T(\mathbf{k})\, 2k^0 L^3 \delta_{\mathbf{k},\mathbf{l}}, \\ [a_m(\mathbf{k}), a_n(\mathbf{l})] &= [a_m^\dagger(\mathbf{k}), a_n^\dagger(\mathbf{l})] = 0. \end{aligned} \tag{7.53}$$

For example,

$$\begin{aligned} [a_m(\mathbf{k}), a_n^\dagger(\mathbf{l})] &= \int d^3x d^3y\, e^{-i\mathbf{kx}+i\mathbf{ly}}\,[k^0 A_m(\mathbf{x}) + i\Pi_m(\mathbf{x}), l^0 A_m(\mathbf{y}) - i\Pi_m(\mathbf{y})] \\ &= (k^0 + l^0)\int d^3x d^3y\, e^{-i\mathbf{kx}+i\mathbf{ly}}\, P_{mn}^T(\mathbf{x},\mathbf{y}) \\ &= (k^0 + l^0) P_{mn}^T(\mathbf{l}) \int d^3y\, e^{i(\mathbf{l}-\mathbf{k})\mathbf{y}} \\ &= 2k^0 P_{mn}^T(\mathbf{k})\, L^3 \delta_{\mathbf{k},\mathbf{l}}. \end{aligned} \tag{7.54}$$

It follows that

$$[a(\mathbf{k},\lambda), a^\dagger(\mathbf{k}',\lambda')] = 2k^0 L^3 \delta_{\mathbf{k},\mathbf{k}'}\delta_{\lambda,\lambda'}, \tag{7.55}$$

$$[a(\mathbf{k},\lambda), a(\mathbf{l},\lambda')] = [a^\dagger(\mathbf{k},\lambda), a^\dagger(\mathbf{l},\lambda')] = 0. \tag{7.56}$$

Hence, the new a and $a^\dagger$ satisfy the commutation relations of creation and annihilation operators of an infinite set of harmonic oscillators labelled by $(\mathbf{k}, \lambda)$.

Expressing the hamiltonian (7.31) and momentum operator (7.32) in terms of the creation and annihilation operators we find (cf. Problem 1)

$$P^0 = \frac{1}{L^3} \sum_{\mathbf{k},\lambda} \frac{1}{2k^0} a^\dagger(\mathbf{k}, \lambda) a(\mathbf{k}, \lambda) k^0 + E_0, \tag{7.57}$$

$$P_m = \frac{1}{L^3} \sum_{\mathbf{k},\lambda} \frac{1}{2k^0} a^\dagger(\mathbf{k}, \lambda) a(\mathbf{k}, \lambda) k_m, \tag{7.58}$$

$$E_0 = \sum_{\mathbf{k},\lambda} \frac{1}{2} k^0. \tag{7.59}$$

By analogy to the ordinary harmonic oscillator we recognize the number operator $a^\dagger(\mathbf{k}, \lambda) a(\mathbf{k}, \lambda)$ for each mode $(\mathbf{k}, \lambda)$. The ground state (state with lowest energy) is the no-quantum state $|0\rangle$ defined by

$$a(\mathbf{k}, \lambda)|0\rangle = 0, \tag{7.60}$$

with

$$P^0|0\rangle = E_0|0\rangle, \quad \mathbf{P}|0\rangle = 0. \tag{7.61}$$

The excited states are given by

$$|k, \lambda\rangle = a^\dagger(\mathbf{k}, \lambda)|0\rangle, \tag{7.62}$$

$$|k_1\lambda_1, k_2\lambda_2\rangle = a^\dagger(\mathbf{k}_1, \lambda_1) a^\dagger(\mathbf{k}_2, \lambda_2)|0\rangle, \tag{7.63}$$

$$|k_1\lambda_1, k_2\lambda_2, k_3\lambda_3\rangle = a^\dagger(\mathbf{k}_1, \lambda_1) a^\dagger(\mathbf{k}_2, \lambda_2) a^\dagger(\mathbf{k}_3, \lambda_3)|0\rangle, \tag{7.64}$$

etc., with

$$[P^\mu - \delta_{\mu,0} E_0]|k_1\lambda_1 \ldots k_n\lambda_n\rangle = (k_1^\mu + \ldots + k_n^\mu)|k_1\lambda_1 \ldots k_n\lambda_n\rangle, \tag{7.65}$$

The four-momenta k^μ correspond to zero mass, $k^\mu k_\mu = 0$. The excited states are the photons, which are massless particles. The symmetry of the basis vectors $|k_1\lambda_1 \cdots k_n\lambda_n\rangle$ under interchange of labels $((\mathbf{k}_i\lambda_i) \leftrightarrow (\mathbf{k}_j\lambda_j)$ has the consequence that photons follow Bose-Einstein statistics.

In the infinite volume limit the ground state represents the vacuum. In this limit the wave vectors become practically continuous, in the sense that for a continuous function $F(\mathbf{k})$,

$$\frac{1}{L^3} \sum_{\mathbf{k}} F(\mathbf{k}) \to \int \frac{d^3k}{(2\pi)^3} F(\mathbf{k}). \tag{7.66}$$

Furthermore, in the sense of generalized functions

$$L^3 \delta_{\mathbf{k},\mathbf{l}} \to (2\pi)^3 \delta(\mathbf{k} - \mathbf{l}). \tag{7.67}$$

Hence

$$\langle k,\lambda|k',\lambda'\rangle \to 2k^0(2\pi)^3\delta(\mathbf{k}-\mathbf{k}')\delta_{\lambda,\lambda'}, \tag{7.68}$$

and the energy density of the ground state takes the form

$$\frac{E_0}{L^3} \to \sum_\lambda \int \frac{d^3k}{(2\pi)^3}\,\frac{1}{2k^0}\,(k^0)^2, \tag{7.69}$$

where we recognize again the Lorentz invariant volume element.

$$d\omega_k \equiv \frac{d^3k}{(2\pi)^3}\,\frac{1}{2k^0}, \quad k^0 = |\mathbf{k}|. \tag{7.70}$$

The energy density is twice ($\sum_\lambda = 2$) that of a massless scalar field. We shall absorb it in a renormalization of the cosmological constant such that

$$E_0 = 0, \tag{7.71}$$

similar to the discussion for the scalar field in sect. 2.3.

7.3 Lorentz invariance

The quantization in Coulomb gauge is straightforward, but not manifestly Lorentz covariant. One might expect covariance to be manifest in a covariant gauge like the Lorentz gauge $\partial_\mu A^\mu = 0$, but quantization is not so easy in this gauge. Later we will make contact with the Lorentz gauge via the path integral. However, the theory in Coulomb gauge should still be Lorentz invariant, because after quantization the generators of Lorentz transformation $J^{\mu\nu}$ exist as operators and they provide a representation of the Lorentz group in Hilbert space, in the usual form $U = \exp(\frac{1}{2}i\omega_{\mu\nu}J^{\mu\nu})$.

Let us take a closer look at this. The Noether form of the energy-momentum tensor turns out to be

$$T^{\mu\alpha}_{\text{Noether}} = F^{\mu\rho}\partial^\alpha A_\rho + \eta^{\mu\alpha}\mathcal{L}, \tag{7.72}$$

which is not symmetric and also not gauge invariant. This tensor can be improved by adding the current

$$j^{\mu\alpha} = -F^{\mu\rho}\partial_\rho A^\alpha, \tag{7.73}$$

which is 'chargeless' when the free Maxwell equations are satisfied,

$$\int d^3x\, j^{0\alpha} = -\int d^3x\, F^{0r}\partial_r A^\alpha = \int d^3x\,(\partial_r F^{0r})\,A^\alpha = 0, \tag{7.74}$$

and conserved

$$\partial_\mu j^{\mu\alpha} = \partial_\mu F^{\mu\rho}\,\partial_\rho A^\alpha + F^{\mu\rho}\,\partial_\mu\partial_\rho A^\alpha = 0. \tag{7.75}$$

The improved tensor

$$T^{\mu\alpha} \equiv T^{\mu\nu}_{\text{Noether}} + j^{\mu\alpha} = F^{\mu\rho}F^{\alpha}_{\ \rho} + \eta^{\mu\alpha}\mathcal{L} \tag{7.76}$$

is gauge invariant, symmetric, conserved, with the same 'charges' P^ν as the Noether charges. It is the one entering in Einstein's equations. Because it is symmetric, the currents

$$J^{\mu\alpha\beta} = x^\alpha T^{\mu\beta} - x^\beta T^{\mu\alpha}, \tag{7.77}$$

are also conserved, $\partial_\mu J^{\mu\alpha\beta} = 0$, and we take it for granted that they are satisfactory generalized angular momentum currents of conserved Lorentz generators

$$J^{\alpha\beta} = \int d^3x\, J^{0\alpha\beta}. \tag{7.78}$$

For example, in Coulomb gauge the angular momenta of the free field (for which $A_0 = 0$) take the canonical form

$$J_k = \frac{1}{2}\epsilon_{klm}J_{lm} = \int d^3x\,(-i\Pi_r)\,[-i\epsilon_{klm}x_l\partial_m\delta_{rs} + (S_k)_{rs}]\,A_s, \tag{7.79}$$

with S_k the spin matrices in the vector representation,[2]

$$(S_k)_{lm} = -i\epsilon_{klm} \tag{7.80}$$

Rotations of the operator fields are generated by taking the commutator $[A_r, J_k]$. With the transverse projector P^T in the commutation relations, this commutator takes the form

$$\begin{aligned}
[A_r(\mathbf{x}), J_k] &= \int d^3x'\, P^T_{rr'}(\mathbf{x},\mathbf{x}')\,[-i\epsilon_{klm}x'_l\partial'_m\delta_{r's} + (S_k)_{r's}]\,A_s(\mathbf{x}'), & (7.81)\\
&= [-i\epsilon_{klm}x_l\partial_m\delta_{rs} + (S_k)_{rs}]\,A_s(\mathbf{x}) + \text{gauge transf.}, & (7.82)
\end{aligned}$$

where the gauge transformation, i.e. a term of the form $\partial_r\omega$, comes from the ∂_r term in $P^T_{rr'}$:

$$P^T_{rr'}(\mathbf{x},\mathbf{x}') = \delta_{rr'}\delta(\mathbf{x}-\mathbf{x}') + \partial_r\partial'_{r'}\frac{1}{4\pi|\mathbf{x}-\mathbf{x}'|}. \tag{7.83}$$

This gauge transformation takes care that $\mathbf{A}$ stays in the Coulomb gauge. In fact, for rotations it does not contribute (make a partial integration for $\partial'_{r'}$), because $\nabla\cdot\mathbf{A} = 0$ is a rotationally invariant equation. However, in general one finds that under Lorentz transformations A^μ transforms like a four-vector field, up to a gauge transformation that keeps it in Coulomb gauge:

$$U^\dagger(\ell)A^\mu(x)U(\ell) = \ell^\mu_{\ \nu}A^\nu(\ell^{-1}x) + \partial^\mu\omega(x), \tag{7.84}$$

with ω such that such that $\nabla\cdot\mathbf{A}(x) = 0$, $A^0(x) = 0$.

[2]With $S_{\alpha\beta}$ in the defining representation (1.52) of the Lorentz group, $S_k = \frac{1}{2}\epsilon_{kab}S_{ab}$.

7.4 Photons

We have seen that the mass of the photon is zero,

$$P_\mu P^\mu|k,\lambda\rangle = (\mathbf{k}^2 - k_0^2)|k,\lambda\rangle = 0. \tag{7.85}$$

We shall now determine its possible helicities. The helicity is defined as the eigenvalue of the angular momentum operator $\mathbf{J}$ in the direction of motion,

$$\hat{k}\cdot\mathbf{J}\,|k,\lambda\rangle = \lambda|k,\lambda\rangle, \quad \hat{k} = \frac{\mathbf{k}}{|\mathbf{k}|}. \tag{7.86}$$

To determine the helicities we use

$$\hat{k}\cdot\mathbf{J}|k,\lambda\rangle = [\hat{k}\cdot\mathbf{J}, a^\dagger(\mathbf{k},\lambda)]\,|0\rangle, \tag{7.87}$$

so we first have to calculate the above commutator. The calculation of $a(k,\lambda)$ in terms of A_l and Π_l is conveniently summarized by the pair of equations

$$A_r(x) = \sum_\lambda \int d\omega_k \left[a(k,\lambda)\, e_r(\mathbf{k},\lambda)\, e^{ikx} + a^\dagger(k,\lambda)\, e_r^*(\mathbf{k},\lambda)\, e^{-ikx}\right], \tag{7.88}$$

$$a(k,\lambda) = \int d^3x\, e^{-ikx}\, e_r^*(\mathbf{k},\lambda)\,(i\overrightarrow{\partial_0} - i\overleftarrow{\partial_0})\, A_r(x), \tag{7.89}$$

which is valid for the free Maxwell field ($kx = \mathbf{kx} - k^0x^0 = \mathbf{kx} - |\mathbf{k}|x^0$). We then have using (7.82),

$$[a(k,\lambda), J_l] = \int d^3x\, e^{-ikx}\, e_r^*(\mathbf{k},\lambda)\,[-i\epsilon_{lmn}x_m\partial_n\delta_{rs} + (S_l)_{rs}]\,(i\overrightarrow{\partial_0} - i\overleftarrow{\partial_0})\, A_s(x). \tag{7.90}$$

Making a partial integration the orbital part of the angular momentum operator, $-i\epsilon_{lmn}x_m\partial_n \to \epsilon_{lmn}x_m k_n$. This is transverse to $\mathbf{k}$, so in $\hat{k}\cdot\mathbf{J} = \hat{k}_l J_l$ this term does not contribute. Thus we have

$$[a(k,\lambda), \hat{k}\cdot\mathbf{J}] = \int d^3x\, e^{-ikx}\,(i\overrightarrow{\partial_0} - i\overleftarrow{\partial_0})\, e_r^*(\mathbf{k},\lambda)\,\hat{k}_l(S_l)_{rs}\, A_s(x), \tag{7.91}$$

and conjugating this expression,

$$[\hat{k}\cdot\mathbf{J}, a^\dagger(k,\lambda)] = -\int d^3x\, e^{ikx}\,(i\overrightarrow{\partial_0} - i\overleftarrow{\partial_0})\, A_r(x)\,\hat{k}\cdot\mathbf{S}_{rs}\, e_s(\mathbf{k},\lambda). \tag{7.92}$$

The problem is reduced to calculating the action of the spin matrix $\mathbf{S}$ on the polarization vectors $\mathbf{e}(\mathbf{k},\lambda)$. Recall that in the vector representation the spin matrices S_1, S_2 and S_3 are represented by $(S_l)_{mn} = -i\epsilon_{lmn}$, which satisfy $[S_k, S_l] = i\epsilon_{klm}S_m$ and $\mathbf{S}^2 = s(s+1) = 1(1+1) = 2$. The right and left handed polarization vectors (7.36) were constructed such that they are eigenvectors of S_3 for the special momentum $\mathbf{k} = |\mathbf{k}|\hat{3}$ (in which case $e_m(\mathbf{k},j) = \delta_{mj}$, $j = 1,2,3$):

$$(S_3)_{mn}\, e_n(\mathbf{k},\pm) = \pm e_m(\mathbf{k},\pm), \quad \mathbf{k} = |\mathbf{k}|\hat{3}, \tag{7.93}$$

with the usual phase relations ($\mathbf{k} = |\mathbf{k}|\hat{3}$)

$$\begin{aligned}(S_1 + iS_2)_{mn} e_n(\mathbf{k}, -) &= \sqrt{2}\, e_m(\mathbf{k}, 0), \quad e_n(\mathbf{k}, 0) \equiv e_n(\mathbf{k}, 3), \\ (S_1 + iS_2)_{mn} e_n(\mathbf{k}, 0) &= \sqrt{2}\, e_m(\mathbf{k}, +),\end{aligned} \tag{7.94}$$

where $e_n(\mathbf{k}, 0)$ is the eigenvector of S_3 with eigenvalue 0. For general $\mathbf{k}$ the polarization vectors may be defined by some standard rotation taking $\hat{3}$ into $\hat{k}$. This does not affect the eigenvalues of the rotation invariant combination $\hat{k}\mathbf{S}$, so finally:

$$\hat{k} \cdot \mathbf{S}_{mn}\, e_n(\mathbf{k}, \pm) = \pm e_m(\mathbf{k}, \pm) \Rightarrow \hat{k} \cdot \mathbf{J}\, |k, \pm\rangle = \pm |k, \pm\rangle, \tag{7.95}$$

where we used (7.87), (7.92) and the conjugate of (7.89). The eigenvector $\mathbf{e}(\mathbf{k}, 3) \propto \mathbf{k}$ of $\hat{k}\mathbf{S}$ with eigenvalue 0 does *not* occur among the polarization vectors.

So the photons have helicity ± 1, and there is no helicity-zero state, as might be expected from the vector representation in which the eigenvalues of S_3 are $+1$, 0, -1. The helicity zero polarization vector would be the longitudinal mode $\mathbf{e}(\mathbf{k}, 3) \propto \mathbf{k}$, which is equivalent to a gauge transformation and therefore unphysical. It was eliminated by the Coulomb gauge condition.

In general, massless particles have only two independent spin-states. This can be understood directly from the representation theory of the Poincaré group, see e.g. Weinberg I, Ryder. Under a general Lorentz transformation one finds that the polarization vectors

$$e^\mu(k, \lambda) \equiv (0, \mathbf{e}(\mathbf{k}, \lambda)), \quad k^0 = |\mathbf{k}|, \tag{7.96}$$

transform covariantly up to a gauge transformation

$$\ell^\mu{}_\nu e^\nu(k, \lambda) = \sum_{\lambda'} C_{\lambda'\lambda}(\ell, k) e^\mu(\ell k, \lambda') + \text{terms} \propto k^\mu, \tag{7.97}$$

and correspondingly[3] for the states in Hilbert space:

$$U(\ell)|k\lambda\rangle = \sum_{\lambda'} C_{\lambda'\lambda}(\ell, k)|k\lambda'\rangle. \tag{7.98}$$

The states are gauge invariant, $a(k, \lambda)$ is gauge invariant. The matrix $C_{\lambda'\lambda} \propto \delta_{\lambda'\lambda}$ in case of the helicity polarization vectors: for massless particles the helicity is Lorentz invariant. This is not true for massive particles, for example, massive spin-1 particles. Loosely speaking, a massive particle has a rest frame, therefore we can make a special Lorentz transformation along its momentum vector $\mathbf{k}$ to bring it first to rest, and then boost it again in direction $-\mathbf{k}$, e.g. giving it momentum $-\mathbf{k}$. This transformation does not affect the transverse polarization

[3] This correspondence is also true for the spin 1/2 case, where the transformation property of the spinor $L(\ell)\, u(p, \lambda) = \sum_{\lambda'} C_{\lambda'\lambda}(\ell, p) u(\ell p, \lambda')$, implies for the state in Hilbert space $U(\ell)\, |p, \lambda\rangle = \sum_{\lambda'} C_{\lambda'\lambda}(p, \lambda)|\ell p, \lambda'\rangle$ (as follows from the expansion of $\psi(x)$ in terms of creation and annihilation operators).

vectors, hence also not the eigenvalue of $\hat{k}\cdot\mathbf{S}$, but since the momentum flips sign, so does the helicity.

In conclusion, in Coulomb gauge the quantized Maxwell field has a sound physical interpretation. Furthermore, local observables satisfy locality, see Appendix 7.9). For perturbative calculations this gauge is however rather awkward, and we shall transform to a manifestly covariant formalism using the path integral, to be introduced next.

7.5 Path integral for the Maxwell field

It is straightforward to write down a path integral representation for the electromagnetic field in the oscillator representation (7.20). However, we shall start afresh and make the Ansatz

$$Z[J] \equiv \langle 0,\infty|0,-\infty\rangle[J] \stackrel{?}{=} \int [dA]\, e^{iS[A]+i\int d^4x\, J^\mu A_\mu}, \tag{7.99}$$

where $S[A]$ is the free Maxwell action and the boundary conditions are taken to be $A^\mu(x) = 0$ at $x^0 = \pm\infty$, similar to the case of the scalar field. We have put a question mark above the equality sign because it is not clear at this stage if the above formula makes sense. In fact, if we assume the measure $[dA]$ to be formally given by

$$\int [dA] = \int_{-\infty}^{\infty} \prod_{x\mu} dA_\mu(x), \tag{7.100}$$

it does not make sense. The reason is that the above implies for each A_μ, also integration over all gauge transforms $A_\mu + \partial_\mu\omega$, for which the integrand of the path integral does not vary because of its gauge invariance. So we expect an infinity of the form $\prod_x \int_{-\infty}^{\infty} d\omega(x)$. We assume in this argument that the external current J^μ is conserved such that $\int d^4x\, J_\mu A^\mu$ is gauge invariant.

To see the problem in another way, suppose we want to separate the J dependence in the usual way by 'completing the square' in the exponent and making a change of variables,

$$A^\mu(x) = A'^\mu(x) + \int d^4y\, G^{\mu\nu}(x,y) J_\nu(y), \tag{7.101}$$

such that

$$\frac{Z[J]}{Z[0]} = \exp[i\frac{1}{2}\int d^4x\, d^4y\, J_\mu(x) G^{\mu\nu}(x,y) J^\nu(y)]. \tag{7.102}$$

Here $G^{\mu\nu}(x,y)$ is supposed to be a Green function of the differential operator

$$K_{\mu\nu} = -\partial^2\eta_{\mu\nu} + \partial_\mu\partial_\nu, \tag{7.103}$$

which enters in the action as (making partial integrations)

$$S = -\int d^4x\, \frac{1}{2} A^\mu K_{\mu\nu} A^\nu, \tag{7.104}$$

such that

$$(-\partial^2\eta_{\mu\nu} + \partial_\mu\partial_\nu)\int d^4y\, G^{\nu\rho}(x,y)\, J_\rho(y) = J_\mu(x). \tag{7.105}$$

However, without gauge fixing such a Green function does not exist! The differential operator $K_{\mu\nu}$ has an infinite number of zero modes, i.e. eigenfunctions with zero eigenvalue. These are the 'pure-gauge' configurations

$$A^\mu_{\rm pg} = \partial^\mu\omega, \quad K_{\mu\nu}\, A^\nu_{\rm pg} = 0, \tag{7.106}$$

where ω is arbitrary. This is infinitely more general than the usual zero modes that can be formed from plane wave solutions $\propto \exp(ikx)$, $k^2 = 0$, and which are dealt with by the $i\epsilon$ prescription , e.g. for the Green function of $-\partial^2$ (propagator of a massless scalar field), $G(k) = 1/(k^2 - i\epsilon)$.

So we have[4] to fix the gauge in the path integral (7.99). One way to do this is the Faddeev–Popov procedure, which also generalizes also to nonabelian gauge field theories. We shall begin with the Coulomb gauge. Consider the functional $\Delta_C[A]$ defined by

$$1 = \Delta_C[A]\int [d\omega]\, \delta[\nabla\cdot\mathbf{A}^\omega], \quad A^\omega_\mu \equiv A_\mu + \partial_\mu\omega. \tag{7.107}$$

Here $\delta[\nabla\mathbf{A}]$ is a delta functional enforcing the Coulomb gauge, formally given by

$$\delta[\nabla\cdot\mathbf{A}] = \prod_x \delta(\partial_k A_k(x)). \tag{7.108}$$

The functional $\Delta_C[A]$ is gauge invariant,

$$\Delta_C[A^{\omega_1}] = \Delta_C[A], \tag{7.109}$$

which follows easily by shifting the dummy integration variable $\omega \to \omega - \omega_1$ in (7.107). Explicitly we have[5]

$$\frac{1}{\Delta_C[A]} = \int [d\omega]\, \delta[\nabla\cdot\mathbf{A} + \nabla^2\omega] = \frac{1}{\det[-\nabla^2]}, \tag{7.110}$$

A way to see this is by making the transformation of variables $\omega' = \nabla^2\omega$, $[d\omega] = [d\omega']/|\det[\nabla^2]|$. The argument of the delta functions $\delta(\nabla^2\omega + \nabla\mathbf{A})$ is assumed to vanish only once, which is correct because the solution $\omega = (-\nabla^2)^{-1}\nabla\cdot\mathbf{A}$ is unique, barring zero modes of the laplacian.

[4]There exist lattice regularizations which do not require gauge fixing and which are currently used in solving gauge theories like QCD nonperturbatively by numerical simulation. However, for the usual perturbation expansion gauge fixing is necessary.

[5]As mentioned earlier we exclude zero-momentum modes, so $-\nabla^2\delta^{(4)}(x-x')$ is a positive 'matrix'.

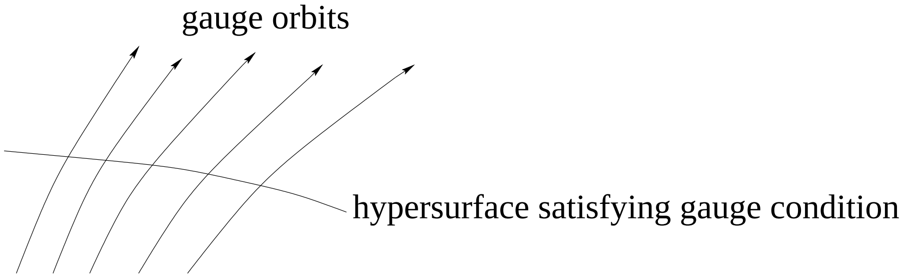

Figure 7.2: Gauge fixing in the path integral. Shown is the space of all gauge field configurations. The directed lines indicate orbits of gauge transformations. Each orbit is supposed to cross the hypersurface of functions satisfying $\nabla \cdot \mathbf{A} = 0$ only once.

So $\Delta_C[A]$ is just a constant, $\det[-\nabla^2]$. We shall keep it in full glory to better present the following reasoning. Inserting the identity (7.107) into the integrand of the Ansatz (7.99) we get

$$Z[J] \stackrel{?}{=} \int [dA]\,[d\omega]\,\delta[\nabla \cdot \mathbf{A}^\omega]\,\Delta_C[A]\,e^{iS[A]+i\int JA} \tag{7.111}$$

$$= \int [dA^{-\omega}]\,[d\omega]\,\delta[\nabla \cdot \mathbf{A}]\,\Delta_C[A^{-\omega}]\,e^{iS[A^{-\omega}]+i\int JA^{-\omega}} \tag{7.112}$$

$$= \int [d\omega] \int [dA]\,\delta[\nabla \cdot \mathbf{A}]\,\Delta_C[A]\,e^{iS[A]+i\int JA}. \tag{7.113}$$

In the second line we made the change of variables $A \to A^{-\omega}$ and in the third line we use the gauge invariance of S, $\int JA$, Δ_C and $[dA]$. The factor $\int [d\omega]$ in front in (7.113) is the divergence anticipated earlier, whereas the remaining path integral over A looks sound. The infinite $\int [d\omega]$ is the volume of the gauge group, a constant, which we drop. So the corrected Ansatz is now

$$Z[J] = \int [dA]\,\delta[\nabla \cdot \mathbf{A}]\,\Delta_C\,e^{iS[A]+i\int d^4x\,J^\mu A_\mu}, \tag{7.114}$$

Fig. 7.2 illustrates what we have done.

We shall now perform the path integral successively over A^0 and $\mathbf{A}$. Because we now only integrate over gauge potentials satisfying the Coulomb gauge condition we can use the form (7.10,7.11) of the action:

$$S = -\int d^4x\,\frac{1}{2}\left[A^0\nabla^2 A^0 + A_k(\partial_0^2 - \nabla^2)A_k\right]. \tag{7.115}$$

The integral over A^0 gives the factor

$$\exp\left[-i\frac{1}{2}\int d^4x\,d^4y\,J^0(x)\frac{\delta(x^0-y^0)}{4\pi|\mathbf{x}-\mathbf{y}|}\,J^0(y)\right] = \exp\left[-i\int dt\,E_C(t)\right], \tag{7.116}$$

where E_C is the Coulomb energy. Since it is a term in the total hamiltonian, we expect indeed a factor $\exp[-i\int dt\,E_C]$ to be present in the matrix element $Z[J] =$

$\langle 0,\infty|0,-\infty\rangle[J]$ of the evolution operator $U(\infty,-\infty) = T\exp[-i\int dt\, H(t)]$. The integral over $\mathbf{A}$ gives

$$\exp\left[\frac{1}{2}\int d^4x\, d^4y\, J_k(x) G_C^{kl}(x,y)\, J_l(y)\right], \tag{7.117}$$

with G_C^{kl} the Coulomb-gauge propagator given by

$$G_C^{kl}(x,y) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik(x-y)}\, \frac{P_{kl}^T(\mathbf{k})}{\mathbf{k}^2 - k_0^2 - i\epsilon}. \tag{7.118}$$

There is the usual $k^2 - i\epsilon$ in the denominator. The transverse projector appears because we only integrate over transverse $\mathbf{A}$. The answer cannot depend on the longitudinal part of $\mathbf{J}$. Defining a complete Coulomb-gauge Green function $G_C^{\mu\nu}(x,y)$ by supplying the other μ,ν combinations as

$$G_C^{00}(x,y) = -\frac{\delta(x^0-y^0)}{4\pi|\mathbf{x}-\mathbf{y}|} = \int \frac{d^4k}{(2\pi)^4}\, e^{ik(x-y)}\, \frac{-1}{\mathbf{k}^2}, \tag{7.119}$$

$$G_C^{0l}(x,y) = G_C^{k0}(x,y) = 0, \tag{7.120}$$

we can summarize the result so far as

$$Z[J] = Z[0]\, \exp\left[i\frac{1}{2}\int d^4x\, d^4y\, J_\mu(x) G_C^{\mu\nu}(x,y)\, J_\nu(y)\right]. \tag{7.121}$$

On the other hand, making the connection with vacuum expectation values of time ordered products of field operators, we should have

$$\frac{Z[J]}{Z[0]} = \frac{\langle 0,\infty|0,-\infty\rangle[J]}{\langle 0,\infty|0,-\infty\rangle[0]} = e^{-i\int dt\, E_C(t)}\, \langle 0|T e^{i\int d^4x\, \mathbf{J}\hat{\mathbf{A}}}|0\rangle. \tag{7.122}$$

As a check we can compare the coefficients quadratic in J with the results obtained by canonical quantization. One finds that (Problem 3)

$$\langle 0|T\hat{A}_k(x)\hat{A}_l(y)|0\rangle = -iG_C^{kl}(x,y), \tag{7.123}$$

indeed. The Faddeev–Popov method has given the same result as canonical quantization. In this check we have to relax the condition $\partial_\mu J^\mu = 0$, otherwise we cannot compare coefficients of *arbitrary* J's. This means that we *define* $Z[J]$ for arbitrary J by (7.114).

