Quantum Field Theory

Wintersemester 2009/10
Sommersemester 2010

Klaus Fredenhagen

II. Institut für Theoretische Physik
Universität Hamburg

Contents

CHAPTER

Introduction

Quantum mechanics is a consistent theory which describes well a large part of physics, in particular atomic and molecular physics. It neglects, however, the effects of special relativity and ignores the quantum nature of force fields, especially that of the electromagnetic field. Moreover, particle number plays a distinguished role in quantum mechanics, such that multiparticle processes and processes with creation and annihilation of particles are not easily described. Each of these points of view leads to an extension of the frame of quantum mechanics which is named quantum field theory. A common, but somewhat misleading equivalent name is "second quantization".

The concept of quantum field theory was developed immediately after the invention of quantum mechanics during the years 1925-1929 by Heisenberg, Jordan, Pauli and Dirac. By first order perturbation theory they were able to explain spontaneous emission of electromagnetic radiation and the Compton effect. For a long time, however, all attempts failed to improve the results by using higher orders of perturbation theory since the computed corrections turned out to be infinite.

The definition of a perturbation series was finally achieved by Tomonaga, Schwinger, Feynman and Dyson in 1947-49 by the method of renormalization. The basic idea was that the observed masses and charges do not coincide with the parameters used in the formulation of the theory (the so-called bare masses, charges etc.). These parameters have to be fixed afterwards ("renormalized") such that the observed parameters obtain the measured values. In the case of electrodynamics the accordingly computed corrections agreed with remarkable precision with the measured values. An impressive example is the magnetic moment of the electron. If μ_0 denotes Bohr's magneton, then one finds for the magnetic moment of the electron the theoretical value

$$(\frac{\mu}{\mu_0})_{\text{theor}} = 1,001159652460(127)(75) \tag{0.1}$$

which has to be compared with the experimental value

$$(\frac{\mu}{\mu_0})_{\text{exper}} = 1,001159652200(40). \tag{0.2}$$

where the error of the theoretical value consists in the uncertainty of the finestructure constant (127) and the numerical error of the calculation

of coefficients of the perturbation series (cf. Nachtmann, Elementary Particles Theory.)

In spite of these great successes it is still unknown whether a consistent theory of quantum electrodynamics exists. Actually, there are plausible arguments which suggest that the formulation of quantum electrodynamics on which the pertubative calculations are based does not correspond to a consistent theory.

After the (partial) success of quantum electrodynamics one tried to describe also the other interactions between elementary particles by quantum field theories. The quanta of the corresponding fields should, analogously to photons, be visible as particles. This consideration led Yukawa to the prediction of mesons as the quanta of nuclear forces, a prediction partially fulfilled by the discovery of the pion. More recently, in a similar way the discovery of W- and Z-bosons as transmitter of weak forces was an impressive justification of this concept.

Presently we have a generally accepted model for the theory of elementary particles, the standard model. It is a combination of the theory of strong interaction called Quantum Chromodynamics (QCD) with the Weinberg-Salam model as the theory of electroweak interactions. There are at present only few facts known which are not in agreement with the standard model: the nonzero mass of neutrinos visible in neutrino oscillations and the presence of large amounts of dark matter in the universe. One hopes that the new accelerator at CERN, the large hadron collider LHC will find other effects so that one may distinguish between the many possible extensions of the standard model.

In spite of its tremendous successes quantum field theory still does not have the status of a consistent theory. The comparison with experiments typically relies on heuristic considerations whose reliability is difficult to judge. An exception are some structural results like the connection between spin and statisics and the PCT theorem which can be derived solely from fundamental properties of the theory.

The plan of these lectures is the following: We will first discuss the multiparticle formalism of quantum mechanics ("2nd quantization"). We will see that this formalism can be interpreted as the quantization of a field theory with the Schrödinger equation as field equation.

We will then treat relativistic single particle systems. This amounts to analyze the unitary representations of the Poincaré group. As a byproduct we will find the relativistic wave equations.

In the 3rd chapter noninteracting relativistic multiparticle systems will be constructed, and it will be shown that they correspond to quantum field theories. Alternatively, we will start from a classical field theory (electrodynamics) and associate to it a quantum field theory, and we will show that this theory describes particles (the photons). Formally, interactions can be introduced and the calculation of cross

sections can be combinatorially described via Feynman diagrams. But only the so called tree graphs give rise to immediately meaningful expressions.

In the 4th chapter interactions will be analyzed in more detail. Scattering theory in terms of LSZ relations and the Haag Ruelle theory will be treated. We will give a thorough definition of the S-matrix, discuss the path integral and define important combinatorial concepts as for instance connected functions and vertex functions.

In the 5th chapter we will treat renormalization. We will discuss several methods, in particular the causal perturbation theory of Epstein and Glaser and the method of flow equations of Polchinsky in the spirit of Wilson, and we will introduce the so-called renormalization group.

The 6th chapter finally is devoted to nonabelian gauge theories.

CHAPTER I

Multiparticle systems in quantum mechanics

1. The n-particle space

According to quantum mechanics a single particle system is described by a wave function $\Phi(\mathbf{x}, m)$ with $\mathbf{x} \in \mathbb{R}^3$ and $m = -s, -s+1, ...s$, where $s = 0, \frac{1}{2}, 1 \ldots$ is the spin of the particle. The wave function has to be square integrable, i.e. the normalization integral

$$||\Phi||^2 := \sum_{m=-s}^{s} \int d^3\mathbf{x} |\Phi(\mathbf{x}, m)|^2 \tag{1.1}$$

has to be finite. These wave functions constitute the single particle Hilbert space $\mathfrak{H}_1$. In order to describe states with 2 particles, one has to couple the single particle systems in an appropriate way. The simplest, but by no means only possibility is the tensor product of the single particle spaces,

$$\mathfrak{H}_2 := \mathfrak{H}_1 \otimes \mathfrak{H}_1 \ \ni \Phi, \Phi = \Phi(\mathbf{x_1}, m_1, \mathbf{x_2}, m_2). \tag{1.2}$$

The wave functions of a 2-particle system can no longer be interpreted as waves in position space. The n-particle wave functions are square integrable functions Φ of n variables $(\mathbf{x_i}, m_i)$, $i = 1, ...n$,

$$\Phi \in \mathfrak{H}_n = \mathfrak{H}_1 \otimes \ldots \otimes \mathfrak{H}_1 \ . \tag{1.3}$$

Observable of the n-particle system are e.g.

- the position operator of the i-th particle

$$(\mathbf{X_i}\Phi_n)(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n) = \mathbf{x_i}\Phi(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n)$$

- the momentum operator of the i-th particle

$$(\mathbf{P_i}\Phi_n)(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n) = \frac{1}{i}\partial_{\mathbf{i}}\Phi(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n)$$

- the kinetic energy of the i-th particle (with mass M_i)

$$(T_i\Phi_n)(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n) = -\frac{1}{2M_i}\Delta_i\Phi_n(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n)$$

- the 3-component of the spin of the i-th particle

$$S_i^{(3)}\Phi_n(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n) = m_i\Phi_n(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n)$$

- the total kinetical energy $T = \sum_i T_i$,
- the total momentum $\mathbf{P} = \sum_i \mathbf{P}_i$,

- and the translation operator $U(\mathbf{x}) = e^{i\mathbf{P}\cdot\mathbf{x}}$,

$$(U(\mathbf{x})\Phi_n)(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n) = \Phi_n(\mathbf{x_1} - \mathbf{x}, m_1, \ldots, \mathbf{x_n} - \mathbf{x}, m_n).$$

Interactions are in typical cases functions $V(\mathbf{X_1}, \ldots, \mathbf{X_n})$ of the position operators. If e.g. $V_{ij}(\mathbf{x_i} - \mathbf{x_j})$ is the interaction potential between the particle i and j, one obtains

$$V = \sum_{i<j} V_{ij}(\mathbf{X_i} - \mathbf{X_j}). \tag{1.4}$$

The Hamiltonian is then given by

$$H = T + V \quad . \tag{1.5}$$

Up to now we assumed that the particles are distinguishable. When the particles are indistinguishable, their numbering is arbitrary. Hence the wave function Φ_n^σ,

$$\Phi_n^\sigma(\mathbf{x_1}, m_1, \ldots, \mathbf{x_n}, m_n) = \Phi_n(\mathbf{x}_{\sigma(\mathbf{1})}, m_{\sigma(1)}, \ldots, \mathbf{x}_{\sigma(\mathbf{n})}, m_{\sigma(1)}) \tag{1.6}$$

describes for any permutation σ the same state as Φ_n. One finds often the argument that therefore $\Phi_n^\sigma = \lambda(\sigma)\Phi_n$ must hold, for a 1-dimensional representation λ of the permutation group S_n. The 1-dimensional representations of the S_n are $\lambda(\sigma) = 1$ (totally symmetric representation) and $\lambda(\sigma) = \text{sign}(\sigma)$ (totally antisymmetric representation). In the first case the wave function is symmetric (Bose statistics), in the second case antisymmetric (Fermi statistics) under permutation of 2 arguments.

Indeed, all known particle have either Bose- or Fermi statistics. But one often considers models where the wave functions have a more complicated behaviour under permutations. In atomic physics, e.g., one often neglects the spin; the position space wave functions then are not necessarily antisymmetric.

Where is the error in the argument? The statement, that Φ_n and Φ_n^σ represent only then the same state, if they differ by a factor with absolut value 1, would be correct, if all operators in $\mathfrak{H}_n$ would be physical observables. But exactly this is not true, when the particles are indistinguishable. The position of the i-th particle, e.g., is no longer an observable. Only those quantities are measurable which do not depend on the numbering of the particles.

We define now a unitary representation of the permutation group on $\mathfrak{H}_n$ by

$$U(\sigma)\Phi = \Phi^\sigma. \tag{1.7}$$

It holds

$$A \text{ observable} \Rightarrow [A, U(\sigma)] = 0 \ \forall \sigma \in S_n. \tag{1.8}$$

We now can decompose the Hilbert space $\mathfrak{H}_n$ in permutation invariant subspaces. Minimal invariant subspaces correspond to irreducible representations of the permutation group. We have the important theorem:

THEOREM I.1. *States which belong to inequivalent irreducible representations of the permutation group cannot coherently be superimposed.*

PROOF. For the proof we consider two minimal invariant subspaces $\mathfrak{H}^a$ and $\mathfrak{H}^b$ with corresponding orthogonal projections E^a, E^b, respectively. According to our assumption these projections as well as all observables A commute with the permutation operators $U(\sigma)$. Thus also E^aAE^b commutes with $U(\sigma)$. When the representations of the permutation group on the 2 spaces are inequivalent, E^aAE^b has to vanish according to Schur's Lemma. Let now $\Phi = \Phi^a + \Phi^b$ with $\Phi^i \in \mathfrak{H}^i$, $i = a, b$ and $\|\Phi\| = 1$. The expectation value of an observable A is

$$\big\langle \Phi, A\Phi \big\rangle = \big\langle \Phi^a, A\Phi^a \big\rangle + \big\langle \Phi^b, A\Phi^b \big\rangle, \tag{1.9}$$

since the mixed terms vanish,

$$\big\langle \Phi^a, A\Phi^b \big\rangle = \big\langle \Phi^a, E^aAE^b\Phi^b \big\rangle = 0. \tag{1.10}$$

The state described by the wave function Φ is thus a mixture of the states Φ^a and Φ^b, i.e. the phase between the components cannot be observed. □

The fact, that states can not always be coherently superimposed, was first observed by Wick, Wightman and Wigner; they named this phenomenon a superselection rule. Sets of states, in which the superposition principle holds without restrictions, are called superselection sectors. Whereas in quantum mechanics with finitely many particles the superselection rules play no important role, they will be important within quantum field theory.

The known elementary particles are either bosons or fermions, according to whether their spin is integer or half-integer. This can be derived within the formalism of relativistic quantum field theory, if the possibility of unobservable inner degrees of freedom is taken into account. But the argument is only valid for theories in a d-dimensional spacetime with $d \geq 4$. In 2- and 3-dimensional theories other cases are possible (anyons, plektons); this possibility may be of importance for the understanding of quasi-1 and 2-dimensional systems (quantum Hall effect, high temperature superconductivity).

2. The bosonic Fock space

If one wants to describe processes, in which the particle number changes, one has to combine the different n-particle spaces. For the case, that the particles are indistinguishable and have Bose statistics, the n-particle space $\mathfrak{H}_n^+$ is the space of symmetric wave functions. One defines now the so-called bosonic Fock space $\mathfrak{H}^+$ over the single particle space $\mathfrak{H}_1$ by forming the direct sum of Hilbert spaces

$$\mathfrak{H}^+ = \bigoplus_{n=0}^{\infty} \mathfrak{H}_n^+ \tag{2.1}$$

with $\mathfrak{H}_0^+ = \mathbb{C}$ and $\mathfrak{H}_1^+ = \mathfrak{H}_1$. Elements $\Phi \in \mathfrak{H}^+$ are sequences

$$\Phi = (\Phi_0, \Phi_1, \ldots) \equiv (\Phi_n)_{n\in\mathbb{N}_0} \tag{2.2}$$

with $\Phi_n \in \mathfrak{H}_n^+$ and

$$||\Phi||^2 = \sum_{n=0}^{\infty} ||\Phi_n||^2 < \infty. \tag{2.3}$$

Their scalar product is given by

$$\big\langle (\Phi_n), (\Psi_n) \big\rangle = \sum_{n=0}^{\infty} \big\langle \Phi_n, \Psi_n \big\rangle \ . \tag{2.4}$$

The vector

$$\Omega = (1, 0, \ldots) \tag{2.5}$$

describes a state with no particle and is called the vacuum vector (often denoted by $|0\rangle$). A unit vector Φ describes a state with not necessarily sharp particle number; $||\Phi_n||^2$ is the probability, to find precisely n particles in this state. Typical observables in $\mathfrak{H}^+$ are

- the particle number operator

$$(N\Phi)_n = n\Phi_n \tag{2.6}$$

with expectation value

$$\langle N \rangle = (\Phi, N\Phi) = \sum_{n=0}^{\infty} n||\Phi_n||^2 \tag{2.7}$$

($\langle N \rangle = \infty$ is possible) and variance

$$\Delta(N)^2 = (\Phi, (N - \langle N \rangle)^2 \Phi) = \sum_{n=0}^{\infty} (n - \langle N \rangle)^2 ||\Phi_n||^2. \tag{2.8}$$

Operators, which map each n-particle space into itself, commute with the particle number operator. Examples are

- the momentum:$(\mathbf{P}\Phi)_n = \mathbf{P}^{(\mathbf{n})}\Phi_n$
- the kinetic energy:$(T\Phi)_n = T^{(n)}\Phi_n$
- the potential energy:$(V\Phi)_n = V^{(n)}\Phi_n$.

The great advantage of the Fock space is that it admits a simple description of the transition between different particle numbers. The procedure may be illustrated by the following example. We may imagine a particle without any degree of freedom, i.e. $\mathfrak{H}_1 = \mathbb{C}$. Then $\mathfrak{H}_n^+ = \mathbb{C}\forall n$ and $\mathfrak{H}^+$ is the space of complex valued square summable sequences l_2. An operator which changes the particle number may be defined by

$$(a\Phi)_n = \sqrt{n+1}\Phi_{n+1} \ . \tag{2.9}$$

It holds

$$[a, N] = a, \tag{2.10}$$

i.e. a lowers the particle number by 1. One calls a the annihilation operator. The adjoint operator a^*, called the creation operator, is characterized by the equation

$$\left\langle \Psi, a^*\Phi \right\rangle = \left\langle a\Psi, \Phi \right\rangle \ . \tag{2.11}$$

One calculates

$$\begin{aligned} \sum_{n=0}^{\infty} \overline{\Psi_n}(a^*\Phi)_n &= \sum_{n=0}^{\infty} (\overline{a\Psi})_n \ \Phi_n \\ = \sum_{n=0}^{\infty} \sqrt{n+1}\ \overline{\Psi_{n+1}}\Phi_n &= \sum_{n=1}^{\infty} \sqrt{n}\ \overline{\Psi_n}\Phi_{n-1}, \end{aligned} \tag{2.12}$$

i.e. $(a^*\Phi)_0 = 0$ and $(a^*\Phi)_n = \sqrt{n}\Phi_{n-1}, n > 0$. It follows

$$a^*a = N, aa^* = N + 1 \quad \text{and} \quad [a, a^*] = 1. \tag{2.13}$$

These relations are known from the treatment of the harmonic oscillator in quantum mechanics. Let

$$H = -\frac{1}{2}\frac{d^2}{dx^2} + \frac{\omega^2}{2}x^2 \tag{2.14}$$

be the Hamiltonian of the harmonic oscillator as an operator on $L^2(\mathbb{R})$. One sets

$$a = \sqrt{\frac{\omega}{2}}x + \sqrt{\frac{1}{2\omega}}\frac{\mathrm{d}}{\mathrm{d}x}, a^* = \sqrt{\frac{\omega}{2}}x - \sqrt{\frac{1}{2\omega}}\frac{\mathrm{d}}{\mathrm{d}x}, \tag{2.15}$$

thus $[a, a^*] = 1$ and

$$a^*a = \frac{\omega}{2}x^2 - \frac{1}{2\omega}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{1}{2} \ . \tag{2.16}$$

Hence $H = \omega(a^*a + \frac{1}{2})$. The n-th energy eigenfunction is, in the Fock space interpretation, the n-particle state. The ground state (the vacuum) can be characterized by

$$a\Omega = 0 \ . \tag{2.17}$$

The corresponding differential equation is

$$(\sqrt{\frac{\omega}{2}}x + \sqrt{\frac{1}{2\omega}}\frac{\mathrm{d}}{\mathrm{d}x})\Omega(x) = 0 \tag{2.18}$$

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}x}\ln\,\Omega(x) = \frac{\frac{\mathrm{d}}{\mathrm{d}x}\Omega(x)}{\Omega(x)} = -\omega x \tag{2.19}$$

with the solution

$$\Omega(x) = e^{-\frac{1}{2}\omega x^2}\mathrm{const.} \tag{2.20}$$

In order to obtain a corresponding definition in the Fock space over an infinite dimensional single particle space, we take into account, that every single particle wave function $f \in \mathfrak{H}_1$ represents one degree

of freedom of the particle. We define the annihilation operator for a particle with wave function f by

$$(a(f)\Phi)_n(\mathbf{x_1},\ldots,\mathbf{x_n}) = \sqrt{n+1}\int d^3\mathbf{x}\overline{f(\mathbf{x})}\Phi_{n+1}(\mathbf{x},\mathbf{x_1},\ldots,\mathbf{x_n}) \quad (2.21)$$

(for the sake of a simpler notation we restrict ourselves to particles with spin 0). The adjoint operator is given by

$$(a(f)^*\Phi)_0 = 0 \quad , \quad (2.22)$$

$$(a(f)^*\Phi)_n(\mathbf{x_1},\ldots,\mathbf{x_n}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} f(\mathbf{x_i})\Phi_{n-1}(\mathbf{x_1},\ldots,\mathbf{x_{i-1}},\mathbf{x_{i+1}},\ldots,\mathbf{x_n}). \quad (2.23)$$

The operators $a(f)$, $a(g)^*$ possess the following commutation relations

$$[a(f),a(g)^*] = \langle f,g\rangle = \int d^3\mathbf{x}\overline{f(\mathbf{x})}g(\mathbf{x}). \quad (2.24)$$

Moreover, we have $[a(f),a(g)] = 0 = [a(f)^*,a(g)^*]$. $a(f)$ annihilates a particle, $a(f)^*$ creates a particle,

$$a(f)N = (N+1)a(f) \quad , \quad a(f)^*N = (N-1)a(f)^* \quad . \quad (2.25)$$

The vacuum can be characterized by the equation

$$a(f)\Omega = 0 \quad \forall f \in \mathfrak{H}_1. \quad (2.26)$$

For $||f|| = 1$ the operator $a(f)^*a(f)$ has the eigenvalues 0,1,2... and the interpretation "number of particles with wave function f". If one chooses an orthonormal basis $(f_i)_{i\in\mathbb{N}}$ in $\mathfrak{H}_1$, then $a_i := a(f_i), i \in \mathbb{N}$ defines a system of independent harmonic oscillators with the commutation relations

$$[a_i,a_j] = 0 = [a_i^*,a_j^*] \quad (2.27)$$

$$[a_i,a_j^*] = \delta_{ij}. \quad (2.28)$$

We now want to investigate the time evolution of the operator $a(f)$. According to Heisenberg's equation we have

$$\frac{\mathrm{d}}{\mathrm{d}t}A(t) = i[H,A(t)] \quad (2.29)$$

for every operator A in $\mathfrak{H}$, if H is the Hamiltonian. Let us first consider the case, in which there are no forces between the particles, but where all particles are subject to an external potential v. The Hamiltonian is then $H = T + V$ with $(V\Phi)_n = V_n\Phi_n$, $V_n(\mathbf{X_1},\ldots\mathbf{X_n}) = \sum_{i=1}^n v(\mathbf{X_i})$. One calculates

$$[H,a(f)] = a(-H_1 f) \quad (2.30)$$

with the single particle Hamiltonian $H_1 = T_1 + V_1$. The solution of the Heisenberg equation (under taking into account, that a depends antilinearly on f) is

$$a(f)(t) = a(e^{iH_1 t}f) \ . \quad (2.31)$$

For each t, $f \mapsto a(f)(t)$ is an antilinear operator valued functional on $\mathfrak{H}_1$. We interpret it as an operator valued distribution and use the symbolic notation

$$a(f)(t) = \int d^3\mathbf{x}\overline{f(\mathbf{x})}a(\mathbf{x},t). \tag{2.32}$$

The simultaneous commutation relations are

$$[a(\mathbf{x},t), a^*(\mathbf{y},t)] = \delta(\mathbf{x}-\mathbf{y}) \quad , \tag{2.33}$$

$$[a(\mathbf{x},t), a(\mathbf{y},t)] = 0 = [a^*(\mathbf{x},t), a^*(\mathbf{y},t)] \tag{2.34}$$

in the sense of operator valued distributions. With the usual definition of derivatives of distributions one obtains the "quantized Schrödinger equation"

$$\boxed{i\frac{\partial}{\partial t}a(\mathbf{x},t) = (-\frac{1}{2M}\Delta + v(\mathbf{x}))a(\mathbf{x},t).} \tag{2.35}$$

The equation of motion of a multiparticle system without internal forces thus can be brought into the same form as the equation of motion of a single particle system, the only difference being that the complex valued wave function has to be replaced by an operator valued distribution. The fact that the Schrödinger equation ("1st quantization") is now interpreted as an operator equation, is the reason for the somewhat misleading name "2nd quantization" for the multiparticle formalism. A better interpretation of the quantized Schrödinger equation is the following one. The Schrödinger equation may be considered as a classical field equation for the matter field (this was the original interpretation of Schrödinger). The transition from a complex valued to an operator valued wave function is then the quantization of this classical field theory. We have thus obtained a first example for the coincidence of multiparticle theory and quantum field theory.