We shall now show that in the physical case in which external current is conserved, the amplitude ratio $Z[J]/Z[0]$ is Lorentz invariant, despite its noncovariant looks. The Coulomb gauge propagator can be written in the form

$$G_C^{\mu\nu}(x,y) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik(x-y)}\, G_C^{\mu\nu}(k), \tag{7.124}$$

$$G_C^{\mu\nu}(k) = \frac{N_C^{\mu\nu}(k)}{k^2 - i\epsilon}, \tag{7.125}$$

$$N_C^{\mu\nu}(k) = \eta^{\mu\nu} - \frac{k^\mu k^\nu + (kn)(k^\mu n^\nu + n^\mu k^\nu)}{k^2 + (kn)^2}, \quad n^\mu = \delta_0^\mu. \tag{7.126}$$

We see that

$$N_C^{\mu\nu}(k) = \eta^{\mu\nu} + \text{terms} \propto k^\mu, \propto k^\nu \text{ and } k^\mu k^\nu. \tag{7.127}$$

Using $k^\mu \tilde{J}_\mu(k) = 0$ ($\tilde{J}_\mu(k)$ is the Fourier transform of $J_\mu(x)$), it follows that

$$\int d^4x\, d^4y\, J_\mu(x) G_C^{\mu\nu}(x,y)\, J_\nu(y) = \int \frac{d^4k}{(2\pi)^4}\, \tilde{J}_\mu(k)^* G_C^{\mu\nu}(k)\, \tilde{J}_\nu(k) \tag{7.128}$$

$$= \int \frac{d^4k}{(2\pi)^4}\, \tilde{J}_\mu(k)^* \frac{\eta^{\mu\nu}}{k^2 - i\epsilon}\, \tilde{J}_\nu(k), \tag{7.129}$$

which is Lorentz invariant.

7.6 Generalized covariant gauge

Although the Coulomb gauge leads to Lorentz invariant results for gauge-invariant quantities, it is very awkward in calculations in higher orders of perturbation theory. We shall therefore make a transformation of variables in the path integral such that we get manifestly covariant expressions.

We start with the gauge-invariant quantity $\Delta[A]$ defined by

$$1 = \Delta[A] \int [d\omega]\, e^{iS_{\rm gf}[A^\omega]}, \tag{7.130}$$

where $S_{\rm gf}[A]$ is a 'gauge fixing' action given by

$$S_{\rm gf}[A] = -\frac{1}{2\xi} \int d^4x\, (\partial^\mu A_\mu)^2. \tag{7.131}$$

Here ξ is a so-called gauge parameter. Rewriting the above expression in the form

$$\frac{1}{\Delta[A]} = \int [d\chi]\, e^{-i\frac{1}{2\xi}\int d^4x\, \chi^2} \int [d\omega]\, \delta[\partial^\mu A_\mu^\omega - \chi] \tag{7.132}$$

we see that here we are not really fixing the gauge but taking an 'average' over covariant gauge conditions $\partial^\mu A_\mu = \chi$, with 'weight' $\exp\left[-i\frac{1}{2\xi}\int d^4x\, \chi^2\right]$. Evaluating the integral over ω in the second form (7.132) we get

$$\int [d\omega]\, \delta[\partial^\mu A_\mu + \partial^2\omega - \chi] = \frac{1}{|\det[-\partial^2]|}, \tag{7.133}$$

which is independent of χ. So we find

$$\Delta[A] = |\det[-\partial^2]| \left[\int [d\chi]\, e^{-i\frac{1}{2\xi}\int d^4x\, \chi^2}\right]^{-1}, \tag{7.134}$$

which is also a constant independent of A, like Δ_C. Inserting the identity (7.130) into the integrand of the path integral for $Z[J]$ and making a transformation of variables $A \to A^{-\omega}$ gives

$$Z[J] = \int [dA]\,[d\omega]\,\delta[\nabla\cdot\mathbf{A}]\,\Delta_C\,\Delta\, e^{iS[A]+iS_{\rm gf}[A^\omega]+i\int JA} \tag{7.135}$$

$$= \int [dA]\,[d\omega]\,\delta[\nabla\cdot\mathbf{A}^{-\omega}]\,\Delta_C\,\Delta\, e^{iS[A]+iS_{\rm gf}[A]+i\int JA} \tag{7.136}$$

$$= \int [dA]\,\Delta\, e^{iS[A]+iS_{\rm gf}[A]+i\int JA}. \tag{7.137}$$

We used the gauge invariance of $[dA]$, Δ_C, Δ, S and $\int JA$, as well as (7.107).

The action plus gauge-fixing term can be written as

$$S + S_{\rm gf} = -\int d^4x\,\frac{1}{2}A^\mu\left(-\partial^2\eta_{\mu\nu} + \partial_\mu\partial_\nu - \frac{1}{\xi}\,\partial_\mu\partial_\nu\right)A^\nu. \tag{7.138}$$

The differential operator in parenthesis has a Green function (for $\xi \neq \infty$), which we shall denote by $D^{\mu\nu}$:

$$D^{\mu\nu}(x,y) = \int \frac{d^4k}{(2\pi)^4}\,e^{ik(x-y)}\,D^{\mu\nu}(k), \tag{7.139}$$

$$\delta^\rho_\mu = \left(k^2\eta^{\mu\nu} - k_\mu k_\nu + \frac{1}{\xi}\,k_\mu k_\nu\right)D^{\nu\rho}(k), \tag{7.140}$$

$$D^{\mu\nu}(k) = \left(\eta^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\right)\frac{1}{k^2} + \xi\,\frac{k^\mu k^\nu}{(k^2)^2}. \tag{7.141}$$

This is easy to check with the help of four-dimensional transverse and longitudinal projectors

$$P_T^{\mu\nu}(k) = \eta^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}, \quad P_L^{\mu\nu}(k) = \frac{k^\mu k^\nu}{k^2}, \tag{7.142}$$

with the properties

$$P_T^2 = P_T, \quad P_L^2 = P_L, \quad P_T P_L = 0, \quad P_T + P_L = 1. \tag{7.143}$$

For $\xi = 1$ we get the 'Feynman gauge' propagator

$$D_F^{\mu\nu}(k) = \frac{\eta^{\mu\nu}}{k^2}, \tag{7.144}$$

which is very convenient for calculations, whereas for $\xi \to 0$ we get the 'Landau gauge' (or 'Lorentz gauge') propagator

$$D_L^{\mu\nu}(k) = \left(\eta^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\right)\frac{1}{k^2}. \tag{7.145}$$

We shall assume the $i\epsilon$ prescription

$$k^2 \to k^2 - i\epsilon \tag{7.146}$$

in the denominators of $D^{\mu\nu}(k)$. It amounts to adding the term $-\epsilon \int d^4x\, \eta_{\mu\nu} A^\mu A^\nu$ in the exponent of the integrand in the path integral. Since $\eta_{00} = -1$, the $A^0 A^0$ term has a positive sign, which is *wrong* for improving convergence in the path integral (recall the motivation for introducing ϵ at the end of sect. 3.3). For convergence reasons we could change the sign of ϵ in this $A^0 A^0$ term, but this would to spoil manifest covariance. Within perturbation theory the prescription (7.146) leads to satisfactory results. Nonperturbatively, the lattice regularization has been used mostly within the imaginary time formulation (where $A^0 A^0 \to -A_4 A_4$), and sometimes in real time in the Coulomb gauge, where the A_0^2 problem does not arise.

7.7 Ghosts

The equivalence between the Coulomb gauge and covariant gauges holds for gauge-invariant quantities, and accordingly we assumed the current J^μ to be conserved. Still, having arrived at a manifestly covariant formalism we can relax the condition $\partial_\mu J^\mu = 0$ and study the correlation functions $\langle A_\mu(x) A_\nu(y) \cdots \rangle$ obtained by differentiating with respect to J^μ and setting $J^\mu = 0$ afterwards. These correlation functions are not themselves gauge invariant, but we they can be used to construct gauge invariant observables.

We can now ask the question: is there a representation in terms of operators $\hat{A}_\mu(x)$ such that

$$\langle A_\mu(x) A_\nu(y) \cdots \rangle = \langle 0 | T \hat{A}_\mu(x) \hat{A}_\nu(y) \cdots | 0 \rangle? \tag{7.147}$$

The answer is yes, but the operators act in an extended space with indefinite metric (hence, not a Hilbert space). Consider the case $\xi = 1$, for which, after some partial integrations,

$$S + S_{\rm gf}^{\xi=1} = -\int d^4x\, \frac{1}{2} \eta^{\rho\sigma} \partial_\mu A_\rho \partial^\mu A_\sigma. \tag{7.148}$$

This looks just like the action of four massless scalar fields, except that the contribution $+\frac{1}{2} \int d^4x\, \partial_\mu A_0 \partial^\mu A_0$ has the wrong sign. For a discussion of canonical quantization with such an action see, e.g. Itzykson and Zuber. Here we shall cut the discussion short by appealing to the following plausible outcome.

If we want to reproduce the correlator

$$\langle 0 | T \hat{A}^\mu(x) \hat{A}^\nu(y) | 0 \rangle = -i D^{\mu\nu}(x,y) = \int d\omega_k\, e^{ik(x-y)}\, \eta^{\mu\nu}, \quad x^0 > y^0, \tag{7.149}$$

by inserting intermediate states,

$$\int d\omega_k\, e^{ik(x-y)}\, \eta^{\mu\nu} = \langle 0|\hat{A}^\mu(x)\hat{A}^\nu(y)|0\rangle \stackrel{?}{=} \sum_\lambda \int d\omega_k\, \langle 0|\hat{A}^\mu(x)|k\lambda\rangle\langle k\lambda|\hat{A}^\nu(y)|0\rangle, \tag{7.150}$$

we find that in

$$\langle 0|\hat{A}^\mu(x)|k\lambda\rangle = e^\mu(k,\lambda)\, e^{ikx}, \quad \lambda = 0,1,2,3, \tag{7.151}$$

we have to allow for states with polarization $\lambda = 0$ and 3. This is needed to reproduce the polarization sum

$$-e^\mu(k,0)e^\nu(k,0)^* + \sum_{\lambda=1}^{3} e^\mu(k,\lambda)\, e^\nu(k,\lambda)^* = \eta^{\mu\nu}. \tag{7.152}$$

Notice that we have used a minus sign in front of the $\lambda = 0$ term in (7.152). Without this sign the left-hand side would constitute a positive-definite matrix, in contradiction with the indefinite $\eta^{\mu\nu}$ on the right hand side. An explicit realization of the $e^\mu(k,\lambda)$ is given by

$$e^\mu(k,\lambda) = (0, \mathbf{e}(\mathbf{k},\lambda)), \quad \lambda = 1,2, \tag{7.153}$$

$$e^\mu(k,3) = (0, \hat{k}), \tag{7.154}$$

$$e^\mu(k,0) = (1, \mathbf{0}). \tag{7.155}$$

The sum over intermediate states in the right of (7.150) has to be re-written as

$$\int d\omega_k \sum_{\lambda=0}^{3} \eta^{\lambda\lambda}\, \langle 0|\hat{A}^\mu(x)|k,\lambda\rangle\langle k,\lambda|\hat{A}^\nu(y)|0\rangle. \tag{7.156}$$

Together with (7.151), $\langle k\lambda|\hat{A}^\mu(x)|0\rangle = \langle 0|\hat{A}^\mu(x)|k\lambda\rangle^*$ and (7.152), this reproduces the expression on the left of (7.150). The minus sign in the decomposition of unity

$$\hat{1} = |0\rangle\langle 0| + \int d\omega_k \left(-|k,0\rangle\langle k,0| + \sum_{\lambda=1}^{3} |k,\lambda\rangle\langle k,\lambda| \right) + \cdots \tag{7.157}$$

reflects an indefinite metric in 'ket space' (it should not be called Hilbert space as this has positive metric),

$$\langle k,\lambda|k',\lambda'\rangle = \eta_{\lambda\lambda'}\, 2k^0\, (2\pi)^3 \delta(\mathbf{k}-\mathbf{k}'), \tag{7.158}$$

such that

$$\hat{1}\,|k,\lambda\rangle = |k,\lambda\rangle \tag{7.159}$$

(with $\hat{1}$ given by (7.157)), also for $\lambda = 0$.

The extra particle-like degrees of freedom with $\lambda = 0$ and 3 are called respectively 'timelike' and 'longitudinal photons'. These unphysical 'particles' are called *ghosts*. The timelike photon states have negative norm. Further analysis reveals that their contributions to gauge-invariant quantities cancel against the longitudinal photons, and only the physical photons corresponding to $\lambda = 1, 2$ remain. The physical-photon states are have positive norm, as needed for the probability interpretation of quantum mechanics. The physical Hilbert space is a subspace of a 'ket space' that has indefinite metric.

There are also other ghosts, named after Faddeev and Popov. They correspond to the determinant in $\Delta[A] = |\det[-\partial^2]|$. We may ignore the absolute value symbols, since Δ is just an A-independent constant that cannot change sign. Using fermionic integration this determinant can be written (up to a factor) as

$$\det[-\partial^2] = \int [d\xi^* d\xi]\, e^{-i\int d^4x\, \partial_\mu \xi^* \partial^\mu \xi}, \tag{7.160}$$

which reveals that these ghosts are scalar fermions! For the electromagnetic field Δ does not depend on A and the Faddeev–Popov ghosts do not interact. But in nonabelian gauge theories $\Delta[A]$ does depend on A and these ghosts have to be explicitly taken into account. However, even in the electromagnetic case we have to keep them in mind when calculating the partition function Z in the imaginary-time formalism. The Faddeev–Popov determinant Δ *does* contribute to Z and compensates, for example, for the unphysical contribution of the longitudinal and timelike photons to the specific heat.

We shall not use the operator language for gauge theories in covariant gauges. Transition amplitudes $\langle \text{out}|\text{in}\rangle$ and matrix elements of operators will be obtained directly from the covariant correlation functions. For the free electromagnetic field these are given by the 'Wick formula'

$$\begin{aligned} \langle A^{\mu_1}(x_1) A^{\mu_2}(x_2) \cdots A^{\mu_n}(x_n)\rangle &= (-i)^{n/2}\, D^{\mu_1\mu_2}(x_1, x_2) \cdots D^{\mu_{n-1}\mu_n}(x_{n-1}, x_n) \\ &\quad + \text{permutations}. \end{aligned} \tag{7.161}$$

To get amplitudes out of these we define the creation and annihilation 'symbols'

$$\begin{aligned} a(k, \lambda, x^0) &= A^{(+)}(k, \lambda, x^0) = \int d^3x\, e^{-ikx}\, e^*_\mu(k, \lambda)\, (i \overrightarrow{\partial_0} - i \overleftarrow{\partial_0})\, A^\mu(x), \\ a(k, \lambda, x^0)^* &= A^{(-)}(k, \lambda, x^0) = \int d^3x\, A^\mu(x)\, (i \overrightarrow{\partial_0} - i \overleftarrow{\partial_0})\, e_\mu(k, \lambda)\, e^{ikx}. \end{aligned} \tag{7.162}$$

where $k^0 = |\mathbf{k}|$, $\lambda = 1, 2$, or $\lambda = \pm$. These are gauge invariant because they are constructed out of the physical polarization vectors which satify[6]

$$k^\mu\, e_\mu(k, \lambda) = 0, \quad k^0 = |\mathbf{k}|, \quad \lambda = 1, 2, \text{ or } \lambda = \pm, \tag{7.163}$$

[6]In fact, k could be arbitrary here (not necessarily 'on shell' $k^2 = 0$). This would do some violence to the notation because originally the k in $e^\mu(k, \lambda)$ was supposed to satisfy $k^0 = |\mathbf{k}|$. But the explicit form $e^\mu(k, \lambda) = (0, \mathbf{e}(\mathbf{k}, \lambda))$ (which really depends only on $\mathbf{k}/|\mathbf{k}|$) shows that the extension to arbitrary k is possible. However, for $A^{(\pm)}(k, \lambda, x^0)$ we only need $k^0 = |\mathbf{k}|$.

So we can interpret correlation functions of them (possibly with other gauge invariant quantities) as being expectation values of operators in Coulomb gauge. However, being gauge invariant they can be directly used in a covariant gauge. A simple example in the free theory is (cf. Problem 4)

$$\begin{aligned}\langle a(k,\lambda,x^0)\, a^*(k',\lambda',x'^0)\rangle &= \int d^3x\, d^3x'\, e^{-ikx}\, e^*_\mu(k,\lambda)\,(i\,\overrightarrow{\partial_0} - i\,\overleftarrow{\partial_0})\,\langle A^\mu(x)A^\nu(x')\rangle \\ &\quad (i\,\overrightarrow{\partial'_0} - i\,\overleftarrow{\partial'_0})\, e_\nu(k',\lambda')\, e^{ik'x'} \qquad (7.164)\\ &= 2k^0\,(2\pi)^3\delta(\mathbf{k}-\mathbf{k}')\,\delta_{\lambda\lambda'}, \quad x^0 > x'^0, \\ &= \langle k\lambda|k'\lambda'\rangle.\end{aligned}$$

The answer is independent of the gauge parameter ξ, at it should be.

7.8 Summary

The gauge invariance of the Maxwell action leads to complications with canonical quantization, which can be solved by choosing the Coulomb gauge. In this gauge the time component A_0 of the gauge field is not a dynamical variable and one more degree of freedom is 'lost' by the vector potential satisfying the condition $\nabla\cdot\mathbf{A} = 0$. The resulting interpretation for the quantized free Maxwell field leads to photons: massless 'spin-one' particles with only two spin polarization degrees of freedom. The formulation is not manifestly Lorentz covariant, but Lorentz invariance is guaranteed by the existence of the unitary representation $U(\ell)$ in Hilbert space, with conserved Noether generators $J^{\mu\nu}$.

Path-integral quantization is also complicated by gauge invariance, which is taken care of with the Faddeev–Popov method. Choosing the Coulomb gauge, the path integral is found to produce the same correlation functions as found with the canonical operator formalism.

The Faddeev–Popov method allows for a very convenient transformation from the non-covariant Coulomb gauge to a manifestly covariant gauge. A corresponding operator description is possible, at the expense of introducing ghosts, unphysical particles, and a space of state vectors with indefinite (positive and negative) metric of which the physical Hilbert space is a subspace. The equivalence to the Coulomb gauge with its (positive metric) Hilbert space guarantees that physical probability is still conserved.

7.9 Appendix: Locality in the Coulomb gauge

We started from an action S which has nice invariance properties and is local: it has the form $S = \int d^4x\,\mathcal{L}(x)$ where $\mathcal{L}(x)$ is a Lorentz scalar which depends on the fields at x and in the immediate neighborhood of x (through the derivatives).

This leads to covariant and local classical equations of motion. No signals can travel faster than the velocity of light. Upon quantization in the Coulomb gauge we have ended up with non-Lorentz and nongauge invariant expressions which furthermore look terribly nonlocal: the projector

$$P^T_{mn}(\mathbf{x}-\mathbf{y}) = \delta_{mn}\delta(\mathbf{x}-\mathbf{y}) + \partial_m\partial_n \frac{1}{4\pi|\mathbf{x}-\mathbf{y}|} \tag{7.165}$$

drops off only like a power of $|\mathbf{x}-\mathbf{y}|$ for large separations.

An important expression of locality and Lorentz invariance is the following. Two observables $O_{1,2}$ associated with compact spacetime regions $R_{1,2}$ ('local observables') commute, when all points $x_1 \in R_1$ are spacelike to all points $x_2 \in R_2$. In the standard lore of quantum mechanics observables correspond to measurements, and measurements in spacelike separated regions should not be able to influence each other. Observables have to be gauge invariant. An example is given by the field strength $F_{\mu\nu}(x)$. Locality is expressed by

$$[F_{\kappa\lambda}(x), F_{\mu\nu}(y)] = 0, \quad (x-y)^2 > 0. \tag{7.166}$$

This is indeed the case as will now be shown for the case of vanishing external current.

Using the expansion

$$A^\mu(\mathbf{x},t) = \sum_\lambda \int d\omega_k [e^{ikx}\, e^\mu(k,\lambda) a(\mathbf{k},\lambda) + h.c.] \tag{7.167}$$

(recall that in Coulomb gauge $e^\mu(k,\lambda)$ has no time component and λ is restricted to the values 1,2), gives

$$[A^\lambda(x), A^\nu(y)] = \int d\omega_k\, (e^{ik(x-y)} - e^{-ik(x-y)}) P_C^{\lambda\nu}(k), \tag{7.168}$$

where (recall $n^\mu = \delta_{\mu,0}$)

$$\begin{aligned} P_C^{\mu\nu}(k) &= \sum_\lambda e^\mu(k,\lambda) e^\nu(k,\lambda)^* \\ &= \eta^{\mu\nu} - \frac{k^\mu k^\nu + (kn)(k^\mu n^\nu + n^\mu k^\nu)}{(kn)^2} \qquad (7.169) \\ &= N_C^{\mu\nu}(k), \quad k^2 = 0. \qquad (7.170) \end{aligned}$$

Working out the derivatives in $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ we get

$$\begin{aligned} [F^{\kappa\lambda}(x), F^{\mu\nu}(y)] &= \int d\omega_k\, \{e^{ik(x-y)}[k^\kappa k^\mu P_C^{\lambda\nu}(k) - k^\lambda k^\mu P_C^{\kappa\nu}(k) \\ &\quad - k^\kappa k^\nu P_C^{\lambda\mu}(k) + k^\lambda k^\nu P_C^{\kappa\mu}(k)] - (k \to -k)\}. \qquad (7.171) \end{aligned}$$

Now the operation of the curl in $F^{\kappa\lambda}$ projects to zero any 4D longitudinal part $\propto k^\lambda$ in $P_C^{\lambda\nu}$, such that only the $\eta^{\lambda\nu}$ part of $P_C^{\lambda\nu}$ contributes. In position space we can then write

$$\begin{aligned}
[F_{\kappa\lambda}(x), F_{\mu\nu}(y)] &= (\partial_\kappa\partial_\mu\eta_{\lambda\nu} - \partial_\lambda\partial_\mu\eta_{\kappa\nu} \\
&\quad - \partial_\kappa\partial_\nu\eta_{\lambda\mu} + \partial_\lambda\partial_\nu\eta_{\kappa\mu})\, i\Delta(x-y), &(7.172)\\
\Delta(x-y) &= i\int d\omega_k\,(e^{ik(x-y)} - e^{-ik(x-y)}). &(7.173)
\end{aligned}$$

The (generalized) function $\Delta(x)$ (which we have met before in the scalar field case) has the following properties:

- $\Delta(x)$ is Lorentz invariant, $\Delta(\ell x) = \Delta(x)$,
- $\Delta(x) = 0$ for $x^0 = 0$, $\mathbf{x} \neq 0$.

Since $x = (\mathbf{x}, 0)$ is spacelike and $\Delta(x)$ is Lorentz invariant it follows that it vanishes for general spacelike distances,

$$\Delta(x-y) = 0, \quad (x-y)^2 > 0. \tag{7.174}$$

(It is also interesting to note that $\Delta(x)$ is the solution of $\partial^2\Delta(x) = 0$ with initial conditions $\Delta(x) = 0$, $\partial_0\Delta(x) = \delta(\mathbf{x})$ at $x^0 = 0$.)

Consequently the field strengths and all local observables that can be made out of these have the locality property (7.166).

7.10 Problems

1. *Expressing P^μ in terms of a's and $a^\dagger$'s.*

   To obtain the expressions (7.57) for the hamiltonian, we insert (7.43) into (7.31), using (7.47):

$$\begin{aligned}
H &= \frac{1}{V^2}\sum_{\mathbf{k},\mathbf{l}} \frac{1}{4k^0 l^0}\int d^3x\,\frac{1}{2}\,[\\
&\quad (-ik^0 e^{i\mathbf{kx}}\,a_m(\mathbf{k}) + ik^0 e^{-i\mathbf{kx}}\,a_m(\mathbf{k})^\dagger)(-il^0 e^{i\mathbf{lx}}\,a_m(\mathbf{l}) + il^0 e^{-i\mathbf{lx}}\,a_m(\mathbf{l})^\dagger) \\
&\quad + (e^{i\mathbf{kx}}\,a_m(\mathbf{k}) + e^{-i\mathbf{kx}}\,a_m(\mathbf{k})^\dagger)\,\mathbf{l}^2\,(e^{i\mathbf{lx}}\,a_m(\mathbf{l}) + e^{-i\mathbf{lx}}\,a_m(\mathbf{l})^\dagger)]. \quad (7.175)
\end{aligned}$$

   The integration sets $\mathbf{l} = \pm\mathbf{k}$ and the aa and $a^\dagger a^\dagger$ terms cancel ($k^0 = |\mathbf{k}|$), leaving

$$H = \frac{1}{V}\sum_{\mathbf{k}} \frac{1}{2k^0}\,k^0\,\frac{1}{2}\,[a_m(\mathbf{k})a_m(\mathbf{k})^\dagger + a_m(\mathbf{k})^\dagger a_m(\mathbf{k})]. \tag{7.176}$$

We can now convert to $a(\mathbf{k},\lambda)$ or use the commutation relation (7.53) directly with $\sum_m P^T_{mm}(\mathbf{k}) = 2 = \sum_\lambda$ to put $a^\dagger$ to the left of a,

$$a_m(\mathbf{k})a_m(\mathbf{k})^\dagger = a_m(\mathbf{k})^\dagger a_m(\mathbf{k}) + 2k^0 V \sum_\lambda . \tag{7.177}$$

This gives (7.31) after converting to $a(\mathbf{k},\lambda)$. This calculation of the hamiltonian is basically the same as for the one dimensional harmonic oscillator. The calculation of the momentum operator (7.32) proceeds in similar fashion,

$$\begin{aligned}
\mathbf{P} &= -\frac{1}{V^2}\sum_{\mathbf{k},\mathbf{l}} \frac{1}{4k^0 l^0}\int d^3x\, [\\
&\quad (-ik^0 e^{i\mathbf{kx}}\, a_m(\mathbf{k}) + ik^0 e^{-i\mathbf{kx}}\, a_m(\mathbf{k})^\dagger)\, i\mathbf{l}\, (e^{i\mathbf{lx}}\, a_m(\mathbf{l}) - e^{-i\mathbf{lx}}\, a_m(\mathbf{l})^\dagger] \\
&= \frac{1}{V}\sum_{\mathbf{k}} \frac{1}{2k^0}\frac{1}{2}\mathbf{k}\, [a_m(\mathbf{k})a_m(-\mathbf{k}) + a_m(\mathbf{k})^\dagger a_m(-\mathbf{k})^\dagger \\
&\quad + a_m(\mathbf{k})a_m(\mathbf{k})^\dagger + a_m(\mathbf{k})^\dagger a_m(\mathbf{k})], \qquad (7.178) \\
&= \frac{1}{V}\sum_{\mathbf{k}} \frac{1}{2k^0}\,\mathbf{k}\, a_m(\mathbf{k})^\dagger a_m(\mathbf{k}) \qquad (7.179)
\end{aligned}$$

($\mathbf{k}$ is odd under $\mathbf{k} \to -\mathbf{k}$, such that accompanying factors even under $\mathbf{k} \to -\mathbf{k}$ do not contribute). Expressing the result in terms of $a(\mathbf{k},\lambda)$ gives (7.32).

2. *Coulomb gauge Green function*

   Verify that $G_C^{\mu\nu}(k)$ in the form (7.126) satisfies the Coulomb gauge condition $k_m G_C^{m\nu}(k) = 0$, and for a conserved current

   $$K_{\mu\nu}(k)G_C^{\nu\rho}(k)\,\tilde{J}_\rho(k) = \tilde{J}_\mu(k). \tag{7.180}$$

   Hence, eq. (7.105) is satisfied as well and although $G_C^{\mu\nu}(x,y)$ is not the inverse of $K_{\mu\nu}$ it is a good Green function in this restricted sense.

3. *Checking (7.123)*

   Verify (7.123) by expanding the left and right hand sides up to terms quadratic in J. Choose time orderings and evaluate the right hand side by contour integration over k^0. On the left hand side insert the resolution of unity

   $$\hat{1} = |0\rangle\langle 0| + \sum_\lambda \int d\omega_k\, |k\lambda\rangle\langle k\lambda| + \cdots, \tag{7.181}$$

   and use the formulas obtained by canonical quantization.

4. *Amplitudes from correlation functions in covariant gauge.*

   Using the Wick formula (7.161), verify (7.164).

Chapter 8

QED

Two concepts have proved very successful in the development of quantum field theory of elementary particles (in addition to locality and Poincaré invariance): gauge invariance and renormalizability.[1] With gauge invariance we mean local symmetries, i.e. symmetries in which the transformations may vary from point to point in space-time. Renormalizability – the property that ultraviolet infinities can be absorbed in the bare parameters of the model – gives strong restrictions on the possible actions: they have to be polynomial in the fields,[2] with coupling constants that can only have dimension ≥ 0 in units of mass. Quantum electrodynamics (QED) is the simplest example of a nontrivial gauge theory. In a wide sense it is the theory of electromagnetic interactions of all things, but we understand it here to be 'spinor electrodynamics' (an even simpler version is 'scalar electrodynamics').

8.1 Gauge invariance

We have seen that the Lagrange density for the Dirac field,

$$\mathcal{L}(x) = -\bar{\psi}(x)(m + \gamma^{\mu} \overleftrightarrow{\partial}_{\mu})\psi(x), \tag{8.1}$$

has a U(1) invariance,

$$\psi'(x) = e^{i\omega}\psi(x), \quad \bar{\psi}'(x) = e^{-i\omega}\,\bar{\psi}(x). \tag{8.2}$$

[1]This does not mean that nonrenormalizable models are not useful. Such models have a regulatorization scale that cannot be removed in perturbation theory, indicating the need of new physical input at that scale. When we increase the regulatorization scale, its influence on the physics diminishes and only renormalizable interactions remain important. Renormalizable models come in universality classes much like critical phenomena.

[2]This holds in the formal continuum presentation. With an explicit regularization such as the lattice the action can be an infinite series with increasing powers of the fields.

In terms of the Majorana fields $\psi_{1,2}$ defined by

$$\psi = \frac{1}{\sqrt{2}}(\psi_1 - i\psi_2), \tag{8.3}$$

this is an SO(2) rotation

$$\begin{pmatrix} \psi_1'(x) \\ \psi_2'(x) \end{pmatrix} = \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix}, \tag{8.4}$$

according to the equivalence U(1)$\simeq$SO(2). We can interpret these transformations also as passive transformations of the coordinate system in an internal two-dimensional space.