The operator valued distribution $a(\mathbf{x}) \equiv a(\mathbf{x},0)$ can be defined as an operator on vectors Φ with finite particle number (i.e. $\Phi_n \neq 0$ only for finitely many n) and wave functions from Schwartz' test function space $\mathcal{S}(\mathbb{R}^{3n})$ (the set of these vectors is a dense subspace $\mathfrak{D} \subset \mathfrak{H}$) by

$$(a(\mathbf{x})\Phi)_n(\mathbf{x_1},\ldots,\mathbf{x_n}) = \sqrt{n+1}\Phi_{n+1}(\mathbf{x},\mathbf{x_1},\ldots,\mathbf{x_n}) \ . \tag{2.36}$$

$a(\mathbf{x})$ has, however no densely defined adjoint and is therefore not closable. So-called normal ordered products,

$$a^*(\mathbf{x_1})\ldots a^*(\mathbf{x_n})a(\mathbf{y_k})\ldots a(\mathbf{y_1}), \tag{2.37}$$

where the creation operators are always on the left side from the annihilation operators, can be defined as sesquilinear forms on $\mathfrak{D}$:

$$\left\langle \Phi, a^*(\mathbf{x_1})\ldots a^*(\mathbf{x_n})a(\mathbf{y_k})\ldots a(\mathbf{y_1})\Psi \right\rangle \tag{2.38}$$

$$:= \left\langle a(\mathbf{x_n})\ldots a(\mathbf{x_1})\Phi, a(\mathbf{y_k})\ldots a(\mathbf{y_1})\Psi \right\rangle \ .$$

In this way one finds

$$\big\langle \Phi, H\Psi \big\rangle = \int d^3\mathbf{x} \big\langle \Phi, h(\mathbf{x})\Psi \big\rangle \tag{2.39}$$

with the energy density

$$h(\mathbf{x}) = \frac{1}{2M} \partial a^*(\mathbf{x}) \cdot \partial a(\mathbf{x}) + a^*(\mathbf{x}) v(\mathbf{x}) a(\mathbf{x}). \tag{2.40}$$

Analogously one obtains for the particle number

$$\big\langle \Phi, N\Psi \big\rangle = \int d^3\mathbf{x} \big\langle \Phi, n(\mathbf{x})\Psi \big\rangle \tag{2.41}$$

with the particle number density

$$n(\mathbf{x}) = a^*(\mathbf{x}) a(\mathbf{x}). \tag{2.42}$$

In the case of 2-particle forces with a potential V the interaction part of the Hamiltonian has the form

$$(H_I \Phi)_n = H_I^{(n)} \Phi_n, \;\; H_I^{(n)} = \sum_{i<j} V(\mathbf{x_i} - \mathbf{x_j}). \tag{2.43}$$

In terms of creation and annihilation operators it may be written as

$$H_I = \frac{1}{2} \int d^3\mathbf{x} d^3\mathbf{y} a^*(\mathbf{x}) a^*(\mathbf{y}) V(\mathbf{x}-\mathbf{y}) a(\mathbf{y}) a(\mathbf{x}). \tag{2.44}$$

For the field $a(\mathbf{x}, t)$ one obtains the nonlinear Schrödinger equation

$$\boxed{i\frac{\partial}{\partial t} a(\mathbf{x}, t) = -\frac{1}{2M} \Delta a(\mathbf{x}, t) + \int d^3\mathbf{y} V(\mathbf{x}-\mathbf{y}) a^*(\mathbf{y}, t) a(\mathbf{y}, t) a(\mathbf{x}, t).} \tag{2.45}$$

As a simple model for a source which emits and absorbs particles, we consider the interaction operator

$$H_I = a(f) + a(f)^* \quad , f \in \mathfrak{H}_1. \tag{2.46}$$

To simulate the effort for the creation of a particle, we add a multiple of the particle number operator to the Hamiltonian

$$H = T + \mu N + H_I \tag{2.47}$$

with the "chemical potential" $\mu > 0$.

In order to get an idea for the treatment of this Hamiltonian, we first look at an analogous problem for the harmonic oscillator under the action of a constant force,

$$H = \omega a^* a + \lambda(a + a^*) = \frac{\omega^2}{2} x^2 - \frac{1}{2} \frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{1}{2}\omega + \lambda\sqrt{2\omega} x \tag{2.48}$$

As it is well known, this force leads to a shift of the oscillating system by a constant length c with $c = \lambda\sqrt{2}\omega^{-\frac{3}{2}}$. It follows that

$$H = e^{ipc}\left(\frac{\omega^2}{2} x^2 - \frac{1}{2} \frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{1}{2}\omega\right) e^{-ipc} - \frac{\lambda^2}{\omega} \tag{2.49}$$

with $p = \frac{1}{i}\frac{\mathrm{d}}{\mathrm{d}x} = \frac{1}{i}\sqrt{\frac{\omega}{2}}(a - a^*)$. Up to the additive constant $-\frac{\lambda^2}{\omega}$ the perturbed Hamiltonian H can be transformed into the unperturbed Hamiltonian by the unitary operator $e^{ipc} = e^{\frac{\lambda}{\omega}(a-a^*)}$.

In an analogous way we search for the Hamiltonian

$$H = T + \mu N + H_I \quad , H_I = a(f) + a(f)^* \quad , f \in \mathfrak{H}_1 \tag{2.50}$$

some $g \in \mathfrak{H}_1$ with the property

$$e^{a(g)-a(g)^*}(T + \mu N)e^{-a(g)+a(g)^*} = H + \text{const.} \tag{2.51}$$

It holds

$$e^A B e^{-A} = \sum_{n=0}^{\infty} \frac{1}{n!}[A, \dots, [A, B] \dots]. \tag{2.52}$$

With $[a(g), T+\mu N] = a((T_1+\mu)g)$ and $[a(g)^*, T+\mu N] = -a((T_1+\mu)g)^*$ and

$$[a(g), a((T_1 + \mu)g)^*] = \big\langle g, (T_1 + \mu)g \big\rangle \tag{2.53}$$

one obtains

$$e^{a(g)-a(g)^*}(T + \mu N)e^{-a(g)+a(g)^*} \tag{2.54}$$

$$= T + \mu N + a((T_1 + \mu)g) + a((T_1 + \mu)g)^* + \big\langle g, (T_1 + \mu)g \big\rangle \tag{2.55}$$

Thus $g = (T_1 + \mu)^{-1} f$ solves the problem. In Fourier space we find

$$\tilde{g}(\mathbf{p}) = (\frac{|\mathbf{p}|^2}{2M} + \mu)^{-1} \tilde{f}(\mathbf{p}) \tag{2.56}$$

which may be transported to position space and yields

$$g(\mathbf{x}) = \frac{M}{2\pi} \int d^3\mathbf{y} \frac{e^{-\sqrt{2M\mu}|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x} - \mathbf{y}|} f(\mathbf{y}). \tag{2.57}$$

The ground state of H is thus

$$\Omega_g = e^{a(g)-a(g)^*}\Omega \tag{2.58}$$

with the ground state energy

$$E_0 = -\big\langle f, T_1+\mu)^{-1} f \big\rangle = -\int d^3\mathbf{p} \frac{|\tilde{f}(\mathbf{p})|^2}{\frac{|\mathbf{p}|^2}{2M} + \mu} = -\frac{M}{2\pi} \int \mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{y} f(\mathbf{x}) \frac{e^{-\sqrt{2M\mu}|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x} - \mathbf{y}|} f(\mathbf{y}). \tag{2.59}$$

Ω_g is a so-called coherent state. It has nonvanishing components for all particle numbers.

According to the Baker-Campbell-Hausdorff formula, for two $(n \times n)$-matrices A, B with $[[A, B], A] = 0 = [[A, B], B]$ the formula

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]} \tag{2.60}$$

holds. If we apply the same formula to $e^{a(g)-a(g)^*}$ where questions of convergence are, for the time being, ignored, one obtains

$$e^{a(g)-a(g)^*} = e^{-a(g)^*} e^{a(g)} e^{-\frac{1}{2}||g||^2} \tag{2.61}$$

and hence (because of $a(g)\Omega = 0$)

$$\Omega_g = e^{-\frac{1}{2}||g||^2} e^{-a(g)^*}\Omega \tag{2.62}$$

i.e.

$$\begin{aligned}(\Omega_g)_n(\mathbf{x_1},\ldots,\mathbf{x_n}) &= \frac{e^{-\frac{1}{2}||g||^2}}{n!}((-a(g)^*)^n\Omega)_n(\mathbf{x_1},\ldots,\mathbf{x_n}) \\ &= \frac{e^{-\frac{1}{2}||g||^2}}{\sqrt{n!}}(-1)^n g(\mathbf{x}_1)\ldots g(\mathbf{x}_n).\end{aligned} \tag{2.63}$$

The probability that Ω_g contains exactly n particles is

$$||(\Omega_g)_n||^2 = \frac{e^{-||g||^2}}{n!}||g||^{2n}. \tag{2.64}$$

The particle number is therefore distributed according to the Poisson distribution with mean particle number

$$\langle N\rangle = ||g||^2 = \int d^3\mathbf{p}\frac{|\tilde{f}(\mathbf{p})|^2}{(\frac{|\mathbf{p}|^2}{2M}+\mu)^2} \tag{2.65}$$

and with mean deviation $\Delta(N)^2 = \langle N\rangle$. In particular, the mean particle number tends for $\mu \to 0$ to ∞, if $\tilde{f}$ is continuous at 0 and does not vanish there. The ground state energy, however, remains finite in this limit.

A further remarkable property of coherent states is that they are eigenvectors of the annihilation operators,

$$a(\mathbf{x})e^{a(g)^*}\Omega = g(\mathbf{x})e^{a(g)^*}\Omega. \tag{2.66}$$

Hence the expectation value of a Hamiltonian in a coherent state is simply obtained by replacing $a(\mathbf{x})$ by $g(\mathbf{x})$ and $a(\mathbf{x})^*$ by $\overline{g(\mathbf{x})}$.

We now consider the limit $f(\mathbf{x}) \to c\delta(\mathbf{x}-\mathbf{y})$, where the particle can be created or annihilated only at the point $\mathbf{y}$. For H_I one finds in this limit

$$H_I = \bar{c}a(\mathbf{y}) + ca^*(\mathbf{y})\ . \tag{2.67}$$

A formal calculation of the ground state energy yields a divergent expression;

$$E_0 = -|c|^2\int d^3\mathbf{x}\delta(\mathbf{x}-\mathbf{y})\frac{M}{2\pi}\frac{e^{-\sqrt{2M\mu}|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}. \tag{2.68}$$

But $g(\mathbf{x}) = c\frac{M}{2\pi}\frac{e^{-\sqrt{2M\mu}|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}$ is a normalizable single particle wave function. Thus

$$H_{\text{ren}} := e^{a(g)-a(g)^*}(T+\mu N)e^{-a(g)+a(g)^*} \tag{2.69}$$

is a well defined selfadjoint operator with ground state Ω_g and with mean particle number

$$\langle N\rangle < \infty \tag{2.70}$$

(see exercise 6). In the Heisenberg equation we replace H by H_{ren}, since both term differ only by a (divergent) constant. The prize to pay is an arbitrary fixing of the ground state energy to 0.

The model described above delivers a simple description of nuclear forces by exchange of mesons (Nelson model). We consider a system of n nucleons, realized by wave functions

$$\Phi(\mathbf{x_1}, \ldots, \mathbf{x_n}) \ , \tag{2.71}$$

coupled to a system of spinless mesons. The Hilbert space of the coupled system is the tensor product of the nucleon space with the Fock space of the mesons. The Hamiltonian is the sum of the kinetic energy of the nucleons, the kinetic energy of the mesons including the term with the chemical potential $\mu = M/2$ (this choice of the chemical potential simulates the behaviour in the corresponding relativistic situation) and of the interaction term

$$H_I = c \sum_{i=1}^{n} (a(\mathbf{X}_i) + a^*(\mathbf{X}_i)), c \in \mathbb{R}, \tag{2.72}$$

where $\mathbf{X}_i$ is the position operator of the i-th nucleon. In the Born-Oppenheimer approximation one neglects in the first step the kinetic energy of the nucleons and obtains for every position space configuration of the nucleons a Hamiltonian of the previously considered form with $g = \sum g_i$. We now study the dependence of the ground state energy on the distribution of nucleons. We find

$$H_{\text{ren}} = e^{a(g)-a(g)^*}(T + \frac{1}{2}MN)e^{-a(g)+a(g)^*} + E_0(\mathbf{x_1}, \ldots, \mathbf{x_n}) \tag{2.73}$$

with

$$E_0(\mathbf{x_1}, \ldots, \mathbf{x_n}) = -\sum_{i<j} \frac{|c|^2 M}{\pi} \frac{e^{-M|\mathbf{x}_i - \mathbf{x}_j|}}{|\mathbf{x}_i - \mathbf{x}_j|} \tag{2.74}$$

Here the divergent diagonal terms were omitted.

The ground state energy of the meson system now yields the Yukawa potential for the nucleon interaction.

Let us now investigate the corresponding problem for a time dependent source,

$$H(t) = T + \mu N + H_{I,t} \quad , H_{I,t} = a(f_t) + a(f_t)^* \tag{2.75}$$

with $f_t(\mathbf{x}) = f(\mathbf{x}, t)$, $f \in \mathcal{S}(\mathbb{R}^4)$. The Schrödinger equation with time dependent Hamiltonian is solved by the time evolution operator $U(t, s)$ which is characterized by the following equations:

$$U(t, t) = 1 \tag{2.76}$$

$$i\frac{\mathrm{d}}{\mathrm{d}t}U(t, s) = H(t)U(t, s) \tag{2.77}$$

$$U(t, s)U(s, r) = U(t, r) \tag{2.78}$$

Time dependent interactions can most easily be discussed in the interaction picture. Let $U_0(t) = e^{it(T+\mu N)}$ and

$$V(t,s) = U_0(t)U(t,s)U_0(-s). \tag{2.79}$$

Then V satisfies the equations

$$V(t,t) = 1 \tag{2.80}$$

$$i\frac{\partial}{\partial t}V(t,s) = H_I(t)V(t,s) \tag{2.81}$$

$$V(t,s)V(s,r) = V(t,r) \tag{2.82}$$

with $H_I(t) = U_0(t)H_{I,t}U_0(-t)$. V satisfies the integral equation

$$V(t,s) = 1 - i\int_s^t dt' H_I(t')V(t',s). \tag{2.83}$$

A solution in the sense of formal power series in H_I is obtained by iteration:

$$V(t,s) = 1 + \sum_{n=1}^{\infty}(-i)^n \int_s^t \mathrm{d}t_1 \int_s^{t_1} \mathrm{d}t_2 \ldots \int_s^{t_{n-1}} \mathrm{d}t_n H_I(t_1)\ldots H_I(t_n)\ . \tag{2.84}$$

This formula can be written in an elegant way by the introduction of time ordered products. Let $A : \mathbb{R} \ni t \mapsto A(t)$ be an operator valued function. Then one defines operator valued functions

$$T_A(t_1,\ldots,t_n) \tag{2.85}$$

(the time ordered product of $A(t_1)$,...,$A(t_n)$) by the following equations

$$\begin{aligned} T_A(t_1,\ldots,t_n)) &= A(t_1)\ldots A(t_n) \text{ for } t_1 \geq t_2 \geq \cdots \geq t_n \\ T_A(t_1,\ldots,t_n)) &= T_A(t_{\sigma(1)},\ldots,t_{\sigma(n)})\ \forall \sigma \in S_n \end{aligned} \tag{2.86}$$

It is common to write

$$T_A(t_1,\ldots,t_n) = TA(t_1)\cdots A(t_n)\ . \tag{2.87}$$

But one has to keep in mind that $T_A(t_1,\ldots,t_n))$ depends on the function $t \mapsto A(t)$ and not only on the operators $A(t_1)$,...,$A(t_n)$.

For the time evolution operator in the interaction picture we get the expression

$$\begin{aligned} V(t,s) &= 1 + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int_s^t dt_1 \ldots \int_s^t dt_n T H_I(t_1)\ldots H_I(t_n) \\ &=: Te^{-i\int_s^t dt' H_I(t')}\ . \end{aligned} \tag{2.88}$$

The last line is called the time ordered exponential.

In the example above we have

$$[[H_I(t_1), H_I(t_2)], H_I(t_3)] = 0. \tag{2.89}$$

Therefore the exponential series can be summed. We study first the case that $H_I(t)$ is piecewise constant. Then

$$\begin{aligned} V(t,s) &= V(t,t_1)V(t_1,t_2)\dots V(t_n,s) \\ &= e^{-i\int_{t_1}^{t}\mathrm{d}t' H_I(t')}\dots e^{-i\int_{s}^{t_n}\mathrm{d}t' H_I(t')} \\ &= e^{-i\int_s^t \mathrm{d}t' H_I(t') - \frac{1}{2}\sum_{i>j}\int_{t_i}^{t_{i+1}}\mathrm{d}t'\int_{t_j}^{t_{j+1}}\mathrm{d}t''[H_I(t'),H_I(t'')]} \\ &= e^{-i\int_s^t \mathrm{d}t' H_I(t') - \frac{1}{2}\int_s^t \mathrm{d}t'\int_s^{t'}\mathrm{d}t''[H_I(t'),H_I(t'')]} \end{aligned} \tag{2.90}$$

One now easily verifies that the term in the last line solves the equation for $V(t,s)$ also in the general case.