The transformations are global in the sense that the angle ω does not depend on the spacetime coordinate x. It is natural to ask if the reference system that picks out the real and imaginary parts of ψ, or equivalently its components ψ_1 and ψ_2, has to be globally defined. For example do we have to choose a reference frame[3] in Amsterdam now the same as somewhere in a far away galaxy five years later? This seems unnatural, and so requiring the parameter ω of a symmetry transformation to be the same all over the universe, for all times, seems unphysical. So one is led to consider transformations with ω depending on the spacetime point x. How rapidly $\omega(x)$ may vary as a function of x is not clear at this point. A natural scale of variation is the Compton wavelength $1/m$, which is very small in human perception. At this point it is simplest to assume that $\omega(x)$ should be allowed to vary without restriction.

Hence, we want to construct an action invariant under U(1) transformations in which the angle $\omega(x)$ depends on spacetime. To achieve this we need to compensate the noninvariance of the derivative terms in the Lagrange density (8.1). Under a local transformation

$$\psi'(x) = e^{i\omega(x)}\,\psi(x), \quad \bar{\psi}'(x) = e^{-i\omega(x)}\,\bar{\psi}(x), \tag{8.5}$$

the term $m\bar{\psi}(x)\psi(x)$ is invariant, but the derivative transforms in an inhomogeneous and noncovariant way

$$\partial_\mu\psi'(x) = \partial_\mu[e^{i\omega(x)}\,\psi(x)] = e^{i\omega(x)}\,[\partial_\mu\psi(x) + i\partial_\mu\omega(x)\psi(x)]. \tag{8.6}$$

Instead, a covariant derivative $D_\mu\psi$ transforming as

$$D_\mu'(x)\psi'(x) = e^{i\omega(x)}\,D_\mu(x)\psi(x), \quad D_\mu'(x)\bar{\psi}'(x) = e^{-i\omega(x)}\,D_\mu(x)\psi(x), \tag{8.7}$$

would allow for the construction of an invariant Lagrange density

$$\mathcal{L}(x) = -\bar{\psi}(x)\left(m + \gamma^\mu \overleftrightarrow{D}_\mu\right)\psi(x), \quad \overleftrightarrow{D}_\mu = \frac{1}{2}\overrightarrow{D}_\mu - \frac{1}{2}\overleftarrow{D}_\mu\,. \tag{8.8}$$

[3]The reasoning here is familiar from in the theory of gravity.

The wellknown construction of such a covariant derivative uses the invariance of the electromagnetic field system under the gauge transformation

$$A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\omega(x) \tag{8.9}$$

(e is here an arbitrary constant). For ψ the form

$$D_\mu(x) = \partial_\mu - ieA_\mu(x), \tag{8.10}$$

has the required property: under the combined gauge transformation (8.5), (8.9),

$$\begin{aligned} D'_\mu(x)\,\psi'(x) &= [\partial_\mu - ieA'_\mu(x)]\,\psi'(x) \\ &= [\partial_\mu - ieA_\mu(x) - i\partial_\mu\omega(x)]\,[e^{i\omega(x)}\,\psi(x)] \\ &= e^{i\omega(x)}\,[\partial_\mu - ieA_\mu(x)]\,\psi(x) \\ &= e^{i\omega(x)}\,D_\mu(x)\,\psi(x), \end{aligned} \tag{8.11}$$

For $\bar\psi$ we have by conjugation

$$D_\mu\bar\psi = \partial_\mu\bar\psi + ieA_\mu\bar\psi \equiv \bar\psi\,\overleftarrow{D}_\mu, \quad D'_\mu\bar\psi' = e^{-i\omega}\,D_\mu\bar\psi. \tag{8.12}$$

A derivative involves the comparison of fields at infinitesimally close points in spacetime. The electromagnetic potentials play the role of a *connection*, which is used in comparing ('connecting') the orientations of the internal spaces at these infinitesimally close points.

The classical action for the combined electromagnetic and Dirac field system is now given by

$$S = \int d^4x \left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \bar\psi(m + \gamma^\mu\,\overleftrightarrow{D}_\mu)\psi\right] \tag{8.13}$$

$$= \int d^4x \left[-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \bar\psi(m + \gamma^\mu\,\overleftrightarrow{\partial}_\mu)\psi + e\bar\psi i\gamma^\mu\psi A_\mu\right]. \tag{8.14}$$

We see in (8.14) an interaction term of the form $\int d^4x\, ej^\mu A_\mu$ familiar from classical electrodynamics, with electromagnetic current e times the Noether current (5.79):

$$j^\mu = \bar\psi i\gamma^\mu\psi. \tag{8.15}$$

This current is also a Noether current of the action (8.14) corresponding to its global U(1) invariance. The construction of a gauge invariant action has led to a specific interaction between the electromagnetic and spinor fields. The parameter e parametrizes the strength of an interaction term trilinear in the fields. It is evidently a coupling constant, and it is dimensionless, $[e] = 0$.

As suggested by the notation, e is also the elementary unit of charge. This can be seen by the following argument. Suppose A_μ is not dynamical but a given external electromagnetic field. Then the action is bilinear in the dynamical variables (the fermion fields), so the system is like a free field and its quantum version is relatively simple. Suppose furthermore that the fermions are protons.[4] Consider now the operator for electric charge

$$\int d^3x\, e j^0(x) = eQ, \tag{8.16}$$

$$Q = \sum_\lambda \int d\omega_p \left[b^\dagger(p,\lambda,t) b(p,\lambda,t) - d^\dagger(p,\lambda,t) d(p,\lambda,t)\right], \tag{8.17}$$

where we used (6.42). Because A_μ may depend on time we have kept the time dependence in the creation and annihilation operators (this similar to the case of the scalar field in sect. 2.10). On the other hand the Noether argument still shows that Q is time independent and it is reasonable to interpret it as counting the number of particles minus antiparticles at time t. It follows that the electric charge of an eigenstate of Q is e times an integer, so e is the elementary charge unit, and $e > 0$.

For electrons and positrons we replace $e \to -e$, because conventionally the electrons are the 'particles' and the positrons the 'antiparticles', such that the electromagnetic current is $-e\bar{\psi}i\gamma^\mu\psi$ and the electric charge operator is $-eQ$. Denoting the charge of the 'particles' (proton, electron, ...) in units of e by q, we have in general

$$D_\mu\psi = (\partial_\mu - ieqA_\mu)\,\psi, \quad D_\mu\bar{\psi} = (\partial_\mu + ieqA_\mu)\,\bar{\psi}, \tag{8.18}$$

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \bar{\psi}(m + \gamma^\mu \overleftrightarrow{D}_\mu)\psi \tag{8.19}$$

$$= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \bar{\psi}(m + \gamma^\mu \overleftrightarrow{\partial}_\mu)\psi + eq\bar{\psi}i\gamma^\mu\psi\, A_\mu. \tag{8.20}$$

So $q = 1$ for the protons, $q = -1$ for electrons, $q = 2/3$ for the up-quark, etc.

It is of course possible to write down other gauge invariant interactions, for example

$$\mathcal{L}_a = \kappa\, \bar{\psi} S^{\mu\nu}\psi\, F_{\mu\nu}, \tag{8.21}$$

which corresponds to an explicit anomalous magnetic moment interaction ($S^{\mu\nu} = -i[\gamma^\mu, \gamma^\nu]/4$, cf. sect. 8.4). The coupling constant κ has dimension -1 in mass units, and this interaction is not renormalizable. Assuming renormalizability, we discard it. The theory of quantum electrodynamics described by (8.20) is the most general form (conserving parity) with coupling constants of mass dimensions ≥ 0 and it *is* renormalizable. We take it as the best possible starting point for a fundamental theory of only photons and electrons. The theory then *predicts* the magnetic moment of the electron (as well as a host of other physical properties).

[4]This is an approximate description, we know of course that in a more fundamental theory protons are to be described in terms of quarks and gluons.

Let us also note at this point that the gauge group here is not U(1) but the group of real numbers R: ω in $A_\mu \to A_\mu + \frac{1}{e}\partial_\mu\omega$ can be any real number. An important consequence of the group being R is that q does not have to be an integer: it can be any real number. If the gauge group were $U(1)$, then $U = \exp(i\omega) \to U^q = \exp(iq\omega)$ would only be a representation if q where integer,[5] which would imply charge quantization.[6]

8.2 Spinor electrodynamics

We continue with the theory of photons, electrons and positrons specified by the lagrange density (8.20) ('spinor electrodynamics') and quantize it covariantly using the path integral methods described in the previous chapter. The previous reasoning goes through with minor modifications. Let us present it here again in a slightly modified form. Let $F[A,\psi,\bar\psi]$ be a gauge invariant functional made out of the fields and $\langle F\rangle$ be its path integral 'average' (e.g. F could be the field strength or the electromagnetic current, or $\langle F\rangle$ could be a transition amplitude $\langle\text{out}|\text{in}\rangle$). The line of argument is now given by

$$1 = \Delta[A]\int[d\omega]\, e^{iS_{\rm gf}[A^\omega]}, \quad S_{\rm gf}[A] = -\int d^4x\, \frac{1}{2\xi}\,(\partial_\mu A^\mu)^2, \tag{8.22}$$

$$\langle F\rangle \stackrel{?}{=} \frac{\int[dA\, d\bar\psi d\psi]\, e^{iS[A,\bar\psi,\psi]}\, F[A,\psi,\bar\psi]}{\int[dA\, d\bar\psi d\psi]\, e^{iS[A,\psi,\bar\psi]}} \tag{8.23}$$

$$= \frac{\int[dA\, d\bar\psi d\psi]\,[d\omega]\,\Delta[A]\, e^{iS_{\rm gf}[A^\omega]+iS[A,\psi,\bar\psi]}\, F[A,\psi,\bar\psi]}{\int[dA\, d\bar\psi d\psi]\,[d\omega]\,\Delta[A]\, e^{iS_{\rm gf}[A^\omega]+iS[A,\psi,\bar\psi]]}} \tag{8.24}$$

$$= \frac{\int[d\omega]}{\int[d\omega]}\, \frac{\int[dA\, d\bar\psi d\psi]\,\Delta[A]\, e^{iS_{\rm gf}[A]+iS[A,\psi,\bar\psi]}\, F[A,\psi,\bar\psi]}{\int[dA\, d\bar\psi d\psi]\,\Delta[A]\, e^{iS_{\rm gf}[A]+iS[A,\psi,\bar\psi]}} \tag{8.25}$$

$$\Rightarrow \langle F\rangle \equiv \frac{\int[dA\, d\bar\psi d\psi]\,\Delta[A]\, e^{iS_{\rm gf}[A]+iS[A,\psi,\bar\psi]}\, F[A,\psi,\bar\psi]}{\int[dA\, d\bar\psi d\psi]\,\Delta[A]\, e^{iS_{\rm gf}[A]+iS[A,\psi,\bar\psi]}}. \tag{8.26}$$

We made a transformation of variables looking like the gauge transformation $A^\omega = A'$, $\psi^\omega = \psi'$, $\bar\psi^\omega = \bar\psi'$, and used the gauge invariance of S, F, $[dA\, d\bar\psi d\psi]$, and Δ. The last line is the motivation for the path integral with sources, which is not gauge invariant, but out of which gauge invariant $\langle F\rangle$s can be constructed,

$$Z[J,\eta,\bar\eta] = \int[dA\, d\bar\psi d\psi]\,\Delta\, e^{iS+iS_{\rm gf}+i\int d^4x\,(JA+\bar\eta\psi+\bar\psi\eta)}. \tag{8.27}$$

We recall that Δ is a constant in QED.

[5] Otherwise U^q would not be single valued.

[6] We do indeed see something similar to charge quantization in Nature, which can be understood if the electromagnetic gauge group is a subgroup of a semisimple nonabelian group, e.g. in a Grand Unified Theory.

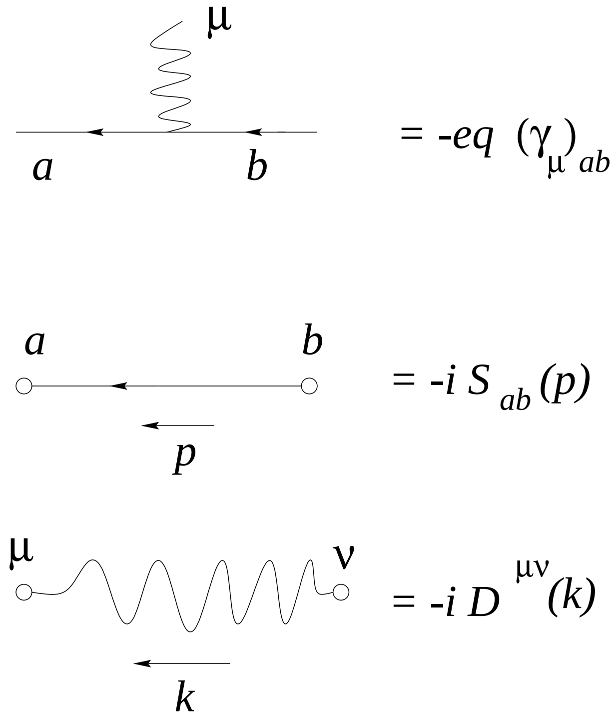

Figure 8.1: Propagators an vertexfunctions in spinor electrodynamics.

Correlation functions can now be defined by differentiation with respect to the sources and they can be calculated in perturbation theory in much the same fashion as in the Yukawa model introduced in section 6.5. The interaction is

$$S_1 = \int d^4x \, eq\bar{\psi} i\gamma^\mu \psi \, A_\mu + \Delta S, \tag{8.28}$$

where ΔS contains counterterms corresponding to mass, charge and wave function renormalization. This will not concern us here as we shall limit ourselves to the tree graph approximation, so $\Delta S \to 0$ in the following.

The ingredients of the diagrams are shown in Fig. 8.1. There is only one bare vertex function, which is in full glory given by

$$S_{\bar{\psi}_a \psi_b A_\mu}(u, v, w) = \int d^4x \, ieq(\gamma_\mu)_{ab} \, \delta^4(u-x) \, \delta^4(v-x) \, \delta^4(w-x), \tag{8.29}$$

such that

$$S_1 = \int d^4u \, d^4v \, d^4w \, \bar{\psi}_a(u) S_{\bar{\psi}_a \psi_b A_\mu}(u, v, w) \psi_b(v) A^\mu(w), \tag{8.30}$$

and a momentum space version which is momentum independent,

$$S_{\bar{\psi}_a \psi_b A_\mu}(p, q, k) = ieq(\gamma_\mu)_{ab}, \tag{8.31}$$

The propagators are given by

$$\begin{aligned} S_{ab}(p) &= \frac{m\,\delta_{ab} - ip_\mu\gamma^\mu_{ab}}{m^2+p^2-i\epsilon}, && (8.32)\\ D^{\mu\nu}(k) &= \frac{\eta^{\mu\nu}}{k^2-i\epsilon} + (\xi-1)\frac{k^\mu k^\nu}{(k^2-i\epsilon)^2}. && (8.33)\end{aligned}$$

We now turn to the calculation of scattering processes in the tree graph approximation. Convenient formulas for obtaining the objects (such as $u(p,\lambda)$, etc.) associated with external fermion or scalar lines in scattering diagrams have already been given in section 6.6. These are also needed for external photon lines. We repeat the expressions for creation and annihilation of photons introduced earlier in (7.162),

$$\begin{aligned} a(k,\lambda,x^0) &= \int d^3x\, e^{-ikx}\, e^*_\mu(k,\lambda)\,(i\overrightarrow{\partial_0} - i\overleftarrow{\partial_0})\,A^\mu(x),\\ a^*(k,\lambda,x^0) &= \int d^3x\, A^\mu(x)\,(i\overrightarrow{\partial_0} - i\overleftarrow{\partial_0})\,e_\mu(k,\lambda)\,e^{ikx}. && (8.34)\end{aligned}$$

where $k^0 = |\mathbf{k}|$, $\lambda = 1, 2$, or $\pm$. In terms of these we have in the free theory (cf. Prob. 7.4)

$$\begin{aligned} \langle A_\mu(x)\,a^*(k,\lambda,-\infty)\rangle_0 &= e_\mu(k,\lambda)\,e^{ikx}, && (8.35)\\ \langle A_\mu(x)\,a^*(k,\lambda,+\infty)\rangle_0 &= 0, && (8.36)\\ \langle a(k,\lambda,-\infty)\,A_\mu(x)\rangle_0 &= 0, && (8.37)\\ \langle a(k,\lambda,+\infty)\,A_\mu(x)\rangle_0 &= e^*_\mu(k,\lambda)\,e^{-ikx}. && (8.38)\end{aligned}$$

The steps given in sect. 6.6 can be followed almost verbatim here as well. Problem 2 deals with photon-electron scattering (called Compton scattering). Note the remarks made in Problem 2.d. Electron-electron scattering is the subject of Problem 3.

Another example is electron-positron scattering, the diagrams of which are shown in Fig. 8.2. The scattering amplitude T is defined in terms of the transition amplitude (for unequal initial and final states) as

$$\begin{aligned} &\langle p_3\lambda_3, \overline{p_4\lambda_4}\,\mathrm{out}|p_1\lambda_1, \overline{p_2\lambda_2}\,\mathrm{in}\rangle\\ \equiv\;& -i(2\pi)^4\delta^4(p_1+p_2-p_3-p_4)\,T(p_3\lambda_3,\overline{p_4\lambda_4};p_1\lambda_1,\overline{p_2\lambda_2})\\ =\;& \langle d(p_4,\lambda_4,\infty)\,b(p_3,\lambda_3,\infty)\,b^*(p_1,\lambda_2,-\infty)\,d^*(p_2,\lambda_2,-\infty)\rangle. && (8.39)\end{aligned}$$

Note the convention of not changing the order of the labels in the bra, compared the ket, and likewise in T, whereas upon conjugation the actual order of b's and d's is interchanged in (8.39). Since we are dealing with fermions, this order matters in getting the right sign. The scattering amplitude is given by

$$\begin{aligned} -iT(p_3\lambda_3,\overline{p_4\lambda_4};p_1\lambda_1,\overline{p_2\lambda_2}) &= -\bar u_3\,e\gamma_\mu\,u_1\;\bar v_2\,e\gamma_\nu\,v_4\;(-i)D^{\mu\nu}(p_1-p_3) \quad (8.40)\\ &\quad + \bar u_3\,e\gamma_\mu\,v_4\;\bar v_2\,e\gamma_\nu\,u_1\;(-i)D^{\mu\nu}(p_1+p_2),\end{aligned}$$

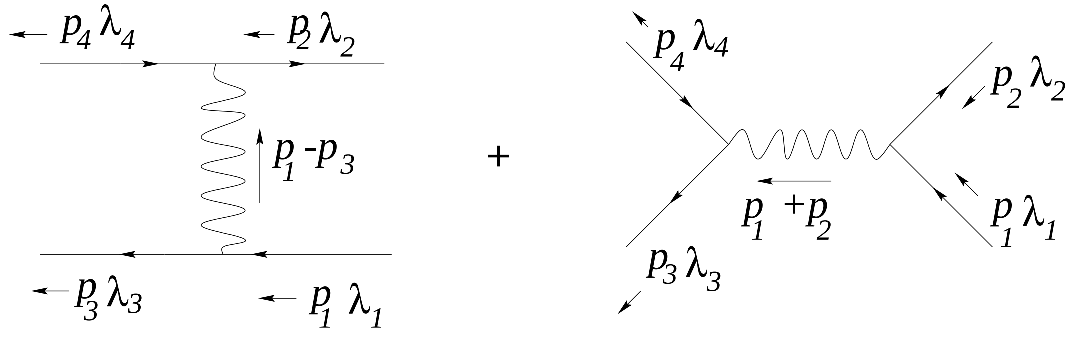

Figure 8.2: Diagram for the process $e_1^- + e_2^+ \rightarrow e_3^- + e_4^+$.

where $\bar{u}_3 = \bar{u}(p_3, \lambda_3)$, etc. The spin averaged squared magnitude $\overline{|T|^2}$ determines the unpolarized cross section. The interference of the two terms makes this calculation less simple than one might wish. In the next section we shall do the calculation for a simpler situation which does not have such interference.

8.3 Example: $e^- + e^+ \rightarrow \mu^- + \mu^+$ scattering

A simple example for working out the unpolarized cross section is the process $e^- + e^+ \rightarrow \mu^- + \mu^+$. The muon mass $M \approx 106$ MeV, much larger than the electron mass $m \approx 0.51$ MeV. For the process to take place, the center-of-mass energy has to be sufficiently high to be able to create the muon-antimuon pair. Similar processes in which quark-antiquarks pairs are produced (which subsequently transform into hadrons) have played an important role in establishing QCD as the theory of the strong interactions, and have led further to the construction of the Standard Model.

We introduce fermion fields for the muon as well as for the electron and the action is given by

$$
\begin{aligned}
S &= \int d^4x \left[-\frac{1}{4} F^2 - \bar{\psi}^{(e)}(m + \gamma\partial)\psi^{(e)} - \bar{\psi}^{(\mu)}(M + \gamma\partial)\psi^{(\mu)} \right. \\
&\quad \left. - e\bar{\psi}^{(e)} i\gamma^\lambda \psi^{(e)} A_\lambda - e\bar{\psi}^{(\mu)} i\gamma^\lambda \psi^{(\mu)} A_\lambda \right]. \qquad (8.41)
\end{aligned}
$$

Because the electron and muon fields are independent there is only one relevant diagram, shown in Fig. 8.3. Then the invariant scattering amplitude is given by

$$
\begin{aligned}
& T(\mu^-(3), \mu^+(4); e^-(1), e^+(2)) \qquad (8.42) \\
&= e^2 \bar{u}(p_3, \lambda_3)\, \gamma_\mu\, v(p_4, \lambda_4)\; \bar{v}(p_2, \lambda_2)\, \gamma_\nu\, u(p_1, \lambda_1) \left[\frac{\eta^{\mu\nu}}{k^2} + (\xi - 1)\frac{k^\mu k^\nu}{(k^2)^2} \right],
\end{aligned}
$$

with $k = p_3 + p_4 = p_1 + p_2$. The gauge terms $\propto k^\mu k^\nu$ in the photon propagator do not contribute because of current conservation (cf. Prob. 3 in chapter 6). For

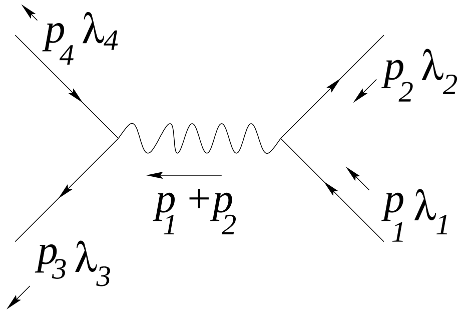

Figure 8.3: Diagram for the process $e_1^- + e_2^+ \to \mu_3^- + \mu_4^+$.

example,

$$\bar{u}(p_3, \lambda_3) i\gamma_\mu v(p_4, \lambda_4)\, (p_3 + p_4)^\mu = 0, \tag{8.43}$$

where we used the fact that the polarization spinors satisfy the (momentum space version of the) Dirac equation,

$$\bar{u}(p_3, \lambda)\, i\gamma p_3 = -M\bar{u}(p_3, \lambda_3), \quad i\gamma p_4\, v(p_4, \lambda_4) = Mv(p_4, \lambda_4). \tag{8.44}$$

Thus we reach the important conclusion that the scattering amplitude is independent of the gauge parameter ξ, as it should be.

To calculate the cross section we need T^*, which leads to

$$(\bar{u}_3\gamma_\rho v_4)^* = v_4^\dagger \gamma_\rho^\dagger \beta u_3 = -\bar{v}_4\gamma_\rho u_3, \tag{8.45}$$

$$(\bar{v}_2\gamma_\sigma u_1)^* = u_1^\dagger \gamma_\sigma^\dagger \beta v_2 = -\bar{u}_1\gamma_\sigma v_2. \tag{8.46}$$

Averaging over initial and final spins gives

$$\overline{|T|^2} = e^4\, \frac{1}{16} \sum_{\lambda_3\lambda_1\lambda_2\lambda_4} \bar{u}_3\gamma_\mu v_4\, \bar{v}_2\gamma_\nu u_1\, \bar{v}_4\gamma_\rho u_3\, \bar{u}_1\gamma_\sigma v_2\, \frac{\eta^{\mu\nu}\eta^{\rho\sigma}}{s^2}, \tag{8.47}$$

where $s = -(p_3 + p_4)^2 = -(p_1 + p_2)^2$ is one of the Mandelstam variables (equal to the total cm energy squared). To evaluate the polarization sums we order the spinor factors in a suggestive way, interpreting $u_a\bar{u}_b$ and $v_a\bar{v}_b$ as matrices and using, for example,

$$\bar{u}_3\gamma_\mu v_4\bar{v}_4\gamma_\rho u_3 = \mathrm{Tr}\,[\gamma_\mu v_4\bar{v}_4\gamma_\rho u_3\bar{u}_3]. \tag{8.48}$$

Then

$$\begin{aligned}\overline{|T|^2} &= e^4\, \frac{\eta^{\mu\nu}\eta^{\rho\sigma}}{s^2}\, \frac{1}{16}\, \mathrm{Tr}\,[\gamma_\mu(\sum_{\lambda_4} v_4\bar{v}_4)\gamma_\rho(\sum_{\lambda_3} u_3\bar{u}_3)] \\ &\quad \mathrm{Tr}\,[\gamma_\nu(\sum_{\lambda_1} u_1\bar{u}_1)\gamma_\sigma(\sum_{\lambda_2} v_2\bar{v}_2)]. \end{aligned} \tag{8.49}$$

We now use the properties

$$\sum_{\lambda_4} v(p_4,\lambda_4)\bar{v}(p_4,\lambda_4) = -(M+i\gamma p_4), \tag{8.50}$$

$$\sum_{\lambda_1} u(p_1,\lambda_1)\bar{u}(p_1,\lambda_1) = m-i\gamma p_1, \tag{8.51}$$

etc. and obtain the form

$$\begin{aligned}\overline{|T|^2} &= e^4\,\frac{\eta^{\mu\nu}\eta^{\rho\sigma}}{s^2}\,\frac{1}{16}\,\mathrm{Tr}\,[\gamma_\mu(M+i\gamma p_4)\gamma_\rho(M-i\gamma p_3)] \\ &\quad \mathrm{Tr}\,[\gamma_\nu(m-i\gamma p_1)\gamma_\sigma(m+i\gamma p_2)]. \end{aligned}\tag{8.52}$$

To evaluate this we use the trace formulas

$$\mathrm{Tr}\,\gamma_\kappa\gamma_\lambda = 4\eta_{\kappa\lambda}, \tag{8.53}$$

$$\mathrm{Tr}\,\gamma_\kappa\gamma_\lambda\gamma_\mu = 0, \tag{8.54}$$

$$\mathrm{Tr}\,\gamma_\kappa\gamma_\lambda\gamma_\mu\gamma_\nu = 4(\eta_{\kappa\lambda}\eta_{\mu\nu}-\eta_{\kappa\mu}g_{\lambda\nu}+\eta_{\kappa\nu}\eta_{\lambda\mu}). \tag{8.55}$$

These follow from the fact that (1) the trace of a product of gamma matrices vanishes unless each γ_0, ..., γ_3 appears an even number of times, (2) $\gamma_0^2=-1$, $\gamma_3^2=\gamma_1^2=\gamma_2^2=1$, (2) the gamma's anticommute and (4) $\mathrm{Tr}\,1=4$. More identities are in Appendix 5.7. The two traces in (8.52) are given by

$$4(M^2\eta_{\mu\rho}+p_{4\mu}p_{3\rho}-\eta_{\mu\rho}p_3p_4+p_{3\mu}p_{4\rho}) \tag{8.56}$$

and

$$4(m^2\eta_{\nu\sigma}+p_{1\nu}p_{2\sigma}-\eta_{\nu\sigma}p_1p_2+p_{2\nu}p_{1\sigma}). \tag{8.57}$$

The evaluation of $\overline{|T|^2}$ is now straightforward and results in a large number of scalar products of the momenta. Using the Mandelstam variables

$$s=-(p_3+p_4)^2=2M^2-2p_3p_4=-(p_1+p_2)^2=2m^2-2p_1p_2, \tag{8.58}$$

$$t=-(p_3-p_1)^2=-m^2-M^2+2p_3p_1=-(p_2-p_4)^2=-m^2-M^2+2p_2p_4, \tag{8.59}$$

$$u=-(p_3-p_2)^2=-m^2-M^2+2p_3p_2=-(p_1-p_4)^2=-m^2-M^2+2p_1p_4, \tag{8.60}$$

the result simplifies to

$$\begin{aligned}\overline{|T|^2} &= \frac{e^4}{s^2}\,[4m^2M^2+M^2(s-2m^2)+m^2(s-2M^2) \\ &\quad +\frac{1}{2}\,(t+m^2+M^2)^2+\frac{1}{2}\,(u+m^2+M^2)^2]. \end{aligned}\tag{8.61}$$

We recall that u can be eliminated in favor of s and t by the relation $s+t+u=2m^2+2M^2$. At high energies we can neglect the electron and muon masses. Then

$$\overline{|T|^2}\approx\frac{e^4}{2s^2}\,(t^2+u^2). \tag{8.62}$$

Under these cricumstances t and u are related to the scattering angle in the centre of mass frame by

$$t \approx -\frac{1}{2}\, s\,(1-\cos\theta), \quad u \approx -\frac{1}{2}\, s\,(1+\cos\theta), \tag{8.63}$$

and we get for the differential crossection at high energies

$$\left[\frac{d\sigma}{d\Omega}\right]_{\rm cm} = \frac{1}{64\pi^2 s}\frac{|\mathbf{p}_3|}{|\mathbf{p}_1|}\, 4\overline{|T|^2} \tag{8.64}$$

$$\approx \frac{\alpha^2}{4s}\,(1+\cos^2\theta), \quad \alpha = \frac{e^2}{4\pi}, \tag{8.65}$$

where α is the finestructure constant. The factor 4 in front of $\overline{|T|^2}$ represents the summation over final spins. The total cross section is given by

$$\sigma = 2\pi\,\frac{\alpha^2}{4s}\int_{-1}^{1} d\cos\theta\,(1+\cos^2\theta) = \frac{4\pi\alpha^2}{3s}. \tag{8.66}$$

It is instructive to rederive these formulas by evaluating first the high energy form of T for given polarization combinations using helicity spinors, and from this $\overline{|T|^2}$. See e.g. Peskin & Schroeder chapter 5.