When the interaction vanishes for sufficiently large times, it makes sense to consider the operator which describes the total influence of the source on the system. We define the S-matrix

$$S = \lim_{t\to\infty, s\to-\infty} V(t,s). \tag{2.91}$$

Let

$$\varphi(f) = \int \mathrm{d}t\mathrm{d}^3\mathbf{x}\left(a(t,\mathbf{x})\overline{f(t,\mathbf{x})} + a^*(t,\mathbf{x})f(t,\mathbf{x})\right) = a(F) + a(F)^* \tag{2.92}$$

with $F = \int \mathrm{d}t e^{i(T_1+\mu)t} f_t$, i.e.

$$\hat{F}(\mathbf{p}) = \sqrt{2\pi}\hat{f}(-(\frac{|\mathbf{p}|^2}{2M}+\mu),\mathbf{p})\ . \tag{2.93}$$

Then one has

$$S = e^{-i\varphi(f)-i\alpha} \tag{2.94}$$

and $\alpha = \frac{1}{2i}\int_{t>s}\mathrm{d}t\mathrm{d}s[H_I(t),H_I(s)]$. It remains to calculate α. We have

$$\alpha = \mathrm{Im}\int_{t>s}\mathrm{d}t\mathrm{d}s(f_t, e^{-i(t-s)(T_1+\mu)}f_s) \tag{2.95}$$

$$= \mathrm{Im}\,(2\pi)^{-1}\int_{t>s}\mathrm{d}t\mathrm{d}s\int\mathrm{d}E\mathrm{d}E'\int\mathrm{d}^3\mathbf{p}\overline{\hat{f}(-E,\mathbf{p})}\hat{f}(-E',\mathbf{p})e^{i(E-E(\mathbf{p}))t}e^{-i(E'-E(\mathbf{p}))s} \tag{2.96}$$

($E(\mathbf{p}) = \frac{|\mathbf{p}|^2}{2M}+\mu$). To be able to perform first the integration over s, we multiply the integrand with $e^{-\varepsilon(t-s)}$, $\varepsilon>0$. In the limit $\varepsilon\to 0$ we get the original integral back. For $\varepsilon>0$ the order of integrations can be interchanged. It holds

$$\int_{s<t}\mathrm{d}s e^{-\varepsilon(t-s)}e^{-i(E'-E(\mathbf{p}))s} = i(E'-E(\mathbf{p})+i\varepsilon)^{-1}e^{-i(E'-E(\mathbf{p}))t} \tag{2.97}$$

and hence

$$\alpha = \frac{1}{2\pi}\lim_{\varepsilon\downarrow 0}\mathrm{Re}\int\mathrm{d}t\int\mathrm{d}^3\mathbf{p}\int\mathrm{d}E\mathrm{d}E'\frac{\overline{\hat{f}(-E,\mathbf{p})}\hat{f}(-E',\mathbf{p})e^{i(E-E')t}}{E'-E(\mathbf{p})+i\varepsilon}\ . \tag{2.98}$$

For performimg the t-integration, we use the formula

$$\int \mathrm{d}t e^{i(E-E')t} = 2\pi\delta(E-E') \ . \tag{2.99}$$

This formula holds in the sense of distributions, i.e. for all test function $h \in \mathcal{S}(\mathbb{R})$ one has

$$\int \mathrm{d}t \int \mathrm{d}E e^{i(E-E')t} h(E) = 2\pi \int \mathrm{d}E \delta(E-E')h(E) = 2\pi h(E') \ . \tag{2.100}$$

(Inversion of the Fourier transformation.) We obtain

$$\alpha = \lim_{\varepsilon \downarrow 0} \int \mathrm{d}^3\mathbf{p} \int \mathrm{d}E |\hat{f}(-E,\mathbf{p})|^2 \frac{E-E(\mathbf{p})}{(E-E(\mathbf{p}))^2+\varepsilon^2} \tag{2.101}$$

$$= \int \mathrm{d}^3\mathbf{p} \int \mathrm{d}E |\hat{f}(-E,\mathbf{p})|^2 P \frac{1}{E-E(\mathbf{p})} \tag{2.102}$$

where the symbol P denotes Cauchy's principal value,

$$\int \mathrm{d}x P\frac{1}{x} h(x) = \lim_{\varepsilon \downarrow 0} \int_{|x|>\varepsilon} \mathrm{d}x \frac{h(x)}{x} \ . \tag{2.103}$$

In case the initial state is the vacuum, the source generates the coherent state

$$S\Omega = e^{-i\alpha}\Omega_{-iF}. \tag{2.104}$$

In case the initial state is already a coherent state Ω_{-iG}, possibly generated by another source which was earlier switched on, one finds

$$S\Omega_{-iG} = c\Omega_{-i(F+G)}, |c| = 1. \tag{2.105}$$

For the expected number of created particles δN one obtains

$$\delta N = ||F+G||^2 - ||G||^2 = ||F||^2 + 2\mathrm{Re}\,(F,G). \tag{2.106}$$

The interference term $2\mathrm{Re}\,(F,G)$ describes, depending on the sign, absorption or induced emission.

3. The fermionic Fock space

The fermionic Fock space is defined as

$$\mathfrak{H}^- = \bigoplus_{n=0}^{\infty} \mathfrak{H}_n^-, \tag{3.1}$$

where $\mathfrak{H}_n^-$ is formed by the totally antisymmetric n-particle wave functions. The annihilation operators are defined as in the bosonic Fock space:

$$(a(f)\Phi)_n(\mathbf{x_1},\ldots,\mathbf{x_n}) = \sqrt{n+1} \int d^3\mathbf{x}\overline{f(\mathbf{x})}\Phi_{n+1}(\mathbf{x},\mathbf{x_1},\ldots,\mathbf{x_n}). \tag{3.2}$$

The creation operator is again defined as the adjoint operator and turns out to be given by

$$(a(f)^*\Phi)_0 = 0 \quad , \tag{3.3}$$

$$(a(f)^*\Phi)_n(\mathbf{x_1},\ldots,\mathbf{x_n}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(-1)^{i+1}f(\mathbf{x_i})\Phi_{n-1}(\mathbf{x_1},\ldots,\mathbf{x_{i-1}},\mathbf{x_{i+1}},\ldots,\mathbf{x_n}). \tag{3.4}$$

One easily verifies the anticommutation relations

$$[a(f),a(g)^*]_+ = \langle f,g\rangle \tag{3.5}$$

$$[a(f),a(g)]_+ = 0 = [a(f)^*,a(g)^*]_+ \tag{3.6}$$

(canonical anticommutation relations (CAR)). As in the Bose case one introduces operator valued distributions $a(\mathbf{x},t)$ and $a^*(\mathbf{x},t)$ by

$$\int d^3\mathbf{x}\overline{f(\mathbf{x})}a(\mathbf{x},t) = a(f)(t) \tag{3.7}$$

$$\int d^3\mathbf{x}f(\mathbf{x})a^*(\mathbf{x},t) = a(f)^*(t) \tag{3.8}$$

with the simultaneous anticommutation relations

$$[a(\mathbf{x},t),a^*(\mathbf{y},t)]_+ = \delta(\mathbf{x}-\mathbf{y}) \quad , \tag{3.9}$$

$$[a(\mathbf{x},t),a(\mathbf{y},t)]_+ = 0 = [a^*(\mathbf{x},t),a^*(\mathbf{y},t)]_+ \ . \tag{3.10}$$

The operator $a(f)^*a(f)$ has, for $||f|| = 1$, as in the Bose case the interpretation "number of particles with wave function f". Because of Fermi statistics $a(f)^*a(f)$ has only the eigenvalues 0 and 1:

$$\begin{gathered}(a(f)^*a(f))^2 = a(f)^*a(f)a(f)^*a(f)\\ = a(f)^*[a(f),a(f)^*]_+a(f) - a(f)^*a(f)^*a(f)a(f) = a(f)^*a(f)\end{gathered} \tag{3.11}$$

since $[a(f),a(f)^*]_+ = ||f||^2 = 1$ and $a(f)^2 = \frac{1}{2}[a(f),a(f)]_+ = 0$. Thus $a(f)^*a(f)$ is a projection. In particular $a(f)$ is a bounded operator with norm

$$||a(f)|| = ||f||. \tag{3.12}$$

In the Bose case, however, we have only

$$\langle\Phi,a(f)^*a(f)\Phi\rangle \le ||f||^2\langle\Phi,N\Phi\rangle \tag{3.13}$$

and therefore, on the orthogonal complement of the vacuum,

$$||a(f)N^{-\frac{1}{2}}|| \le ||f||. \tag{3.14}$$

An important speciality of fermion systems is their ability to fill states. Let E be the projection on some finite dimensional subspace of $\mathfrak{H}_1$, and let $\{f_1,\ldots,f_n\}$ be an orthonormal basis of this space. We consider the vector

$$\Phi_E = a(f_1)^*\ldots a(f_n)^*\Omega. \tag{3.15}$$

(Φ_E is known as the Slater determinant of the single particle wave functions $f_1,\ldots,f_n$.) Φ_E is annihilated by all operators of the form $a(Ef)^*$ and $a((1-E)f)$, $f\in\mathfrak{H}_1$. This property characterizes Φ_E uniquely up to a phase.

The operator

$$b(f) = a(Ef)^* + a((1-E)f) \tag{3.16}$$

can be interpreted as an annihilation operator for a quasi-particle. If $f \in E\mathfrak{H}_1$, then the quasi-particle is a "hole", if $f \in (1-E)\mathfrak{H}_1$, it is a particle of the original sort. The operators $b(f)$ and $b(g)^*$ satisfy the following commutation relations

$$[b(f), b(g)^*]_+ = \big\langle f, (1-E)g\big\rangle + \big\langle g, Ef\big\rangle, \tag{3.17}$$

$$[b(f), b(g)]_+ = 0 = [b(f)^*, b(g)^*]_+ \ . \tag{3.18}$$

The vector Φ_E describes the vacuum state for the quasi-particles.

Operators b and b^* are also meaningful for projections E on infinite dimensional subspace. But then the Fock space does not contain the vacuum for the quasi-particles. It holds the following theorem:

THEOREM I.2. *Let $\Phi \in \mathfrak{H}^-$ with $a(f)^*\Phi = 0 \forall f \in \mathfrak{K}$, where $\mathfrak{K}$ is an infinite dimensional subspace of $\mathfrak{H}_1$. Then $\Phi = 0$.*

PROOF. Let $f \in \mathfrak{K}$ with $||f|| = 1$. Then

$$\Phi = [a(f), a(f)^*]_+\Phi = a(f)^*a(f)\Phi. \tag{3.19}$$

If $g \perp f$, then $a(g)$ commutes with $a(f)^*a(f)$. Let $(f_i)_{i\in\mathbb{N}}$ be an orthonormal basis of $\mathfrak{K}$. Then for all $k \in \mathbb{N}$

$$\Phi = a(f_1)^* \dots a(f_k)^* a(f_k) \dots a(f_1)\Phi. \tag{3.20}$$

Let now $\Psi \in \mathfrak{H}^-$ be a state with bounded particle number, i.e. $\exists n_0 \in \mathbb{N}$ with $\Psi_n = 0$ for $n > n_0$. Then

$$a(f_k) \dots a(f_1)\Psi = 0 \tag{3.21}$$

for $k > n_0$. Hence for $k > n_0$

$$\begin{aligned} \big\langle \Psi, \Phi\big\rangle &= \big\langle \Psi, a(f_1)^* \dots a(f_k)^* a(f_k) \dots a(f_1)\Phi\big\rangle \\ &= \big\langle a(f_k) \dots a(f_1)\Psi, a(f_k) \dots a(f_1)\Phi\big\rangle = 0. \end{aligned} \tag{3.22}$$

Φ is thus orthogonal to all vectors with bounded particle number, i.e. $\Phi_n = 0 \forall n$, thus $\Phi = 0$. □

One may now introduce a Fock space for quasi-particles,

$$\mathfrak{H}^{E,-} = \bigoplus_{n=0}^{\infty} \mathfrak{H}_n^{E,-} \tag{3.23}$$

with the n-quasi particle spaces

$$\mathfrak{H}_n^{E,-} = \{\Phi \in \underbrace{\mathfrak{H}_1^E \otimes \dots \otimes \mathfrak{H}_1^E}_{n}, \ U(\sigma)\Phi = \mathrm{sign}\sigma\Phi, \sigma \in S_n\} \tag{3.24}$$

and the single quasi particle space

$$\mathfrak{H}_1^E = (1-E)\mathfrak{H}_1 \oplus \overline{E\mathfrak{H}_1} \ . \tag{3.25}$$

$\mathfrak{H}_1^E$ coincides with $\mathfrak{H}_1$ as a real Hilbert space. The difference consists in the fact that the operator of multiplication with i is replaced by $i(1-E)-iE$ and thus the scalar product is now defined by

$$\big\langle f,g\big\rangle_E = \big\langle f,(1-E)g\big\rangle + \big\langle g,Ef\big\rangle\ . \tag{3.26}$$

On the new Fock space, the new creation and annihilation operators act in an analogous way. Since the new n-particle spaces are not immediately given as spaces of functions, it is convenient to characterize the operators b and b^* abstractly:

$$b(f)\sum_{(i)} c_{(i)}\Phi_{i_1}\otimes\cdots\otimes\Phi_{i_n} = \sqrt{n}\sum_{(i)} c_{(i)}(f,\Phi_{i_1})\Phi_{i_2}\otimes\cdots\otimes\Phi_{i_n} \tag{3.27}$$

$$b(f)^*\sum_{(i)}\Phi_{i_1}\otimes\cdots\otimes\Phi_{i_n} = \frac{1}{\sqrt{n+1}}\sum_{k=0}^{n}(-1)^n\Phi_{i_1}\otimes\cdots\otimes\Phi_{i_k}\otimes f\otimes\Phi_{i_{k+1}}\cdots\Phi_{i_n} \tag{3.28}$$

Φ_E is the vacuum vector in the quasi particle Fock space $\mathfrak{H}^{E,-}$.

An important example for such a projection is the projection on energies below the Fermi energy $\mu>0$,

$$\hat{Ef}(\mathbf{p}) = \begin{cases} \hat{f}(\mathbf{p}) & , \ \frac{|\mathbf{p}|^2}{2M}<\mu \\ 0 & , \ \frac{|\mathbf{p}|^2}{2M}\geq\mu \end{cases}\ . \tag{3.29}$$

Then Φ_E describes the state, in which all single particle states with energy below μ are occupied. Another important example is the so-called Dirac sea where all states with negative energies are occupied.

CHAPTER II

Relativistic single particle systems

1. The Poincaré group

According to the principles of special relativiy, physical systems which differ only by a uniform relative motion have identical properties. Furthermore there is a limiting velocity for the propagation of signals, the velocity of light c which is equal in all uniformly moving systems. It is customary in relativistic physics to use a system of units where $c = 1$ holds.

Points x in spacetime are described by elements of $\mathbb{R} \times \mathbb{R}^3 = \mathbb{R}^4$,

$$x = (t, \mathbf{x}) = (x^0, x^1, x^2, x^3) \tag{1.1}$$

with $x^0 = t$ and $(x^1, x^2, x^3) = \mathbf{x}$. A uniform motion with velocity $\mathbf{v}$ is represented in this space by a straight line,

$$x(\tau) = (1, \mathbf{v})\tau + a \ , \ \tau \in \mathbb{R}, \ a \in \mathbb{R}^4 \ . \tag{1.2}$$

For any motion which is connected with the transmission of a signal one has $|\mathbf{v}| \leq 1$. Spacetime points x which can be reached in finite time from the point 0 by a signal with $|\mathbf{v}| < 1$, form the forward lightcone

$$V_+ = \{x \in \mathbb{R}^4, \ x^0 > |\mathbf{x}|\} \ , \tag{1.3}$$

those from which the point 0 can be reached form the backward lightcone

$$V_- = \{x \in \mathbb{R}^4, -x^0 > |\mathbf{x}|\} = -V_+ \ . \tag{1.4}$$

$V_+ \cup V_- = \{x \in \mathbb{R}^4, (x^0)^2 > |\mathbf{x}|^2\}$ is called the set of timelike points. The boundary of the forward lightcone,

$$\partial V_+ = \{x \in \mathbb{R}^4, \ x^0 = |\mathbf{x}|\} \tag{1.5}$$

is the set of points which can be reached from the origin by a light signal. Points with $|x^0| = |\mathbf{x}|$ are called lightlike, those with $|x^0| < |\mathbf{x}|$ spacelike.

The structure of spacetime can be described by an (indefinite) scalar product

$$\langle x|y\rangle \equiv xy := x^0 y^0 - \mathbf{x} \cdot \mathbf{y} \ . \tag{1.6}$$

Other notations are

$$xy = x^\mu y_\mu = x^\mu g_{\mu\nu} y^\nu \ . \tag{1.7}$$

Here always the summation convention is used that all indices which occur as an upper and a lower index have to be summed over, from 0 to 3. $g_{\mu\nu}$ is the metric tensor with components

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.8}$$

and $y_\mu = g_{\mu\nu} y^\nu$. $\mathbb{R}^4$ endowed with this scalar product is called Minkowski space $\mathbb{M}$.

The group of motions of Minkowski space is the set of all maps L with the property

$$(Lx - Ly)^2 = (x - y)^2 \ , \ x, y \in \mathbb{M} \ , \tag{1.9}$$

and is called the Poincaré group (notation $\mathcal{P}$). The maps $\Lambda \in \mathcal{P}$ which leave the origin fixed are called Lorentz transformations. They form a subgroup $\mathcal{L}$, the Lorentz group.

Every Poincaré transformation can be written as a product of a Lorentz transformation Λ and a translation a,

$$Lx = a + \Lambda x \ , \ L = (a, \Lambda) \ . \tag{1.10}$$

The Lorentz transformations of Minkowski space are analogous to the rotations of euclidean space and are likewise linear maps,

$$(\Lambda x)^\nu = \Lambda^\mu{}_\nu x^\nu \tag{1.11}$$

with $\Lambda^\mu{}_\nu \in \mathbb{R}$.

The group law in $\mathcal{P}$ is

$$(a_1, \Lambda_1)(a_2, \Lambda_2) = (a_1 + \Lambda_1 a_2, \Lambda_1 \Lambda_2) \ . \tag{1.12}$$

Lorentz transformations are volume preserving (i.e. $|\det \Lambda| = 1$), but may change the orientation. E.g., the parity

$$I_S(x^0, \mathbf{x}) = (x^0, -\mathbf{x}) \tag{1.13}$$

is a Lorentz transformation with $\det \Lambda = -1$.

A further property of Lorentz transformations is that they either leave the forward lightcone invariant or map it into the backward lightcone.

PROOF. Let $x, y \in V_+$. Then $x^0 > |\mathbf{x}|$, $y^0 > |\mathbf{y}|$ and thus

$$xy = x^0 y^0 - \mathbf{x} \cdot \mathbf{y} > |\mathbf{x}||\mathbf{y}| - \mathbf{x} \cdot \mathbf{y} \geq 0 \ . \tag{1.14}$$

Λx and Λy are timelike, hence belong to $V_+ \cup V_-$. If one of these points would be in V_+ and the other in V_-, we would have

$$\Lambda x \Lambda y < 0 < xy \tag{1.15}$$

in contradiction to the fact that the Lorentz transformations preserve the scalar product of Minkowski space. □

For the matrix entries of Λ the property above means that $|\Lambda^0{}_0| \geq 1$,

$$1 = \Lambda(1, \mathbf{0})^2 = (\Lambda^0{}_0)^2 - \sum_{i=1}^{3} (\Lambda^i{}_0)^2 \ . \tag{1.16}$$

The Lorentz group decomposes into the following connected components:

$\mathcal{L}_+^\uparrow$ (proper orthochronous Lorentz group): $\det \Lambda = 1$, $\Lambda^0{}_0 \geq 1$. These Lorentz transformations preserve space and time orientation (left and right, past and future are not exchanged). The identical transformation 1 belongs to this component.

$\mathcal{L}_+^\downarrow$: $\det \Lambda = 1$, $\Lambda^0{}_0 \leq -1$, e.g. -1.

$\mathcal{L}_-^\uparrow$: $\det \Lambda = -1$, $\Lambda^0{}_0 \geq 1$, e.g. parity I_S.

$\mathcal{L}_-^\downarrow$: $\det \Lambda = -1$, $\Lambda^0{}_0 \leq -1$, e.g. time reflection I_T.

The Poincaré -group $\mathcal{P}$ decomposes accordingly into the connected components $\mathcal{P}_+^\uparrow, \mathcal{P}_+^\downarrow, \mathcal{P}_-^\uparrow, \mathcal{P}_-^\downarrow$. Only the elements of the connected component of the unit correspond to physically realizable transformations. Therefore $\mathcal{P}_+^\uparrow$ is considered as the relativistic invariance group. Indeed parity and time reflection are not symmetries of nature.

The group $\mathcal{P}_+^\uparrow$ is two-fold connected, i.e. the set of closed curves in $\mathcal{P}_+^\uparrow$ decomposes into two different classes which cannot be continuously transformed into each other. The reason is that $\mathcal{P}_+^\uparrow$ contains the likewise twofold connected rotation group. The simply connected covering group of $\mathcal{P}_+^\uparrow$ is the so-called inhomogeneous $\mathrm{SL}(2, \mathbb{C})$ which is denoted by $\mathcal{P}^c$ in what follows. This group consists of pairs (a, A) with translations a and complex (2×2)-matrices A with determinant 1. $\mathrm{SL}(2, \mathbb{C})$ is the covering group of the proper orthochronous Lorentz group $\mathcal{L}_+^\uparrow$. The covering homomorphism

$$\Lambda : \left\{ \begin{array}{ccc} \mathrm{SL}(2, \mathbb{C}) & \to & \mathcal{L}_+^\uparrow \\ A & \mapsto & \Lambda(A) \end{array} \right. \tag{1.17}$$

is defined in the following way: By

$$x \mapsto \underset{\sim}{x} = x^0 \mathbf{1} + \mathbf{x} \cdot \vec{\sigma} = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} \tag{1.18}$$

a bijective linear map from Minkowski space into the space of hermitean 2×2-matrices is introduced. It holds

$$\det \underset{\sim}{x} = x^2 \ . \tag{1.19}$$

One now sets

$$\underset{\sim}{\Lambda(A)x} = A \underset{\sim}{x} A^* \ . \tag{1.20}$$

One easily verifies that $\Lambda(A)$ is a Lorentz transformation. We have $\Lambda(A) = \Lambda(B) \Longrightarrow A = \pm B$ und $\Lambda(\mathrm{SL}(2,\mathbb{C})) = \mathcal{L}_+^\uparrow$.

Example: (i) Let

$$A = e^{i\theta/2}\frac{1}{2}(1 + \mathbf{n}\cdot\vec{\sigma}) + e^{-i\theta/2}\frac{1}{2}(1 - \mathbf{n}\cdot\vec{\sigma}) \ , \tag{1.21}$$

$|\mathbf{n}| = 1$, $\theta \in \mathbb{R}$. Then $\Lambda(A)$ is a rotation with rotation axis $\mathbf{n}$ and rotation angle $-\theta$.

(ii) Let

$$A = e^{\theta/2}\frac{1}{2}(1 + \mathbf{n}\cdot\vec{\sigma}) + e^{-\theta/2}\frac{1}{2}(1 - \mathbf{n}\cdot\vec{\sigma}) \ . \tag{1.22}$$

$|\mathbf{n}| = 1$, $\theta \in \mathbb{R}$. Then $\Lambda(A)$ is a Lorentz boost with velocity $\mathbf{v} = \mathbf{n}\tanh\theta$.

The group multiplication in $\mathcal{P}^c$ is given by

$$(a_1, A_1)(a_2, A_2) = (a_1 + \Lambda(A_1)a_2, A_1A_2) \ . \tag{1.23}$$

Besides of the map $x \mapsto \underset{\sim}{x}$ we will use a further map from Minkowski space into the set of hermitean 2×2-matrices,

$$x \mapsto \widetilde{x} = x^0\mathbf{1} - \mathbf{x}\cdot\vec{\sigma} = \begin{pmatrix} x^0 - x^3 & -x^1 + ix^2 \\ -x^1 - ix^2 & x^0 + x^3 \end{pmatrix} \ . \tag{1.24}$$

It holds

$$\underset{\sim}{x}\widetilde{x} = x^2\mathbf{1} \tag{1.25}$$

and thus

$$\Lambda(\widetilde{A})x = (A^*)^{-1}\widetilde{x}A^{-1} \ . \tag{1.26}$$

2. Poincaré symmetry in quantum mechanics

We assume that the states of a relativistic particle can be described by the rays (one dimensional subspaces) $\hat{\Phi}$ of some Hilbert space $\mathfrak{H}$,

$$\hat{\Phi} = \{\lambda\Phi, \lambda \in \mathbb{C}\} \ , \ \Phi \in \mathfrak{H} \ , \ \Phi \neq 0 \ . \tag{2.1}$$

A Poincaré transformation $L \in \mathcal{P}_+^\uparrow$ transforms single particle states into single particle states,

$$\hat{\Phi} \mapsto \hat{T}_L\hat{\Phi} = \hat{\Psi} \ , \ \Psi \in \mathfrak{H} \ , \ \Psi \neq 0 \ . \tag{2.2}$$

The transformed system shall have the same physical properties as the original system. In particular the transition probabilities (ray products)

$$\left\langle \hat{\Phi}, \hat{\Psi} \right\rangle := \frac{|\langle \Phi, \Psi\rangle|^2}{\|\Phi\|^2\|\Psi\|^2} \tag{2.3}$$

shall not change,

$$\left\langle \hat{T}_L\hat{\Phi}, \hat{T}_L\hat{\Psi} \right\rangle = \left\langle \hat{\Phi}, \hat{\Psi} \right\rangle \ . \tag{2.4}$$

For transformations which leave the ray product invariant, one has the following theorem of Wigner (for a proof see Weinberg):

THEOREM II.1. *Let $\hat{T}$ be an invertible and ray product preserving map of the rays of a Hilbert space into itself. Then there is an invertible, $\mathbb{R}$-linear and isometric map $T : \mathfrak{H} \to \mathfrak{H}$ with the property*

$$\widehat{T\Phi} = \hat{T}\hat{\Phi} \ . \tag{2.5}$$

T is unique up to a factor of modulus 1 and is either unitary or antiunitary.

If $\hat{T}$ can be written as the square of another ray transformation,

$$\hat{T} = \hat{S}^2 \ , \tag{2.6}$$

then $T = \lambda S^2$ is necessarily unitary, since the square of an antiunitary operator is unitary. In the connected component of 1 in the Poincaré group each element can be written as a product of squares, hence T_L is unitary for all $L \in \mathcal{P}_+^\uparrow$.

If one performs two Poincaré transformations L_1 und L_2 one after another, the transformed state shall coincide with that which one would obtain by the transformation $L = L_1 L_2$, i.e.

$$\hat{T}_{L_1}\hat{T}_{L_2} = \hat{T}_{L_1 L_2} \tag{2.7}$$

(ray representation or projective representation). For the operators T_L it follows

$$T_{L_1}T_{L_2} = e^{i\omega(L_1,L_2)}T_{L_1L_2} \tag{2.8}$$

with $\omega(L_1, L_2) \in \mathbb{R}$. If one modifies the definition of T_L,

$$T'_L = e^{i\alpha(L)}T_L \ , \tag{2.9}$$

one obtains

$$T'_{L_1}T'_{L_2} = e^{i\omega'(L_1,L_2)}T'_{L_1L_2} \tag{2.10}$$

with

$$\omega'(L_1, L_2) = \omega(L_1, L_2) + \alpha(L_1) + \alpha(L_2) - \alpha(L_1L_2) \ . \tag{2.11}$$

This suggests the question whether one can find a choice of α such that ω' vanishes. This question leads to the investigation of the group cohomology of $\mathcal{P}_+^\uparrow$.

Let us assume that the ray representation is continuous in the following sense,

$$\left\langle \hat{T}_L\hat{\Phi}, \hat{\Psi} \right\rangle \to \left\langle \hat{\Phi}, \hat{\Psi} \right\rangle \tag{2.12}$$

for $L \to 1$. For these ray representations the problem is solved by the following theorem of Wigner and Bargmann (see Weinberg for a proof).

THEOREM II.2. *For every continuous ray representation $L \to \hat{T}_L$ of the proper orthochronous Poincaré group $\mathcal{P}_+^\uparrow$ there is a strongly continuous unitary representation U of the twofold covering group $\mathcal{P}^c$, such that*

$$\widehat{U(a,A)}\Phi = \hat{T}_{a,\Lambda(A)}\hat{\Phi} \tag{2.13}$$

with the covering map Λ.

For the description of a relativistic single particle system we therefore use the following ansatz which goes back to Wigner: The Hilbert space of the states of a particle is the representation space of an irreducible, continuous, unitary representation of $\mathcal{P}^c$. Our next task is to find these representations.

3. The representations of the Poincaré group

Let U be a strongly continuous, unitary, irreducible representation of $\mathcal{P}^c$ in some Hilbert space $\mathfrak{H}$. We consider first the restriction of U to the subgroup of translations,

$$\mathbb{R}^4 \ni a \mapsto U(a) \ . \tag{3.1}$$

This representation can be characterized by 4 selfadjoint, mutually commuting operators P_μ, $\mu = 0, \ldots, 3$,

$$U(a) = e^{iPa} \ , \ Pa = P_\mu a^\mu \ . \tag{3.2}$$

The joint spectrum of these 4 operators is a closed subset of $\mathbb{R}^4$,

$$\operatorname{sp} P = \{(p_0, p_1, p_2, p_3) \in \mathbb{R}^4 | p_\mu \in \operatorname{sp} P_\mu, \mu = 0, \ldots, 3\} \ . \tag{3.3}$$

where $\mathbb{R}^4$ is interpreted as momentum space, i.e. the dual $\mathbb{M}^*$ of the translation group of Minkowski space (which by choice of an origin is itself identified with the Minkowski space $\mathbb{M}$). In our notation we characterize components of elements of $\mathbb{M}^*$ by lower indices. Lorentz transformations act on $\mathbb{M}^*$ by transposition,

$$(p\Lambda)_\nu = p_\mu \Lambda^\mu{}_\nu \ . \tag{3.4}$$

The relation

$$U(A)U(a)U(A)^{-1} = U(\Lambda(A)a) \ , \ A \in \mathrm{SL}(2,\mathbb{C}) \tag{3.5}$$

follows from the condition that U is a representation and implies

$$U(A)PaU(A)^{-1} = P\Lambda(A)a \ . \tag{3.6}$$

Thus $\operatorname{sp} P$ is Lorentz invariant. Let us first look at the point spectrum. If Φ is a joint eigenvector of Pa, $a \in \mathbb{M}$ with eigenvalues pa, then $U(A)\Phi$ is an eigenvector $P\Lambda(A)a$ with eigenvalues pa, or, equivalently, an eigenvector of Pa with eigenvalues $p\Lambda(A^{-1})a$. But eigenvectors of selfadjoint operators with different eigenvalues are orthogonal. Hence only the eigenvalue $p = 0$ is compatible with the assumed continuity of the representation U.

We obtain a class of irreducible representations by

$$U(a, A) = U(A) \tag{3.7}$$

with a unitary irreducible representation U of $\mathrm{SL}(2, \mathbb{C})$. These representations, however, do not describe particles, since the transition probabilities $\Phi \to U(a)\Psi$ are independent of a. The case $U(A) = 1$ characterizes the transformation properties of the vacuum.

We now investigate the continuous spectrum. We consider for each point $p \in \mathbb{M}^*$ its orbit O_p under Lorentz transformations,

$$O_p = \{q \in \mathbb{M}^*, \exists \Lambda \in \mathcal{L}_+^\uparrow \text{ with } p\Lambda = q\} \ . \tag{3.8}$$

In an irreducible representation, $\mathrm{sp}\, P$ is the closure of a single orbit. The orbits can be classified as follows

$$\begin{aligned} H_m^+ &= \{p \in \mathbb{M}^* | p^2 = m^2, \ p_0 > 0\} \\ H_m^- &= \{p \in \mathbb{M}^* | p^2 = m^2, \ p_0 < 0\} \\ \partial V_+ &= \{p \in \mathbb{M}^* | p^2 = 0, \ p_0 > 0\} \\ \partial V_- &= \{p \in \mathbb{M}^* | p^2 = 0, \ p_0 < 0\} \\ H_{im}^+ &= \{p \in \mathbb{M}^* | p^2 = -m^2\} \\ &\{0\} \end{aligned} \tag{3.9}$$

with $m > 0$, and by abuse of notation, $p^2 := p_\mu p^\mu$ with $p_\mu = g_{\mu\nu} p^\nu$. In an irreducible representation, P^2 is a multiple of the unit. If $P^2 = m^2 \mathbf{1}$, $m^2 > 0$, then $\mathrm{sp}\, P = H_m^+$ or $\mathrm{sp}\, P = H_m^-$. If $P^2 = 0$, then $\mathrm{sp}\, P = \{0\}$, $\mathrm{sp}\, P = \overline{\partial V_+}$ or $\mathrm{sp}\, P = \overline{\partial V_-}$; in the latter cases the spectrum contains not only a single orbit, but (being a closed set) in addition the origin. 0 is, however, not an eigenvalue, but belongs to the continuous spectrum.

In analogy to nonrelativistic quantum mechanics we interpret $P_0 = P^0 \equiv H$ as energy and $\mathbf{P} = (P^1, P^2, P^3)$ as the spatial momentum of the particle. We will justify this interpretation later on. For the moment it leads us to accept only the orbits with $p_0 > 0$ as spectra of physical particles.

Let us turn to the case $\mathrm{sp}\, P = H_m^+$. We choose a representation in which the momenta are diagonal. Because of Poincaré invariance this representation can be determined very explicitly.

We first convince ourselves that our Hilbert space contains a dense subspace $\mathfrak{D}$ on which the expectation values of spatial translations vanish fast at infinity,

$$\lim_{|x| \to \infty} |\mathbf{x}|^n \big\langle \Phi, U(\mathbf{x})\Psi \big\rangle = 0 \ \forall n \in \mathbb{N} \ . \tag{3.10}$$

We set

$$\mathfrak{D} = \{\Phi \in \mathfrak{H}, \mathrm{SL}(2, \mathbb{C}) \ni A \mapsto U(A)\Phi \text{ infinitely often differentiable}\} \ . \tag{3.11}$$

PROOF. Let $\mathbf{x} \neq 0$ and let $(A_\theta)_\theta$ be the one-parameter group in $\mathrm{SL}(2,\mathbb{C})$ which describes Lorentz boosts in the direction of $\mathbf{x}$. Then

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\Lambda(A_\theta)(0,\mathbf{x})|_{\theta=0} = (|\mathbf{x}|,0) \ . \tag{3.12}$$

Thus we find for $\Phi, \Psi \in \mathfrak{D}$

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\big\langle U(A_\theta)\Phi, U(\mathbf{x})U(A_\theta)\Psi\big\rangle|_{\theta=0} = -i|\mathbf{x}|(\Phi, P_0 U(\mathbf{x})\Psi) \tag{3.13}$$

But because of the condition on $\mathrm{sp}\, P$, with Φ also $P_0^{-1}\Phi \in \mathfrak{D}$. We therefore can replace Φ by $P_0^{-1}\Phi$ in this equation. On the left hand side we can perform the differentiation of vector valued functions inside the scalar product and obtain finitely many vectors $\Phi_i, \Psi_i \in \mathfrak{D}$ with

$$|\mathbf{x}|\big\langle \Phi, U(\mathbf{x})\Psi\big\rangle = \sum_i \big\langle \Phi_i, U(\mathbf{x})\Psi_i\big\rangle \ . \tag{3.14}$$

We know that the matrix elements of translations are bounded. But then it follows from this equation that they have to decay faster than any power. (Theorem of Hepp and Jost, see R.Jost: General Theory of Quantum Fields.) □

We therefore can define the Fourier transform of these matrix elements

$$\big\langle \Phi, \Psi\big\rangle_{\mathbf{p}} := (2\pi)^{-3}\int \mathrm{d}^3\mathbf{x}\big\langle \Phi, U(\mathbf{x})\Psi\big\rangle e^{i\mathbf{p}\cdot\mathbf{x}} \ . \tag{3.15}$$

From the formula for the inverse Fourier transform we get

$$\big\langle \Phi, U(\mathbf{x})\Psi\big\rangle = \int \mathrm{d}^3\mathbf{p}\big\langle \Phi, \Psi\big\rangle_{\mathbf{p}} e^{-i\mathbf{p}\cdot\mathbf{x}} \ . \tag{3.16}$$

We now convince ourselves that $\big\langle \cdot, \cdot\big\rangle_{\mathbf{p}}$ is a positive semidefinite sesquilinearform on $\mathfrak{D}$.

PROOF. Let f be a continuous function on $\mathbb{R}^3$ with compact support. Then the integral

$$\Phi_f := \int d^3\mathbf{x} f(\mathbf{x}) e^{i\mathbf{p}\cdot\mathbf{x})} U(\mathbf{x})\Phi \ , \ \Phi \in \mathfrak{D} \tag{3.17}$$

is a well defined element of $\mathfrak{H}$ as a limit of Riemann sums. Thus we have the inequality

$$0 \leq \big\langle \Phi_f, \Phi_f\big\rangle = \int d^3\mathbf{x}\int d^2\mathbf{y}\overline{f(\mathbf{x})}f(\mathbf{y})\big\langle \Phi, U(\mathbf{y}-\mathbf{x})\Phi\big\rangle e^{i\mathbf{p}\cdot(\mathbf{y}-\mathbf{x}))} \tag{3.18}$$

We now substitute $\mathbf{y}$ by $\mathbf{z} = \mathbf{y} - \mathbf{x}$, replace f by f_λ,

$$f_\lambda(\mathbf{x}) = \lambda^{-\frac{3}{2}} f(\lambda\mathbf{x}) \tag{3.19}$$

substitute $\mathbf{x}$ by $\lambda\mathbf{x}$ and obtain

$$0 \leq \int d^3\mathbf{x}\int d^3\mathbf{z}\overline{f(\mathbf{x})}f(\mathbf{x}+\lambda\mathbf{z})\big\langle \Phi, U(\mathbf{z})\big\rangle e^{i\mathbf{p}\mathbf{z}} \tag{3.20}$$

The integrand is bounded by an integrable function uniformly in λ and converges pointwise to $|f(\mathbf{x})|^2\langle\Phi, U(\mathbf{z})\rangle e^{i\mathbf{pz}}$ for $\lambda \to 0$. Hence, according to Lebesgues' theorem of dominated convergence, limit and integration can be interchanged, and we obtain

$$\langle\Phi, \Phi\rangle_{\mathbf{p}} = (2\pi)^{-3} \int d^3\mathbf{x}\langle\Phi, U(\mathbf{x})\Phi\rangle e^{i\mathbf{p}\cdot\mathbf{x}} \geq 0 \ . \tag{3.21}$$

□

As usual we take the quotient with respect to the subspace of vectors with length 0,

$$N_{\mathbf{p}} = \{\Phi \in \mathfrak{D} | \langle\Phi, \Phi\rangle_{\mathbf{p}} = 0\} \ , \tag{3.22}$$

and complete the arising pre-Hilbert space. We obtain a Hilbert space $\mathfrak{H}_{\mathbf{p}}$. In this way every $\Phi \in \mathfrak{D}$ yields a family of vectors $\Phi(\mathbf{p}) := \Phi + N_{\mathbf{p}} \in \mathfrak{H}_{\mathbf{p}}$ with

$$\|\Phi(\mathbf{p})\|^2 = \|\Phi\|^2_{\mathbf{p}} \ . \tag{3.23}$$

The vectors of the original Hilbert space become thus sections of a vector bundle over the mass shell H^+_m. The elements of the bundle might be interpreted as (improper) momentum eigenstates. A smooth structure on the bundle could be implicitly introduced by identifying the elements of $\mathfrak{D}$ with the smooth sections. Actually, the bundle will turn out to be trivial, so that the sections can be replaced by functions on the mass shell with values in some fixed Hilbert space.

The construction described above distinguishes a Lorentz frame. It is, however, possible to modify it a little bit such that it becomes explicitly Lorentz covariant. For this purpose we set

$$\langle\Phi, \Psi\rangle_p = c \int \mathrm{d}^4x \delta(px)\langle\Phi, U(x)\Psi\rangle \tag{3.24}$$

with a factor $c > 0$ to be determined. For $p = (m, 0)$ the new scalar product coincides with the scalar product defined above up to a factor. In the general case we find (with $\omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2}$)

$$\langle\Phi, U(x)\Psi\rangle = \int \mathrm{d}^3\mathbf{q} e^{i\omega(\mathbf{q})x^0 - i\mathbf{q}\cdot\mathbf{x}}\langle\Phi, \Psi\rangle_{\mathbf{q}} \tag{3.25}$$

and thus

$$\begin{aligned} &\int \mathrm{d}^4x \delta(px)\langle\Phi, U(x)\Psi\rangle \\ &= \int \mathrm{d}^3\mathbf{x}\mathrm{d}^3\mathbf{q} \ \omega(\mathbf{p})^{-1} e^{-i\mathbf{x}\cdot(\mathbf{q} - \mathbf{p}\omega(\mathbf{q})/\omega(\mathbf{p}))}\langle\Phi, \Psi\rangle_{\mathbf{q}} \end{aligned} \tag{3.26}$$

We introduce as a new momentum variable $\mathbf{k} = \mathbf{q} - \mathbf{p}\omega(\mathbf{q})/\omega(\mathbf{p})$ with the Jacobian

$$|\det \frac{\partial \mathbf{k}}{\partial \mathbf{q}}| = 1 - \frac{\mathbf{p}\cdot\mathbf{q}}{\omega(\mathbf{p})\omega(\mathbf{q})} \ . \tag{3.27}$$

Integration over $\mathbf{x}$ yields the delta function in $\mathbf{k}$, multiplied by $(2\pi)^3$. The evaluation at $\mathbf{k} = 0$ corresponds to $\mathbf{q} = \mathbf{p}$, hence we obtain the

following relation between the covariant and the non-covariant scalar product

$$\big\langle \Phi, \Psi \big\rangle_p = c(2\pi)^3 \frac{\omega(\mathbf{p})}{m^2} \big\langle \Phi, \Psi \big\rangle_{\mathbf{p}} \ , \ p^0 = \omega(\mathbf{p}), (p^1, p^2, p^3) = \mathbf{p} \ . \quad (3.28)$$

In particular, the null spaces N_p and $N_{\mathbf{p}}$ of both scalar products coincide. We set $c = (2\pi)^{-3} 2m^2$ and arrive at the Lorentz covariant momentum space decomposition of the scalar product in $\mathfrak{H}$,

$$\big\langle \Phi, \Psi \big\rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \big\langle \Phi, \Psi \big\rangle_p \ . \quad (3.29)$$

Here we exploit the fact that $\frac{d^3\mathbf{p}}{2\omega(\mathbf{p})}$ is a Lorentz invariant measure. This may be understood from the somewhat formal relation

$$\int \frac{d^3\mathbf{p}}{2\omega(\mathbf{p})} f(\omega(p), \mathbf{p}) = \int d^4p \delta(p^2 - m^2) \Theta(p_0) f(p) \quad (3.30)$$

with the Heavyside function

$$\Theta(x) = \begin{cases} 1 & , \quad x > 0 \\ 0 & , \quad x \leq 0 \end{cases}$$

The scalar products with different momenta p are related by Lorentz transformations. We have

$$\begin{aligned} \big\langle U(A)\Phi, U(A)\Phi \big\rangle_p =& c \int d^4x \delta(px) \big\langle U(A)\Phi, U(x)U(A)\Phi \big\rangle \\ =& c \int d^4x \delta(px) \big\langle \Phi, U(\Lambda(A)^{-1}x)\Phi \big\rangle \end{aligned} \quad (3.31)$$

hence substituting $y = \Lambda(A)^{-1}x$ and using the fact that Lorentz transformations are volume preserving we find

$$\big\langle U(A)\Phi, U(A)\Phi \big\rangle_p = \big\langle \Phi, \Phi \big\rangle_{p\Lambda(A)} \ . \quad (3.32)$$

In particular, $U(A)\Phi \in N_p$ iff $\Phi \in N_{p\Lambda(A)}$. Thus the unitaries $U(A)$ induce unitary maps between the spaces with sharp momentum,

$$\underline{U}(A) : \mathfrak{H}_{p\Lambda(A)} \to \mathfrak{H}_p \quad (3.33)$$

by

$$\underline{U}(A)(\Phi + N_{p\Lambda(A)}) = U(A)\Phi + N_p \ . \quad (3.34)$$

For each momentum $p \in H_m^+$ there is a subgroup G_p of $\mathrm{SL}(2, \mathbb{C})$, the so-called little group, whose elements do not change p,

$$G_p = \{A \in \mathrm{SL}(2, \mathbb{C}), p\Lambda(A) = p\} \ . \quad (3.35)$$

Then $\underline{U} \restriction_{G_p}$ is an irreducible representation of the little group on $\mathfrak{H}_p$.