We shall now briefly review the relation with hadron production. The total cross section $\sigma(e^+ + e^- \to \text{hadrons})$ goes via the intermediate production of quark-antiquark pairs, which subsequently transform into hadrons. The quarks come in three colors for every flavor $f = u,\, d,\, s,\, c,\, b,\, t$ (i.e. up, down, strange, charm, bottom, top). Their *effective* masses are of the order of $m_u \approx m_d \approx 0.3$ GeV, $m_s \approx 0.4$ GeV, $m_c \approx 1.4$ GeV, $m_b \approx 4.1$ GeV and $m_t \approx 175$ GeV. At energies much higher than the thresholds for production of the quark-antiquark bound states and resonances, the ratio of the cross section into hadrons is simply given by

$$R \equiv \frac{\sigma(e^+ + e^- \to \text{hadrons})}{\sigma(e^+ + e^- \to \mu^+ + \mu_-)} = 3\sum_f q_f^2 + \text{corrections}, \tag{8.67}$$

where the sum is over the contributing flavors and the factor 3 is the number of colors. The electric charges of the quarks are given by

$$q_u = q_c = q_t = \frac{2}{3}, \quad q_d = q_s = q_b = -\frac{1}{3}. \tag{8.68}$$

The corrections (mainly due to QCD interactions) are reasonably small at energies above a few GeV; at higher energies electroweak radiative corrections become also important.

Experimentally one finds (cf. Fig. 8.4) in the region leading up to the charm-anticharm threshold at ≈ 3 GeV, modulo the corrections,

$$R \approx 3\left(\frac{4}{9} + \frac{1}{9} + \frac{1}{9}\right) = 2,$$

10 ***37. Plots of cross sections and related quantities***

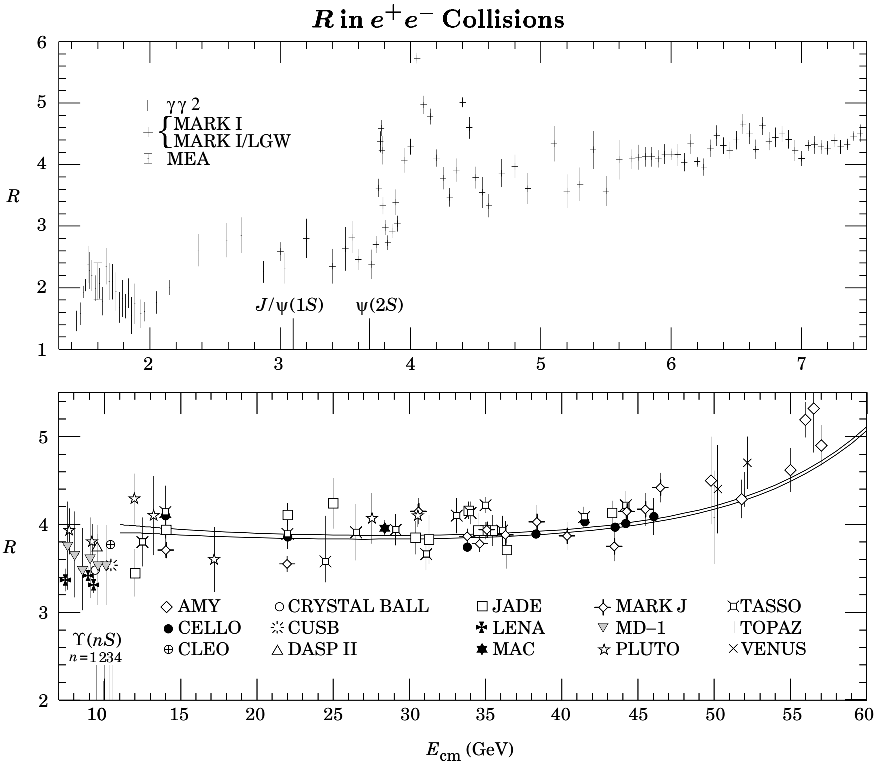

Figure 37.16: Selected measurements of $R \equiv \sigma(e^+e^- \to \text{hadrons})/\sigma(e^+e^- \to \mu^+\mu^-)$, where the annihilation in the numerator proceeds via one photon or via the Z. Measurements in the vicinity of the Z mass are shown in the following figure. The denominator is the calculated QED single-photon process; see the section on Cross-Section Formulae for Specific Processes. Radiative corrections and, where important, corrections for two-photon processes and τ production have been made. Note that the ADONE data ($\gamma\gamma2$ and MEA) is for ≥ 3 hadrons. The points in the $\psi(3770)$ region are from the MARK I—Lead Glass Wall experiment. To preserve clarity only a representative subset of the available measurements is shown—references to additional data are included below. Also for clarity, some points have been combined or shifted slightly ($< 4\%$) in $E_{\rm cm}$, and some points with low statistical significance have been omitted. Systematic normalization errors are not included; they range from ~5–20%, depending on experiment. We caution that especially the older experiments tend to have large normalization uncertainties. Note the suppressed zero. The horizontal extent of the plot symbols has no significance. The positions of the $J/\psi(1S)$, $\psi(2S)$, and the four lowest Υ vector-meson resonances are indicated. Two curves are overlaid for $E_{\rm cm} > 11$ GeV, showing the theoretical prediction for R, including higher order QCD [M. Dine and J. Sapirstein, Phys. Rev. Lett. **43**, 668 (1979)] and electroweak corrections. The Λ values are for 5 flavors in the $\overline{\rm MS}$ scheme and are $\Lambda^{(5)}_{\overline{\rm MS}} = 60$ MeV (lower curve) and $\Lambda^{(5)}_{\overline{\rm MS}} = 250$ MeV (upper curve). (Courtesy of F. Porter, 1992.) References (including several references to data not appearing in the figure and some references to preliminary data):

AMY: T. Mori *et al.*, Phys. Lett. **B218**, 499 (1989);
CELLO: H.-J. Behrend *et al.*, Phys. Lett. **144B**, 297 (1984); and H.-J. Behrend *et al.*, Phys. Lett. **183B**, 400 (1987);
CLEO: R. Giles *et al.*, Phys. Rev. **D29**, 1285 (1984); and D. Besson *et al.*, Phys. Rev. Lett. **54**, 381 (1985);
CUSB: E. Rice *et al.*, Phys. Rev. Lett. **48**, 906 (1982);
CRYSTAL BALL: A. Osterheld *et al.*, SLAC-PUB-4160; and Z. Jakubowski *et al.*, Z. Phys. **C40**, 49 (1988);
DASP: R. Brandelik *et al.*, Phys. Lett. **76B**, 361 (1978);
DASP II: Phys. Lett. **116B**, 383 (1982);
DCI: G. Cosme *et al.*, Nucl. Phys. **B152**, 215 (1979);
DHHM: P. Bock *et al.* (DESY-Hamburg-Heidelberg-MPI München Collab.), Z. Phys. **C6**, 125 (1980);
γγ2: C. Bacci *et al.*, Phys. Lett. **86B**, 234 (1979);
HRS: D. Bender *et al.*, Phys. Rev. **D31**, 1 (1985);
JADE: W. Bartel *et al.*, Phys. Lett. **129B**, 145 (1983); and W. Bartel *et al.*, Phys. Lett. **160B**, 337 (1985);
LENA: B. Niczyporuk *et al.*, Z. Phys. **C15**, 299 (1982).
MAC: E. Fernandez *et al.*, Phys. Rev. **D31**, 1537 (1985);
MARK J: B. Adeva *et al.*, Phys. Rev. Lett. **50**, 799 (1983); and B. Adeva *et al.*, Phys. Rev. **D34**, 681 (1986);
MARK I: J.L. Siegrist *et al.*, Phys. Rev. **D26**, 969 (1982);
MARK I + Lead Glass Wall: P.A. Rapidis *et al.*, Phys. Rev. Lett. **39**, 526 (1977); and P.A. Rapidis, thesis, SLAC-Report-220 (1979);
MARK II: J. Patrick, Ph.D. thesis, LBL-14585 (1982);
MD-1: A.E. Blinov *et al.*, Z. Phys. **C70**, 31 (1996);
MEA: B. Esposito *et al.*, Lett. Nuovo Cimento **19**, 21 (1977);
PLUTO: A. Bäcker, thesis Gesamthochschule Siegen, DESY F33-77/03 (1977); C. Gerke, thesis, Hamburg Univ. (1979); Ch. Berger *et al.*, Phys. Lett. **81B**, 410 (1979); and W. Lackas, thesis, RWTH Aachen, DESY Pluto-81/11 (1981);
TASSO: R. Brandelik *et al.*, Phys. Lett. **113B**, 499 (1982); and M. Althoff *et al.*, Phys. Lett. **138B**, 441 (1984);
TOPAZ: I. Adachi *et al.*, Phys. Rev. Lett. **60**, 97 (1988); and
VENUS: H. Yoshida *et al.*, Phys. Lett. **198B**, 570 (1987).

Figure 8.4: Experimental results for R (from the web site of the Particle Data Group, see http://pdg.lbl.gov/2000/contents_plots.html).

then, after passing through the $c\bar{c}$ resonance region leading up to the bottom-antibottom threshold at ≈ 9 GeV,

$$R \approx 2 + 3\frac{4}{9} = \frac{10}{3}.$$

Beyond the upsilon resonance region R settles at about

$$R \approx \frac{10}{3} + 3\frac{1}{9} = \frac{11}{3},$$

presumably untill the top-antitop The ratio R has given important confirmation for the color degree of freedom of the hadron constituents (i.e. the quarks): without it theory and experiment would disagree by about a factor 3. For further discussion see Peskin and Schroeder chapter 5, De Wit & Smith ch. 6, Brown sect. 8.2,

8.4 Magnetic moment of the electron

In non-relativistic quantum mechanics, an electron in an external static electromagnetic potential is described by the hamiltonian

$$\hat{H} = \frac{\hat{\mathbf{p}}^2 + e[\hat{\mathbf{p}} \cdot \mathbf{A}(\hat{\mathbf{x}}) + \mathbf{A}(\hat{\mathbf{x}}) \cdot \hat{\mathbf{p}}] + e^2\mathbf{A}^2(\hat{\mathbf{x}})}{2m} - eA^0(\hat{\mathbf{x}}) + \frac{eg}{2m}\mathbf{S} \cdot \hat{\mathbf{B}}(\hat{\mathbf{x}}), \quad (8.69)$$

where $\hat{\mathbf{x}}$ is the position operator, $\hat{\mathbf{p}}$ the momentum operator, $\hat{\mathbf{S}}$ the spin operator, g is the Landé g-factor (gyromagnetic ratio) and $\mathbf{B} = \nabla \times \mathbf{A}$ is the magnetic field. The terms $\hat{\mathbf{p}} \cdot \mathbf{A}(\hat{\mathbf{x}}) + \mathbf{A}(\hat{\mathbf{x}}) \cdot \hat{\mathbf{p}} + e^2\mathbf{A}^2(\hat{\mathbf{x}})$ come from the 'minimal substitution' $\hat{\mathbf{p}}^2 \to (\hat{\mathbf{p}} + e\mathbf{A}(\hat{\mathbf{x}}))^2$. It will be shown in this section that in the approximation where (8.69) is valid, spinor electrodynamics predicts $g = 2$.

We first derive the form of the non-relativistic H in the momentum representation and then identify the corresponding form in spinor electrodynamics. Using momentum states with the normalization

$$\langle \mathbf{p}'\lambda'|\mathbf{p}\lambda\rangle = (2\pi)^3\,\delta(\mathbf{p}' - \mathbf{p})\,\delta_{\lambda'\lambda}, \quad (8.70)$$

we have

$$\langle \mathbf{x}\lambda'|\mathbf{p}\lambda\rangle = \delta_{\lambda'\lambda}\,e^{i\mathbf{p}\mathbf{x}}. \quad (8.71)$$

The momentum representation of H is given by

$$\begin{aligned} \langle \mathbf{p}'\lambda'|\hat{H}|\mathbf{p}\lambda\rangle \;=\; & \frac{\mathbf{p}^2}{2m}(2\pi)^3\,\delta(\mathbf{p}' - \mathbf{p})\,\delta_{\lambda'\lambda} + \frac{e}{2m}(\mathbf{p}' + \mathbf{p}) \cdot \tilde{\mathbf{A}}(\mathbf{p}' - \mathbf{p})\,\delta_{\lambda'\lambda} \\ & + \frac{e^2}{2m}\langle \mathbf{p}'\lambda'|\mathbf{A}^2(\hat{\mathbf{x}})|\mathbf{p}\lambda\rangle \\ & - e\tilde{A}^0(\mathbf{p}' - \mathbf{p})\,\delta_{\lambda'\lambda} + \frac{eg}{4m}\boldsymbol{\sigma}_{\lambda'\lambda} \cdot \tilde{\mathbf{B}}(\mathbf{p}' - \mathbf{p}), \end{aligned} \quad (8.72)$$

where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the vector of Pauli matrices and we used

$$\begin{aligned} \langle \mathbf{p}'\lambda'|A^\mu(\hat{\mathbf{x}})|\mathbf{p}\lambda\rangle &= \sum_{\lambda''}\int d^3x\, \langle \mathbf{p}'\lambda'|\mathbf{x}\lambda''\rangle\, \langle \mathbf{x}\lambda''|\mathbf{p}\lambda\rangle\, A^\mu(\mathbf{x}) && (8.73)\\ &= \delta_{\lambda'\lambda}\, \tilde{A}^\mu(\mathbf{p}'-\mathbf{p}), && (8.74)\\ \tilde{A}^\mu(\mathbf{k}) &= \int d^3x\, e^{-i\mathbf{kx}}\, A^\mu(\mathbf{x}). && (8.75)\end{aligned}$$

The Fourier transform of the magnetic field is related to that of the vector potential by

$$\tilde{B}_k(\mathbf{k}) = i\epsilon_{lmn}\, k_m \tilde{A}_n(\mathbf{k}). \qquad (8.76)$$

In spinor electrodynamics, the approximation where (8.72) is valid, is the non-relativistic approximation in which radiation effects corresponding to the quantized photon field are neglected. So we consider the electron field in an external static electromagnetic potential $A^\mu(\mathbf{x})$. The hamiltonian of this system can be derived by the Noether argument and is given by

$$\begin{aligned} \hat{H} &= \int d^3x\, \hat{\psi}^\dagger(\mathbf{x}) \left[m\beta - eA^0(\mathbf{x}) + \boldsymbol{\alpha}\cdot(-i\nabla + ie\mathbf{A}(\mathbf{x}))\right] \hat{\psi}(\mathbf{x}) && (8.77)\\ &= \hat{H}_0 - \int d^3x\, (-e)\hat{j}^\mu(x)\, A_\mu(\mathbf{x}), && (8.78)\end{aligned}$$

with $\hat{H}_0$ the free fermion hamiltonian. Using our usual relativistic normalization, the one-particle eigenstates of the free hamiltonian satisfy

$$\hat{H}_0|p\lambda\rangle = p^0|p\lambda\rangle, \quad p^0 = \sqrt{m^2+\mathbf{p}^2}, \quad \langle p'\lambda'|p\lambda\rangle = 2p^0(2\pi)^3\delta(\mathbf{p}'-\mathbf{p})\delta_{\lambda'\lambda}. \quad (8.79)$$

The matrix elements of $\hat{H}$ in the one-particle subspace are given by

$$\langle p'\lambda'|H|p\lambda\rangle = p^0\, \langle p'\lambda'|p\lambda\rangle + e\int d^3x\, \langle p'\lambda'|\hat{j}^\mu(x)|p\lambda\rangle\, A_\mu(\mathbf{x}). \qquad (8.80)$$

Using the results derived in Problem 6.3 we have

$$\langle p'\lambda'|\hat{j}^\mu(x)|p\lambda\rangle = \bar{u}(p',\lambda')\, i\gamma^\mu\, u(p,\lambda)\, e^{i(p-p')x}, \qquad (8.81)$$

$$\bar{u}(p',\lambda')\, i\gamma^\mu\, u(p,\lambda) = \frac{1}{2m}\, \bar{u}(p',\lambda')\, [(p+p')^\mu + 2i(p-p')_\nu\, S^{\mu\nu}]\, u(p,\lambda). \qquad (8.82)$$

In the non-relativistic approximation the spinor-matrix elements become

$$\begin{aligned} \bar{u}(p',\lambda')\, u(p,\lambda) &= 2m\, [\delta_{\lambda'\lambda} + O(\mathbf{p}^2/m^2)], && (8.83)\\ \bar{u}(p',\lambda')\, S^{0n}\, u(p,\lambda) &= 2m\, [O(|\mathbf{p}|/m)], && (8.84)\\ \bar{u}(p',\lambda')\, S^{mn}\, u(p,\lambda) &= 2m\, [\epsilon_{mnl}\, (\sigma_l/2)_{\lambda'\lambda} + O(|\mathbf{p}|/m)]. && (8.85)\end{aligned}$$

In the chiral representation $S^{0n} = \gamma_5\sigma_n/2$ and γ_5 anticommutes with β. So at zero momentum, $\beta u = u$, $\bar{u}\beta = \bar{u}$, and it follows that $\bar{u}'\, S^{0n}\, u = 0$. Substitution in (8.80) and deviding by the relativistic normalization factor $1/2p^0 \to 1/2m$ to get the same normalization as in (8.70), reproduces the magnetic moment coupling in (8.72) with $g = 2$. The rest energy m is omitted in the usual non-relativistic expressions. The $\mathbf{A}^2$ term in (8.72) is not reproduced this way; this will be clarified in section 8.5.

The g factor has been calculated in higher orders in perturbation theory (including the quantized photon field) to great accuracy, for the electron and the muon. The theoretical calculations agree with stunningly precise experimental results,

$$(g_e - 2)/2 = (1159.652193 \pm 0.000010) \times 10^{-6}, \tag{8.86}$$

$$(g_\mu - 2)/2 = (1165.9230 \pm 0.0084) \times 10^{-6}, \tag{8.87}$$

to eight significant digits in the case of the electron. For more comments on precision results in QED, see for example Peskin & Schroeder, end of section 6.3.

8.5 Gauge-invariant non-relativistic reduction

In the non-relativistic regime the energies are so small that particle creation is not possible becasue of energy conservation. We may then expect the number of particles and antiparticles to be separately conserved to a very good approximation, and indeed, the non-relativistic field theory derived in section 6.4 has a corresponding U(1) symmetry. Similarly, a one-particle wave function in the momentum representation,

$$\Phi_\lambda(\mathbf{p}, t) = \langle \mathbf{p}\lambda|\Phi, t\rangle, \quad |\mathbf{p}\lambda\rangle \equiv |p\lambda\rangle/\sqrt{2p^0}, \tag{8.88}$$

may be expected to satisfy a Schrödinger equation

$$i\frac{d}{dt}\Phi_\lambda(\mathbf{p}, t) = \sum_{\lambda'} \int \frac{d^3p'}{(2\pi)^3} \langle \mathbf{p}\lambda|\hat{H}|\mathbf{p}'\lambda'\rangle\, \Phi_{\lambda'}(\mathbf{p}', t) + \cdots, \tag{8.89}$$

for which the contribution $\cdots$ of other intermediate states (e.g. multi-particle states) is negligible in the non-relativistic limit.

However, the one-particle matrix element $\langle \mathbf{p}\lambda|\hat{H}|\mathbf{p}'\lambda'\rangle$ is linear in $\mathbf{A}$ (see (8.80)), and the $\mathbf{A}^2$ term of (8.72) is lacking. This quadratic term is needed for gauge-covariance of the non-relativistic Schrödinger equation.[7] Evidently, the $\cdots$ contribution in (8.89) must contain the $\mathbf{A}^2$ term somehow, but bringing this to the fore could be cumbersome. Instead, let us give a heuristic derivation of the non-relativistic theory along the lines of section 6.4.

[7] Although the original hamiltonian $\hat{H}$ is gauge invariant, its matrix elements $\langle \mathbf{p}\lambda|\hat{H}|\mathbf{p}'\lambda'\rangle$ do not have this property, because free hamiltonian $\hat{H}_0$ and its eigenstates $|\mathbf{p}\lambda\rangle$ are not gauge invariant.

Consider the path integral for the electron field coupled to an external electromagnetic potential, with sources η and $\bar{\eta}$,

$$\begin{aligned} Z[\eta,\bar{\eta}] &= \int [d\bar{\psi}d\psi] \exp\left\{i\int d^4x \, [-\bar{\psi}(\gamma D+m)\psi + \bar{\eta}\psi + \bar{\psi}\eta]\right\} \\ &= Z[0,0] \exp\left\{i\int d^4x\, d^4y\, \bar{\eta}(x)G(x,y,A)\eta(y)\right\}, \end{aligned} \tag{8.90}$$

where $D_\mu = \partial_\mu + ieA_\mu$ and $G(x,y,A)$ is the fermion propagator in the given A_μ field. It satisfies

$$(m+\gamma^\mu D_\mu)_{ab}\, G_{bc}(x,y,A) = \delta_{ac}\delta^4(x-y), \tag{8.91}$$

with appropriate boundary conditions, or in condensed notation (suppressing the $i\epsilon$),

$$G = (m+\gamma D)^{-1}. \tag{8.92}$$

We now first try to write G in a form similar to that of the free propagator: $(m+\gamma\partial)^{-1} = (m-\gamma\partial)(m^2-\partial^2)^{-1}$, which in momentum space is the familiar $(m+i\gamma p)^{-1} = (m-i\gamma p)(m^2+p^2)^{-1}$. Using $\gamma^\mu\gamma^\nu = \eta^{\mu\nu} + 2iS^{\mu\nu}$ and $[D_\mu, D_\nu] = ieF_{\mu\nu}$, with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ as usual, we have

$$\begin{aligned} (m+\gamma D)(m-\gamma D) &= m^2 - \gamma^\mu\gamma^\nu D_\mu D_\nu = m^2 - D^\mu D_\mu + eS^{\mu\nu}F_{\mu\nu} \\ &\equiv m^2 - D^2 + eSF. \end{aligned} \tag{8.93}$$

Hence,

$$G = (m-\gamma D)(m^2-D^2+eSF)^{-1} = (m^2-D^2+eSF)^{-1}(m-\gamma D). \tag{8.94}$$

We now mimic the approximations made in section 6.4, treating eA_0 of the order of the kinetic energy $\mathbf{p}^2/2m$ and $e\mathbf{A}$ of the order of the momentum $|\mathbf{p}|$. We start with the analog of (6.117):

$$m-\gamma D \to m - \gamma^0\partial_0 = m + \beta\, i\partial_0 \to 2mP_+, \quad P_+ = (1+\beta)/2, \tag{8.95}$$

which leads again to the conclusion that the non-relativistic sources (containing only frequencies near $\pm m$) have only components in the $\beta = +1$ subspace. In the present Dirac case (instead of the Majorana case of sect. 6.4), this is implemented by

$$\eta(x) \to e^{-imx^0}\,\zeta(x), \quad \bar{\eta}(x) \to e^{imx^0}\,\zeta^\dagger(x), \quad \zeta = P_+\zeta, \quad \zeta^\dagger = \zeta^\dagger P_+, \tag{8.96}$$

where ζ and ζ^* possess only frequencies with magnitude $\ll m$, and wave vectors $|\mathbf{p}| \ll m$.

The time components S^{0n} of the matrices $S^{\mu\nu}$ appearing in the denominator in (8.94) may give transitions between the subspaces with $\beta = +1$ and -1, but this is a higher order effect: we have[8]

$$\begin{aligned}
& P_+(m^2 - D^2 + eSF)^{-1}P_+ \\
= \; & P_+(m^2 - D^2)^{-1} - (m^2 - D^2)^{-1}P_+ eSF\, P_+\,(m^2 - D^2)^{-1} + \cdots \\
= \; & P_+\left[(m^2 - D^2)^{-1} - (m^2 - D^2)^{-1} 2e\mathbf{S}\cdot\mathbf{B}(m^2 - D^2)^{-1}\right] + \cdots \\
\approx \; & P_+(m^2 - B^2 + 2e\mathbf{S}\cdot\mathbf{B})^{-1},
\end{aligned} \tag{8.97}$$

where we used the fact that S^{0n} anticommutes with β, whereas the spin matrices $S_{mn} = \epsilon_{mnp}S_p$ commute with β, so $P_+S^{0n}P_+ = 0$, and

$$P_+SF\,P_+ = P_+(2S^{0n}F_{0n} + S^{mn}F_{mn})P_+ = 2\mathbf{S}\cdot\mathbf{B}\,P_+. \tag{8.98}$$

The analog of (6.116), combined with (8.94) – (8.97) now leads to

$$\begin{aligned}
\bar{\eta}G\eta &= \bar{\eta}(m - \gamma D)(m^2 - D^2 + eSF)^{-1}\eta \\
&\approx \zeta^\dagger \left(-\mathbf{D}^2/2m + 2e\mathbf{S}\cdot\mathbf{B}/2m - 2m(iD_0 + i\epsilon)\right)^{-1}\zeta.
\end{aligned} \tag{8.99}$$

The corresponding path integral in terms of two-component non-relativistic spinor fields is given by

$$Z[\zeta,\zeta^*] = \int [d\psi^* d\psi]\,\exp\left\{i\int d^4x\,\psi^\dagger\left[iD_0 - \frac{-\mathbf{D}^2}{2m} - \frac{2e}{2m}\mathbf{S}\cdot\mathbf{B}\right]\psi + \zeta^\dagger\psi + \psi^\dagger\zeta\right\}. \tag{8.100}$$

We see again the magnetic-moment coupling with g-factor 2, and the gauge invariance is manifest.

8.6 Summary

Gauge invariance and renormalizability are guiding principles in constructing covariant actions for interacting fields. QED is a prime example, which predicts a wealth of phenomena which have been verified experimentally.

Scattering amplitudes can be shown to be independent of the gauge parameter specifying a covariant gauge. Spin-averaged amplitudes are Lorentz invariant and can be evaluated with the trace technique. The Landé g-factor, which

[8]A useful expansion formula is

$$\begin{aligned}
(A+B)^{-1} &= [(1+BA^{-1})A]^{-1} = A^{-1}(1+BA^{-1})^{-1} = A^{-1}\left[1 + \sum_{n=1}^{\infty}(-BA^{-1})^n\right] \\
&= A^{-1} - A^{-1}BA^{-1} + A^{-1}BA^{-1}BA^{-1} + \cdots.
\end{aligned}$$

parametrizes the non-relativistic magnetic-moment coupling of the electron to an external magnetic field, is 2, with calculable corrections that have been verified by experiment to extremely high precision.

The road is now open for generalization to theories with non-abelian gauge fields, such as the Standard Model.

8.7 Problems

1. *Gauge invariant measure*

   Verify formally (using the ill defined $\prod_x$) the gauge invariance of $[dA]$ and $[d\bar{\psi}d\psi]$.

2. *Compton scattering*

   The elastic scattering of a photon and an electron is an example of Compton scattering. Consider the process $\gamma(k\kappa)+e^-(p\lambda) \to \gamma(k'\kappa')+e^-(p'\lambda')$, where $k,\ldots,p'$ denote the four momenta and $\kappa,\ldots,\lambda'$ the spin polarizations of the particles. The amplitude can be calculated by expanding[9]

$$\begin{aligned}
&\langle k'\kappa', p'\lambda'\,\mathrm{out}|k\kappa, p\lambda\,\mathrm{in}\rangle \\
&= \langle A^{(+)}(k'\kappa',\infty)\psi^{(+)}(p'\lambda',\infty)\bar{\psi}^{(-)}(p\lambda,-\infty)A^{(-)}(k\kappa,-\infty)\rangle \\
&= \frac{\langle A^{(+)}(k'\kappa',\infty)\psi^{(+)}(p'\lambda',\infty)\bar{\psi}^{(-)}(p\lambda,-\infty)A^{(-)}(k\kappa,-\infty)\, e^{iS_1}\rangle_0}{\langle e^{iS_1}\rangle_0}
\end{aligned}$$

   in terms of

$$S_1 = \int d^4x\, eq\bar{\psi}i\gamma^\mu\psi A_\mu + \Delta S. \tag{8.101}$$

   Recall that $q = -1$ for the electron field; the counterterms ΔS can be neglected in the tree-graph approximation. The first nontrivial scattering occurs at second order and given by

$$\begin{aligned}
\langle k'\kappa', p'\lambda'\,\mathrm{out}|k\kappa, p\lambda\,\mathrm{in}\rangle \;=\; & \frac{(ieq)^2}{2!}\int d^4x\, d^4x'\, \langle A^{(+)}(k'\kappa',\infty)\psi^{(+)}(p'\lambda',\infty) \\
& \bar{\psi}(x)i\gamma^\mu\psi(x)A_\mu(x)\bar{\psi}(x')i\gamma^\nu\psi(x')A_\nu(x') \\
& \bar{\psi}^{(-)}(p\lambda,-\infty)A^{(-)}(k\kappa,-\infty)\rangle_0 + \cdots
\end{aligned} \tag{8.102}$$

   In short-hand, we evaluate the contractions

$$\begin{aligned}
\langle A^{(+)}A_\mu A'_\nu A^{(-)}\rangle_0 \;=\; & \langle A^{(+)}A_\mu\rangle_0\langle A'_\nu A^{(-)}\rangle_0 + \langle A^{(+)}A'_\nu\rangle_0\langle A_\mu A^{(-)}\rangle_0 \\
& + \langle A^{(+)}A^{(-)}\rangle_0\langle A_\mu A'_\nu\rangle_0
\end{aligned} \tag{8.103}$$

[9]For a change, we use the notation $A^{(+)}(k,\kappa,t) = a(k,\kappa,t)$, etc.

(the last contraction leads to loop diagrams) and

$$\begin{aligned}\langle\psi^{(+)}\bar{\psi}i\gamma^{\mu}\psi\bar{\psi}'i\gamma^{\nu}\psi'\bar{\psi}^{(-)}\rangle_0 &= \langle\psi^{(+)}\bar{\psi}\rangle_0 i\gamma^{\mu}\langle\psi\bar{\psi}'\rangle_0 i\gamma^{\nu}\langle\psi'\bar{\psi}^{(-)}\rangle_0 \\ &\quad + x\leftrightarrow x',\quad \mu\leftrightarrow\nu \qquad (8.104)\\ &\quad \text{plus contractions leading to loop diagrams.}\end{aligned}$$

a. Verify that the above leads to the following scattering contribution to (8.102):

$$2\frac{(ieq)^2}{2!}\int d^4x\, d^4x'\, e^{-ip'x}\,\bar{u}(p'\lambda')i\gamma^{\mu}(-i)S(x-x')i\gamma^{\nu}u(p\lambda)e^{ipx'}$$

$$\left[e^{-ik'x}\, e^*_{\mu}(k'\kappa')e_{\nu}(k\kappa)e^{ikx'} + e^{-ik'x'}\, e^*_{\nu}(k'\kappa')e_{\mu}(k\kappa)e^{ikx}\right] \qquad (8.105)$$

$$= -i(2\pi)^4\delta(p'+k'-p-k)T, \qquad (8.106)$$

with

$$-iT = \qquad (8.107)$$

$$e^*_{\mu}(k'\kappa')\bar{u}(p'\lambda')(-eq\gamma^{\mu})(-i)\frac{m-i\gamma(p+k)}{m^2+(p+k)^2}(-eq\gamma^{\nu})u(p\lambda)e_{\nu}(k\kappa)+$$

$$e^*_{\nu}(k'\kappa')\bar{u}(p'\lambda')(-eq\gamma^{\mu})(-i)\frac{m-i\gamma(p-k')}{m^2+(p-k')^2}(-eq\gamma^{\nu})u(p\lambda)e_{\mu}(k\kappa)$$

b. Write down the corresponding diagrams, include all the relevant labels and establish a 1-1 correspondence between the diagrams and the expression for T.

c. Verify that the other contributions lead to loop diagrams

d. The scattering amplitude has the form

$$T = e^*_{\mu}(k'\kappa')T^{\mu\nu}(p',k',p,k)e_{\nu}(k\kappa), \qquad (8.108)$$

which defines $T^{\mu\nu}(p',k',p,k)$ (of course only three four-momenta are independent because $p'+k'=p+k$). Using the Ward-Takahasi identities

$$i\gamma k = S^{-1}(p+k) - S^{-1}(p) = S^{-1}(p') - S^{-1}(p'-k) \qquad (8.109)$$

and the Dirac equation for the spinors $u(p\lambda)$ and $\bar{u}(p'\lambda')$, verify that

$$k_{\nu}T^{\mu\nu}(p',k',p,k) = 0, \qquad (8.110)$$

and similarly $k'_{\mu}T^{\mu\nu}(p',k',p,k) = 0$. Using this result we know that on calculating the unpolarized cross section we may use eq. (7.169) and replace

$$\sum_{\kappa} e_{\mu}(k\kappa)e^*_{\rho}(k\kappa) \to \eta_{\mu\rho}. \qquad (8.111)$$

The property (8.110) can be understood as implementing invariance with respect to gauge tranformations on the photon polarization vectors, $e_\nu(k,\kappa) \to e_\nu(k,\kappa) + \propto k_\mu$, which occur in Lorentz transformations in Coulomb gauge. The replacement (8.111) implements Lorentz invariance in this respect. Note that $T^{\mu\nu}$ has the form $T^{\mu\nu} = \bar{u}'\tilde{T}^{\mu\nu}u$, with $\tilde{T}^{\mu\nu}$ transforming as a tensor-matrix, and that the polarization sum $\sum_{\rm spins} \bar{u}'\tilde{T}_{\mu\nu}u\,(\bar{u}'\tilde{T}_{\rho\sigma}u)^*$ results in a Lorentz tensor.

e. Further evaluation of the differential cross section is described for example in the book by Weinberg (part I sec. 8.7), or by Bjorken & Drell (part I), Sterman, etc. We quote the Klein-Nishina formula which holds in the laboratory frame:

$$\left[\frac{d\sigma}{d\Omega}\right]_{\rm lab} = \left(\frac{e^2}{4\pi}\right)^2 \frac{1}{2m^2}\left(\frac{k'^0}{k^0}\right)^2 \left(\frac{k^0}{k'^0} + \frac{k'^0}{k^0} - \sin^2\theta\right). \tag{8.112}$$

Express k'^0 in terms of m, k^0 and θ.

3. *Electron-electron scattering*

In this problem we shall work out the unpolarized cross section for the process $e^- + e^- \to e^- + e^-$. Consider the amplitude for $e_1^- + e_2^- \to e_3^- + e_4^-$,

$$T(34;12) = -e^2\left[\bar{u}_3\gamma_\mu u_1\,\bar{u}_4\gamma_\nu u_2\,D^{\mu\nu}(k) - \bar{u}_3\gamma_\nu u_2\,\bar{u}_4\gamma_\mu u_1\,D^{\mu\nu}(l)\right], \tag{8.113}$$

where $k = p_3 - p_1$ and $l = p_4 - p_1$.

a. Derive this.

b. Show using the Dirac equation in momentum space that $\bar{u}_3\gamma_\mu u_1\,k^\mu = 0$, $\bar{u}_3\gamma_\nu u_2\,l^\nu = 0$. This corresponds to current conservation. Consequently the amplitude can be simplified to

$$T = -e^2\left[\bar{u}_3\gamma_\mu u_1\,\bar{u}_4\gamma^\mu u_2\,\frac{1}{k^2} - \bar{u}_3\gamma_\nu u_2\,\bar{u}_4\gamma^\nu u_1\,\frac{1}{l^2}\right]. \tag{8.114}$$

c. When the incoming electrons are unpolarized and we do not measure the spins of the outgoing electrons, we have to average $|T|^2$ over initial spin polarizations and sum over final spin polarizations. We use the shorthand notation

$$\overline{|T|^2} = \frac{1}{16}\sum_{\lambda_1\cdots\lambda_4}|T|^2. \tag{8.115}$$

Derive along similar lines as in sect. 8.3, that

$$\overline{|T|^2} = \frac{e^4}{16}\left[\frac{T_1}{((p_1-p_3)^2)^2} - \frac{T_2}{(p_1-p_3)^2(p_1-p_4)^2} + (p_3 \leftrightarrow p_4)\right], \tag{8.116}$$

where, using the convenient 'slash' notation $\not{p} = p_\mu \gamma^\mu$,

$$T_1 = \mathrm{Tr}\,[\gamma_\mu(m - i\not{p}_1)\gamma_\nu(m - i\not{p}_3)]\,\mathrm{Tr}\,[\gamma^\mu(m - i\not{p}_2)\gamma^\nu(m - i\not{p}_4)], \quad (8.117)$$

$$T_2 = \mathrm{Tr}\,[\gamma_\mu(m - i\not{p}_1)\gamma_\nu(m - i\not{p}_4)\gamma^\mu(m - i\not{p}_2)\gamma^\nu(m - i\not{p}_3)]. \quad (8.118)$$

d. Using the 'trace identities' and of course momentum conservation $p_1 + p_2 = p_3 + p_4$ and $p_i^2 = -m^2$, show that

$$T_1 = 32[2m^4 + 2m^2 p_1 p_3 + (p_1 p_2)^2 + (p_1 p_4)^2], \quad (8.119)$$

$$T_2 = -32[2m^2 p_1 p_2 + (p_1 p_2)^2]. \quad (8.120)$$

e. These expressions can be evaluated in the center of mass frame. Let θ be the scattering angle between particles 1 and 3, $p_1 p_3 = -m^2 - |\mathbf{p}|^2(1 - \cos\theta)$. From now on we use the notation $p \equiv |\mathbf{p}|$. Verify that

$$T_1 = 64\left[m^4 + 4m^2 p^2 \cos^2\frac{\theta}{2} + 2p^4(1 + \cos^4\frac{\theta}{2})\right], \quad (8.121)$$

$$T_2 = -32(-m^4 + 4p^4), \quad (8.122)$$

$$\overline{|T|^2} = \frac{e^4}{256p^4}\left[\frac{T_1}{\sin^4\frac{\theta}{2}} - \frac{T_2}{\cos^2\frac{\theta}{2}\sin^2\frac{\theta}{2}} + (\theta \to \pi - \theta)\right]. \quad (8.123)$$

f. Under ultra-relativistic conditions we may neglect the electron mass m. Verify that in the center of mass frame

$$\left[\frac{d\sigma}{d\Omega}\right]_{\text{cm, ur}} = \frac{\alpha^2}{8p^2}\left[\frac{1 + \cos^4\frac{\theta}{2}}{\sin^4\frac{\theta}{2}} + \frac{2}{\cos^2\frac{\theta}{2}\sin^2\frac{\theta}{2}} + \frac{1 + \sin^4\frac{\theta}{2}}{\cos^4\frac{\theta}{2}}\right]. \quad (8.124)$$

g. Under non-relativistic conditions we may neglect p compared to m. Verify that

$$\left[\frac{d\sigma}{d\Omega}\right]_{\text{cm, nr}} = \frac{\alpha^2 m^2}{16p^4}\left[\frac{1}{\sin^4\frac{\theta}{2}} - \frac{1}{\cos^2\frac{\theta}{2}\sin^2\frac{\theta}{2}} + \frac{1}{\cos^4\frac{\theta}{2}}\right]. \quad (8.125)$$

The middle term is due to the interference of the two diagrams contributing to the amplitude. The first term goes over in the Rutherford formula for Coulomb scattering off a heavy target, upon expressing it in terms of the reduced mass $m_{\text{red}} = mm/(m + m) = m/2$.

The total cross section is infinite because the integration over angles diverges at $\theta = 0$. This can be attributed to the infinite range of the Coulomb potential.

Chapter 9

Scattering

Having seen how vacuum expectation values of time-ordered products of fields can be calculated in perturbation theory, we derive here how these can be used for the construction of scattering amplitudes. The discussion will be quite general, without recourse to perturbation theory, such that it also applies to composite fields and bound states as found in QCD. Results of scattering experiments are expressed in terms of cross sections. We start with the relation between cross sections, transition amplitudes $\langle \text{out}|\text{in}\rangle$ and invariant-scattering amplitudes T (often denoted by $\mathcal{M}$). After deriving so-called spectral representations for two-point functions, which exhibit the connection between the occurrence of poles in the complex frequency plane and the existence of particles, we derive a relation between connected n-point functions and scattering amplitudes. This relation is often called 'the LSZ formula', after the work of Lehmann, Symanzik and Zimmermann. We shall not follow their derivation, but use an intuitive reasoning inspired by Schwinger's Source Theory. We end with a heuristic derivation of formulas for unstable particle decay.

9.1 Cross section

In a colliding beam experiment, two bunches of particles approach each other such that they overlap for a short period, during which the particles scatter. The bunches 1,2 may be described by a four-current density $j^{\mu}_{1,2}(x)$. We assume that these currents are slowly varying on the scale of the Compton wavelengths $m^{-1}_{1,2}$, and that the spread in momentum of the particles is small, such that to a good approximation the current is just the average velocity times the density,

$$\mathbf{j}_i(x) = \mathbf{v}_i\, j^0_i(x), \quad \mathbf{v}_i = \mathbf{p}_i/p^0_i, \quad i = 1, 2. \tag{9.1}$$

Here p^{μ}_i is the average four-momentum of the particles in bunch i. The number of scattering events N_{Δ} of a kind specified by Δ, in a frame in which the bunch velocities are aligned, $\mathbf{v}_1 \propto \mathbf{v}_2$, is given by

$$N_\Delta = \sigma_\Delta v_{12} \int d^4x\, j_1^0(x)\, j_2^0(x) \tag{9.2}$$

where σ_Δ is the corresponding cross section and $v_{12} = |\mathbf{v}_1 - \mathbf{v}_2|$ is the relative velocity. Multiple scattering is neglected. The number of events is Lorentz invariant, and the cross section is *defined* to be invariant, by generalizing the above formula to

$$\begin{aligned} N_\Delta &= \int d^4x\, \sigma_\Delta \sqrt{[j_1^\mu(x) j_{\mu\,2}(x)]^2 - j_1^\mu(x) j_{\mu\,1}(x)\, j_2^\nu(x) j_{\nu\,2}(x)} && (9.3)\\ &= \sigma_\Delta \frac{\sqrt{(p_1p_2)^2 - p_1^2 p_2^2}}{p_1^0 p_2^0} \int d^4x\, j_1^0(x)\, j_2^0(x) && (9.4)\\ &= \sigma_\Delta \sqrt{(\mathbf{v}_1 - \mathbf{v}_2)^2 - (\mathbf{v}_1 \times \mathbf{v}_2)^2} \int d^4x\, j_1^0(x)\, j_2^0(x). && (9.5) \end{aligned}$$

Specializing to the case of aligned momenta $\mathbf{p}_1 \propto \mathbf{p}_2$, for which

$$\sqrt{(p_1p_2)^2 - m_1^2 m_2^2} = p_1^0 p_2^0 v_{12}, \tag{9.6}$$

we get back (9.2).

To make contact with the elementary discussion in chapter 2, in which we used the concept of event *rate*, consider very big uniform ‘bunches’. Then the integration over time in (9.5) will be proportional to the total overlap time of the bunches, and we identify the event rate

$$dN_\Delta/dt = \sigma_\Delta\, L(t), \tag{9.7}$$

with

$$L(x^0) = \sqrt{(\mathbf{v}_1 - \mathbf{v}_2)^2 - (\mathbf{v}_1 \times \mathbf{v}_2)^2} \int d^3x\, j_1^0(x)\, j_2^0(x), \tag{9.8}$$

the *luminosity*. However, in the following we shall use (9.5) to identify the scattering amplitude.

If we redefine the currents by dividing by the total number of particles $N_{1,2}$:

$$\frac{1}{N_{1,2}}\, j_{1,2}^\mu(x) \to j_{1,2}^\mu(x), \qquad \int d^3x\, j_{1,2}^0(x) = 1, \tag{9.9}$$

then

$$N_\Delta \to P_\Delta, \tag{9.10}$$

with P_Δ the *probability* for events of type Δ. Possible specifications of Δ have been discussed in Problem 3 in chapter 2. Spin is ignored above; in case the particles have spin, we may assume that it is not specified in the initial state and not observed in the final state.

9.2 Scattering amplitude

We now consider the scattering from the quantum point of view. The particles in the bunches are assumed to be effectively noninteracting, so their specification is like single particles. Let

$$|p\lambda\rangle, \quad \langle p\lambda|p'\lambda'\rangle = \delta_{\lambda\lambda'}\, 2p^0\, (2\pi)^3\, \delta(\mathbf{p}-\mathbf{p}'), \tag{9.11}$$

denote a covariant basis state for a single massive particle with four-momentum p and spin index λ. The index λ may be taken to be the eigenvalue of J_3 in a rest frame of the particle, where J_3 is the third component of the angular momentum. Then $\lambda = -s, -s+1, \cdots, s$, with s the spin of the particle and $s(s+1)$ the eigenvalue of $\mathbf{J}^2$ in the rest frame. Another possibility is to use the helicity, i.e. the eigenvalue of $\mathbf{J}$ in the direction of momentum, which also takes $2s+1$ values. For massless particles there is no rest frame, and we shall assume λ to be the helicity, which takes only two values in case it is non-zero. We have seen this to be true for the photon and the same holds for the graviton; a general discussion is in Weinberg I, and in Ryder. Massless particles require special treatment in scattering theory, but in a first approach it is useful to simply generalize formulas derived for massive particles to the massless case.

For spinless particles we have seen in (2.119) how $|p\rangle$ transforms under Lorentz transformations. In case of spin we have

$$U(\ell)|p,\lambda\rangle = \sum_{\lambda'} C_{\lambda'\lambda}(\ell,p)|\ell p,\lambda'\rangle, \tag{9.12}$$

with $C_{\lambda'\lambda}(\ell,p)$ a unitary matrix depending on $\mathbf{p}$ and ℓ,

$$\sum_{\lambda'} C^*_{\lambda'\lambda}(\ell,p)\, C_{\lambda'\lambda''}(\ell,p) = \delta_{\lambda\lambda''}, \quad C^{-1}(\ell,p) = C^\dagger(\ell,p) = C(\ell^{-1},p). \tag{9.13}$$

This unitarity is a consequence of the unitarity of $U(\ell)$. The explicit form of C is complicated, see e.g. Weinberg I or Ryder, but we do not need this.

In terms of these basis states we define wave-packet states

$$|f\rangle = \sum_\lambda \int d\omega_q\, f(q,\lambda)\, |q\lambda\rangle, \quad \langle f|f\rangle = \sum_\lambda \int d\omega_q\, |f(q,\lambda)|^2 = 1, \tag{9.14}$$

and wave functions

$$f(x,\lambda) = \int d\omega_q\, f(q,\lambda)\, e^{iqx}, \tag{9.15}$$

similar to the spinless case discussed in Sect. 2.8. Under Lorentz transformations $f(q,\lambda)$ transforms as

$$f'(q,\lambda) = \sum_{\lambda'} C_{\lambda\lambda'}(\ell,q)\, f(\ell^{-1}q,\lambda'). \tag{9.16}$$

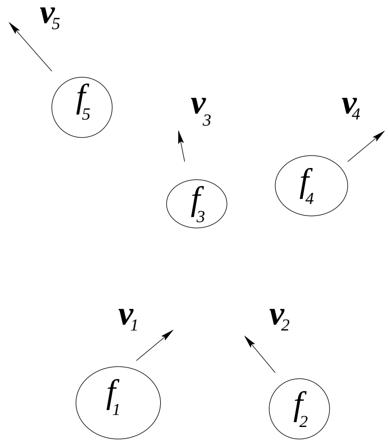

Figure 9.1: In and outgoing wave packet states. The ingoing states are specified in the far past, the outgoing in the far future

Generalizing the discussion in section 2.8 to particles with spin, let $f(q, \lambda)$ be suitably sharply peaked about a mean momentum $\mathbf{p}$. Then

$$\begin{aligned} j^\mu(x) &\equiv \sum_\lambda \left[-f^*(x,\lambda)\, i\partial^\mu f(x,\lambda) + i\partial^\mu f^*(x,\lambda)\, f(x,\lambda)\right] \\ &\simeq 2p^\mu \sum_\lambda |f(x,\lambda)|^2 \end{aligned} \tag{9.17}$$

is a good probability current for finding the particle at spacetime point x. Note that $f'(x,\lambda) \simeq \sum_{\lambda'} C_{\lambda\lambda'}(\ell, \ell p) f(\ell^{-1}x, \lambda')$ and so $\sum_\lambda |f(x,\lambda)|^2$ is a scalar field under Lorentz transformations, because of the unitarity of the matrix C. We identify the current (9.17) with the normalized classical current (9.9). More generally, one can consider a spin density matrix $j^0(x; \lambda, \lambda')$,

$$j^0(x; \lambda, \lambda') = 2p^0 f(x,\lambda) f^*(x,\lambda'), \tag{9.18}$$

such that, for instance, the expectation value of the spin in the particle rest frame is given by $\mathrm{Tr}\, j^0(x)\boldsymbol{S}$, where $\boldsymbol{S}_{\lambda\lambda'}$ is the spin matrix (e.g. for spin 1/2, $\boldsymbol{S} = \boldsymbol{\sigma}/2$, with $\boldsymbol{\sigma}$ the Pauli matrices).

Consider now an incoming state of two widely separated particles converging to a scattering region as in Fig. 9.1:

$$|f_1, f_2\, \mathrm{in}\rangle. \tag{9.19}$$

We are using the Heisenberg picture in which the states in Hilbert space are time-independent. Quantum numbers other than momentum and spin (e.g. distinguishing an electron from a proton) are suppressed. How to make such a state in a quantum field theory with interactions will be discussed later in this chapter. Similarly, consider states representing widely separated particles at late times, which may or may not have emerged from the scattering region, as in Fig. 9.1:

$$|f_3, \cdots, f_n \,\text{out}\rangle. \tag{9.20}$$

The probability amplitude for the process $f_1 + f_2 \to f_3 + \cdots + f_n$ is

$$\langle \text{out}\, f_3, \cdots, f_n | f_1, f_2 \,\text{in}\rangle. \tag{9.21}$$

It is assumed that there are momentum-basis states $|q_1\lambda_1, q_2\lambda_2 \,\text{in}\rangle$, in terms of which we can write

$$|f_1, f_2 \,\text{in}\rangle = \sum_{\lambda_1\lambda_2} \int d\omega_{q_1}\, d\omega_{q_2}\, f_1(q_1, \lambda_1)\, f_2(q_2, \lambda_2)\, |q_1\lambda_1, q_2\lambda_2 \,\text{in}\rangle, \tag{9.22}$$

and similar for the out states. Note that these sharp momentum-states are certainly overlapping in position space (being plane-wave states). They are therefore not merely tensor products of free-particle states, and the labels 'in' and 'out' are important. Only single-particle states are free states[1] in the sense that the labels 'in' and 'out' may be omitted: $|p, \lambda \text{ in}\rangle = |p, \lambda \text{ out}\rangle = |p, \lambda\rangle$. The probability for ending up in the final state specified by a momentum region Δ is now

$$P_\Delta = \sum_{\lambda_3\cdots\lambda_n} \int_\Delta d\omega_{p_3} \cdots d\omega_{p_n}\, |\langle \text{out}\, p_3\lambda_3 \cdots p_n\lambda_n | f_1, f_2 \,\text{in}\rangle|^2. \tag{9.23}$$

We are going to rewrite this expression into the form (9.2).
First, making the wave functions explicit we have

$$\begin{aligned} |\langle \text{out}\, p_3\lambda_3 \cdots p_n\lambda_n | f_1, f_2 \,\text{in}\rangle|^2 \;=\; & \sum_{\lambda_1\lambda_2\lambda_1'\lambda_2'} \int d\omega_{q_1}\, d\omega_{q_2}\, d\omega_{q_1'}\, d\omega_{q_2'} \\ & f_1(q_1, \lambda_1)\, f_2(q_2, \lambda_2)\, f_1^*(q_1', \lambda_1')\, f_2^*(q_2', \lambda_2') \\ & \langle \text{out}\, p_3\lambda_3 \cdots p_n\lambda_n | q_1\lambda_1, q_2\lambda_2 \,\text{in}\rangle \\ & \langle \text{out}\, p_3\lambda_3 \cdots p_n\lambda_n | q_1'\lambda_1', q_2'\lambda_2' \,\text{in}\rangle^*. \end{aligned} \tag{9.24}$$

Second, the in and out states are specified in the far past and far future. Therefore we expect energy-momentum conservation to hold, such that we may write schematically,[2]

$$\langle \text{out}\, \alpha | \beta \,\text{in}\rangle = \delta_{\alpha,\beta} - i(2\pi)^4 \delta^4(p_\alpha - p_\beta)\, T(\alpha, \beta). \tag{9.25}$$

Here the first term on the right hand side represents the possibility of no scattering, while the second term represents scattering, with energy-momentum conservation made explicit. The occurrence of the Dirac delta function $\delta^4(p_\alpha - p_\beta)$

[1] Here these states are assumed to be eigenstates of the *complete* hamiltonian, including interaction terms.

[2] The matrix $S_{\alpha,\beta} \equiv \langle \text{out}\, \alpha | \beta \,\text{in}\rangle$ is called the scattering matrix, or S-matrix for short.

will be evident later on (we have seen it emerge in lowest order calculations); it is also implicitly present in the $\delta_{\alpha\beta}$-term. More explicitly, (9.25) reads, in case the final state is different from the initial,

$$\langle \text{out}\, p_3\lambda_3\cdots p_n\lambda_n|q_1\lambda_1, q_2\lambda_2\,\text{in}\rangle \tag{9.26}$$
$$= -i(2\pi)^4\delta^4(p_3+\cdots+p_n-q_1-q_2)\,T(p_3\lambda_3,\cdots,p_n\lambda_n;q_1\lambda_1,q_2\lambda_2). \tag{9.27}$$

The function $T(p_3\lambda_3,\cdots)$ is the called the invariant amplitude, which is often denoted by $\mathcal{M}$. Assuming the final states contributing in Δ to be different from the initial state, only the T-terms contribute.
Third, we have the identity

$$\delta^4(P-q_1-q_2)\,\delta^4(P-q_1'-q_2') = \delta^4(P-q_1-q_2)\,\delta^4(q_1+q_2-q_1'-q_2'). \tag{9.28}$$

Fourth, the delta functions occur in integrals with functions $f_{1,2}(q,\lambda)$ which are sharply peaked about $p_{1,2}$, so we may make the replacement

$$\delta^4(P-q_1-q_2)\to\delta^4(P-p_1-p_2). \tag{9.29}$$

Fifth, using

$$(2\pi)^4\delta^4(q_1+q_2-q_1'-q_2') = \int d^4x\, e^{i(q_1+q_2-q_1'-q_2')x}, \tag{9.30}$$

the factor in (9.24)

$$\int d\omega_{q_1}\,d\omega_{q_2}\,d\omega_{q_1'}\,d\omega_{q_2'}\,(2\pi)^4\delta^4(q_1+q_2-q_1'-q_2')$$
$$f_1(q_1,\lambda_1)\,f_2(q_2,\lambda_2)\,f_1^*(q_1',\lambda_1')\,f_2^*(q_2',\lambda_2') \tag{9.31}$$

equals approximately

$$\frac{1}{4p_1^0p_2^0}\int d^4x\, j_1^0(x;\lambda_1,\lambda_1')\,j_2^0(x;\lambda_2,\lambda_2'), \tag{9.32}$$

with

$$j_i^0(x;\lambda,\lambda') = 2p_i^0 f_i(x,\lambda)f_i^*(x,\lambda') \tag{9.33}$$

the density matrix for particle $i=1,2$, as in (9.18).

Putting things together we can rewrite (9.23) in the form

$$P_\Delta = \sum_{\lambda_1\cdots\lambda_2'}\sigma_\Delta(\lambda_1'\lambda_2',\lambda_1\lambda_2)v_{12}\int d^4x\, j_1^0(x;\lambda_1,\lambda_1')\,j_2^0(x;\lambda_2,\lambda_2') \tag{9.34}$$

with

$$\sigma_\Delta(\lambda_1'\lambda_2',\lambda_1\lambda_2) = \frac{1}{4p_1^0p_2^0v_{12}}\sum_{\lambda_3\cdots\lambda_n}\int_\Delta d\omega_{p_3}\cdots d\omega_{p_n}\,(2\pi)^4\delta^4(p_1+p_2-p_3\cdots-p_n)$$
$$T(p_3\lambda_3,\cdots;p_1\lambda_1,p_2\lambda_2)\,T(p_3\lambda_3,\cdots;p_1\lambda_1',p_2\lambda_2')^*. \tag{9.35}$$

The initial density matrices have here a pure-state form. In an experimental setup they have to be replaced by the appropriate density matrices. In case one is interested in measuring spin dependence these would correspond to a spin-polarized initial state and the spin dependence of the final state would also be analyzed in more detail. Here we shall assume that spin is not monitored, unless mentioned otherwise. This means that we should replace the pure-state forms by

$$j_i^0(x;\lambda,\lambda') \to j_i^0(x)\,\frac{\delta_{\lambda\lambda'}}{2s_i+1}, \quad i=1,2, \tag{9.36}$$

with $s_{1,2}$ the spin of the particle and $j^0(x)$ defined in (9.17). Correspondingly we introduce the short-hand notation

$$\overline{|T|^2} = \frac{1}{\prod_{i=1}^n (2s_i+1)} \sum_{\lambda_1\cdots\lambda_n} |T(p_3\lambda_3,\cdots;p_1\lambda_1,p_2\lambda_2)|^2, \tag{9.37}$$

which is a Lorentz-invariant function of the momenta (as follows from (9.12) and (9.26)). Recalling (9.6) gives finally the formula for the unpolarized cross section

$$\sigma_\Delta = \frac{(2s_3+1)\cdots(2s_n+1)}{4\sqrt{(p_1p_2)^2-m_1^2m_2^2}} \int_\Delta d\omega_{p_3}\cdots d\omega_{p_n}\,(2\pi)^4\delta^4(p_1+p_2-p_3\cdots-p_n)\,\overline{|T|^2}. \tag{9.38}$$

Problem 3 in chapter 2 elaborates on how to use this formula for the case $1+2\to 3+4$.

9.3 Fields, particles and poles

In free field theory the eigenstates of the energy-momentum operator P^μ are found by application of the creation operators $a^\dagger(p)$ on the vacuum state $|0\rangle$. In an interacting field theory the creation and annihilation operators can be defined similarly at some time t, with a corresponding empty state $|\emptyset,t\rangle$, $a(p,t)|\emptyset,t\rangle=0$, as we did e.g. in sect. 2.5. However the states $a^\dagger(p_1,t)\cdots a^\dagger(p_n,t)|\emptyset,t\rangle$ are not eigenstates of P^μ anymore; in perturbation theory they are only eigenstates at lowest order. They are called bare particle states. The true, 'dressed', eigenstates are complicated linear superposition of the bare states. We can try to calculate these in perturbation theory, but this is cumbersome, and it would not work in case of bound states, which are non-perturbative.[3] We have learned to focus on the probability *amplitudes* $\langle\text{out}|\text{in}\rangle$, rather than on states themselves. These amplitudes can be expressed in terms of vacuum expectation values of field operators. We shall keep the discussion general, such that it applies also to the case

[3]For example, the ground state of hydrogen cannot be found by perturbing a state with a free electron and a free proton. In QCD all eigenstates of the hamiltonian are bound states of quarks and gluons.

of bound states, proceed heuristically and appeal to intuition in avoiding difficult proofs.

Suppose $\phi^a(x)$ is a hermitian operator field transforming in a representation L of the Lorentz group,

$$U(\ell)^\dagger \phi^a(x) U(\ell) = D_{ab}(\ell)\phi^b(\ell^{-1}x). \tag{9.39}$$

It may be an elementary field, i.e. a basic variable appearing in the action, or a composite field contructed out of products of elementary fields, e.g. a current $j^\mu(x)$. Suppose furthermore that $J_a(x)$ is a function with support in a compact spacetime region, with the help of which we 'smear out' the field ϕ^a,

$$\phi(J) \equiv \int d^4x\, J_a(x)\phi^a(x). \tag{9.40}$$

Intuitively one expects $\phi(J)$ to create a particle out of the vacuum $|0\rangle$ in the spacetime region J, i.e. the amplitude

$$\langle p\lambda|\phi(J)|0\rangle \tag{9.41}$$

is nonzero.[4]

Quantum numbers other than p, λ (such as electric charge, lepton number, ...) have to correspond to those of $\phi(J)$, otherwise the above amplitude would be zero. However, because of interactions the state $\phi(J)|0\rangle$ will have many more components in Hilbert space, multiparticle states and possibly even the vacuum itself: for a scalar field $\langle 0|\phi(x)|0\rangle$ may be nonzero. In the following we shall assume that the vacuum expectation value of ϕ^a has been subtracted from the field:

$$\langle 0|\phi^a(x)|0\rangle = 0. \tag{9.42}$$

To see how to construct transition amplitudes with a given number of particles we start with the simplest of all: $\langle p\lambda|p'\lambda'\rangle$, which can be made out of the two point function

$$\langle 0|\phi^a(x)\phi^b(y)|0\rangle. \tag{9.43}$$

First we derive a *spectral representation.*[5] We are going to insert a complete set of intermediate states between the two fields above. To contribute, the intermediate states must have the same quantum numbers as the field, because of Schur's lemma (we suppress these quantum numbers in our notation). For the eigenvalues of the energy-momentum operator we make some reasonable assumptions:

[4] The annihilation and creation operators $a(p,t)$ etc., defined in sect. 6.6 etc., can also be written in the general form $\phi(J)$, with suitable $J(x)$ containing $\delta(x^0 - t)$ and its derivative. The notation J suggests it to be a source for ϕ and this aspect is indeed and essential element in the methods employed by Schwinger in his 'Source Theory' .

[5] In this context also called a Källén-Lehmann representation.