The little groups on the same orbit are conjugated within $\mathrm{SL}(2, \mathbb{C})$,

$$q = p\Lambda(A) \Longrightarrow G_p = AG_qA^{-1} \ , \quad (3.36)$$

and the representations of the little groups on the respective Hilbert spaces with sharp momentum satisfy the equation

$$\underline{U}(ARA^{-1}) = \underline{U}(A)\underline{U}(R)\underline{U}(A)^{-1} \ . \tag{3.37}$$

with $R \in G_q$ and $q = p\Lambda(A)$.

The representation U is thus fixed if the representation of some little group is known. In the case of the mass shell H_m^+ we consider the little group for the momentum $(m,0)$. Since $\underset{\sim}{(m,0)} = m1$, the little group at this point is

$$G_{(m,0)} = \{R \in \mathrm{SL}(2,\mathbb{C}) | R^* R = 1\} = \mathrm{SU}(2) \ . \tag{3.38}$$

In the rest system of the particle the little group is thus the covering group of the rotation group.

The irreducible representations of SU(2) can be parametrized by the spin quantum number $s = 0, \frac{1}{2}, 1, \ldots$. The spin, relative to the rest system, characterizes thus the unitary irreducible representations of $\mathcal{P}^c$ with $\mathrm{sp}(P) = H_m^+$.

The existence of the unitary maps $\underline{U}(A)$ allows us to identify the Hilbert spaces $\mathfrak{H}_p$ with a reference space. In the present case one can choose the space $\mathfrak{H}_{(m,0)}$ with vanishing spatial momentum. We then search a family of Lorentz transformations A_p with

$$(m,0)\Lambda(A_p) = p \ . \tag{3.39}$$

As for spacetime vectors x we set for momenta p

$$\underset{\sim}{p} = p^0 \mathbf{1} - \mathbf{p} \cdot \vec{\sigma} = p_0 \mathbf{1} + \sum p_i \sigma_i \ . \tag{3.40}$$

Then

$$px \equiv p_\mu x^\mu = \frac{1}{2} \operatorname{tr} \underset{\sim}{p} x \tag{3.41}$$

Thus

$$p\Lambda(A)x = \frac{1}{2} \operatorname{tr} \underset{\sim}{p} A x A^* = \frac{1}{2} \operatorname{tr} A^* \underset{\sim}{p} A x \ , \tag{3.42}$$

i.e.

$$\underset{\sim}{(p\Lambda(A))} = A^* \underset{\sim}{p} A \ . \tag{3.43}$$

We choose A_p positive definite, i.e. as a pure boost. Then

$$A_p = \sqrt{\frac{\underset{\sim}{p}}{m}} \ . \tag{3.44}$$

We now can identify the Hilbert space $\mathfrak{H}$ with a space $\mathfrak{K}$ of wave functions with values in the space $\mathfrak{H}_{(m,0)}$,

$$\mathfrak{K} = L^2(H_m^+, \mathfrak{H}_{(m,0)}, \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})}) \ .$$

For this purpose we introduce the unitary operator $V : \mathfrak{H} \to \mathfrak{K}$,

$$(V\Phi)(p) = \underline{U}(A_p)\Phi(p) \ . \tag{3.45}$$

The representation U of $\mathcal{P}^c$ on $\mathfrak{H}$ then is equivalent to the representation

$$U'(L) = VU(L)V^{-1} \ , \ L \in \mathcal{P}^c \tag{3.46}$$

on the space $\mathfrak{K}$. The representation U' is explicitly given by

$$(U'(a,A)\Phi)(p) = e^{ipa}\underline{U}(R(p,A))\Phi(p\Lambda(A)) \ . \tag{3.47}$$

Here $R(p,A) = A_p A A^{-1}_{p\Lambda(A)}$ is an element of the little group $G_{(m,0)} = \mathrm{SU}(2)$, the so-called Wigner rotation. Each irreducible representation of $\mathrm{SU}(2)$ thus induces an irreducible representation of the $\mathcal{P}^c$, and all representations of $\mathcal{P}^c$ with $\mathrm{sp}\, P = H_m^+$ are of this form.

For what follows it is useful to describe the irreducible representations of $\mathrm{SU}(2)$ in the following way. Let $\mathfrak{H}^{(s)}$, $s \in \mathbb{N}_0/2$, be the $2s$-fold symmetrical tensor product of $\mathbb{C}^2$. $\mathfrak{H}^{(s)}$ is the linear span of the vectors

$$\underbrace{\xi \otimes \cdots \otimes \xi}_{2s} \ , \ \xi \in \mathbb{C}^2 \ . \tag{3.48}$$

$\mathfrak{H}^{(s)}$ has the dimension $2s+1$. A representation V_s of $\mathrm{SU}(2)$ on $\mathfrak{H}^{(s)}$ is defined by

$$V_s(R)\xi \otimes \cdots \otimes \xi = R\xi \otimes \cdots \otimes R\xi \ . \tag{3.49}$$

One easily verifies that this is the representation with spin s. We now use the fact that this representation has a natural extension to $\mathrm{SL}(2,\mathbb{C})$,

$$V_{s0}(A)\xi \otimes \cdots \otimes \xi = A\xi \otimes \cdots \otimes A\xi \ . \tag{3.50}$$

Using this extension, the Wigner rotation can be factorized in the representation V_{s0},

$$V_{s0}(R(p,A)) = V_{s0}(A_p)V_{s0}(A)V_{s0}(A^{-1}_{p\Lambda(A)}) \ . \tag{3.51}$$

We now set

$$\Psi(p) = V_{s0}(A_p^{-1})\Phi(p) \ . \tag{3.52}$$

Ψ transforms under $(a,A) \in \mathcal{P}^c$ according to

$$(U''(a,A)\Psi)(p) = e^{ipa}V_{s0}(A)\Psi(p\Lambda(A)) \ . \tag{3.53}$$

For U'' to become unitary, the norm of Ψ has to be chosen in such a way that it coincides with the norm of Φ. One obtains

$$\|\Psi\|^2 = \|\Phi\|^2 = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})}\left\langle \Psi(p), V_{s0}\left(\frac{\widetilde{p}}{m}\right)\Psi(p)\right\rangle \ . \tag{3.54}$$

We interpret Ψ as momentum space wavefunction of a particle with mass m and spin s. By Fourier transformation, the corresponding position space wavefunction turns out to be

$$\Psi(x) = (2\pi)^{-3/2}\int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})}e^{-ipx}\Psi(p) \ . \tag{3.55}$$

It transforms under Poincaré transformations according to

$$(U(a,A)\Psi)(x) = V_{s0}(A)\Psi(\Lambda(A)^{-1}(x-a)) \ . \tag{3.56}$$

In the case of mass zero the analysis has to be modified. We first choose some momentum $q \in \partial V_+$, e.g. $q = (\frac{1}{2}, 0, 0, \frac{1}{2})$. Then

$$\widetilde{q} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} . \tag{3.57}$$

The little group is

$$G_q = \{R \in \mathrm{SL}(2,\mathbb{C}), R^* \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} R = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\} . \tag{3.58}$$

We find

$$R = \begin{pmatrix} e^{i\varphi} & 0 \\ a & e^{-i\varphi} \end{pmatrix} , \ \varphi \in \mathbb{R}, a \in \mathbb{C} . \tag{3.59}$$

G_q is the two-fold covering group $\widetilde{\mathrm{E}(2)}$ of the group of motions $\mathrm{E}(2)$ of the euclidean plane. Thereby, the case $a = 0$ corresponds to the rotation by 2φ and the case $\varphi = 0$ to the translation by $(\mathrm{Re}\, a, \mathrm{Im}\, a)$.

In the first step we have to determine representations of the $\widetilde{\mathrm{E}(2)}$. Here one can proceed in the same way as for the Poincaré group. The representations of the subgroup of translations in $\widetilde{\mathrm{E}(2)}$ is determined by the spectrum of their generators K_1, K_2,

$$\underline{U} \begin{pmatrix} 1 & 0 \\ a_1 + ia_2 & 1 \end{pmatrix} = e^{i(K_1 a_1 + K_2 a_2)} . \tag{3.60}$$

In an irreducible representation of $\widetilde{\mathrm{E}(2)}$ the spectrum has to be a circle in $\mathbb{R}^2$ with the origin as center.

If the radius of the circle vanishes, the translations are trivial, and the corresponding representations of the $\widetilde{\mathrm{E}(2)}$ are given by

$$\underline{U} \begin{pmatrix} e^{i\varphi} & 0 \\ a & e^{-i\varphi} \end{pmatrix} = e^{i\varphi n} , \ n \in \mathbb{Z} . \tag{3.61}$$

$n/2$ is called helicity, the associated representation of $\mathcal{P}^c$ is named helicity representation.

In case the radius of the circle does not vanish, the little groups for each spectral value of (K_1, K_2) is $\{\pm\mathbf{1}\} = \mathbb{Z}_2$. This group has the two irreducible representations

$$\pm\mathbf{1} \to 1 , \ \pm\mathbf{1} \to \pm 1 . \tag{3.62}$$

$\mathfrak{H}_q$ can be identified with $L^2(0, 2\pi)$. The representation is

$$\left(\underline{U} \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} \Phi\right)(\alpha) = \Phi(\alpha - 2\varphi) \tag{3.63}$$

with $\Phi(\alpha + 2n\pi) = (\pm 1)^n \Phi(\alpha)$. The associated representations of the $\mathcal{P}^c$ could not yet be used for the description of particles.

In order to present an explicit description of the helicity representations we look for a family $B_p \in \mathrm{SL}(2,\mathbb{C})$, $p \in \partial V_+$ with the property

$$B_p^* \widetilde{q} B_p = \widetilde{p} . \tag{3.64}$$

A possible choice is

$$B_p = (p_0 + p_3)^{-\frac{1}{2}} \begin{pmatrix} p_0 + p_3 & p_1 - ip_2 \\ 0 & 1 \end{pmatrix} . \tag{3.65}$$

Then we may identify $\mathfrak{H}$ with the space of wave functions on ∂V_+ with the scalar product

$$\big\langle \Phi, \Psi \big\rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|} \overline{\Phi(p)} \Psi(p) . \tag{3.66}$$

The representation of the $\mathcal{P}^c$ on this space is given by

$$(U(a, A)\Phi)(p) = e^{ipa} h(p, A)^n \Phi(p\Lambda(A)) \tag{3.67}$$

with the helicity phase

$$h(p, A) = (B_p A B^{-1}_{p\Lambda(A)})_{11} . \tag{3.68}$$

One can obtain the helicity representations with helicity $s \in \mathbb{N}_0/2$ as limit of representations with mass m and spin s in the limit $m \to 0$. For this purpose we consider the one dimensional space $\mathfrak{H}_q$ as that subspace of the $(2s + 1)$-dimensional representation of the SU(2) in which the 3-component of the spin has the eigenvalue s,

$$\mathfrak{H}_q = V_{s0}(\widetilde{q})\mathfrak{H}^{(s)} \tag{3.69}$$

with

$$V_{s0}(A)\xi \otimes \cdots \otimes \xi = A\xi \otimes \cdots \otimes A\xi \tag{3.70}$$

for all 2×2-matrices A. Because of

$$\widetilde{q} B_p A B^{-1}_{p\Lambda(A)} = \widetilde{q} h(p, A) , \ A \in \mathrm{SL}(2, \mathbb{C}) \tag{3.71}$$

the representation of the $\mathrm{SL}(2, \mathbb{C})$ has the form

$$(U(A)V_{s0}(\widetilde{q})\Phi)(p) = V_{s0}(\widetilde{q} B_p A B^{-1}_{p\Lambda(A)})\Phi(p\Lambda(A)) . \tag{3.72}$$

We now set, similar to our proceeding in the case of representations with positive mass

$$\Psi(p) = V_{s0}(B_p^{-1})\Phi(p) . \tag{3.73}$$

Ψ transforms as in the massive case

$$(U''(A)\Psi)(p) = V_{s0}(A)\Psi(p\Lambda(A)) . \tag{3.74}$$

The scalar product is

$$\|\Psi\|^2 = \int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|} \big\langle \Psi(p), V_{s0}(\widetilde{p})\Psi(p) \big\rangle \tag{3.75}$$

We thus obtain essentially the same description as in the massive case, with the difference that before the limit $m \to 0$ the scalar product was multiplied by m^{2s}. Thereby only the component with $s_3 = s$ survives in the limit. Similar constructions are possible for the case of negative helicities.

There remain the representations with imaginary mass ("tachyons") and those with negative energy and positive mass. They can be treated by the same method but do not correspond to physical particles.

We now want to justify the interpretation of the operators P^μ as energy-momentum operators. For this purpose we consider the transition probabilities

$$|\langle \Phi, U(t, \mathbf{v}t)\Phi \rangle|^2 \ . \tag{3.76}$$

In the limit $t \to \infty$ only the component of Φ which moves with the velocity $\mathbf{v}$, will contribute to the transition probability.

Let Φ be smooth with compact support. The integrand in

$$\langle \Phi, U(t, \mathbf{v}t)\Phi \rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} e^{i(\omega(\mathbf{p})-\mathbf{p}\cdot\mathbf{v})t} \|\Phi(p)\|^2 \tag{3.77}$$

oscillates strongly for large t with the exception of the stationary points of the exponent

$$0 = \nabla_{\mathbf{p}}(\omega(\mathbf{p}) - \mathbf{p}\cdot\mathbf{v}) = \frac{\mathbf{p}}{\omega(\mathbf{p})} - \mathbf{v} \ . \tag{3.78}$$

Thus

$$\frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \|\Phi(\mathbf{p})\|^2 \tag{3.79}$$

can be viewed as the probability that the velocity $\mathbf{v}$ is concentrated in the region $\frac{m^2}{2(1-|\mathbf{v}|^2)^2}\mathrm{d}^3\mathbf{v}$ around the point $\mathbf{v} = \frac{\mathbf{p}}{\omega(\mathbf{p})}$. This is precisely the relativistic connection between momentum and velocity. In the case of imaginary mass one finds superluminal velocities $|\mathbf{v}| > 1$, in the case $p_0 < 0$ velocity and momentum point into opposite directions, corresponding to negative energy.

4. Relativistic wave equations

The position space wave function Ψ, obtained by Fourier transformation from the momentum space wave function in irreducible representations of $\mathcal{P}^c$, transforms under Poincaré transformations as expected. Ψ is uniquely determined by its initial values at time $t = 0$,

$$\Psi(t, \mathbf{x}) = 2i \int \mathrm{d}^3\mathbf{y} \frac{\partial}{\partial t} \Delta_+(t, \mathbf{x} - \mathbf{y})\Psi(0, \mathbf{y}) \ . \tag{4.1}$$

with

$$\Delta_+(t, \mathbf{x}) = (2\pi)^{-3} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} e^{-i(\omega(\mathbf{p})t - \mathbf{p}\cdot\mathbf{x})} \ . \tag{4.2}$$

But surprisingly it is not possible to localize the wave function strictly. If the wave function is localized at a given instant of time t in a compact region G,

$$\Psi(t, \mathbf{x}) = 0 \ \forall \mathbf{x} \notin G \ , \tag{4.3}$$

then it expands in an arbitrary short time over the whole space. This behaviour seems to contradict the principle of Einstein causality which

says that signals can propagate at most with the velocity of light. We have the following theorem:

THEOREM II.3. *Let $\Psi \neq 0$ be the position space wave function of a relativistic particle. Then Ψ cannot vanish on an open nonempty subset of Minkowski space.*

PROOF. The momentum space wave function Φ has support in $\overline{V_+}$. Thus its Laplace transform

$$\Psi(x - iy) = (\pi)^{-3/2} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \Phi(p) e^{-ip(x-iy)} \tag{4.4}$$

with $y \in V_+$ is analytic in the so-called tube $\mathcal{T} = \mathbb{M} - iV_+$. The position space wave functions result as boundary values for $y \to 0$. According to the "Edge of the Wedge"-Theorem (a multidimensional version of Schwarz' reflection principle) Ψ is identical zero if its boundary values vanish on a nonempty open set. □

The normalization integral in position space turns out to be

$$\|\Psi\|^2 = 2i^{(2s+1)} \int \mathrm{d}^3\mathbf{x} \big\langle \Psi(t,\mathbf{x}), \frac{\partial}{\partial t} V_{s0}(m^{-1}\widetilde{\partial})\Psi(t,\mathbf{x}) \big\rangle_{\mathfrak{H}^{(s)}} \tag{4.5}$$

independent of t. Here $\partial = (\frac{\partial}{\partial t}, \nabla_{\mathbf{x}})$.

In the case $s = 1/2$ the normalization integral can be simplified. It holds

$$\mathbf{1}\frac{\partial}{\partial t} = \frac{1}{2}(\widetilde{\partial} + \underset{\sim}{\partial}) \ , \ \widetilde{\partial}\underset{\sim}{\partial}\Psi = -m^2\Psi \ . \tag{4.6}$$

We set $\chi = \frac{i}{m}\widetilde{\partial}\Psi$, define a 4-component wave function

$$\psi = \begin{pmatrix} \Psi \\ \chi \end{pmatrix} \tag{4.7}$$

and find

$$\|\Psi\|^2 = m \int \mathrm{d}^3\mathbf{x} \|\psi(t,\mathbf{x})\|^2 \tag{4.8}$$

with the standard scalar product on $\mathbb{C}^4$.