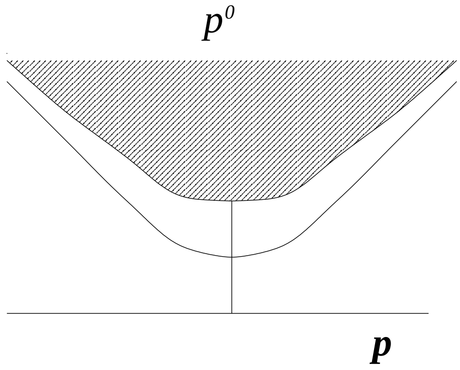

Figure 9.2: Energy-momentum spectrum of contributing states. The single curve represents the mass shell $p^2 = -m^2$, $p^0 > 0$ of a single particle state. The hatched area represents $-p^2 \geq (m + m')^2$ of the multiparticle states.

- The groundstate (vacuum state) $|0\rangle$ is a non-degenerate Lorentz-invariant eigenvector of P^μ and $J^{\mu\nu}$ with eigenvalue zero,

$$P^\mu|0\rangle = 0, \quad J^{\mu\nu}|0\rangle = 0. \tag{9.44}$$

- The contributing single-particle states are stable. The single-particle nature implies that p is on a mass shell as illustrated in Fig. 9.2,

$$P^\mu|p\lambda\rangle = p^\mu|p\lambda\rangle, \quad p^0 = \sqrt{m^2 + \mathbf{p}^2}. \tag{9.45}$$

- The infinite remainder of the basis states are multi-particle states (e.g. in-going *or* out-going),

$$P^\mu|p,\Lambda\rangle = p^\mu|p,\Lambda\rangle, \quad -p^2 > m^2. \tag{9.46}$$

  Here p^μ is the total four-momentum of the states. Other quantum numbers (such as relative momentum, spins) are lumped into the symbol Λ. Unstable particles decay eventually in terms of stable particles, so unstable particles are to be somehow represented by the states $|p,\Lambda\rangle$.

In writing $-p^2 > m^2$ we have assumed that there are no massless particles. For example, suppose the smallest mass other than m is m'. Then, for a two-particle state with four-momenta p and p', $p^2 = -m^2$, $p'^2 = -m'^2$, the total invariant mass is given in the center of mass frame by $-(p+p')^2 = m^2 + 2\sqrt{m^2 + \mathbf{p}^2}\sqrt{m'^2 + \mathbf{p}^2} + m'^2 \geq (m+m')^2$, which is larger than m^2 if $m' \neq 0$. The spectrum of contributing states is illustrated in Fig. 9.2.

Inserting the complete set of contributing states we write[6]

$$\langle 0|\phi^a(x)\phi^b(y)|0\rangle = \sum_\lambda \int d\omega_p \, \langle 0|\phi^a(x)|p\lambda\rangle\langle p\lambda|\phi^b(y)|0\rangle \tag{9.47}$$
$$+ \sum_\Lambda \int_{m+m'}^{\infty} dM \int d\omega_p(M)\, \langle 0|\phi^a(x)|p\Lambda\rangle\langle p\Lambda|\phi^b(y)|0\rangle,$$
$$d\omega_p = \frac{d^3p}{(2\pi)^3 2\sqrt{m^2+\mathbf{p}^2}}, \quad d\omega_p(M) \equiv \frac{d^3p}{(2\pi)^3 2\sqrt{M^2+\mathbf{p}^2}}, \tag{9.48}$$

where we have explicitly indicated the invariant mass $-p^2 \equiv M^2$ of the multi-particle states. Note that the shorthand $\sum_\Lambda$ includes also integrations over the continuous variables buried in Λ. The above expression can be simplified further using Poincaré symmetry. Translation invariance gives

$$\langle 0|\phi^a(x)|p,\lambda\rangle = \langle 0|e^{-iPx}\,\phi^a(0)e^{iPx}\,|p,\lambda\rangle = \langle 0|\phi^a(0)|p,\lambda\rangle\, e^{ipx}. \tag{9.49}$$

Lorentz invariance gives

$$\langle 0|U(\ell)^\dagger\phi^a(0)U(\ell)|p,\lambda\rangle = D_{ab}(\ell)\langle 0|\phi^b(0)|p,\lambda\rangle \tag{9.50}$$
$$= C_{\lambda'\lambda}(\ell,p)\langle 0|\phi^a(0)|\ell p,\lambda'\rangle, \tag{9.51}$$

where we used (9.12) and (9.39) (and used the summation convention for the repeated λ'). This shows that $U_a(p,\lambda)$, defined up to a constant $\sqrt{Z}$ by[7]

$$\sqrt{Z}\, U_a(p,\lambda) = \langle 0|\phi^a(0)|p,\lambda\rangle, \tag{9.52}$$

transforms just like the basis vector $|p,\lambda\rangle$:

$$D_{ab}(\ell)U_b(p,\lambda) = C_{\lambda'\lambda}(\ell,p)\, U_a(\ell p,\lambda') \tag{9.53}$$

The $U_a(p,\lambda)$ are polarization tensors or spinors, depending on the representation L. They can be normalized in some convenient covariant way. The constant $\sqrt{Z}$ depends on ϕ^a. It can we any complex number, including zero, but we assume it here to be nonzero. In case ϕ^a is an elementary field, the phases of the basis states $|p,\lambda\rangle$ are usually chosen such that $\sqrt{Z}$ is positive, and Z is the wave function renormalization constant first introduced in chapter 4. For a scalar field and a spin zero state, eq. (9.53) reads $U(p) = U(\ell p)$, so $U(p)$ is a Lorentz invariant function of p. For the on-shell case here (i.e. $p^0 = \sqrt{m^2+\mathbf{p}^2}$) this means it is just a constant, which we choose to be unity, i.e. $U(p) = 1$ for a scalar field.

[6]We have written this as if there is only one type of particle contributing to the one-particle intermediate states. If there are more, all should be included. For instance, both particles and antiparticles may contribute if $\phi^a \to$ a hermitian combination of charged currents. In such a case one often uses non-hermitian ϕ^a for which only one type of particle contribute.

[7]This $U_a(p,\lambda)$ should of course not be confused with the representation $U(a,\ell)$ of Poincaré transformations in Hilbert space.

For a spinor field $U_a(p,\lambda) = u_a(p,\lambda)$, the spinor in the Majorana representation (recall ϕ^a is hermitian). Problem 1 deals the case of the vector representation and massive spin-one particles. The photon case is special and will be discussed later.

In (9.47) we need the summation over spins

$$P_{ab}(p) = \sum_{\lambda} U_a(p,\lambda)\, U_b^*(p,\lambda). \tag{9.54}$$

Because of the unitarity of $C_{\lambda\lambda'}(\ell, p)$, this is an invariant tensor or spinor matrix,

$$D_{aa'}(\ell) D_{bb'}(\ell) P_{a'b'}(p) = P_{ab}(\ell p). \tag{9.55}$$

For example, in the spinor case, using a Majorana representation for the γ-matrices, $P_{ab}(p) = [(m - i\gamma p)\beta]_{ab}$. For the vector case $P_{ab}(p) \to P_{\mu\nu}(p) = \eta_{\mu\nu} + p_\mu p_\nu/m^2$, see Problem 1.

We conclude that translation and Lorentz invariance restrict the form of the single particle contribution to

$$\sum_{\lambda} \langle 0|\phi^a(x)|p,\lambda\rangle\langle p,\lambda|\phi^b(y)|0\rangle = Z\left[P_{ab}(p)\, e^{ip(x-y)}\right]_{p^0=\sqrt{\mathbf{p}^2+m^2}}, \quad Z \equiv \sqrt{Z}\left(\sqrt{Z}\right)^*. \tag{9.56}$$

The constant Z above is evidently positive. This is a direct consequence of our implicit assumption that 'ket space' is a Hilbert space with positive metric. Within the same assumption $P_{ab}(p)$ is a positive matrix (i.e. for any c^a, $c^{a*} P_{ab}(p) c^b \geq 0$). In a perturbative treatment of gauge theories in covariant gauges, the metric is not positive, and the summation over intermediate states has to be modified accordingly, as in (7.156,7.157). Then we cannot conclude anymore that Z is positive. For simplicity we shall continue under the assumption of a positive metric Hilbert space. (This is actually the situation in nonperturbative treatments of QCD with the lattice regularization.)

Similarly, we can write for the combination appearing in the multiple particle contribution,

$$\sum_{\Lambda} \langle 0|\phi^a(x)|p,\Lambda\rangle\langle p,\Lambda|\phi^b(y)|0\rangle = \rho(M)\,\left[P_{ab}(p)\, e^{ip(x-y)}\right]_{p^0=\sqrt{\mathbf{p}^2+M^2}}. \tag{9.57}$$

We have separated a kinematical part in the form of $P_{ab}(p)\exp[ip(x-y)]$. The dynamical (interaction dependent) part is in the positive function $\rho(M)$, which is the analogue of Z for the multiple particle contribution.

Summarizing, we have derived the *spectral representation*

$$\begin{aligned} \langle 0|\phi^a(x)\phi^b(y)|0\rangle &= Z\int d\omega_p\,\left[P_{ab}(p)\, e^{ip(x-y)}\right]_{p^0=\sqrt{\mathbf{p}^2+m^2}} \\ &\quad + \int_{m+m'}^{\infty} dM\,\int d\omega_p(M)\,\rho(M)\,\left[P_{ab}(p)\, e^{ip(x-y)}\right]_{p^0=\sqrt{\mathbf{p}^2+M^2}}. \end{aligned} \tag{9.58}$$

Calculations are usually carried out in terms of time-ordered products of fields. The ordinary T product, when defined in terms of Heaviside step functions $\theta(x^0 - y^0)$, is not necessarily covariant, except for scalar and spinor fields. Manifestly covariant path-integral techniques lead to expressions $\langle\phi^a(x)\phi^b(y)\cdots\rangle$ which represent *covariant* time ordered products. These are denoted by T^*, and still satisfy the usual relations

$$\begin{aligned}\langle\phi^a(x)\phi^b(y)\rangle = \langle 0|T^*\phi^a(x)\phi^b(y)|0\rangle &= \langle 0|\phi^a(x)\phi^b(y)|0\rangle, \quad x^0 > y^0 \\ &= \pm\langle 0|\phi^b(y)\phi^a(x)|0\rangle, \quad x^0 < y^0,\end{aligned} \tag{9.59}$$

with the upper/lower sign for Bose/Fermi fields. The difference between the T product and the T^* product is concentrated at $x^0 = y^0$, i.e. it contains the Dirac delta function $\delta(x^0 - y^0)$ and possibly also its derivatives. The spectral representation for the covariant time ordered product can be found as follows.

First one observes that the invariant tensor/spinor $P_{ab}(p)$ has an extension off the mass shell. This means that there is a quantity $N_{ab}(p,m)$, a polynomial in p (hence defined for all p, not only those with $p^0 = \sqrt{\mathbf{p}^2 + m^2}$ or $\sqrt{\mathbf{p}^2 + M^2}$), such that on-shell

$$N_{ab}(p,m) = P_{ab}(p), \quad p^0 = \sqrt{\mathbf{p}^2 + m^2}, \tag{9.60}$$

$$N_{ab}(-p,m) = \pm P_{ba}(p), \quad p^0 = \sqrt{\mathbf{p}^2 + m^2}, \tag{9.61}$$

where we may also read M instead of m. Examples are the Majorana spinor case

$$N_{ab}(p,m) = [(m - i\gamma^\mu p_\mu)\beta]_{ab}, \tag{9.62}$$

and the vector case

$$N_{\mu\nu}(p,m) = \eta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2} \tag{9.63}$$

(cf. in Prob. 1).

In terms of N_{ab} the spectral representation for the covariant time ordered product of the vacuum expectation value of ϕ^a is given by[8]

$$\begin{aligned}\langle 0|T^*\phi^a(x)\phi^b(y)|0\rangle &= -i\int \frac{d^4p}{(2\pi)^4}\, e^{ip(x-y)}\, G^{ab}(p), \\ G^{ab}(p) &= Z\,\frac{N_{ab}(p,m)}{m^2 + p^2 - i\epsilon} + \int_{m+m'}^{\infty} dM\,\rho(M)\,\frac{N_{ab}(p,M)}{M^2 + p^2 - i\epsilon}.\end{aligned} \tag{9.64}$$

It is straightforward to check that (9.64) leads to (9.58), by performing the integral over p^0 for $x^0 > y^0$ or $x^0 < y^0$ by contour integration in the usual way, provided that one ignores possible nonconvergence of the integral at $p^0 = \pm\infty$. This happens with massive particles of spin $> 1/2$. The distribution (generalised

[8]In this chapter we do not distinguish dressed propators (G') from bare ones (G).

function) aspects come more to the forefront in the higher spin case. Such matters are to be resolved in actual calculations using the path integral.

The analytic structure of the full ϕ^a propagator $G^{ab}(p)$ is transparant: it has poles at $p^0 = \pm\sqrt{\mathbf{p}^2 + m^2}$, the positions of which are determined by the mass of the particle created out of the vacuum, and branch points[9] at $p^0 = \pm\sqrt{(m+m')^2 + \mathbf{p}^2}$. Accordingly, the function $G(p)$ is rendered unique by cutting the complex p^0 plane along the intervals $p^0 = (m+m', \infty)$ and $(-\infty, -m-m')$. If there is no particle with the quantum numbers of ϕ^a, then $Z = 0$ and there is no pole but only a branch cut.

This is the *particle-pole connection*. It gives us a very important method for calculating particle masses: construct a field ϕ^a with well defined quantum numbers and look for the position of a pole in its full propagator. If there is a pole, the particle exists; if there is no pole, it does not.

9.4 Scattering amplitudes from correlation functions

We now continue our derivation of scattering amplitudes from correlation functions (also called connected Green functions) $\langle \phi^a(x)\phi^b(y)\cdots\rangle^{\rm conn}$. First we want to recover the free particle amplitude $\langle p\lambda|p'\lambda'\rangle$ from the fully interacting two point function.

Consider

$$\begin{aligned} \langle 0|\phi(J_1)\phi(J_2)|0\rangle &= \int d^4x\, d^4y\, J_{1a}(x) J_{2b}(y) \langle 0|\phi^a(x)\phi^b(y)|0\rangle & (9.65)\\ &= Z\sum_\lambda \int d\omega_p\, J_1^*(p,\lambda) J_2(p,\lambda) & (9.66)\\ &\quad + \int_{m+m'}^\infty dM\,\rho(M) \int d\omega_p(M) J_{1a}^*(p) P_{ab}(p) J_{2b}(p). \end{aligned}$$

where we used the spectral representation (9.58), with (9.54) and the definition

$$J(p,\lambda) = \tilde{J}_a(p)\, U_a^*(p,\lambda), \quad p^0 = \sqrt{m^2 + \mathbf{p}^2}, \tag{9.67}$$

in which $\tilde{J}_a(p)$ is the Fourier transform

$$\tilde{J}_a(p) = \int d^4x\, e^{-ipx}\, J_a(x). \tag{9.68}$$

Note that this Fourier transform is defined for arbitrary p whereas in (9.67) only on-shell values of p are used. The first term in (9.66) is the single particle

[9] A continous superposition of poles leads to a discontinuity along a branch cut, ending at branch points, as in the example $\int_0^1 dx\, 1/(x + y - i\epsilon) = \ln(1 + y - i\epsilon/y - i\epsilon)$, where the cut runs from 0 to 1.

contribution. It has the desired form of an inner product of wave packet states,

$$\sum_\lambda \int d\omega_p \, J_1^*(p,\lambda) J_2(p,\lambda) = \langle J_1 | J_2 \rangle. \tag{9.69}$$

The second term is the multi-particle contribution which we do not want. To get rid of it we take J_1 to the far future and J_2 to the far past, i.e. the support (the region where it is non-zero) of $J_1(x)$ moves to the far future of that of $J_2(x)$. Intuitively, the multi-particles then disperse and the overlap $\langle 0|\phi(J_1)|\text{multi}\rangle\langle\text{multi}|\phi(J_2)|0\rangle$ goes to zero. Mathematically, the product of Fourier transforms $\tilde{J}_1^*(p)\tilde{J}_2(p)$ picks up a phase factor $\exp(ip^0 t)$, where t is the time difference between the supports of J_1 and J_2, causing the integral in (9.66) to die out as $t \to \infty$ because of the Riemann-Lebesque lemma, In doing so we do not want to change the on-shell ($p^2 = -m^2$) values of $\tilde{J}_a(p)$, since these determine the wave functions $J(p,\lambda)$ of our particles.[10]

Of course, in a practical experimental setup we do not actually take limits. At finite but sufficiently large separation the multiparticle contribution is negligible. Sufficiently large means much larger than the typical relevant time scale, which is here set by m^{-1}. This is generally a very short time compared to the experimental situation, e.g. of order of 10^{-10} cm for $m = 1$ MeV. For an estimate of how fast the multiparticle contribution vanishes, see e.g. Brown sect. 6.2.

So we know how to make $\langle J_1|J_2\rangle$ from the two point function:

$$\begin{aligned} \langle J_1|J_2\rangle &= \frac{1}{Z} \lim \langle 0|\phi(J_1)\phi(J_2)|0\rangle &(9.70)\\ &= \frac{1}{Z} \lim \int d^4x\, d^4y\, J_{1a}(x)\,(-i)G^{ab}(x,y)\, J_{2b}(y), &(9.71)\end{aligned}$$

where we have replaced $\langle 0|\phi^a(x)\phi^b(y)|0\rangle$ by the time ordered product $-iG^{ab}(x,y)$, as the time ordering is just right for the limit we are taking. As a corollary we note the following useful formulas which follow along the same lines,

$$\begin{aligned} \frac{1}{\sqrt{Z}} \lim_{J\to\text{future}} \int d^4x\, J_a(x)(-i)G^{ab}(x,y) &= \sqrt{Z}^* \sum_\lambda \int d\omega_p\, J^*(p,\lambda)\, U_b^*(p,\lambda)\, e^{-ipy} \\ &= \langle J|\phi^b(y)|0\rangle, &(9.72)\\ \frac{1}{\sqrt{Z}^*} \lim_{J\to\text{past}} \int d^4y\, (-i)G^{ab}(x,y) J_b(y) &= \sqrt{Z} \sum_\lambda \int d\omega_p\, J(p,\lambda)\, U_a(p,\lambda)\, e^{ipx} \\ &= \langle 0|\phi^a(x)|J\rangle. &(9.73)\end{aligned}$$

[10]To see this more explicitly, consider a translation in a time-like direction v^μ, $v^0 > 0$, $v^2 = -1$ (v can be taken along the average four-velocity of the wave packet): $J(x) \to J(x,t) = J(x - vt)$. In Fourier space this means $\tilde{J}(p,t) = \tilde{J}(p)\exp(-it\,pv)$. We can keep the on-shell values of $\tilde{J}$ fixed by multiplication with $\exp[it\,\text{sign}(pv)\sqrt{m^2+p_\perp^2}]$, $p_\perp^\mu = p^\mu + v^\mu\, pv$, because $pv - \text{sign}(pv)\sqrt{m^2+p_\perp^2} = 0$ for $p^2 = -m^2$. (This factor causes $J(x,t)$ to spread as t gets large.) Using now the form $\tilde{J}(p,t) = \tilde{J}(p)\, \exp\left[-it\left\{pv - \text{sign}(pv)\sqrt{m^2+p_\perp^2}\right\}\right]$, for J_1 and J_2, the multiparticle contribution $\int_{m+m'}^\infty dM\,\rho(M)\int d\omega_p(M)\tilde{J}_{1a}^*(p,t_1)P_{ab}(p)\tilde{J}_{2b}(p,t_2)$ goes to zero for $t_1 \to \infty$, $t_2 \to -\infty$ by the Riemann-Lebesque lemma.

Consider now two sources (smearing functions) J_1, J_2 in the past and an arbitrary number J_3, ..., J_n in the future as in Fig. 9.1. If the initial sources do not overlap and are outside each other's light cone it is plausible that the particles are created independently. Similarly when the final sources are space-like with respect to each other will particles be absorbed independently . Thus, for large separations of the sources,[11]

$$\begin{aligned}\langle J_3 \cdots J_n \,\text{out}|J_1 J_2 \,\text{in}\rangle^{\text{conn}} &= \frac{1}{\sqrt{Z_1^* \cdots Z_n}} \lim \langle 0|T^* \phi(J_n) \cdots \phi(J_3)\phi(J_1)\phi(J_2)|0\rangle^{\text{conn}} \\ &= \frac{1}{\sqrt{Z_1^* \cdots Z_n}} \lim \int d^4x_1 \cdots d^4x_n \, J_{1a_1}(x_1) \cdots J_{na_n}(x_n) \\ &\quad \langle 0|T^* \phi^{a_1}(x_1) \cdots \phi^{a_n}(x_n)|0\rangle^{\text{conn}}. \qquad (9.74)\end{aligned}$$

The limit is $J_{1,2} \to$ far past, $J_{3\cdots n} \to$ far future, e.g. along the average four-velocities of the wave packets, again without changing the on-shell values of the $\tilde{J}$. This gives us the desired probability amplitude for scattering. Taking the connected part eliminates the no scattering contribution.

We now continue with case that the fields ϕ^a are the dynamical variables that appear in the lagrangian model. Then we can proceed using the fact that in a diagrammatical analysis, the connected n-point functions (also called connected Green functions) can decomposed in two-point functions and 1PI vertex functions, and there is always a two-point function G' at each external leg. This implies that they can be written in the form

$$\begin{aligned}\langle 0|T^* \phi^{a_1}(x_1) \cdots \phi^{a_n}(x_n)|0\rangle^{\text{conn}} &\equiv (-i)^{n-1} G^{a_1 \cdots a_n}(x_1, \cdots, x_n) \qquad (9.75) \\ &= \int d^4y_1 \cdots d^4y_n \, (-i) G^{a_1 b_1}(x_1, y_1) \cdots \\ &\quad (-i) G^{a_n b_n}(x_n . y_n) \, i H_{b_1 \cdots b_n}(y_1, \cdots, y_n),\end{aligned}$$

which defines the function H. Using the first lines in (9.72) and (9.73), eq. (9.74) can be written in the form

$$\begin{aligned}\langle J_3 \cdots J_n \,\text{out}|J_1 J_2 \,\text{in}\rangle^{\text{conn}} &= \sqrt{Z_1 Z_2 Z_3^* \cdots Z_n^*} \sum_{\lambda_1 \cdots \lambda_n} \int d\omega_{p_1} \cdots d\omega_{p_n} \\ &\quad (2\pi)^4 \delta(p_1 + p_2 - p_3 \cdots - p_n) \\ &\quad J_3^*(p_3, \lambda_3) \cdots J_n^*(p_n, \lambda_n) \, U_{a_3}^*(p_3, \lambda_3) \cdots U_{a_n}^*(p_n, \lambda_n) \\ &\quad i H_{a_n \cdots a_3 a_1 a_2}(p_n, \cdots, p_3, -p_1, -p_2) \\ &\quad U_{a_1}(p_1, \lambda_1) \, U_{a_2}(p_2, \lambda_2) \, J_1(p_1, \lambda_1) \, J_2(p_2, \lambda_2) \qquad (9.76)\end{aligned}$$

Here the momentum-conserving delta function is due to translation invariance; recall that in our notation such delta functions are extracted from the definition (4.27) of the Fourier transform. Because the J's are arbitray (within the

[11]For notational simplicity we assume $\left(\sqrt{Z}\right)^* = \sqrt{Z^*}$.

constraint of describing reasonable wave packet states) we conclude that for the sharp-momentum basis states

$$\begin{aligned}
\langle p_3\lambda_2\cdots p_n\lambda_n\,\mathrm{out}|p_1\lambda_1,p_2\lambda_2\,\mathrm{in}\rangle^{\mathrm{conn}} &= -i(2\pi)^4\delta(p_1+p_2-p_3\cdots-p_n)\,T,\\
T(p_3\lambda_3\cdots p_n\lambda_n;p_1\lambda_1,p_2\lambda_2) &= -\sqrt{Z_1Z_2Z_3^*\cdots Z_n^*}\\
&\quad U^*_{a_3}(p_3,\lambda_3)\cdots U^*_{a_n}(p_n,\lambda_n)\\
&\quad H_{a_n\cdots a_3a_1a_2}(p_n,\cdots,p_3,-p_1,-p_2)\\
&\quad U_{a_1}(p_1,\lambda_1)\,U_{a_2}(p_2,\lambda_2), \qquad (9.77)
\end{aligned}$$

where T is the scattering amplitude.

So the recipe for obtaining the scattering amplitude is quite simple:

1. take the full connected n-point function in momentum space and make explicit the two-point functions on the external lines:

$$G^{a_1\cdots a_n}(p_1,\cdots,p_n)=G^{a_1b_1}(p_1)\cdots G^{a_nb_n}(p_n)\,H_{b_1\cdots b_n}(p_1,\cdots,p_n); \qquad (9.78)$$

2. omit the external line G''s so as to obtain

$$H_{a_1\cdots a_n}(p_1,\cdots,p_n), \qquad (9.79)$$

which is sometimes called 'the amputated Green function';

3. put H on-shell and contract with the polarization tensors/spinors,

$$\begin{aligned}
&U^*_{a_3}(p_3,\lambda_3)\cdots U^*_{a_n}(p_n,\lambda_n)\,H_{a_n\cdots a_3a_1a_2}(p_n,\cdots,p_3,-p_1,-p_2)\\
&U_{a_1}(p_1,\lambda_1)\,U_{a_2}(p_2,\lambda_2); \qquad (9.80)
\end{aligned}$$

4. multiply by the external-line wave-function renormalization factors to obtain $-T$, as in (9.77).

   The minus sign can be avoided by changing the sign in the definition of T, a convention that is also followed in the literature.

This recipe generalizes straightforwardly to the case in which (some of) the ϕ^a are composite fields, as is typically the case for hadrons in QCD (see section 9.6). The off-shell form of H is then somewhat arbitrary, since it depends on the choice of field ϕ^a used to create and absorb a particle (usually there is more than one natural candidate). However, this arbitrariness is compensated for by the $\sqrt{Z}$ factors and the contraction with the Us, and the scattering amplitude T is unique. The same is true in the case in which the ϕ^a are dynamical variables: then T is invariant under a transformations of variables. This uniqueness property of T is sometimes called 'the arbitrariness of the interpolating field'.

For simplicity of notation we emphasized only the properties of ϕ^a under Lorentz transformations and suppressed the type of particle, e.g. is it a W boson or a proton? This information can be added to the case at hand. Alternatively, it is sometimes useful to think of ϕ^a as running over a set of fields and use the U^a select a particle.

In practise we often do not use real fields but complex fields, for example Dirac fields. With the experience obtained in the previous chapters it is easy to adapt the formulas above for this case. Alternatively, we can consistently work with real fields, which leads to quite a transparent formalism (see e.g. the books by J. Schwinger, 'Particles, Sources and Fields'), but which is not generally used.

Let us now comment on the effect of massless particles. the typical example being the photon. Fundamentally, their effect invalidates our derivation, since the pole in spectral representations is not separated from the cut if $m' = 0$. Single particle contributions involving massless particles cannot be separated from multiparticle contributions. However, we are partially saved by the fact that with each photon emission or absorption comes a factor of $\alpha \approx 1/137$. It is customary to talk of 'hard' and 'soft' photons, the first having a sizable energy and the second having arbitrarily small energy. The hard photons are limited in number by energy conservation, and since α is so small, we can treat them perturbatively. For example, in electron-electron scattering we can neglect them altogether in a first approximation. But the soft photons can be unlimited in number, and when there are more than $1/\alpha$ of them their contribution falls outside perturbation theory. The difficulty is resolved by a careful analysis that takes into account the fact that we cannot distinguish experimentally a particle from a particle with a cloud of soft photons (Bloch-Nordsieck resolution of the 'infrared catastrophe'). Massless fermions are less dangerous because the Pauli principle forbids them piling up into nonperturbative numbers, and their effect can usually be treated in perturbation theory. Still, also in this case the analysis of what is measurable has to be refined. In practise, one can sometimes ignore these complications.

Finally, some comments on gauge invariance and using the Coulomb gauge. The wave function renormalization constants Z depend on the gauge. In a general covariant gauge they depend on the gauge parameter ξ. However, this dependence is just what is needed to make the transition amplitudes gauge *in*-dependent. In the Coulomb gauge 'ket space' has positive metric, but we have lost manifest Lorentz covariance. This means, for example, that the wave function renormalization constants Z depend on $|\mathbf{p}|$ (there is still manifest rotation invariance), which causes explicit demonstrations of equivalence to covariant gauges to be quite involved.

9.5 Decay revisited

The reasoning in the previous sections can also be extended heuristically to the decay of unstable particles. Let us briefly present a derivation for the example of the simple scalar field model of sect. 2.6, with interaction $S_1 = -\int d^4x\, \frac{1}{2} g\chi\varphi^2$. This new derivation complements that given in sect. 2.6 by expressing the decay amplitude in terms of the full three point function $\langle\varphi(x_1)\varphi(x_2)\chi(y)\rangle$. For example, it allows for arbitrarily strong interactions of the particles in the final state.

We start from a normalized wave packet state for the unstable χ particle,

$$|f\rangle = \int d\omega_q\, f(q)\,|q\rangle, \quad d\omega_q = \frac{d^3q}{(2\pi)^3 2q^0}, \quad q^0 = \sqrt{M^2 + \mathbf{q}^2}, \tag{9.81}$$

with M the mass of χ. We assume that $f(q)$ is sharply peaked about $\mathbf{q} = \mathbf{P}$. The probability amplitude for decay into two φ particles described by wave functions $f_1(p_1)$ and $f_2(p_2)$ is, following the same heuristic reasoning as in the previous sections, given by

$$\begin{aligned} \langle f_1 f_2\, \mathrm{out}|f\rangle &= \lim \langle 0|\varphi(J_1)\,\varphi(J_2)\,\chi(J)|0\rangle, \\ &= \lim \int d^4x_1 d^4x_2 d^4y\, J_1(x_1) J_2(x_2) J(y) \\ &\quad \langle 0|T\varphi(x_1)\varphi(x_2)\chi(y)|0\rangle, \end{aligned} \tag{9.82}$$

$$f(q) = \tilde{J}(q), \quad q^0 = \sqrt{M^2 + \mathbf{q}^2}, \tag{9.83}$$

$$f_{1,2}(p_{1,2}) = \tilde{J}^*_{1,2}(p_{1,2}), \quad p^0_{1,2} = \sqrt{m^2 + \mathbf{p}^2_{1,2}}, \tag{9.84}$$

where the limit consists of taking $J_{1,2}$ to the far future; the source J for the χ particle stays put. The wave packet state $|f\rangle$ is supposed to be given, somehow, at some time, say $t = 0$. It is *not* made by a limiting process in which the source J is moved to the far past. This would not make sense because of the instability of the χ particle. (Imagine the time reversed amplitude $\langle f|J_1 J_2\, \mathrm{in}\rangle$: it would not make sense to move the source J for the decaying χ particle to the far future.)

The total probability P for decay is tentatively given by

$$\mathcal{P} = \frac{1}{2}\int d\omega_{p_1} d\omega_{p_2}\, |\langle p_1 p_2\, \mathrm{out}|f\rangle|^2, \tag{9.85}$$

where the factor 1/2 takes care of the fact that the two φ particles in the final state are identical. Note that the mass hidden in the $d\omega_p$'s is the φ-mass m and not M.

Using the reasoning in the previous sections, the above formula can be expressed in terms of an invariant decay-amplitude T, which can be computed in terms of the connected three-point function with external lines removed, H,

$$\langle p_1 p_2\, \mathrm{out}|q\rangle = -i(2\pi)^4 \delta(p_1 + p_2 - q)\, T(p_1, p_2; q), \tag{9.86}$$

$$T(p_1, p_2; q) = -\frac{1}{\sqrt{Z_\chi Z_\varphi^2}}\, H(p_1, p_2, -q). \tag{9.87}$$

The amplitude $T(p_1, p_2; q)$ is Lorentz invariant, so it may depend on scalar products of the momenta, which leaves apparently one non-trivial invariant, $(p_1 - p_2)q$. However, this vanishes in the rest frame of the decaying particle. So T is just a constant.

Using the same type of approximations as in the derivation of the scattering cross section we have

$$\begin{aligned} \mathcal{P} &\approx \frac{1}{2}\int d\omega_{p_1} d\omega_{p_2}\, d\omega_q d\omega_{\bar{q}} f(q) f(\bar{q})^* (2\pi)^4 \delta(p_1 + p_2 - P)(2\pi)^4 \delta(q - \bar{q})\, |T|^2 \\ &\approx |T|^2 \frac{1}{2} I(P) \int d^4x\, |f(x)|^2, \qquad (9.88) \\ I(P) &= \int d\omega_{p_1} d\omega_{p_2}\, (2\pi)^4 \delta(p_1 + p_2 - P), \qquad (9.89) \\ f(x) &= \int d\omega_q\, e^{iqx}\, f(q). \qquad (9.90) \end{aligned}$$

For a wave packet $f(q)$ sharply peaked about $\mathbf{P}$ we have

$$\int d^4x\, |f(x)|^2 = \int dx^0 \int d\omega_q \frac{1}{2q^0} |f(q)|^2 \approx \frac{1}{2P^0} \int dx^0, \quad P^0 = \sqrt{M^2 + \mathbf{P}^2}, \tag{9.91}$$

and we discover that we have obtained a result that diverges at large times: $\int dx^0 \to \infty$.

This divergence is due to the fact that we have neglected higher order contributions. These will be such that, if we start with a large number N of χ particles, this number will decay exponentially,

$$N(t) = N(0) e^{-\Gamma t}, \tag{9.92}$$

with Γ the decay rate. Furthermore, the number of φ particles will be given by (assuming no other decay modes)

$$N_\varphi(t) = 2N(0)(1 - e^{-\Gamma t}), \tag{9.93}$$

or, interpreting this as a probability:

$$\mathcal{P}(t) = \frac{N_\varphi(t)/2}{N(0)} = 1 - e^{-\Gamma t}. \tag{9.94}$$

Now if we expand $\mathcal{P}(t)$ in Γ, $\mathcal{P}(t) = \Gamma t + \cdots$, we get nonsense for large t if we throw away the higher order terms. We can make sense out of our result (9.88)–(9.91) for $\mathcal{P}$ by identifying $\int dx^0$ in (9.91) with t:

$$\mathcal{P} = \Gamma t + \cdots \approx \Gamma \int dx^0, \tag{9.95}$$

with

$$\Gamma = \frac{1}{2} I(P) \frac{|T|^2}{2P^0}. \tag{9.96}$$

We expect that the result to all orders will look like

$$\mathcal{P} = \int_0^\infty dx^0 \, \Gamma \, e^{-\Gamma x^0} = 1, \tag{9.97}$$

where we have replaced t_0 by 0. It is to be expected that the particle decays with probability 1 after infinite time.

A more rigorous treatment of decaying particles can be given, which is somewhat involved, since it should deal with infinite orders in perturbation theory. Some of this is covered in Brown section 6.3. It turns out that the 'exponential decay law' itself is an approximation. However, eq. (9.96) is a useful and correct result if the rate Γ is not too large, or equivalently, if the lifetime $\tau \equiv \Gamma^{-1}$ is not too small, on the relevant time scales involved.

For a particle at rest ($\mathbf{P} = 0$) we have already verified in Problem 2.3 that

$$I(P) = \frac{p}{4\pi M}, \quad p \equiv |\mathbf{p}_1|, \tag{9.98}$$

hence,

$$\Gamma = \frac{p}{16\pi M^2} \, |T|^2, \tag{9.99}$$

where we can express p in terms of m and M. Recall that for a moving particle the rate behaves $\propto M/P^0$, in accordance with relativistic time dilatation.

In a similar way it is possible to derive a formula for the total rate for a decay $1 \to 2 + \cdots + n$:

$$\begin{aligned} \Gamma &= \frac{S}{2p_1^0(2s_1+1)} \sum_{\lambda_1 \cdots \lambda_n} \int d\omega_{p_2} \cdots d\omega_{pn} \, (2\pi)^4 \delta(p_1 - p_2 \cdots - p_n) \\ & \quad |T(p_2\lambda_2, \cdots, p_n\lambda_n; p_1\lambda_1)|^2, \end{aligned} \tag{9.100}$$

where S is a statistical factor taking care of identical particles in the final state. The invariant decay amplitude is then given by the analog of (9.77),

$$\begin{aligned} T(p_2\lambda_2 \cdots p_n\lambda_n; p_1\lambda_1) &= -\sqrt{Z_1 Z_2^* \cdots Z_n^*} \, U_{a_2}^*(p_2,\lambda_2) \cdots U_{a_n}^*(p_n,\lambda_n) \\ & \quad H_{a_n \cdots a_2 a_1}(p_n, \cdots, p_2, -p_1) \, U_{a_1}(p_1,\lambda_1), \end{aligned} \tag{9.101}$$

Unstable particles really require a much more involved discussion, especially in case the decay rate Γ is a sizable fraction of the mass of the unstable particle. See the books for more information, e.g. Brown, sect. 6.3, Peskin and Schroeder, sect. 7.3, De Wit and Smith, sect. 3.6.

9.6 Example: the decays $\pi^- \to e + \bar{\nu}_e$ and $\pi^- \to \mu^- + \bar{\nu}_\mu$

In the following we consider the decay process $\pi^- \to \mu^- + \bar{\nu}_\mu$ to illustrate how to the formalism of this chapter can be used to deal with bound states.

QCD possesses the remarkable property of *confinement*: quarks and gluons cannot occur as free particles, they form bound states, called *hadrons*, which *do* occur as free particles (to a good approximation). Well-known hadrons of low mass are the proton, neutron, the charged pions $\pi^\pm$, and the neutral pion π^0. When the electroweak interactions are neglected, these particles are stable, but in the Standard Model only the proton is stable (up to a negligible decay probability related to so-called sphaleron processes).

For the description the $\pi^\pm$ we use composite pseudo-scalar fields,

$$\phi(x) \equiv \bar{\psi}_d(x) i\gamma_5 \psi_u(x), \quad \phi^\dagger(x) = \bar{\psi}_u(x) i\gamma_5 \psi_d(x), \tag{9.102}$$

where the $\psi_{u,d}$ fields are the 'up' and 'down' quark fields. The computation of correlation functions such as

$$\langle 0|T\phi(x)\phi^\dagger(y)|0\rangle \equiv \langle \phi(x)\phi^\dagger(y)\rangle \equiv G_\phi(x,y) \tag{9.103}$$

is a complicated problem in non-perturbative QCD, which can be done using computers and the lattice regularization ('lattice QCD'). The Fourier transform of this correlator has a pole at $p^2 = -m_{\pi^+}^2 = -m_{\pi^-}^2$, corresponding to the matrix elements

$$\langle p,\pi^+|\phi^\dagger(x)|0\rangle = \sqrt{Z}\, e^{ipx}, \quad \langle p,\pi^-|\phi^\dagger(x)|0\rangle = 0, \tag{9.104}$$

$$\langle p,\pi^-|\phi(x)|0\rangle = \sqrt{Z}\, e^{ipx}, \quad \langle p,\pi^+|\phi(x)|0\rangle = 0 \tag{9.105}$$

(the zeros follow e.g. from conservation of electric charge, $Q_u = 2/3$, $Q_d = -1/3$, $Q_{\pi^+} = 1$)[12]. We assumed $\sqrt{Z}$ to be positive. Hence,

$$\langle 0|T\phi(x)\phi^\dagger(y)|0\rangle \overset{x^0\geq y^0}{=} \int d\omega_p\, \langle 0|\phi(x)|p,\pi^+\rangle\langle p,\pi^+|\phi^\dagger(y)|0\rangle + \text{m.p.} \tag{9.106}$$

$$\overset{x^0\leq y^0}{=} \int d\omega_p\, \langle 0|\phi^\dagger(y)|p,\pi^-\rangle\langle p,\phi^-|\phi(x)|0\rangle + \text{m.p.} \tag{9.107}$$

$$= -i\int \frac{d^4p}{(2\pi)^4}\, e^{ip(x-y)}\, \frac{Z}{m_\pi^2 + p^2 - i\epsilon} + \text{m.p.}, \tag{9.108}$$

$$G_\phi(p) = \frac{Z}{m_\pi^2 + p^2 - i\epsilon} + \text{m.p.}, \tag{9.109}$$

where m.p. denotes the contribution from multi-particle intermediate states. The value of Z can be obtained with the methods of lattice QCD, but we shall not need this in the following.

For the decays, let us use an *effective action* that can be derived as an approximation to the Standard Model. It is given by

$$S = \int d^4x\, \left(\mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{lepton}}^{\text{free}} + \mathcal{L}_1\right) = S_0 + S_1, \tag{9.110}$$

[12] Formally: $[Q,\phi] = -\phi \to Q\phi|0\rangle = [Q,\phi]|0\rangle = -\phi|0\rangle$, hence $\phi|0\rangle$ has charge $Q = -1$.

$$\mathcal{L}_0 = \mathcal{L}_{\rm QCD} + \mathcal{L}^{\rm free}_{\rm lepton}, \tag{9.111}$$

$$\mathcal{L}^{\rm free}_{\rm lepton} = -\bar\psi_e(\gamma^\mu\partial_\mu + m_e)\psi_e - \bar\psi_{\nu_e}\gamma^\mu\partial_\mu\psi_{\nu_e} + e\to\mu, \tag{9.112}$$

$$\mathcal{L}_1 = \frac{G_{\rm F}\cos\theta}{\sqrt{2}}\,\bar\psi_e i\gamma^\mu(1-\gamma_5)\psi_{\nu_e}\, j^\dagger_\mu + h.c. + e\to\mu, \tag{9.113}$$

$$j^\dagger_\mu = \bar\psi_u i\gamma_\mu(1-\gamma_5)\psi_d. \tag{9.114}$$

Here $\mathcal{L}_{\rm QCD}$ is the QCD lagrangian in terms of quark and gluon fields, which will not be specified further; $\mathcal{L}^{\rm free}_{\rm lepton}$ is the free lagrangian for the electron and the muon and the corresponding neutrinos, which are taken to be massless; $\mathcal{L}_1$ is the relevant part of an effective lagrangian for the weak interactions, the *four-Fermi interaction*, with $G_{\rm F}$ the Fermi constant and θ the Cabibbo angle. The electromagnetic interactions are neglected.

For the decay $\pi^-(p)\to e(k)+\bar\nu_e(k')$ we need the correlation function

$$\langle 0|T\psi_e(x)\,\bar\psi_{\nu_e}(y)\,\phi(z)|0\rangle \equiv \langle\psi_e(x)\,\bar\psi_{\nu_e}(y)\,\phi(z)\rangle \equiv (-i)^2 G_{e\bar\nu_e\phi}(x,y,z), \tag{9.115}$$

Perturbing in S_1 we have

$$\langle\psi_e(x)\,\bar\psi_{\nu_e}(y)\,\phi(z)\rangle = \frac{\langle e^{iS_1}\,\psi_e(x)\,\bar\psi_{\nu_e}(y)\,\phi(z)\rangle_0}{\langle e^{iS_1}\rangle_0} \tag{9.116}$$

$$= i\int d^4u\,\langle\mathcal{L}_1(u)\,\psi_e(x)\,\bar\psi_{\nu_e}(y)\,\phi(z)\rangle_0 + \ldots \tag{9.117}$$

where the $\ldots$ denote terms of higher order in G_F and $\langle\cdot\rangle_0$ is the path-integral 'average' over the full QCD variables and the free lepton variables,

$$\langle F\rangle_0 = \frac{\int_{\rm QCD,lepton} e^{iS_0}\,F}{\int_{\rm QCD,lepton} e^{iS_0}}. \tag{9.118}$$

Performing the contractions of the lepton fields gives

$$i\langle\mathcal{L}_1(u)\psi_e(x)\bar\psi_{\nu_e}(y)\phi(z)\rangle_0 = \frac{iG_{\rm F}\cos\theta}{\sqrt{2}}(-i)G_e(x,u)i\gamma^\mu(1-\gamma_5)(-i)G_{\nu_e}(u,y) \times\langle j^\dagger_\mu(u)\phi(z)\rangle_0 \tag{9.119}$$

The second factor above is a correlator within QCD:

$$\langle j^\dagger_\mu(u)\,\phi(z)\rangle_0 = \langle j^\dagger_\mu(u)\,\phi(z)\rangle_{\rm QCD} \equiv -iG_{j^\dagger_\mu\phi}(u,z), \tag{9.120}$$

since the path integral over the lepton variables is just a factor that cancels between numerator and denominator when F in (9.118) contains only QCD fields. Putting things together and taking the Fourier transform, the three-point correlator becomes

$$(-i)^2G_{e\bar\nu_e\phi}(k,k',p) = \frac{iG_{\rm F}\cos\theta}{\sqrt{2}}\,(-i)G_e(k)\,i\gamma^\mu(1-\gamma_5)(-i)G_{\nu_e}(-k')\,(-i)G_{j^\dagger_\mu\phi}(-p), \tag{9.121}$$

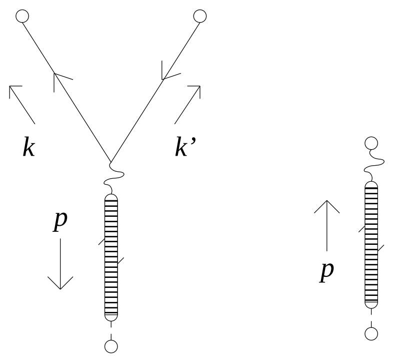

Figure 9.3: Left: Diagram for the three-point function $(-i)^2 G_{e\bar{\nu}_e\phi}(k, k', p)$. It consists of the vertex function $(iG_{\rm F}\cos\theta/\sqrt{2})i\gamma^\mu(1-\gamma_5)$, the electron propagator $-iG_e(k)$, the neutrino propagator $-iG_{\nu_e}(-k')$, and the correlator $-iG_{j^\dagger_\mu\phi}(-p)$. In the approximation used here, the four-Fermi interaction, the wiggly line merely indicates a vector index. In the Standard Model it would represent the W-propagator. Right: the correlator $-iG_{j^\dagger_\mu\phi}(p)$.

whith the diagrammatical representation in figure 9.3.

Consider next the $j^\dagger_\mu\phi$ correlator. Choosing $u^0 > z^0$ and inserting intermediate states gives

$$\langle 0|T j^\dagger_\mu(u)\,\phi(z)|0\rangle \stackrel{u^0\geq z^0}{=} \int d\omega_p\,\langle 0|j^\dagger_\mu(u)|p,\pi^-\rangle\langle p,\pi^-|\phi(z)|0\rangle + \text{m.p.}, \tag{9.122}$$

where we exhibited the one-pion intermediate-state contribution. The matrix element of the current can be parametrized, on grounds of Poincaré invariance, in terms of only one constant, f_π, which has the dimension of mass,

$$\langle 0|j^\dagger_\mu(u)(u)|p,\pi^-\rangle = ip_\mu f_\pi\sqrt{2}\,e^{ipu} \tag{9.123}$$

(only the pseudo-vector part of j_μ contributes, cf. problem 10.2). Hence,

$$\langle 0|T j^\dagger_\mu(u)\,\phi(z)|0\rangle \stackrel{u^0\geq z^0}{=} \int d\omega_p\,ip_\mu f_\pi\sqrt{2Z}\,e^{ip(u-z)} + \text{m.p.}. \tag{9.124}$$

The other time-ordering gives

$$\langle 0|T j^\dagger_\mu(u)\,\phi(z)|0\rangle \stackrel{u^0\leq z^0}{=} \int d\omega_p\,(-ip_\mu)f_\pi\sqrt{2Z}\,e^{-ip(u-z)} + \text{m.p.}, \tag{9.125}$$

and so the correlator has the form

$$\langle 0|T j^\dagger_\mu(u)\,\phi(z)|0\rangle = -i\int \frac{d^4p}{(2\pi)^4}\,e^{ip(u-z)}\,\frac{ip_\mu f_\pi\sqrt{2Z}}{m_\pi^2 + p^2 - i\epsilon} + \text{m.p.}, \tag{9.126}$$

$$G_{j^\dagger_\mu\phi}(p) = \frac{ip_\mu f_\pi\sqrt{2Z}}{m_\pi^2 + p^2 - i\epsilon} + \text{m.p.}. \tag{9.127}$$

To obtain the H function, we need to strip off the external-line propagators (correlators) from the three-point $e\bar{\nu}_u\phi$ correlator. There is no explicit ϕ correlator in (9.121), so we simply multiply by the inverse ϕ correlator:

$$iH_{e\bar{\nu}_e\phi}(k,k',-p) = \frac{iG_{\rm F}\cos\theta}{\sqrt{2}}\, i\gamma^\mu(1-\gamma_5)(-i)G_{j^\dagger_\mu\phi}(p)\, iG_\phi(p)^{-1} \tag{9.128}$$

Going on-shell,

$$G_\phi(p)^{-1} \stackrel{p^2\to -m_\pi^2}{\to} Z^{-1}(m_\pi^2+p^2), \tag{9.129}$$

which selects the pole part of $G_{j^\dagger_\mu\phi}(p)$. Hence,

$$H(k',k',-p) \stackrel{\rm o.s.}{=} \frac{G_{\rm F}\cos\theta}{\sqrt{2}}\, i\gamma^\mu(1-\gamma_5)\,\frac{ip_\mu f_\pi\sqrt{2}}{\sqrt{Z}}, \tag{9.130}$$

and the decay amplitude becomes

$$\begin{aligned} T(k,\lambda;k'\lambda';p) &= -\sqrt{Z}\,\bar{u}(k,\lambda)H(k,k',-p)v(k',\lambda') &(9.131)\\ &= G_{\rm F}\cos\theta f_\pi p_\mu \bar{u}(k,\lambda)\gamma^\mu(1-\gamma_5)v(k',\lambda'). &(9.132)\end{aligned}$$

Alternatively, we could have used the divergence of an axial current to obtain a pseudoscalar field

$$\phi'(x) = \partial_\mu\left[\bar{\psi}_d(x)i\gamma^\mu\gamma_5\psi_u(x)\right]. \tag{9.133}$$

In this case the corresponding Z' in known in terms of f_π and m_π. The resulting expression for T is the same.

9.7 Problems

1. *Spin-one polarization vectors*

   Let $\phi^a \to V^\mu$ be a vector field and $|p,\lambda\rangle$ basis vectors for spin-one particles. Then the $e^\mu(p,\lambda)$ in

   $$\langle 0|V^\mu(x)|p,\lambda\rangle = \sqrt{Z}\,e^\mu(p,\lambda)\,e^{ipx} \tag{9.134}$$

   are the polarization vectors for this case. For a particle at rest we may assume

   $$e^0(\bar{p},\lambda)=0, \quad e^k(\bar{p},\lambda)=\delta_{k\lambda}, \quad \lambda=1,2,3. \tag{9.135}$$

   where

   $$\bar{p}=(m,\mathbf{0}). \tag{9.136}$$

   For general p the polarization vectors can be defined by applying the special Lorentz boost ℓ_p which transforms $\bar{p}$ into p:

   $$e^\mu(p,\lambda) = e^\mu(\ell_p\,\bar{p},\lambda) = \ell_p{}^\mu{}_\nu\, e^\nu(\bar{p},\lambda). \tag{9.137}$$

a. Derive for a boost in the 3-direction that

$$e^0(p,3) = \frac{|\mathbf{p}|}{m}, \quad e^k(p,3) = \frac{p^0}{k}\delta_{k,3}, \tag{9.138}$$

where $p^0 = \sqrt{\mathbf{p}^2 + m^2}$. With an eye on completing the two vectors $\mathbf{e}(\mathbf{p},\lambda)$ used for the photon we may generalize this to

$$e^0(p,3) = \frac{|\mathbf{p}|}{m}, \quad \mathbf{e}(p,3) = \hat{p}\,\frac{p^0}{m}, \tag{9.139}$$

$$e^0(p,\lambda) = 0, \quad \mathbf{e}(p,\lambda) = \text{as for photon, } \lambda = 1,2 \tag{9.140}$$

b. Verify

$$e^*_\mu(p,\lambda)e^\mu(p,\lambda') = \delta_{\lambda\lambda'}. \tag{9.141}$$

c. Verify

$$p_\mu e^\mu(p,\lambda) = 0, \quad \lambda = 1,2,3. \tag{9.142}$$

d. Verify

$$P^{\mu\nu}(p) \equiv \sum_\lambda e^\mu(p,\lambda)\, e^{*\nu}(p,\lambda) = \eta^{\mu\nu} + \frac{p^\mu p^\nu}{m^2}. \tag{9.143}$$

(This can also be obtained directly by transformation of the polarization sum at rest.)

e. Verify that

$$N^{\mu\nu}(p,m) = \eta^{\mu\nu} + \frac{p^\mu p^\nu}{m^2} \tag{9.144}$$

satisfies (9.61).

2. *Effective action for massive spin-one field*

In this problem we look for an effective action for the field theoretic description of massive spin-one particles. We choose to describe the particles with a vector field V^μ. Suppose the vector field of the previous problem is normalized such that $Z = 1$. If the field is free, then the multiparticle contribution in its spectral representation wil vanish, and the propagator for the free system is given by

$$G^{\mu\nu}(p) = \frac{N^{\mu\nu}(p,m)}{m^2 + p^2 - i\epsilon} = \frac{\eta^{\mu\nu} + p^\mu p^\nu/m^2}{m^2 + p^2 - i\epsilon}. \tag{9.145}$$

The inverse $G^{-1}_{\mu\nu}(p)$ is hopefully a polynomial in p, such that we get a differential operator in position space ($p_\mu \to -i\partial_\mu$). This is indeed the case:

a. Verify

$$[\eta_{\mu\nu}(m^2 - \partial^2) + \partial_\mu\partial_\nu]G^{\nu\rho}(x-y) = \delta^\rho_\mu \delta^4(x-y), \tag{9.146}$$

with $G^{\nu\rho}(x-y)$ the Fourier transform of $G^{\mu\nu}(p)$.

b. Show using path integral quantization that the effective action

$$S = -\int d^4x \left(\frac{1}{4}V_{\mu\nu}V^{\mu\nu} + \frac{1}{2}m^2 V_\mu V^\mu\right), \quad V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu, \tag{9.147}$$

leads to the above propagator.

The above effective action should in first instance only be used for tree diagram calculations. If we go beyond this and use it as a starting point for a fundamental theory of interacting spin-one particles, then the fact that the above propagator $G^{\mu\nu}(p)$ does *not* go to zero nicely for $p \to \infty$ causes a lot of trouble with divergencies in loop diagrams, especially in non-abelian gauge theories. These problems are resolved by invoking the Higgs mechanism as the underlying reason for the mass of spin-one particles like W and Z in the Standard Model. For their contribution to the theory of renormalization of non-abelian gauge theories, 't Hooft and Veltman received the Nobel price in 1999. See, e.g. Veltman's book for more information on the relation between gauge invariance, renormalizability and 'unitarity'.

3. *Calculation of matrix elements*

   The construction of transition amplitudes given in this chapter can be extended to matrix elements of local operators, e.g. the single particle matrix element $\langle f|T^{\mu\nu}(x)|f\rangle$. In free scalar field theory, derive (2.133) using the Wick formula.

4. *The decays $\pi^- \to \mu^- + \bar{\nu}_\mu$ and $\pi^- \to e^- + \bar{\nu}_e$*

   The charged pions $\pi^\mp$ are unstable and decay mainly into muons $\mu^\mp$ and muon neutrinos $(\bar{\nu}_\mu)\nu_\mu$, with a life time of 2.60×10^{-8} s, or $\Gamma^{-1} = 780$ cm. There is a corresponding decay into electrons $e^\mp$ and electron neutrinos $(\bar{\nu}_e)\nu_e$, with a much smaller rate. These processes can be described by an effective action of the form $S = S_0 + S_1$, where S_0 is the sum of the actions for the free pions,[13] muons, electrons, muon neutrinos and electron neutrinos,

$$S_0 = S_\pi + S_\mu + S_e + S_{\nu_\mu} + S_{\nu_e}, \tag{9.148}$$

$$S_\pi = -\int d^4x\, (\partial_\mu\varphi^*\partial^\mu\varphi + m_\pi^2\varphi^*\varphi), \tag{9.149}$$

$$S_\mu = -\int d^4x\, \bar{\psi}_\mu(\gamma^\kappa\partial_\kappa + m_\mu)\psi_\mu, \tag{9.150}$$

   and similar for e, ν_μ and ν_e with $m_{\nu_\mu} = m_{\nu_e} = 0$, and S_1 is the interaction

$$S_1 = c\int d^4x\, [\partial_\kappa\varphi^*\, \bar{\psi}_\mu i\gamma^\kappa(1-\gamma_5)\psi_{\nu_\mu} + \partial_\kappa\varphi\, \bar{\psi}_{\nu_\mu} i\gamma^\kappa(1-\gamma_5)\psi_\mu + (\mu \to e)]. \tag{9.151}$$

[13] The $\pi^\pm$ are described by the *complex* scalar field φ.

The constant c is given by

$$c = f_\pi \, G_F \, \cos\theta_C, \qquad (9.152)$$

with f_π the pion decay constant, G_F the Fermi weak interaction constant and θ_C the Cabibbo angle.

Notice that the interaction S_1 does not conserve parity P, as it is the sum of terms odd and even under parity.

a. Draw the diagram for the decay $\pi^-(p) \to \mu^-(k,\lambda) + \bar{\nu}_\mu(k',\lambda')$ and verify that the decay amplitude is given by

$$\begin{aligned} \langle k\lambda, \overline{k'\lambda'}\,\mathrm{out}|p\,\mathrm{in}\rangle &= -i(2\pi)^4\delta(p-k-k')\,T(k,\lambda,k',\lambda';p), \quad (9.153)\\ T(k,\lambda,k',\lambda';p) &= c\bar{u}(k,\lambda)\gamma p(1-\gamma_5)v(k',\lambda'). \quad (9.154)\end{aligned}$$

b. Verify the polarization sum

$$\overline{|T|^2} = \frac{1}{2}c^2\mathrm{Tr}\,[\gamma p(1-\gamma_5)(i\gamma k')\gamma p(1-\gamma_5)(m_\mu - i\gamma k)]. \qquad (9.155)$$

(The neutrino has only one spin polarization.)

c. Using the anticommutation relations of the gamma matrices, the properties of the right and lefthanded projectors $P_{R,L} = (1 \pm \gamma_5)/2)$ and the 'trace identities', show that

$$\overline{|T|^2} = 4c^2[2(pk)(pk') - p^2kk']. \qquad (9.156)$$

d. In the rest frame of the pion, verify

$$|\mathbf{k}| = \frac{m_\pi^2 - m_\mu^2}{2m_\pi}, \quad k^0 - |\mathbf{k}| = \frac{m_\mu^2}{m_\pi}, \qquad (9.157)$$

and

$$\overline{|T|^2} = 2c^2(m_\pi^2 - m_\mu^2)m_\mu^2, \qquad (9.158)$$

and

$$\Gamma(\pi^- \to \mu^- + \bar{\nu}_\mu) = \frac{c^2}{4\pi}\frac{(m_\pi^2 - m_\mu^2)^2 m_\mu^2}{m_\pi^3}. \qquad (9.159)$$

e. The masses of the particles are given by $m_{\pi^\pm} = 139.6$ MeV, $m_\mu = 105.7$ MeV, $m_e = 0.5110$ Mev (the neutrino masses were already assumed to be zero). Using $G_F \approx 1.17 \times 10^{-5}$ GeV^{-2}, $\theta_C \approx 13°$, and the fact that π^- decays for 99.988% into $\mu^- + \bar{\nu}_\mu$, verify that $f_\pi \approx 93$ MeV from the rate $\Gamma = (780\mathrm{cm})^{-1}$.

f. Calculate the branching ratio

$$\frac{\Gamma(\pi^- \to e^- + \bar{\nu}_e)}{\Gamma(\pi^- \to \mu^- + \bar{\nu}_\mu)} \tag{9.160}$$

and compare this with the experimental value 1.22×10^{-4}.

The striking smallness of the above branching ratio is a consequence of the combination $\gamma^\kappa(1-\gamma_5)$ in the interaction S_1. The interaction conserves chirality: $1-\gamma_5$ projects on to chirality –1, in the neutrino fields as well as in the electron or muon fields (recall that $\bar{\psi}$ contains $\beta = i\gamma^0$ and γ_5 commutes with $i\gamma^0\gamma^\kappa$). For the massless antineutrinos, chirality –1 means helicity +1/2. For the electron and muon, chirality –1 would mean helicity –1/2 if these particles were massless. However, angular momentum conservation requires that the muon or electron have the same helicity (+1/2) as the antineutrino, since the pion at rest has angular momentum zero. Hence, if m_μ and m_e would be zero, the decay amplitude would vanish ($1-\gamma_5$ acting on a massless helicity +1/2 *particle* spinor gives zero). So we may expect that the decay amplitude is proportional to $m_{\mu,e}$ as $m_{\mu,e}$ goes to zero. In fact, it can be shown using helicity spinors that the decay amplitude is given by

$$T = 2icm_\pi \sqrt{|\mathbf{k}|}\, (\sqrt{k^0 + m} - \sqrt{k^0 - m})\, \delta_{\lambda,\lambda'}\, \delta_{\lambda',+}, \tag{9.161}$$

with $k^0 = \sqrt{\mathbf{k}^2 + m^2}$ and $m = m_\mu$ or m_e. In this way we can understand why the above branching ratio $\propto m_e^2/m_\mu^2$ is so small.