The relativistic energy-momentum relation

$$P_0 = \sqrt{|\mathbf{P}|^2 + m^2} \tag{4.9}$$

is non-polynomial. Therefore it does not yield a differential equation in position space; this is the origin of the nonlocal behaviour of its solutions. It is suggestive to use instead the squared equation

$$P_0^2 = |\mathbf{P}|^2 + m^2 \ . \tag{4.10}$$

This equation has in addition solutions with negative energy. In position space it becomes the Klein-Gordon equation

$$(\Box + m^2)\Psi = 0 \ , \ \Box = \partial^\mu \partial_\mu \ . \tag{4.11}$$

The general solution of the Klein-Gordon equation is a superposition of a solution with positive energy and a solution with negative energy,

$$\varphi(x) = (2\pi)^{-3/2} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \big(a_+(p)e^{-ipx} + a_-(p)e^{ipx}\big) \ . \tag{4.12}$$

For the determination of $\varphi(t,\mathbf{x})$ one needs in addition to $\varphi(0,\mathbf{x})$ also $\frac{\partial}{\partial t}\varphi(0,\mathbf{x})$ as initial conditions (Cauchy data). With

$$\Delta(x) := 2\mathrm{Im}\,\Delta_+(x) \tag{4.13}$$

we obtain

$$\varphi(t,\mathbf{x}) = -\int \mathrm{d}^3\mathbf{y}\big(\frac{\partial}{\partial t}\Delta(t,\mathbf{x}-\mathbf{y})\varphi(0,\mathbf{y}) + \Delta(t,\mathbf{x}-\mathbf{y})\frac{\partial}{\partial t}\varphi(0,\mathbf{y})\big) \ . \tag{4.14}$$

Δ is antisymmetric

$$\Delta(-x) = -\Delta(x) \tag{4.15}$$

and invariant under proper orthochronous Lorentz transformations

$$\Delta(\Lambda x) = \Delta(x) \ , \ \Lambda \in \mathcal{L}_+^\uparrow \ . \tag{4.16}$$

This implies that $\Delta(x)$ vanishes for spacelike x ($x^2 < 0$). Namely, for purely spatial $x = (0,\mathbf{x})$ there is a Lorentz transformation $\Lambda \in \mathcal{L}_+^\uparrow$ with $\Lambda x = -x$, e.g. the rotation by π around an axis orthogonal to $\mathbf{x}$. Since each spacelike point can be brought into the $t = 0$ hyperplane by a suitable Lorentz transformation this property holds for all spacelike points x. Thus for $x^2 < 0$ and a suitable $\Lambda \in \mathcal{L}_+^\uparrow$

$$\Delta(x) = \Delta(\Lambda x) = \Delta(-x) = -\Delta(x) \ , \tag{4.17}$$

i.e. $\Delta(x) = 0$.

This property implies that solutions of the Klein-Gordon-Gleichung propagate in a causal way. Namely, if the Cauchy data at time $t = 0$ are concentrated within a ball of radius r, then, at time t, they are concentrated within a ball of radius $r + |t|$. But because of the nonvanishing component with negative energy the localized solutions cannot be considered as physical particles.

Δ itself is a distributional solution of the Klein-Gordon equation with the Cauchy data

$$\Delta(0,\mathbf{x}) = 0 \ , \ \frac{\partial}{\partial t}\Delta(0,\mathbf{x}) = -\delta(\mathbf{x}) \ . \tag{4.18}$$

For particles with spin $\frac{1}{2}$ one usually uses instead of the Klein-Gordon equation the Dirac equation. Thereby one defines in the massive case a second 2-component wave function

$$\chi = \frac{i}{m}\widetilde{\partial}\varphi \tag{4.19}$$

and combines the two spinors φ und χ to a 4-component bispinor ψ,

$$\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} . \tag{4.20}$$

ψ satisfies the equation

$$\begin{pmatrix} 0 & \partial \\ \widetilde{\partial} & 0 \end{pmatrix} \psi = -im\psi . \tag{4.21}$$

This is the Dirac equation

$$(i\gamma\partial - m)\psi = 0 \tag{4.22}$$

with the γ-matrices

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} , \ \gamma^i = \begin{pmatrix} 0 & -\sigma_i \\ +\sigma_i & 0 \end{pmatrix} . \tag{4.23}$$

The equation remains meaningful in the massless case. In this case ψ decomposes into the two chiral components φ and χ which satisfy the Weyl equations

$$\widetilde{\partial}\varphi = 0 , \ \underset{\sim}{\partial}\chi = 0 . \tag{4.24}$$

While φ transforms under Lorentz transformations according to the defining representation $A \mapsto A$ of $\mathrm{SL}(2,\mathbb{C})$, for χ one gets

$$(U(A)\chi)(x) = (A^*)^{-1}\chi(\Lambda(A)^{-1}x) . \tag{4.25}$$

The Dirac spinor ψ transforms therefore according to the reducible representation

$$A \mapsto \begin{pmatrix} A & 0 \\ 0 & (A^*)^{-1} \end{pmatrix} . \tag{4.26}$$

CHAPTER III

Free fields and Feynman diagrams

1. The scalar field

Let

$$\mathfrak{H}_1 = \{\Phi : H_m^+ \to \mathbb{C} | \|\Phi\|^2 = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} |\Phi(p)|^2 < \infty\} \tag{1.1}$$

be the single particle space of a relativistic particle with spin 0 and mass $m > 0$. As in the nonrelativistic case, we define the bosonic Fock space

$$\mathfrak{H}^+ = \bigoplus_{n=0}^{\infty} \mathfrak{H}_n^+ \tag{1.2}$$

with the n-particle spaces

$$\begin{aligned} \mathfrak{H}_n^+ =& \{\Phi : H_m^+ \times \cdots \times H_m^+ \to \mathbb{C} \text{ symmetric } | \\ & \|\Phi\|^2 = \int \frac{\mathrm{d}^3\mathbf{p}_1}{2\omega(\mathbf{p}_1)} \cdots \int \frac{\mathrm{d}^3\mathbf{p}_n}{2\omega(\mathbf{p}_n)} |\Phi(p_1, \ldots, p_n)|^2 < \infty\} \ . \end{aligned} \tag{1.3}$$

The representation of the Poincaré group $\mathcal{P}^c$ is fixed by the representation on the single particle space,

$$(U(x, A)\Phi)_n(p_1, \ldots, p_n) = e^{i\sum p_k x} \Phi_n(p_1\Lambda(A), \ldots, p_n\Lambda(A)) \ . \tag{1.4}$$

Creation and annihilation operators are declared as in the nonrelativistic case,

$$(a(f)\Phi)_n(p_1 \ldots, p_n) = \sqrt{n+1} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \overline{f(p)} \Phi_{n+1}(p, p_1, \ldots, p_n) \ ,$$

$$(a(f)^*\Phi)_n(p, p_1, \ldots, p_n) = \begin{cases} 0 & , \ n = 0 \\ \frac{1}{\sqrt{n}} \sum f(p_k) \Phi_{n-1}(p_1 \ldots, p_{k-1}, p_{k+1} \ldots, p_n) & , \ n > 0 \end{cases} \tag{1.5}$$

with $f \in \mathfrak{H}_1$. Accordingly we obtain the commutation relations

$$[a(f), a(g)^*] = \langle f, g \rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \overline{f(\mathbf{p})} g(\mathbf{p}) \ , \tag{1.6}$$

$$[a(f), a(g)] = 0 = [a(f)^*, a(g)^*] \ . \tag{1.7}$$

Creation and annihilation operators with sharp momentum are introduced as operator valued distributions by

$$a(f) = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \overline{f(p)} a(p) \ , \tag{1.8}$$

$$a(f)^* = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} a^*(p) f(p) \ . \tag{1.9}$$

They obey the commutation relations

$$[a(p), a^*(q)] = 2\omega(\mathbf{p})\delta(\mathbf{p}-\mathbf{q}) \ , \ [a(p), a(q)] = 0 = [a^*(p), a^*(q)] \ . \tag{1.10}$$

In the sense of operator valued distributions we can also introduce creation and annihilation operators in position space,

$$a(x) = (2\pi)^{-3/2} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} e^{-ipx} a(p) \ , \tag{1.11}$$

$$a^*(x) = (2\pi)^{-3/2} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} a^*(p) e^{ipx} \ . \tag{1.12}$$

These operator valued distributions are solutions of the Klein-Gordon equation. They have the commutation relations

$$[a(x), a^*(y)] = \Delta_+(x-y) \ , \ [a(x), a(y)] = 0 = [a^*(x), a^*(y)] \ . \tag{1.13}$$

We search now for operators which can be interpreted as local measurements (local observables). When $A(x)$ describes a measurable quantity at the point x, the translated observable

$$A(y) = U(y-x) A(x) U(x-y) \tag{1.14}$$

should describe the corresponding quantity at the point y. Einstein's causality principle requires that signals cannot be transmitted with superluminal velocity, thus measurements at spacelike separated points must be mutually compatible, i.e. the corresponding operators have to commute,

$$[A(x), A(y)] = 0 \text{ if } (x-y)^2 < 0 \ . \tag{1.15}$$

An example for such a quantity is the scalar field

$$\varphi(x) = a(x) + a^*(x) \ . \tag{1.16}$$

It is a solution of the Klein-Gordon equation, transforms under Poincaré transformations according to

$$U(x,\Lambda)\varphi(y)U(x,\Lambda)^{-1} = \varphi(\Lambda y + x) \tag{1.17}$$

and fulfils the commutation relation

$$\boxed{[\varphi(x), \varphi(y)] = i\Delta(x-y) \ .} \tag{1.18}$$

Since Δ vanishes for spacelike points, $\varphi(x)$ can be interpreted as a measurement at the point x. From the Cauchy data for Δ we obtain the following commutation relations for φ at coinciding times,

$$[\varphi(t,\mathbf{x}), \varphi(t,\mathbf{y})] = 0 = [\dot\varphi(t,\mathbf{x}), \dot\varphi(t,\mathbf{y})] \ , \tag{1.19}$$

$$[\varphi(t,\mathbf{x}), \dot\varphi(t,\mathbf{y})] = i\delta(\mathbf{x}-\mathbf{y}) \ . \tag{1.20}$$

These relations are the continuous analog of the canonical commutation relations of quantum mechanics,

$$[q_k, q_l] = 0 = [p_k, p_l] \ , \ [q_k, p_l] = i\delta_{kl} \ , \tag{1.21}$$

where the position $\mathbf{x}$ plays the role of the component k and $\dot{\varphi}$ is the canonical conjugated momentum to φ.

The fact that measurements at spacelike separated points have to be compatible does not at all mean that they have to be uncorrelated. Correlations are present, in contrast to the non-relativistic case, already in the vacuum,

$$\begin{aligned}\big\langle\Omega,\varphi(x)\varphi(y)\Omega\big\rangle &= \big\langle\Omega,(a(x)+a^*(x))(a(y)+a^*(y))\Omega\big\rangle \\ &= [a(x),a^*(y)] = \Delta_+(x-y)\ .\end{aligned} \tag{1.22}$$

The vacuum expectation values of products of fields are called Wightman functions,

$$\boxed{W_n(x_1,\ldots,x_n) = \big\langle\Omega,\varphi(x_1)\cdots\varphi(x_n)\Omega\big\rangle\ .} \tag{1.23}$$

Wightman functions are tempered distributions. They can be easily computed by using the commutation relations. We split the field into its annihilation and creation part,

$$\begin{aligned}&\big\langle\Omega,\varphi(x_1)\cdots\varphi(x_n)\Omega\big\rangle \\ =&\big\langle\Omega,(a(x_1)+a^*(x_1))\varphi(x_2)\cdots\varphi(x_n)\Omega\big\rangle\ .\end{aligned} \tag{1.24}$$

The creation operator is brought to the left hand side of the scalar product. There it acts as an annihilation operator and annihilates the vacuum. The annihilation part of the field is permuted with the product of the other fields and annihilates the vacuum on the right hand side. It remains the expectation value of the commutator,

$$\begin{aligned}&[a(x_1),\varphi(x_2)\cdots\varphi(x_n)] \\ =&\sum_{k=2}^{n}\Delta_+(x_1-x_k)\varphi(x_2)\cdots\varphi(x_{k-1})\varphi(x_{k+1})\cdots\varphi(x_n)\ .\end{aligned} \tag{1.25}$$

This implies the recursion relation

$$W_n(x_1,\ldots,x_n) = \sum_{k=2}^{n}\Delta_+(x_1-x_k)W_{n-2}(x_2,\ldots,\hat{x}_k,\ldots,x_n)\ , \tag{1.26}$$

where the symbol $\hat{y}$ means that the argument y is cancelled. Together with the initial conditions $W_0 = 1$ and $W_1 = 0$, this relation determines all Wightman functions.

It has turned out to be very efficient to formulate the resulting combinatorial formulas in terms of graphs. Let $\mathcal{G}_+(n)$ denote the set of all directed graphs with n vertices $\{1,\ldots,n\}$ such that every vertex is connected with exactly one other vertex by a line where the initial point ("source") $s(l)$ of a line l has always a smaller index than the final point $r(l)$ ("range"). For odd n this set is obviously empty. It holds

$$W_n(x_1,\ldots,x_n) = \sum_{G\in\mathcal{G}_+(n)}\prod_l \Delta_+(x_{s(l)}-x_{r(l)})\ . \tag{1.27}$$

It is often useful to describe combinatorial formulas in terms of generating functions. Let f be a real valued test function. According to the Baker-Campbell-Hausdorff formula we have

$$e^{i\int \mathrm{d}^4x\varphi(x)f(x)} = e^{i\int \mathrm{d}^4x a^*(x)f(x)} e^{i\int \mathrm{d}^4x a(x)f(x)} e^{-\frac{1}{2}\int \mathrm{d}^4x\int \mathrm{d}^4y f(x)\Delta_+(x-y)f(y)} . \tag{1.28}$$

Hence

$$\left\langle \Omega, e^{i\int \mathrm{d}^4x\varphi(x)f(x)}\Omega\right\rangle = e^{-\frac{1}{2}\int \mathrm{d}^4x\int \mathrm{d}^4y f(x)\Delta_+(x-y)f(y)} . \tag{1.29}$$

Thus the symmetrized Wightman functions can be obtained as functional derivatives of the generating function

$$\begin{aligned} &i^n W_n(x_1,\ldots,x_n)_{\mathrm{symm}} \\ &= \frac{\delta^n}{\delta f(x_1)\cdots\delta f(x_n)} e^{-\frac{1}{2}\int \mathrm{d}^4x\int \mathrm{d}^4y f(x)\Delta_+(x-y)f(y)} \restriction_{f=0} . \end{aligned} \tag{1.30}$$

Because of the singularities of the 2-point function Δ_+, powers of the field are not well-defined. If one try to compute, e.g., the vacuum expectation value of $\varphi(x)^2$, then one would obtain

$$\left\langle \Omega, \varphi(x)^2\Omega\right\rangle = \Delta_+(0) = (2\pi)^{-3}\int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} = \infty . \tag{1.31}$$

Instead we define normal ordered products (Wick products) by splitting the field into creation and annihilation operators and putting the annihilation operators on the right hand side of the creation operators,

$$:\varphi(x_1)\cdots\varphi(x_n):= \sum_{I\subset\{1,\ldots,n\}} \prod_{i\in I} a^*(x_i) \prod_{j\notin I} a(x_j) . \tag{1.32}$$

In terms of generating functions we find for the normal ordered products

$$i^n :\varphi(x_1)\cdots\varphi(x_n):= \frac{\delta^n}{\delta f(x_1)\cdots\delta f(x_n)} :e^{i\int \mathrm{d}^4x\varphi(x)f(x)}:\restriction_{f=0} \tag{1.33}$$

with

$$:e^{i\int \mathrm{d}^4x\varphi(x)f(x)}:= e^{i\int \mathrm{d}^4x\varphi(x)f(x)} e^{\frac{1}{2}\int \mathrm{d}^4x\int \mathrm{d}^4y f(x)\Delta_+(x-y)f(y)} . \tag{1.34}$$

The normal ordered products can be restricted to coinciding points, and one obtains the so-called Wick powers $:\varphi(x)^n:$. They are themselves quantum fields; their Wightman functions are obtained from the Wightman functions of φ by setting all arguments belonging to the same Wick power to the same value and canceling all graphs which contain lines which connect coinciding points ("tadpoles"),

$$\left\langle \Omega, :\varphi(x_1)^{n_1}:\cdots:\varphi(x_k)^{n_k}:\Omega\right\rangle = \sum_{G\in\mathcal{G}_+(n_1,\ldots,n_k)} c_G \prod_l \Delta_+(x_{s(l)}-x_{r(l)}) . \tag{1.35}$$

Here $\mathcal{G}_+(n_1,\ldots,n_k)$ denotes the set of all directed graphs with k vertices $1,\ldots,k$ with n_i lines at the i-th vertex. The lines always connect two different vertices and are directed to the vertex with the larger

index. The factor c_G counts the number of graphs $G_0 \in \mathcal{G}_+(\sum n_i)$ which by contraction of the first n_1 vertices to the vertex 1, the next n_2 vertices to a vertex 2, etc. yield the graph G.

For calculating the combinatorial factor c_G we replace the vertices $i \in \{1, \dots, k\}$ of the graph G by sets of n_i points such that the n_i lines attached to i are attached to different points. Let l_{ij} be the number of lines between the vertices i and j. There are $\frac{n_i!}{\prod_j l_{ij}!}$ possibilities to decompose the set of points associated to the vertex i into subsets V_{ij} of points with l_{ij} elements. For all vertices we get

$$\frac{\prod n_i!}{\prod_{i \neq j}(l_{ij}!)} \tag{1.36}$$

possibilities. Moreover, there are $l_{ij}!$ different possibilities to choose the lines which connect the subsets V_{ij} and V_{ji}. Altogether we obtain

$$c_G = \frac{\prod_i n_i!}{\prod_{i<j} l_{ij}!} \ . \tag{1.37}$$

The formula above for Wightman functions of Wick powers can also be written in the form

$$\boxed{\left\langle \Omega, \prod_i \frac{:\varphi(x_i)^{n_i}:}{n_i!} \Omega \right\rangle = \sum_G \prod_{i<j} \frac{\Delta_+(x_i - x_j)^{l_{ij}}}{l_{ij}!} \ .} \tag{1.38}$$

It is crucial that the products of Δ_+ occuring in the formula are well-defined distributions (see the exercises).

As an example for a more general Wick polynomial we consider the energy density

$$h(x) = \frac{m^2}{2} :\varphi(x)^2: + \frac{1}{2} :\dot{\varphi}(x)^2: + \frac{1}{2} :\nabla\varphi(x)^2: \ . \tag{1.39}$$

In terms of creation and annihilation operators at sharp momenta we get

$$h(x) =$$
$$\frac{1}{2}(2\pi)^{-3} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \int \frac{\mathrm{d}^3\mathbf{q}}{2\omega(\mathbf{q})} \big(2(m^2 + \omega(\mathbf{p})\omega(\mathbf{q}) + \mathbf{p}\cdot\mathbf{q}) a^*(p)a(q)e^{i(p-q)x}$$
$$+ (m^2 - \omega(\mathbf{p})\omega(\mathbf{q}) - \mathbf{p}\cdot\mathbf{q})(a(p)a(q)e^{-i(p+q)x} + a^*(p)a^*(q)e^{i(p+q)x})\big) \ . \tag{1.40}$$

We see that the energy density has a term which preserves the number of particles, a term which creates and another term which annihilates a pair of particles. After integration over $\mathbf{x}$ at fixed time the terms which change the particle number get a factor $(2\pi)^3\delta(\mathbf{p}+\mathbf{q})$ and the term which preserves the particle number get a factor $(2\pi)^3\delta(\mathbf{p}-\mathbf{q})$.

The term $m^2 - \omega(\mathbf{p})\omega(\mathbf{q}) - \mathbf{p}\cdot\mathbf{q}$ vanishes at $\mathbf{p} = -\mathbf{q}$, hence only the particle number preserving term contributes to the integral. We find

$$\int \mathrm{d}^3\mathbf{x}\, h(t,\mathbf{x}) = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})}\omega(\mathbf{p})a^*(p)a(p) = H \ . \tag{1.41}$$

This justifies the interpretation of h as energy density.

The classical energy density of the Klein-Gordon field has the same form, up to the disappearance of normal ordering. Contrary to the classical energy density the energy density of the quantum field is not always positive. This follows from the fact that the expected energy density in the vacuum vanishes, but not its fluctuations. Formally we have

$$\big\langle \Omega, h(x)^2\Omega\big\rangle = \infty \ . \tag{1.42}$$

Thus the vacuum is not an eigenstate of the energy density, and in a decomposition into eigenstates the contributions of negative and positive eigenvalues have to cancel.

The existence of states with negative energy density has consequences for General Relativity. Many results of this theory rely on the assumption that the energy density satisfies some positivity condition. If this assumption is violated, exotic spacetimes with worm holes etc. are possible. In the science fiction literature this possibility is often exploited. A careful analysis, however, shows that negative energy densities are possible only under very restrictive conditions. For instance the integral of the energy density over time at a given spatial point must be nonnegative, thus a negative energy density at some time has to be compensated at other times, actually with some additional amount of positive energy ("quantum interest").

As in the nonrelativistic case, we want now to investigate the effect of a time dependent source on our system. Let the time dependent Hamiltonian be

$$H(t) = H_0 - \int \mathrm{d}^3\mathbf{x}\, \varphi(0,\mathbf{x})f(t,\mathbf{x}) \tag{1.43}$$

with a real valued test function $f \in \mathcal{D}(\mathbb{M})$. The calculations are completely analogous to those in chapter I. The time evolution operator in the interaction picture is

$$V(t,s) = Te^{i\int_{s<x^0<t}\mathrm{d}^4x\varphi(x)f(x)} \ , \tag{1.44}$$

and for the S-matrix we find

$$S(f) = Te^{i\int \mathrm{d}^4x\varphi(x)f(x)} = e^{i\int \mathrm{d}^4x\varphi(x)f(x)}e^{-i\alpha} \tag{1.45}$$

with $\alpha = \frac{1}{2}\int_{x^0>y^0}\mathrm{d}^8(x,y)f(x)f(y)\Theta(x^0-y^0)\Delta(x-y)$. In terms of normal ordered products this yields

$$S(f) =: e^{i\int \mathrm{d}^4x\varphi(x)f(x)} : \big\langle \Omega, S(f)\Omega\big\rangle \tag{1.46}$$

with

$$\left\langle \Omega, S(f)\Omega \right\rangle = e^{-\frac{1}{2}\int \mathrm{d}^8(x,y) f(x) f(y) (\Delta_+(x-y) + i\Theta(x^0-y^0)\Delta(x-y))} \tag{1.47}$$

Here Θ is the Heaviside function

$$\Theta(x^0) = \begin{cases} 1 & , \quad x^0 > 0 \\ 0 & , \quad x^0 \leq 0 \end{cases} . \tag{1.48}$$

The symmetrized part of the coefficient of $f(x)f(y)$ in the integrand of the exponent is the Feynman propagator

$$\begin{aligned} i\Delta_F(x-y) &= \Theta(x^0-y^0)\Delta_+(x-y)) + \Theta(y^0-x^0)\Delta_+(y-x)) \\ &= \left\langle \Omega, T\varphi(x)\varphi(y)\Omega \right\rangle . \end{aligned} \tag{1.49}$$

One calculates

$$\Delta_F(x) = \lim_{\varepsilon \downarrow 0} (2\pi)^{-4} \int \mathrm{d}^4 p \frac{e^{-ipx}}{p^2 - m^2 + i\varepsilon} . \tag{1.50}$$

The Feynman propagator is invariant under the full Lorentz group (including the time inversion) and is a Green's function of the Klein-Gordon equation,

$$(\Box + m^2)\Delta_F(x) = -\delta(x) . \tag{1.51}$$

The formulas above yield a definition of the time ordered products of the field φ via functional derivatives

$$\begin{aligned} & i^n T\varphi(x_1)\cdots\varphi(x_n) \\ =& \frac{\delta^n}{\delta f(x_1)\cdots\delta f(x_n)} : e^{i\int \mathrm{d}^4 x \varphi(x) f(x)} : e^{-\frac{i}{2}\int \mathrm{d}^8(x,y) f(x) f(y) \Delta_F(x-y)} \restriction_{f=0} \end{aligned} \tag{1.52}$$

The resulting combinatorial formula can be visualized in terms of graphs. Let $\mathcal{G}(n)$ be the set of graphs with vertices $1, \ldots, n$ and undirected lines which either connect two vertices (inner lines $l \in K_1$) or go from a vertex to the exterior (external lines $l \in K_2$). Each vertex v is a boundary point $v \in \partial l$ of precisely one line l. Then

$$T\varphi(x_1)\cdots\varphi(x_n) = \sum_{G \in \mathcal{G}(n)} : \prod_{K_2} \varphi(x_{\partial l}) : \prod_{K_1} \left\langle \Omega, T \prod_{v \in \partial l} \varphi(x_v)\Omega \right\rangle . \tag{1.53}$$

Formally one obtains analogous formulas if one inserts higher Wick powers as interactions. Again the S-matrix can be computed in terms of time ordered products which, as in the case of Wightman functions, arise from the graphs $G \in \mathcal{G}(n_1, \ldots, n_k)$, which originate by merging the appropriate vertices in graphs from $\mathcal{G}(\sum n_i)$. One finds

$$\boxed{T \prod_i \frac{:\varphi(x_i)^{n_i}:}{n_i} = \sum_{G \in \mathcal{G}(n_1,\ldots,n_k)} : \prod_j \frac{\varphi(x_j)^{l_i}}{l_i!} : \prod_{j<m} \frac{(i\Delta_F(x_j - x_m))^{l_{jm}}}{l_{jm}!} .} \tag{1.54}$$

Unfortunately, the products of Feynman propagators occuring in this formula are, in general, ill defined. This is in remarkable contrast to the corresponding formulas in the graphical presentation of Wightman functions and is the origin of ultraviolet divergences in quantum field theory.

The arising graphs can be classified according to the number of independent loops. Graphs without loops are called trees. The corresponding terms in the expansion of time ordered products are well defined, since the arguments of the Feynman propagators are independent differences $y_{jm} = x_j - x_m$, and no higher powers of propagators occur, $l_{jm} = 0, 1$.

As an example we consider the S matrix

$$S = Te^{i\int \mathrm{d}^4x f(x):\varphi(x)^3:} \ . \tag{1.55}$$

up to second order. We find

$$S = 1 + i\int \mathrm{d}^4x f(x) : \varphi(x)^3 : -\frac{1}{2}\int \mathrm{d}^8(x,y) f(x)f(y)T :\varphi(x)^3::\varphi(y)^3: \tag{1.56}$$

with

$$\begin{aligned} T\frac{:\varphi(x)^3::\varphi(y)^3:}{3!\cdot 3!} &= \frac{:\varphi(x)^3\varphi(y)^3:}{3!\cdot 3!} + \frac{:\varphi(x)^2\varphi(y)^2:}{2!\cdot 2!} i\Delta_F(x-y) \\ &+ :\varphi(x)\varphi(y): \frac{(i\Delta_F(x-y))^2}{2!} + \frac{(i\Delta_F(x-y))^3}{3!} \ . \end{aligned} \tag{1.57}$$

We now expand the Wick polynomials into sums of products of creation and annihilation operators. For the part which maps 2-particle states into 2-particle states we obtain, e.g.

$$S_{2\to 2} = 1-$$
$$9i\int \mathrm{d}^8(x,y) f(x)f(y)(a^*(x)^2a(y)^2 + 2a^*(x)a^*(y)a(x)a(y))\Delta_F(x-y) \ . \tag{1.58}$$

The first term can be interpreted as describing the annihilation of a pair and the creation of a "virtual" particle at the point y, followed by a decay of the virtual particle into 2 particles at the point x. The second term may be interpreted as the exchange of a virtual particle between a particle at x and another particle at y.

2. Fields with spin; the connection between spin and statistics

Let $\mathfrak{H}_1$ be the state space of a particle with spin s and mass $m > 0$. For the sake of a simpler notation we set $m \equiv 1$ in this section. This means that momenta and energies are measured in units of m.

According to chapter II, $\mathfrak{H}_1$ is the representation space of $\mathcal{P}^c$ for the irreducible representation labeled by the pair (m, s). We realize it

in the form

$$\mathfrak{H}_1 = \{\Phi : H_1^+ \to \mathfrak{H}^{(s)} | \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \big\langle \Phi(p), V_{s0}(\widetilde{p})\Phi(p)\big\rangle < \infty\} \ . \tag{2.1}$$

with the notations from chapter II. The representation of $\mathrm{SL}(2,\mathbb{C})$ on $\mathfrak{H}_1$ is given by

$$(U(A)\Phi)(p) = V_{s0}(A)\Phi(p\Lambda(A)) \ . \tag{2.2}$$

For the following it is important that $\mathfrak{H}^{(s)}$ has a natural complex conjugation such that

$$\overline{V_{s0}(A)\Phi} = V_{s0}(\overline{A})\overline{\Phi} \ , \ \Phi \in \mathfrak{H}^{(s)} \ . \tag{2.3}$$

We choose a real orthonormal basis in $\mathfrak{H}^{(s)}$. The elements of $\mathfrak{H}_1$ can then be represented as functions with $(2s+1)$ components .

The symmetrical and anti-symmetrical n-particle spaces $\mathfrak{H}_n^{\pm}$ consist of functions of n momenta with $(2s+1)^n$ components with the respective symmetry. The scalar product is

$$\|\Phi\|^2 = \int \frac{\mathrm{d}^3\mathbf{p}_1}{2\omega(\mathbf{p}_1)} \cdots \frac{\mathrm{d}^3\mathbf{p}_n}{2\omega(\mathbf{p}_n)} \sum_{k_1,\ldots,k_n,j_1,\ldots,j_n} \tag{2.4}$$

$$\overline{\Phi_{k_1,\ldots,k_n}(p_1,\ldots,p_n)} V_{s0}(\widetilde{p})_{k_1 j_1} \cdots V_{s0}(\widetilde{p})_{k_n j_n} \Phi_{j_1,\ldots,j_n}(p_1,\ldots,p_n)$$

Bosonic and fermionic Fock spaces $\mathfrak{H}^{\pm}$ can be constructed as before. Annihilation operators are defined in terms of the scalar product on $\mathfrak{H}_1$,

$$(a(f)\Phi)_{n;k_1,\ldots,k_n}(p_1,\ldots,p_n) =$$

$$\sqrt{n+1} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \sum_{kj} \overline{f_k(p)} V_{s0}(\widetilde{p})_{kj} \Phi_{n+1;j,k_1,\ldots,k_n}(p,p_1,\ldots,p_n) \ . \tag{2.5}$$

The representation U of the Poincaré group on $\mathfrak{H}_{\pm}$ arises naturally from the single particle representation. In particular we have

$$U(x,A)a(f)U(x,A)^* = a(U(x,A)f) \ . \tag{2.6}$$

The commutation relations are

$$a(f)a(g) = \varepsilon a(g)a(f) \ , \ a(f)a(g)^* = \varepsilon a(g)^* a(f) + \big\langle f, g\big\rangle \ , \tag{2.7}$$

with $\varepsilon = 1$ for Bose and $\varepsilon = -1$ for Fermi statistics.

As in the case $s = 0$, we now look for fields with local commutation relations,

$$\varphi_i(x)\varphi_j(y) = \varepsilon\varphi_j(y)\varphi_i(x) \ , \ (x-y)^2 < 0 \ . \tag{2.8}$$

These fields should transform covariantly under Poincaré transformations,

$$U(y,A)\varphi_i(x)U(y,A)^{-1} = \sum_j \varphi_j(\Lambda(A)x + y)S(A)_{ji} \ , \tag{2.9}$$

with a finite dimensional representation S of $\mathrm{SL}(2,\mathbb{C})$ with $\overline{S(A)_{ji}} = S(\overline{A})_{ji}$.

The fields are operator valued distributions. Therefore we consider the smeared fields

$$\varphi(f) = \int \mathrm{d}^4x \sum \varphi_i(x) f_i(x) \tag{2.10}$$

where f is test function with values in the representation space of S. On the space $\mathcal{D}$ of these test functions, $\mathrm{SL}(2,\mathbb{C})$ acts by

$$(V(A)f)(x) = S(A)f(\Lambda(A)^{-1}x) \ . \tag{2.11}$$

The covariance condition on φ is

$$U(A)\varphi(f)U(A)^{-1} = \varphi(V(A)f) \ . \tag{2.12}$$

We first search for a linear combination of creation and annihilation operators which transforms covariantly,

$$\varphi(f) = a(Q(f)) + a(R(f))^* \ , \tag{2.13}$$

with maps R and Q from the test function space $\mathcal{D}$ into the single particle space. R is linear and Q is antilinear. Q and R should intertwine the representations V and U,

$$QV(A) = U(A)Q \ , RV(A) = U(A)R \ . \tag{2.14}$$

The condition on R can be easily satisfied. We choose $S = V_{s0}$ and set

$$R(f)(p) = (2\pi)^{-\frac{3}{2}} \int \mathrm{d}^4x e^{ipx} f(x) \ . \tag{2.15}$$

But in the case of Q antilinearity creates problems. We therefore search for an operator C ("charge conjugation") in $\mathfrak{H}_1$ with $U(A)C\overline{\Phi} = C\overline{U(A)\Phi}$. An operator with this property is

$$(C\Phi)(p) = V_{s0}(\underset{\sim}{p\zeta})\Phi(p) \tag{2.16}$$

with $\zeta = i\sigma_2$. Namely, with $(\Lambda = \Lambda(A))$

$$(\underset{\sim}{p\Lambda}) = A^{-1}\underset{\sim}{p}(A^*)^{-1} \ , \ (A^*)^{-1}\zeta = \zeta\overline{A} \tag{2.17}$$

and

$$\begin{aligned} (U(A)C\overline{\Phi})(p) &= V_{s0}(A(\underset{\sim}{p\Lambda})\zeta)\overline{\Phi}(p\Lambda) \ , \\ (CU(\overline{A})\overline{\Phi})(p) &= V_{s0}(\underset{\sim}{p\zeta\overline{A}})\overline{\Phi}(p\Lambda) \ . \end{aligned}$$

the assertion follows from the equation

$$\underset{\sim}{p\zeta}\overline{A} = A(\underset{\sim}{p\Lambda})\zeta \ . \tag{2.18}$$

We now choose

$$Q(f) = CR(\overline{f}) \ . \tag{2.19}$$

We check whether the field defined in this way has local commutation relations. We have

$$\varphi(f)\varphi(g) - \varepsilon\varphi(g)\varphi(f) = \big\langle Q(f), R(g)\big\rangle - \varepsilon\big\langle Q(g), R(f)\big\rangle \ . \tag{2.20}$$

A direct calculation yields

$$\begin{aligned}\big\langle Q(f), R(g)\big\rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})}\big\langle V_{s0}(\zeta)R(\overline{f})(p), R(g)(p)\big\rangle = \\ \int \mathrm{d}^8(x,y)\Delta_+(x-y)\sum_{jk} V_{s0}(\zeta)_{kj}f_j(x)g_k(y)\ .\end{aligned} \tag{2.21}$$

Since $V_{s0}(\zeta)_{kj} = (-1)^{2s}V_{s0}(\zeta)_{jk}$ we finally obtain

$$\begin{aligned}\varphi(f)\varphi(g) - \varepsilon\varphi(g)\varphi(f) = \\ \int \mathrm{d}^8(x,y)\sum_{jk} V_{s0}(\zeta)_{kj}f_j(x)g_k(y)\left(\Delta_+(x-y) - \varepsilon(-1)^{2s}\Delta_+(y-x)\right)\ .\end{aligned} \tag{2.22}$$

If $\varepsilon = (-1)^{2s}$, i.e. if the spin statistics connection holds, then we find in the last bracket the scalar commutator function $i\Delta$ which we know from the scalar theory. In this case the covariant fields satisfy local commutation relations

$$[\varphi_j(x), \varphi_k(y)]_\pm = i\Delta(x-y)V_{s0}(\zeta)_{kj}\ . \tag{2.23}$$

In case of a wrong connection of spin and statistics local commutation relations are not valid since Δ_+ does not vanish for spacelike arguments.

The fields constructed above are non hermitean, for $s > 0$. The hermitean conjugated fields can be represented as linear combinations of the fields φ_i. We consider the operator

$$\varphi(f)^* = a(Q(f))^* + a(R(f))\ . \tag{2.24}$$

We find

$$Q(f) = CR(\overline{f}) = R(V_{s0}(-i\underset{\sim}{\partial}\zeta)\overline{f}) = R(\mathcal{C}\overline{f})\ . \tag{2.25}$$

with $\mathcal{C} = V_{s0}(i\underset{\sim}{\partial}\zeta)$. as well as

$$Q(\mathcal{C}\overline{f}) = CR(\overline{\mathcal{C}}f) = R(\mathcal{C}\overline{\mathcal{C}}f)\ . \tag{2.26}$$

with $\overline{\mathcal{C}} = V_{s0}(-i\zeta\widetilde{\partial})$. Since $\zeta^2 = -1$, $\underset{\sim}{\partial}\widetilde{\partial} = \Box$, $R\circ\Box^{2s} = (-m^2)^{2s}R$ and $m = 1$ it follows $R\circ\mathcal{C}\overline{\mathcal{C}} = R$, hence

$$Q(\mathcal{C}\overline{f}) = R(f)\ . \tag{2.27}$$

We thus obtain the Majorana condition

$$\varphi(f)^* = \varphi(\mathcal{C}\overline{f})\ . \tag{2.28}$$

The hermitean conjugated fields, implicitly defined by

$$\varphi(f)^* = \int \mathrm{d}^4x \sum_j \varphi_j^*(x)\overline{f_j(x)} \tag{2.29}$$

are given by

$$\varphi_k^*(x) = V_{s0}(i\widetilde{\partial}\zeta)_{kj}\varphi_j(x) \tag{2.30}$$

The physical content of the Majorana condition is that the particle is its own antiparticle. In case particle and antiparticle are different, the single particle space is the direct sum of two irreducible representation spaces with the same mass and spin. Let a denote the annihilation operator for the first particle and b that for the second particle. The local fields can be defined by

$$\varphi(f) = a(Q(f)) + b(R(f))^* \ . \tag{2.31}$$

In this case the fields φ_i commute with each other. For the (anti-)commutator of φ and φ^* we find as in the Majorana case (with the right spin and statistics connection)

$$[\varphi(f)^*, \varphi(g)]_\pm = i^{2s+1} \int \mathrm{d}^8(x,y) \overline{f_k(x)} g_j(y) V_{s0} (\widetilde{\partial})_{kj} \Delta(x-y) \ . \tag{2.32}$$

hence

$$[\varphi_k^*(x), \varphi_j(y)]_\pm = i^{2s+1} V_{s0} (\widetilde{\partial})_{kj} \Delta(x-y) \ . \tag{2.33}$$

3. The free Dirac field

We specialize now to the case $s = \frac{1}{2}$. The commutation relations at coinciding times assume the simple form

$$\{\varphi_k^*(t,\mathbf{x}), \varphi_l(t,\mathbf{y})\} = \frac{1}{m} \delta(\mathbf{x}-\mathbf{y}) \delta_{kl} \ , \ k,l = 1,2. \tag{3.1}$$

Here we used the facts that $\Delta(0,\mathbf{x}) = 0$ and $\partial_t \Delta(0,\mathbf{x}) = -\delta(\mathbf{x})$.

φ fulfils the Klein-Gordon equation and transforms under Poincaré-Transformationen according to

$$U(x,A) \varphi_k(y) U(x,A)^{-1} = \sum_{j=1}^{2} \varphi_j(\Lambda(A)y + x) A_{jk} \ . \tag{3.2}$$

We now introduce a second field χ by

$$\chi_k(x) = \frac{i}{m} \sum_j (\widetilde{\partial}\zeta)_{kj} \varphi_j(x) \ . \tag{3.3}$$

χ also fulfils the Klein-Gordon equation. In the Majorana case we have $\chi = \varphi^*$. χ transforms under Poincaré transformations according to the conjugated representation of $\mathrm{SL}(2,\mathbb{C})$,

$$U(x,A) \chi_k(y) U(x,A)^{-1} = \sum_{j=1}^{2} \chi_j(\Lambda(A)y + x) \overline{A_{jk}} \ . \tag{3.4}$$

We combine the fields φ and χ to a 4-component field ψ,

$$\psi = \sqrt{m} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} \ . \tag{3.5}$$

ψ transforms under Poincaré transformations according to

$$U(x,A)\psi_\alpha(y)U(x,A)^{-1} = \sum_{\beta=1}^{4} \psi_\beta(\Lambda(A)y+x)S(A)_{\beta\alpha} \tag{3.6}$$

with $S(A) = \begin{pmatrix} A & 0 \\ 0 & \tilde{A} \end{pmatrix}$ and fulfils the Dirac equation

$$(i\partial\!\!\!/ - m)\psi = 0 \tag{3.7}$$

with

$$\partial\!\!\!/ = \begin{pmatrix} 0 & -\zeta\overset{\sim}{\partial} \\ \overset{\sim}{\partial}\zeta & 0 \end{pmatrix} = \partial_\mu\gamma^\mu \ . \tag{3.8}$$

Here the γ matrices were chosen in the form

$$\gamma^\mu = \begin{pmatrix} 0 & -\zeta\sigma_\mu \\ \sigma^\mu\zeta & 0 \end{pmatrix} \ . \tag{3.9}$$

with $\sigma_0 = 1$ and $\sigma^\mu = g^{\mu\nu}\sigma_\nu$. This form arises by the similarity transformation with the unitary matrix $\begin{pmatrix} \zeta & 0 \\ 0 & 1 \end{pmatrix}$ from the form given in Chapter II. They fulfil also the anti commutation relations

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \ , \tag{3.10}$$

and transform under Lorentz transformations according to

$$S(A)^T\gamma^\mu S(A^{-1})^T = \Lambda(A)^\mu{}_\nu\gamma^\nu \ . \tag{3.11}$$

Moreover, the representation S satisfies the pseudo unitarity condition

$$S(A)^*\gamma^0 S(A) = \gamma^0 \ . \tag{3.12}$$

The arising commutation relations of the Dirac field are

$$\{\psi_\alpha(x), \psi^*_\beta(y)\} = i(S(x-y)\gamma^0)_{\alpha\beta} \tag{3.13}$$

with the 4×4-matrix valued anticommutator function

$$S(x-y) = (i\partial\!\!\!/ + m)\Delta(x-y) \ . \tag{3.14}$$

In particular one finds for the simultaneous anticommutation relations

$$\{\psi_\alpha(t,\mathbf{x}), \psi^*_\beta(t,\mathbf{y})\} = \delta_{\alpha\beta}\delta(\mathbf{x}-\mathbf{y}) \ . \tag{3.15}$$

The 2-point Wightman function is

$$\left\langle \Omega, \psi_\alpha(x)\psi^*_\beta(y)\Omega \right\rangle = (S_+(x-y)\gamma^0)_{\alpha\beta} \tag{3.16}$$

with

$$S_+(x-y) = (i\partial\!\!\!/ + m)\Delta_+(x-y) \ . \tag{3.17}$$

For the definition of time ordered products one has to take into account the Fermi statistics. One sets

$$T\psi_\alpha(x)\psi^*_\beta(y) = \Theta(x^0-y^0)\psi_\alpha(x)\psi^*_\beta(y) - \Theta(y^0-x^0)\psi^*_\beta(y)\psi_\alpha(x) \ . \tag{3.18}$$

This yields the Feynman propagator of the Dirac field

$$\left\langle \Omega, T\psi_\alpha(x)\psi^*_\beta(y)\Omega \right\rangle = i(S_F(x-y)\gamma^0)_{\alpha\beta} \tag{3.19}$$

with

$$S_F(x) = (i\partial\!\!\!/ + m)\Delta_F(x) = (2\pi)^{-4}\int \mathrm{d}^4p\, e^{-ipx}(p\!\!\!/ - m + i\varepsilon)^{-1} \ . \tag{3.20}$$

Normal ordered products are defined similar as in the Bose case. The difference is that each term is multiplied with the sign of a permutation which brings the creation operators to the left hand side of the annihilation operators (different permutations with this property have the same sign). Again, normal ordered products can be defined on coinciding points; one obtains as in the case of the Bose field the Wick polynomials as new local fields.