Chapter 10

C, P, and T

In addition to Lorentz invariance, the discrete symmetry transformations C (charge conjugation), P (parity) and T (time reversal) are of crucial importance in particle physics and cosmology. We shall give a brief introduction here which can only scratch the surface, as a proper presentation of the important experimental implications and results would lead much too far. For more information, see e.g. the books by Weinberg, Perkins, Branco et al., Sanda & Bigi.

10.1 Charge conjugation

We have seen that the QED action is invariant under global (and of course local) U(1) transformations. In Majorana language this means SO(2) invariance, in which the real and imaginary parts $\psi_{1,2}$ of the Dirac field get rotated into each other, as in (8.3). In fact, the free action is invariant under the full O(2) group, obtained by adding reflections,

$$\psi_1 \to \psi_1, \quad \psi_2 \to -\psi_2, \tag{10.1}$$

which are orthogonal transformations in internal space with determinant minus one. For the Dirac field this means $\psi \leftrightarrow \psi^*$, in a Majorana representation of the γ-matrices. In a general representation we have to be a bit careful, because ψ^* does not transform like ψ under Lorentz transformations. However, as seen in Appendix 5.5, the charge-conjugate spinors

$$\psi^{(c)} = \beta C \psi^* = (\bar{\psi} C)^T, \quad \bar{\psi}^{(c)} = -(C^\dagger \psi)^T \tag{10.2}$$

do transform in the same way. Recall that C is unitary and antisymmetric with

$$\gamma^{\mu T} = -C^\dagger \, \gamma^\mu \, C, \quad \beta C \, \gamma^{\mu *} \, C^\dagger \beta = \gamma^\mu, \tag{10.3}$$

and recall also our convention $\beta C = 1$ in a Majorana representation.

It is straightforward to check that the transformation C defined by

$$C: \quad \psi \to (\bar{\psi}C)^T, \quad \bar{\psi} \to -(C^\dagger\psi)^T, \tag{10.4}$$

leaves the free Dirac action invariant. We can extend this symmetry to QED if we adopt as the transformation rule for the electromagnetic field:

$$C: \quad A_\mu \to -A_\mu. \tag{10.5}$$

Let us check this for the interaction term in the Lagrange density:

$$\begin{aligned} \mathcal{L}_1 &= ej^\mu A_\mu, \quad j^\mu = \bar{\psi}i\gamma^\mu\psi, && (10.6)\\ j^\mu &\to -(C^\dagger\psi)^T i\gamma^\mu(\bar{\psi}C)^T = \bar{\psi}Ci\gamma^{\mu T}C^\dagger\psi = -\bar{\psi}i\gamma^\mu\psi = -j^\mu, && (10.7)\\ \mathcal{L}_1 &\to e(-j^\mu)(-A_\mu) = \mathcal{L}_1. && (10.8) \end{aligned}$$

In the second line we interchanged ψ and $\bar{\psi}$, which gives a minus sign by their anticommuting property, and then used (10.3). This symmetry of the action under C implies, via the path integral, a unitary operator U_C in the quantum Hilbert space (or in the extended negative metric space, in case of a covariant gauge). The derivation of this is cumbersome (we also have not discussed yet fermionic coherent states) and it is simpler to guess the answer directly. Let the unitary operator C have the properties

$$\begin{aligned} U_C\psi U_C^\dagger &= (\bar{\psi}C)^T, \quad U_C\bar{\psi}U_C^\dagger = -(C^\dagger\psi)^T, && (10.9)\\ U_C A_\mu U_C^\dagger &= -A_\mu, && (10.10) \end{aligned}$$

It is straightforward to check that U_C leaves the canonical (anti)commutation relations invariant (otherwise it could not be unitary), as well as the total hamiltonian,

$$\begin{aligned} U_C H U_C^\dagger &= H, \quad [U_C, H] = 0, && (10.11)\\ H &= \int d^3x \left(\frac{1}{2}E^2 + \frac{1}{2}B^2 + \psi^\dagger H_D\psi + ej^0A^0 - e\mathbf{jA}\right), && (10.12) \end{aligned}$$

where H_D is the Dirac hamiltonian (5.85). So U_C is time-independent and we may assume that it leaves the vacuum invariant,

$$U_C|0\rangle = |0\rangle. \tag{10.13}$$

We end this section with miscellaneous remarks:

- Products of fermion fields in operators should be antisymmetrized, such that they may be freely interchanged as anticommuting (the delta functions from the canonical anticommutation relations cancel). An example is the current operator

$$j^\mu = \frac{1}{2}\bar{\psi}i\gamma^\mu\psi - \frac{1}{2}\psi^T i\gamma^{\mu T}\bar{\psi}^T, \tag{10.14}$$

which satisfies

$$U_C j^\mu U_C^\dagger = -j^\mu \tag{10.15}$$

also in operator language.

- C interchanges particles and antiparticles:

$$U_C\, b(p,\lambda) U_C^\dagger = d(p,\lambda), \quad U_C\, d(p,\lambda) U_C^\dagger = b(p,\lambda). \tag{10.16}$$

- Neutral states have a definite C-parity:

$$U_C\, a^\dagger(k,\lambda)|0\rangle = -a^\dagger(k,\lambda)|0\rangle, \tag{10.17}$$
$$U_C\, b^\dagger(p,\lambda) d^\dagger(p',\lambda')|0\rangle = -b^\dagger(p,\lambda) d^\dagger(p',\lambda')|0\rangle. \tag{10.18}$$

The photon has negative negative C-parity.

- The C-parity has to be the same in initial and final states.[1]

- The Yukawa model with pseudoscalar coupling is an example of an effective theory for neutral pions and protons. The interaction lagrangian

$$\mathcal{L}_1 = g\bar{\psi} i\gamma_5 \psi \phi \tag{10.19}$$

is invariant under C if

$$U_C \phi U_C^\dagger = +\phi. \tag{10.20}$$

The neutral pion described by the ϕ field has positive C-parity. It decays primarily electromagnetically into two photons and cannot decay into an odd number of photons, because of C-parity conservation.

- QED and QCD are C-invariant, but the full Standard Model violates C.

- For nonabelian gauge fields, as in QCD, the coupling to fermions can be written in the form

$$g\bar{\psi} i\gamma^\mu A_\mu \psi, \tag{10.21}$$

where g is a coupling constant and A_μ is a matrix acting on additional ('color' or 'flavor') indices of the fermion fields. The charge conjugation transformation for A_μ is now given by

$$U_C A_\mu U_C^\dagger = -A_\mu^T. \tag{10.22}$$

[1]Neglecting the C-violation caused by the weak interactions.

- We may make an additional global U(1) transformation in the definition of U_C,
$$U_C\,\psi U_C^\dagger = \eta(\bar\psi C)^T, \quad U_C\,\bar\psi U_C^\dagger = -\eta^*(C^\dagger\psi)^T, \tag{10.23}$$
with $|\eta| = 1$. This puts phase factors in (10.16) but does not affect the minus sign in (10.18).

  As a generalization of this, it is always possible to change the definition of a discrete transformation like U_C by making an additional transformation of variables, which may or may not be a symmetry transformation (the same holds for P and T). This may turn an apparently noninvariant model into an invariant one. It is clearest to make such transformations of variables first, putting the action into some convenient form. This is commmon practise in the Standard Model, e.g. for the Cabibbo-Kobayashi-Maskawa matrix.

10.2 Parity

We have encountered parity already several times. It represents the 'improper' Lorentz transformation $(x^0, \mathbf{x}) \to (x^0, -\mathbf{x})$,

$$x' = \ell_P\,x, \quad \ell^\mu_{P\nu} = \mathrm{diag}(1,-1,-1,-1), \tag{10.24}$$
$$D_P = D(\ell_P) = \gamma^0, \quad D_P^{-1} = -\gamma^0 = D_P^\dagger, \tag{10.25}$$

where D_P is a spinor representation. In terms of a unitary operator U_P the transformation rules read in operator language

$$U_P\psi(x)U_P^\dagger = D_P\psi(\ell_P x), \quad U_P\bar\psi(x)U_P^\dagger = \bar\psi(\ell_P x)D_P^{-1}, \tag{10.26}$$
$$U_P A^\mu(x)U_P^\dagger = \ell^\mu_{P\nu}A^\nu(\ell_P x) \tag{10.27}$$

(with identical transformation rules for Majorana fields.) Note that $\ell_P = \ell_P^{-1}$ (but for our choice, $D_P \neq D_P^{-1}$).

Miscellaneous remarks:

- The QED hamiltonian (10.12) is parity invariant. Experimentally, QED and QCD are parity invariant, e.g. there is no indication for non-invariant terms like $\epsilon^{\kappa\lambda\mu\nu}F_{\kappa\lambda}F_{\mu\nu}$ in the lagrangian.

- It follows from the group properites of ℓ and ℓ_P that momentum and angular momentum are respectively odd and even under parity:
$$U_P\,\mathbf{P}U_P^\dagger = -\mathbf{P}, \quad U_P\,\mathbf{J}U_P^\dagger = +\mathbf{J}. \tag{10.28}$$

- Hence, a particle at rest can be in an eigenstate of U_P. The eigenvalue is called the (intrinsic) parity of the particle. For single fermions with a charge, we can choose

$$U_P\, b^\dagger(\bar{p},\lambda)|0\rangle = ib^\dagger(\bar{p},\lambda)|0\rangle, \quad \bar{p}=(m,\vec{0}) \tag{10.29}$$
$$U_P\, d^\dagger(\bar{p},\lambda)|0\rangle = id^\dagger(\bar{p},\lambda)|0\rangle, \tag{10.30}$$

this is convention dependent, because we can always put an additional phase in the parity transformation, similar to (10.23) for C. However, these phases cancel in a fermion-antifermion state, for which the resulting minus sign is convention *in*dependent. A fermion-antifermion in a spatially symmetric bound state with total momentum zero has negative parity. The neutral pion, a quark-antiquark state, is an example. Its effective field transforms as $\phi(x) \to -\phi(\ell_P x)$ under parity. The effective Yukawa model (10.19) illustrates the properties of π^0 under both C and P.

A neutral fermion (describable by a Majorana field) can have parity $\pm i$.

- The following interaction cannot be invariant under C or P,

$$\mathcal{L}_1 = \bar{\psi} i\gamma^\mu (g_V + g_A\,\gamma_5)\psi\, A_\mu, \tag{10.31}$$

if the coupling constants g_V and g_A are both non-zero, but it is invariant under the product CP. This illustrates the situation in the Standard Model.

10.3 Time reversal

The time reflection is defined by $(x^0,\mathbf{x}) \to (-x^0,\mathbf{x})$, or

$$x' = \ell_T\, x, \quad \ell^\mu_{T\nu} = \mathrm{diag}(-1,1,1,1), \tag{10.32}$$
$$D_T = D(\ell_T) = i\gamma^0\gamma_5, \quad D_P^{-1} = -i\gamma^0\gamma_5 = D_P^\dagger, \tag{10.33}$$

where D_T is a spinor representation. It turns out that in Hilbert space this symmetry is represented by an *antiunitary* operator U_T, i.e. a unitary operator that is antilinear.[2] To understand this we first turn to non-relativistic quantum mechanics.

Consider a lagrangian of the conventional type $L(q,\dot{q}) = \frac{1}{2}m\dot{q}^2 - V(q)$. It is invariant under the transformation $q'(t) = q(-t)$. The equations of motion are invariant under time reversal. If we have a solution $q(t)$ satisfying $q(t_1) = q_1$, $q(t_2) = q_2$, then $q'(t) = q(-t)$ is also a solution, with boundary values $q'(-t_1) = q_1$ and $q'(-t_2) = q_2$. If the velocity $\dot{q}$ happens to be positive throughout the whole trajectory, then $\dot{q}'$ is negative.[3] The symmetry of L has consequences in the path integral. To cut the discussion short, consider the semiclassical approximation

$$\langle q_1,t_1|q_2,t_2\rangle \propto \exp\left(i\int_{t_2}^{t_1} dt\, L[q,\dot{q}]\right), \tag{10.34}$$

[2]Recall that for an antilinear operator A and complex number λ, $A\lambda|\psi\rangle = \lambda^* A|\psi\rangle$.

[3]A time-honoured German name for T is *Bewegungsumkehr* (motion-reversal).

where $q(t)$ is a classical solution as mentioned earlier. Recall that $|q,t\rangle = U^\dagger(t,0)|q\rangle$ is the eigenvector of the Heisenberg operator $\hat{q}(t)$ with eigenvalue q. Substituting $q(t) = q'(-t)$ we see that the right-hand side of (10.34) turns into

$$\exp\left(i\int_{-t_1}^{-t_2} dt\, L[q',\dot{q}']\right) \propto \langle q_2,-t_2|q_1,-t_1\rangle. \tag{10.35}$$

The consequence of T-invariance is evidently

$$\langle q_1,t_1|q_2,t_2\rangle = \langle q_2,-t_2|q_1,-t_1\rangle. \tag{10.36}$$

The interchange of initial and final states is intuitively appealing.

Next we introduce U_T by

$$U_T\,|q,t\rangle = |q,-t\rangle, \quad U_T\, e^{iHt}\,|q\rangle = e^{-iHt}\,|q\rangle. \tag{10.37}$$

We want it to represent the property (10.36), i.e.

$$\langle q_1,t_1|q_2,t_2\rangle = (\langle q_2,t_2|U_T^\dagger)\,(U_T\,|q_1,t_1\rangle). \tag{10.38}$$

Using $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$ on the right-hand side this gives the relation

$$\langle q_1,t_1|q_2,t_2\rangle = [(\langle q_1,t_1|U_T^\dagger)\,(U_T\,|q_2,t_2\rangle)]^*. \tag{10.39}$$

Note the parenthesis, which are needed here in the bra-ket notation. Usually we are dealing with linear operators, for which it does not matter if we interpret the action of an operator in the bra-ket notation to act to the right or the left: $\langle\psi|\,(O|\phi\rangle) = (\langle\psi|O)\,|\phi\rangle$. However, this cannot be true for U_T. We shall now show that it has to be antiunitary.

First, choosing $t_1 = t_2 = 0$ the amplitude $\langle q_1,0|q_2,0\rangle = \langle q_1|q_2\rangle$ is real. For any state $|\psi\rangle = \int dq\,\psi(q)|q\rangle$, the norm $\langle\psi|\psi\rangle^* = \langle\psi|\psi\rangle$ is real, and we get $(\langle\psi|U_T^\dagger)(U_T|\psi\rangle) = \langle\psi|(U_T^\dagger U_T|\psi\rangle) = \langle\psi|(U_T^\dagger U_T)|\psi\rangle$, where it does not matter if U_T is linear or antilinear. Since $|\psi\rangle$ is arbitrary this means that U_T is unitary,

$$U_T^\dagger\, U_T = 1. \tag{10.40}$$

Note that the product $U_T^\dagger U_T$ is linear. Secondly, for general $t_{1,2}$, the amplitude (10.39) is complex, and letting $U_T^\dagger$ act to the right and using $U_T^\dagger\, U_T = 1$ has to produce a complex conjugate, such that it equals the left hand side of (10.39). So U_T should satisfy

$$(\langle\psi|U_T)\,|\phi\rangle = [\langle\psi|\,(U_T\,|\phi\rangle)]^*, \tag{10.41}$$

for arbitrary ψ and ϕ (from which $(\langle\phi|U_T^\dagger)\,|\psi\rangle = [\langle\phi|\,(U_T^\dagger\,|\psi\rangle)]^*$ follows by conjugation). From (10.41) follows that U_T is antilinear: let λ be a complex number

$$\langle\psi|(U_T\,\lambda|\phi\rangle) = [(\langle\psi|U_T)\,\lambda|\phi\rangle]^* = \lambda^*[(\langle\psi|U_T)\,|\phi\rangle]^* = \lambda^*\,\langle\psi|(U_T\,|\phi\rangle), \tag{10.42}$$

so

$$U_T\,\lambda|\phi\rangle = \lambda^* U_T|\phi\rangle, \quad U_T\,\lambda = \lambda^*\,U_T. \tag{10.43}$$

For $t = 0$ we see from (10.37) that U_T acts as the identity on the coordinate-basis $|q\rangle$,

$$U_T\,|q\rangle = |q\rangle. \tag{10.44}$$

So on this basis it is just the operator of complex conjugation. It follows that in the momentum basis

$$U_T\,|p\rangle = U_T \int dq\, e^{ipq}\,|q\rangle = \int dq\, e^{-ipq}\,|q\rangle = |-p\rangle. \tag{10.45}$$

For general t the relation (10.37) and the unitarity of U_T gives

$$U_T\,\exp[iHt]\,|q\rangle = \exp[-it\,U_T\,HU_T^\dagger]\,|q\rangle = \exp[-iHt]\,|q\rangle, \tag{10.46}$$

so the hamiltonian is invariant,

$$U_T\,HU_T^\dagger = H, \quad [U_T\,, H] = 0. \tag{10.47}$$

For the canonical Heisenberg operators we get

$$U_T\,q(t)U_T^\dagger = q(-t), \quad U_T\,p(t)U_T^\dagger = -p(-t). \tag{10.48}$$

Having familiarized ourselves with the antilinear nature of U_T we turn to QED, using operator notation. For Majorana fields with a Majorana representation of the Dirac matrices it is natural to define the T transformation as

$$U_{\tilde{T}}\,\psi(x)U_{\tilde{T}}^\dagger = D_T\psi(\ell_T x), \quad U_{\tilde{T}}\,\psi^T(x)\beta U_{\tilde{T}}^\dagger = \psi^T D_T^T\beta^* = \psi^T\beta D_T^{-1}. \tag{10.49}$$

For the complex Dirac field ψ build out of two Majorana fields $\psi_{1,2}$ according to $\psi = (\psi_1 - i\psi_2)/\sqrt{2}$, this $U_{\tilde{T}}$ would imply interchanging ψ and $\psi^\dagger$, as in charge conjugation. This looks intruiging, reminding us of the idea that charged particles running backwards in time are just antiparticles running forward (cf. Feynman). However, the conventional definition of T for Dirac fields does not interchange ψ and $\psi^\dagger$, so let us undo this by following the transformation (10.49) by charge conjugation. Thus defining T we get

$$U_T\,\psi(x)U_T^\dagger = C^\dagger\beta D_T\psi(\ell_T x), \quad U_T\,\bar{\psi}(x)U_T^\dagger = \bar{\psi}(\ell_T x)D_T^{-1}\beta C \tag{10.50}$$

$$U_T\,A^\mu(x)U_T^\dagger = -\ell^\mu_{T\nu}\,A^\nu, \tag{10.51}$$

where we also included the transformation for the electromagnetic field. To see what this does, consider the transformation of the current (suppressing the anti-symmetrization as in (10.14)),

$$\begin{aligned} U_T\,j^\mu(x)U_T^\dagger &= U_T\,\bar{\psi}(x)i\gamma^\mu\psi(x)U_T^\dagger = \bar{\psi}(\ell_T x)D_T^{-1}\beta C\,(-i\gamma^{\mu *})\,C^\dagger\beta D_T\psi(\ell_T x) \\ &= -\ell^\mu_{T\nu}\,j^\nu(\ell_T x), \end{aligned} \tag{10.52}$$

where we used (10.3). With this new definition of T the charge does not flip sign, but the current gets inverted: $j^0(\mathbf{x},t) \to j^0(\mathbf{x},-t)$, $\mathbf{j}(\mathbf{x},t) \to -\mathbf{j}(\mathbf{x},-t)$. The total QED hamiltonian (10.12) is T-invariant, $U_T H(t) U_T^\dagger = H(-t) = H(t)$, because it is time-independent.

Miscellaneous remarks:

- U_T changes the signs of both momentum and angular-momentum operators.

- Eigenvalues of U_T have no special significance, their phase can be changed by a phase transformation of the state vector. However, U_T^2 is linear and its eigenvalues have physical significance. For a fermion state $|p,\lambda\rangle$ it is -1.

- U_T changes in-states into out-states (cf. chapter 9),

$$U_T|p_1,\lambda_1,\cdots,p_n,\lambda_n\ \text{in}\rangle \;\propto\; |\tilde{p}_1,\tilde{\lambda}_1,\cdots,\tilde{p}_n,\tilde{\lambda}_n\ \text{out}\rangle, \qquad (10.53)$$

$$U_T|p_1,\lambda_1,\cdots,p_n,\lambda_n\ \text{out}\rangle \;\propto\; |\tilde{p}_1,\tilde{\lambda}_1,\cdots,\tilde{p}_n,\tilde{\lambda}_n\ \text{in}\rangle, \qquad (10.54)$$

  where $\tilde{p} = (p^0, -\mathbf{p})$, and the constant of proportionality is a phase factor.

- The experimental detection of T-violation is difficult, because the product CPT seems to be an exact symmetry of Nature (see next section) and CP-violation turns out to be effectively very small.

10.4 CPT

The CPT theorem states that in a local, Lorentz invariant and unitary (hermitian hamiltonian) theory, all interactions are invariant under the product transformation CPT. This is just our original $U_{\tilde{T}}$ for Majorana fields in (10.49) times U_P. We shall demonstrate CPT invariance here by showing that any lagrange density $\mathcal{L}$ satisfying the assumptions, keeping for simplicity only scalar, spinor and vector fields, transforms as $\mathcal{L}(x) \to \mathcal{L}(-x)$. This then leads to a CPT invariant hamiltonian.

Let us use Majorana language (Majorana fermion fields and a Majorana representation for the Dirac matrices, and real bosonic fields), which is most general, as any Dirac structure can be translated into this form. It also helps avoiding overlooking less familiar possibilities, such as so-called Majorana interactions which do not allow any charge conservation. Consider a theory with K real Majorana fields ψ_α, L real scalar fields, ϕ_k, and M real vector fields A_r^μ, $\alpha = 1,\ldots,K$, $k = 1,\ldots,L$, $r = 1,\ldots,M$. Under $U_\theta \equiv U_P\, U_{\tilde{T}}$ these fields transform simply as

$$U_\theta\, \phi_k(x) U_\theta^\dagger = \phi_k(-x), \quad U_\theta\, A_r^\mu(x) U_\theta^\dagger = -A_r^\mu(-x), \quad U_\theta\, \psi_\alpha(x) U_\theta^\dagger = i\gamma_5 \psi_\alpha(-x), \qquad (10.55)$$

where we used $\ell_T \ell_P = -1$, $D_T D_P = i\gamma_5$ and did not introduce any other phase (sign) factors. For the fermion bilinears it follows that

$$U_\theta\, \psi^T_\alpha(x)\beta\gamma^{\mu_1}\cdots\gamma^{\mu_n}\psi_{\alpha'}(x)U^\dagger_\theta = (-1)^n\, \psi^T_\alpha(-x)\beta\gamma^{\mu_1}\cdots\gamma^{\mu_n}\psi_{\alpha'}(-x). \tag{10.56}$$

Hence, these transform just like tensors build out of the Bose fields, e.g. $U_\theta\, \partial_\nu A^\mu_r(x)U^\dagger_\theta = (-1)^2[\partial/\partial(-x^\nu)]A^\mu_r(-x)$. The fermion bilinears are to be antisymmetrized, $\psi_{a\alpha}\psi_{b\beta} \to (\psi_{a\alpha}\psi_{b\beta} - \psi_{b\beta}\psi_{a\alpha})/2$, if necessary, such that the field operators can be freely anticommuted as if they were Grassmann numbers. Then they are hermitian, and either symmetric or antisymmetric under exchange of the labels α and α'. Corresponding coupling constants (or mass matrices) have to have the same symmetry. They have to be real in order that $\mathcal{L}$ be hermitian. Hence, they are not affected by the antilinearity of U_θ. It now follows that any Lorentz invariant multinomial in the fields with real coupling constants is CPT-covariant. For example in

$$\begin{aligned}\mathcal{L}_1 \;=\; & g_{\alpha\alpha' r}\frac{1}{4}\left(\psi^T_\alpha\beta\gamma^\mu\psi_{\alpha'} - \psi^T_{\alpha'}\gamma^{\mu T}\beta^T\psi_\alpha\right)A_{r\mu} + g'_{\alpha\alpha' r}\frac{1}{2}\psi^T_\alpha\beta\gamma^\mu i\gamma_5\psi_{\alpha'}A_{r\mu} \\ & + h_{\alpha\alpha' k}\frac{1}{2}\psi^T_\alpha\beta\psi_{\alpha'}\phi_k + h'_{\alpha\alpha' k}\frac{1}{2}\psi^T_\alpha\beta i\gamma_5\psi_{\alpha'}\phi_k.\end{aligned} \tag{10.57}$$

the couplings have to satisfy

$$g_{\alpha\alpha' r} = -g_{\alpha'\alpha r} = g^*_{\alpha\alpha' r}, \qquad g'_{\alpha\alpha' r} = g'_{\alpha'\alpha r} = g'^*_{\alpha\alpha' r}, \tag{10.58}$$

$$h_{\alpha\alpha' k} = h_{\alpha'\alpha k} = h^*_{\alpha\alpha' k}, \qquad h'_{\alpha\alpha' k} = h'_{\alpha'\alpha k} = h'^*_{\alpha\alpha' k}, \tag{10.59}$$

and $\mathcal{L}_1$ is hermitian and covariant under CPT, $U_\theta\, \mathcal{L}_1(x)U^\dagger_\theta = \mathcal{L}_1(-x)$. Note that terms involving the epsilon tensor $\epsilon_{\kappa\lambda\mu\nu}$ are also invariant, because $\det(\ell_P\ell_T) = 1$, similarly the presence of $\gamma_5 = i\epsilon_{\kappa\lambda\mu\nu}\gamma^\kappa\gamma^\lambda\gamma^\mu\gamma^\nu/4!$ does not affect CPT invariance.

Finally, we note that the somewhat mysterious presence of C in the usual treatments of the CPT transformation, which is essentially just $x \to \ell_P\ell_T x = -x$, has disappeared in the Majorana formulation used here: there is no trace of C in the natural $U_\theta = U_P\, U_{\tilde{T}}$.

An important consequence of CPT-invariance is:

- particles and antiparticles have the same mass and lifetime, even when C, P and T are all violated (as they are in Nature).

10.5 Problems

1. *Decay of the neutral pion*

   The neutral pion decays primarily electromagnetically into two photons: $\pi^0(p) \to \gamma(k,\lambda) + \gamma(k',\lambda')$. This process can be described in terms of an effective action, the form of which depends on the parity of π^0. In the following we assume not to know the parity of π^0 and intend to investigate how it might be determined. Candidate gauge invariant interactions are given by

$$S_1 = \frac{\alpha C}{m} \int d^4x\, \phi(x) \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu}(x) F_{\rho\sigma}(x), \tag{10.60}$$

and

$$S_2 = \frac{\alpha C}{m} \int d^4x\, \phi(x) F^{\mu\nu}(x) F_{\mu\nu}(x). \tag{10.61}$$

Here $\alpha = e^2/4\pi$ is the fine structure constant, m the π^0 mass, C a constant and $\epsilon^{\mu\nu\rho\sigma}$ the Levi-Civita tensor with $\epsilon_{0123} = 1$.

a. How should ϕ transform for S_1 to be invariant under C, P, and T?

b. How should ϕ transform for S_2 to be invariant under C, P, and T?

c. Compare your results for **a**,**b** with the transformation properties of the composite fields $\bar\psi\psi$ and $\bar\psi i\gamma_5\psi$.

d. Calculate the vertex function

$$S_{\mu\mu'}(u, v, v') \equiv \frac{\delta^3 S}{\delta\phi(u)\delta A^\mu(v)\delta A^{\mu'}(v')}, \tag{10.62}$$

for S_1 as well as for S_2.

e. Determine the corresponding vertex functions in momentum space, $S_{\mu\mu'}(p, k, k')$.

f. Draw the Feyman diagrams for the decay process $\pi^0(p) \to \gamma(k, \lambda) + \gamma(k', \lambda')$ in the tree-graph approximation.

g. Write down the corresponding decay amplitude $T(k\lambda, k'\lambda'; p)$ for both cases 1 and 2.

h. Suppose you can observe correlations between the photon polarizations and their momenta. Discuss how you might determine the pion parity from the decay distribution of the photons.

i. Calculate $\overline{|T|^2}$ in both cases. For 1 the identity $\epsilon^{\alpha\beta\mu\nu}\epsilon_{\alpha\beta\rho\sigma} = -2(\delta^\mu_\rho \delta^\nu_\sigma - \delta^\mu_\sigma \delta^\nu_\rho)$ is useful.

j. Calculate the unpolarized distribution $d\Gamma/d\Omega$ and the total rate Γ in the pion rest frame.

2. *The pion decay constant f_π*

The positively charged pion decays primarily through the process $\pi^+ \to \mu^+ + \nu_\mu$. Its decay amplitude depends on the matrix element of the current

$$j_\mu = \bar\psi_d i\gamma_\mu \psi_u - \bar\psi_d i\gamma_\mu\gamma_5\psi_u \equiv V_\mu - A_\mu \tag{10.63}$$

between the vacuum state and the one-pion state in QCD; V_μ is called the (polar-)vector current, A_μ the axial-vector current.

a. Verify that the following parametrization of these matrix elements has the general form compatible with Poincaré invariance:

$$\begin{aligned} \langle 0|A^\mu(x)|p,\pi^+\rangle &= -ip^\mu f_\pi \sqrt{2}\, e^{ipx}, & (10.64)\\ \langle 0|V^\mu(x)|p,\pi^+\rangle &= -ip^\mu \tilde{f}_\pi \sqrt{2}\, e^{ipx}. & (10.65) \end{aligned}$$

Hint: use the Poincaré in variance of the vacuum, (9.12) for the scalar case and (9.39) for the vector case; insert $U(\ell)^\dagger U(\ell) = 1$ into the matrix elements, and similar for translations.

Verify that the constants f_π and $\tilde{f}_\pi$ have the dimension of mass.

b. Parity is a good symmetry in QCD, which implies $U_P|0\rangle = |0\rangle$. The pions are known to be pseudo-scalars, i.e.

$$U_P|p,\pi^+\rangle = -|\ell_P p,\pi^+\rangle. \quad (10.66)$$

Verify that because of parity, $\tilde{f}_\pi = -\tilde{f}_\pi$, hence $\tilde{f}_\pi = 0$, but that there is no such constraint on f_π.

Hint: insert $U_P^\dagger U_P = 1$ into the matrix elements.

c. Charge conjugation is a good symmetry in QCD. Use C to obtain the relations for π^- with the corresponding currents.