An example for such a field is the current

$$j^\mu(x) =: \overline{\psi}(x)\gamma^\mu\psi(x): \ . \tag{3.21}$$

The notation here should be understood in the following way: ψ is a column vector, ψ^* the adjoint row vector. $\overline{\psi}$ is the row vector defined by $\psi^*\gamma^0$. Because of the pseudo-unitarity of the representation of $\mathrm{SL}(2,\mathbb{C})$, $:\overline{\psi}\psi:$ is a scalar field, and j transforms as a Lorentz vector,

$$U(A)j^\mu(x)U(A)^{-1} = \Lambda(A^{-1})^\mu{}_\nu j^\nu(\Lambda(A)x) \ . \tag{3.22}$$

j is a conserved current, $\partial_\mu j^\mu = 0$. Hence the integral

$$Q = \int \mathrm{d}^3\mathbf{x} j^0(t,\mathbf{x}) \ . \tag{3.23}$$

(the charge operator) is a conserved quantity (i.e. independent of t). The integral above exists in the sense of expectation values on a dense domain and coincides with the difference of the number of particles and the number of antiparticles. The spectrum of Q therefore is the set of integers.

The Dirac equation is a first order differential equation Therefore we expect that its solutions are determined by their initial values at a given time. Indeed, one finds

$$\psi(t,\mathbf{x}) = \int \mathrm{d}^3\mathbf{y}\, S(t,\mathbf{x}-\mathbf{y})\psi(0,\mathbf{y}) \ . \tag{3.24}$$

As an example for an interaction we consider the coupling of the Dirac field to an external electromagnetic field with vector potential A_μ. The interacting Dirac field satisfies the equation

$$(i\partial\!\!\!/ - eA\!\!\!/ - m)\psi = 0 \ . \tag{3.25}$$

We identify the interacting field at time $t = 0$ with the free Dirac field and obtain the time dependent Hamiltonian

$$H(t) = H_0 - e\int \mathrm{d}^3\mathbf{x}\, j^\mu(0,\mathbf{x})A_\mu(t,\mathbf{x}) \ . \tag{3.26}$$

This yields the S-matrix

$$S = Te^{i\int \mathrm{d}^4x\, j^\mu(x) eA_\mu(x)} \ . \tag{3.27}$$

Again, as in the case of the scalar field, the time ordered products can be visualized in terms of Feynman graphs. The arising graphs have directed lines, and every vertex is source and range of a line. The internal lines $l \in K_2$ correspond to Feynman propagators iS_F, the incomimg external lines $l \in K_+$ correspond to a factor $\overline{\psi}$, the outgoing external lines $l \in K_-$ correspond to a factor ψ in the normal ordered product. One finds

$$\begin{aligned} Tj^{\mu_1}(x_1)\cdots j^{\mu_n}(x_n) &= \sum_G \varepsilon \sum_{\alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_n} \\ &: \prod_{l\in K_+} \overline{\psi}_{\beta_{\partial l}}(x_{\partial l}) \prod_{l\in K_-} \psi_{\alpha_{\partial l}}(x_{\partial l}) : \prod_{l\in K_2} iS_F(x_{s(l)} - x_{r(l)})_{\alpha_{s(l)}\beta_{r(l)}} \prod_{i=1}^n \gamma^{\mu_i}_{\beta_i\alpha_i} \end{aligned} \tag{3.28}$$

with the sign ε of the permutation of Fermi fields.

Let us consider first the connected graphs. There are two types, the linear graphs and the loop graphs. For the linear graphs with n vertices there are $n!$ possible orderings in which the path composed of the lines of the graph meets the vertices. Let the order be labeled by a permutation $\sigma \in \mathrm{S}_n$. The corresponding graph yields the contribution

$$:\overline{\psi}(x_{\sigma(1)})\gamma^{\mu_{\sigma(1)}} iS_F(x_{\sigma(1)} - x_{\sigma(2)}) \cdots iS_F(x_{\sigma(n-1)} - x_{\sigma(n)})\gamma^{\mu_{\sigma(n)}}\psi(x_{\sigma(n)}): \ . \tag{3.29}$$

We now set

$$\Gamma(x,y) = ie\not{A}(x) iS_F(x-y) \tag{3.30}$$

and consider Γ as an integral operator,

$$(\Gamma f)(x) = \int \mathrm{d}^4y\, \Gamma(x,y) f(y) \ , \ f \in \mathcal{S}(\mathbb{R}^4, \mathbb{C}^4) \ . \tag{3.31}$$

This is justified, if $A \in \mathbf{S}(\mathbb{R}^4, \mathbb{R}^4)$. Then the sum over all linear graphs with n vertices yields after integration

$$n! : \int \mathrm{d}^8(x,y) \overline{\psi}(x) \Gamma^{n-1}(x,y) ie\not{A}(y)\psi(y) :\equiv n! : \int \overline{\psi}\Gamma^{n-1} ie\not{A}\psi : \ . \tag{3.32}$$

For the loop graphs with n vertices there are only $(n-1)!$ different possibilities since the choice of the initial vertex is arbitrary. Again we denote the order of vertices in the graph by a permutation σ and obtain (taking into account the sign arising from the permutation of Fermi fields) the contribution

$$-\operatorname{tr} \gamma^{\mu_{\sigma(1)}} iS_F(x_{\sigma(1)} - x_{\sigma(2)}) \cdots \gamma^{\mu_{\sigma(n)}} iS_F(x_{\sigma(n)} - x_{\sigma(1)}) \ , \tag{3.33}$$

where tr denotes the trace of a 4×4-matrix. Integration finally yields

$$-(n-1)!\,\mathrm{Tr}\,\Gamma^n := (n-1)! \int \mathrm{d}^4x\,\mathrm{tr}\,\Gamma^n(x,x)\ . \tag{3.34}$$

Since every graph can be decomposed into connected components, we now can compute the contributions of all graphs. Let G be a graph with n vertices and connected components $L_1, \dots L_k$, $S_1, \dots S_l$, where the linear graphs L_i have n_i vertices and the loop graphs S_j have m_j vertices. The corresponding contribution is

$$:\prod_{i=1}^{k} \int \overline{\psi}\Gamma^{n_i-1} ie\not{A}\psi : \prod_{j=1}^{l} \mathrm{Tr}\,\Gamma^{m_j}\ . \tag{3.35}$$

There are

$$\frac{n!}{\prod n_i! \prod m_j!} \tag{3.36}$$

different possibilities to decompose a set of n vertices into subsets with $n_1, \dots, n_k, m_1, \dots, m_l$ vertices. But the numbering of the components is arbitrary, hence we have to divide by $k!l!$. The contribution of all these graphs is

$$n! : \frac{1}{k!}\prod_{i=1}^{k} \int \overline{\psi}\Gamma^{n_i-1} ie\not{A}\psi : \frac{1}{l!}\prod_{j=1}^{l}(-\,\mathrm{Tr}\,\frac{\Gamma^{m_j}}{m_j}) \tag{3.37}$$

After summation over n in the exponential series for S the summation over n_i and m_j is only resticted by the conditions $n_i > 0$ and $m_j > 1$. We use

$$\sum_{n=0}^{\infty} \Gamma^n = (1-\Gamma)^{-1} \tag{3.38}$$

(geometric series) and

$$\sum_{n=2}^{\infty} \frac{1}{n}\Gamma^n = -\ln(1-\Gamma) - \Gamma \tag{3.39}$$

(Taylor series of $\ln(1+x)$), sum over k and l and obtain finally for the S-matrix the closed formula

$$S =: e^{\int \overline{\psi}(1-\Gamma)^{-1} ie\not{A}\psi} : e^{\mathrm{Tr}\ln(1-\Gamma)+\Gamma}\ . \tag{3.40}$$

By using the formula

$$e^{\mathrm{Tr}\,B} = \det e^B\ , \tag{3.41}$$

valid for square matrices B, we can write the vacuum expectation value of S as

$$\big\langle \Omega, S\Omega \big\rangle = \det(1-\Gamma)e^{\Gamma}\ . \tag{3.42}$$

The considerations above solve the combinatorial problem associated to the expansion into graphs. It remains to show that the derived expressions are mathematically well defined. For the term in the normal ordered product this can be done for $A \in \mathbf{S}(\mathbb{R}^4, \mathbb{R}^4)$ and e sufficiently

small. The vacuum expectation value of the S-matrix, however, diverges and has to be renormalized.

By similar methods as in the construction of the S-matrix we can construct also the interacting field. Let ψ be the operator valued distribution

$$\psi(f) = \int \mathrm{d}^4x \sum_\alpha \psi_\alpha(x) f_\alpha(x) \ . \tag{3.43}$$

If ψ satisfies the Dirac equation with the vector potential A_μ, then

$$\begin{aligned} 0 &= \int \mathrm{d}^4x \sum_{\alpha\beta} (i\partial\!\!\!/ - eA\!\!\!/ - m)_{\alpha\beta} \psi_\beta(x) f_\alpha(x) \\ &= \int \mathrm{d}^4x \sum_{\alpha\beta} \psi_\beta(x) (-i\partial\!\!\!/ - eA\!\!\!/ - m)_{\alpha\beta} f_\alpha(x) \ . \end{aligned} \tag{3.44}$$

i.e. with $D = i\partial\!\!\!/ - m$ we have

$$\psi((D^t - eA\!\!\!/^t)f) = 0 \ , \tag{3.45}$$

where the exponent t denotes the transposed operator. In the case of an integral operator G in $\mathcal{S}(\mathbb{R}^4, \mathbb{C}^4)$ with integral kernel $G(x, y)$ the integral kernel of the transposed operator is

$$G^t(x, y) = G(y, x)^T \ . \tag{3.46}$$

Here T refers to the transposition of the 4×4-matrix.

The free Dirac field ψ_0 fulfils the equation

$$\psi_0(D^t f) = 0 \ . \tag{3.47}$$

We therefore look for a linear operator

$$W : \mathcal{S}(\mathbb{R}^4, \mathbb{C}^4) \to \mathcal{S}(\mathbb{R}^4, \mathbb{C}^4) \tag{3.48}$$

with

$$D^t = W(D^t - eA\!\!\!/^t) \ . \tag{3.49}$$

Then

$$\psi = \psi_0 \circ W \tag{3.50}$$

solves the Dirac equation with interaction.

Let G be a Green's function of the operator $D - eA\!\!\!/$,

$$(D - eA\!\!\!/)G(x, y) = \delta(x - y)\mathbf{1} \ . \tag{3.51}$$

Then a solution is

$$W = D^t G^t \ . \tag{3.52}$$

Let G_0 be a Green's function of the free operator D. Then one obtains a series representation of G by

$$G = \sum_{n=0}^{\infty} G_0 (eA\!\!\!/ G_0)^n \tag{3.53}$$

As Green's function G_0 we could choose the Feynman propagator S_F. It is more convenient to take the retarded Green's function,

$$G_0(x,y) = S_{\rm ret}(x-y) = \Theta(x^0-y^0)S(x-y) \tag{3.54}$$

In this case the interacting field coincides with the free field ψ_0 in the limit $t \to -\infty$, $\psi_{\rm in} = \psi_0$. In the limit $t \to +\infty$ one obtains another free field, the outgoing free field ψ_{out}. It is related to the incomimg free field by

$$\psi_{\rm in}(f)S = S\psi_{\rm out}(f) \ . \tag{3.55}$$

4. Electrodynamics

Up to now we considered relativistic multiparticle systems and found that they give rise to quantized field theories, under the condition that the correct connection between spin and statistics was chosen. We now want to start from a classical field theory and associate to it a quantum theory. We will see that this delivers directly a particle interpretation of electrodynamics which is known since the work of Planck and Einstein.

The Maxwell equations for the magnetic field $\mathbf{B}$ and the electrical field $\mathbf{E}$ in the absence of charges are

$$\begin{aligned} \operatorname{curl}\mathbf{B} = \frac{\partial}{\partial t}\mathbf{E} \quad &, \quad \operatorname{curl}\mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B} \\ \operatorname{div}\mathbf{E} \quad &, \quad \operatorname{div}\mathbf{B} = 0 \end{aligned}$$

By introducing the electromagnetical field strength tensor $F_{\mu\nu}$ the equations can be written in Lorentz covariant form: we set

$$\mathbf{E} = (F_{01}, F_{02}, F_{03}) \ , \ \mathbf{B} = (F_{32}, F_{13}, F_{21}) \ , \ F_{\mu\nu} = -F_{\nu\mu} \ , \tag{4.1}$$

and find

$$\partial^\mu F_{\mu\nu} = 0 \ , \ \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0 \ . \tag{4.2}$$

In particular, $F_{\mu\nu}$ satisfies the wave equation,

$$\begin{aligned} \Box F_{\mu\nu} = \partial^\rho\partial_\rho F_{\mu\nu} &= -\partial^\rho(\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu}) \\ &= \partial_\mu\partial^\rho F_{\rho\nu} - \partial_\nu\partial^\rho F_{\rho\mu} = 0 \ . \end{aligned} \tag{4.3}$$

We now want to realize the fields $F_{\mu\nu}$ as hermitean (tempered) operator valued distributions in some Hilbertraum $\mathfrak{H}$. The Maxwell equation should hold; furthermore there should be a representation U of the Poincaré group in $\mathfrak{H}$ with the property

$$U(x,\Lambda)F_{\mu\nu}(y)U(x,\Lambda)^{-1} = F_{\rho\sigma}(\Lambda y + x)\Lambda^\rho{}_\mu\Lambda^\sigma_\nu \tag{4.4}$$

(covariance), such that the energy is positive, independent of the Lorentz system

$$U(x) = e^{iPx} \ , \ \operatorname{sp} P \subset \overline{V_+} \ . \tag{4.5}$$

(spectrum condition). Measurements of field strengths at spacelike separated points should be compatible,

$$[F_{\mu\nu}(x), F_{\rho\sigma}(y)] = 0 \ , \ (x-y)^2 < 0 \ . \tag{4.6}$$

(locality). Finally we require the existence of a unique (up to phase) Poincaré invariant unit vector Ω (for the description of the vacuum), and we assume that Ω is cyclic for the fields $F_{\mu\nu}$, i.e. the vectors

$$\sum \prod F(f_i)\Omega \ , \tag{4.7}$$

$F(f) = \frac{1}{2}\int \mathrm{d}^4x F_{\mu\nu}(x) f^{\mu\nu}(x)$, $f^{\mu\nu} \in \mathcal{S}(\mathbb{R}^4)$, should span a dense subspace of $\mathfrak{H}$.

We first look at the 2-point function

$$W_{\mu\nu\sigma\tau}(x,y) = \big\langle \Omega, F_{\mu\nu}(x) F_{\sigma\tau}(y)\Omega \big\rangle \ . \tag{4.8}$$

Due to the translation invariance of Ω and the covariance of $F_{\mu\nu}$ the 2-point function depends only on the difference $x-y$, hence can be written in the form

$$W_{\mu\nu\sigma\tau}(x,y) = \int \mathrm{d}^4p e^{-ip(x-y)} \rho_{\mu\nu\sigma\tau}(p) \tag{4.9}$$

with tempered distributions $\rho_{\mu\nu\sigma\tau}$. Because of the spectrum condition the support of ρ is contained in $\overline{V_+}$. Since $F_{\mu\nu}$ is a solution of the Maxwell equations, we have

$$p^2 \rho_{\mu\nu\sigma\tau}(p) = 0 \ . \tag{4.10}$$

Thus ρ is of the form

$$\rho_{\mu\nu\sigma\tau}(p) = P_{\mu\nu\sigma\tau}(p)\delta(p^2)\Theta(p_0) \ , \tag{4.11}$$

where P is a function on ∂V_+ which transforms covariantly under Lorentz transformations. The only functions with this property are linear combinations of $p_\mu p_\nu p_\sigma p_\tau$, $g_{\mu\nu}p_\sigma p_\tau$, $g_{\mu\nu}g_{\sigma\tau}$ and the functions arising by permuting the indices. Because of the antisymmetry of the electromagnetic field tensor the 4-fold products of momenta cannot occur. The Maxwell equation $\partial^\mu F_{\mu\nu} = 0$ excludes also the momentum independent terms. Hence we obtain

$$P_{\mu\nu\sigma\tau} = -c\big(g_{\mu\sigma}p_\nu p_\tau - g_{\nu\sigma}p_\mu p_\tau - g_{\mu\tau}p_\nu p_\sigma + g_{\nu\tau}p_\mu p_\sigma\big) \tag{4.12}$$

with an up to now undetermined constant $c > 0$.

We conclude that the 2-point function is

$$W_{\mu\nu\sigma\tau}(x,y) = c(2\pi)^3\big(g_{\mu\sigma}\partial_\nu\partial_\tau - g_{\nu\sigma}\partial_\mu\partial_\tau - g_{\mu\tau}\partial_\nu\partial_\sigma + g_{\nu\tau}\partial_\mu\partial_\sigma\big)D_+(x-y) \tag{4.13}$$

with the 2-point function of a massless free scalar field

$$D_+(x) = (2\pi)^{-3}\int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|}e^{-ipx} = \Delta_+(x, m=0) \ . \tag{4.14}$$

D_+ can be explicitly computed: Let f be a test function from the Schwartz space $\mathcal{S}(\mathbb{R}^4)$. As a distribution, D_+ is defined by

$$\int \mathrm{d}^4x D_+(x)f(x) = (2\pi)^{-3}\int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|}\int \mathrm{d}^4x f(x)e^{-ipx} \ . \tag{4.15}$$

We can interchange the integration over the angular part of $\mathbf{p}$ with the integration over x and obtain

$$(2\pi)^{-2}\int_0^\infty \mathrm{d}p \int \mathrm{d}^4x f(x)\frac{e^{-ip(t-r)} - e^{-ip(t+r)}}{2ir} \tag{4.16}$$

with $x = (t, \mathbf{x})$ and $r = |\mathbf{x}|$. We replace t by $t - i\varepsilon$ with $\varepsilon > 0$. In the limit $\varepsilon \to 0$ we reobtain the previous integral. For $\varepsilon > 0$ we can interchange the integration over p with the x integration and get

$$\lim_{\varepsilon\downarrow 0}\int \mathrm{d}^4x f(x)\frac{-1}{4\pi^2\big((t-i\varepsilon)^2 - r^2\big)}, \tag{4.17}$$

hence

$$\begin{aligned} D_+(x) &= \lim_{\varepsilon\downarrow 0}\frac{-1}{4\pi^2\big((t-i\varepsilon)^2 - r^2\big)} \\ &= \frac{-1}{4\pi^2x^2} \text{ for } x^2 \neq 0 \ . \end{aligned} \tag{4.18}$$

One sees explicitly that D_+ is the boundary value of an analytical function in the tube $\mathcal{T}_+ = \mathbb{R}^4 - iV_+$.

We now consider the following subspace $\mathfrak{H}_1$ of $\mathfrak{H}$,

$$\mathfrak{H}_1 = \{F(f)\Omega, f^{\mu\nu} \in \mathcal{S}(\mathbb{R}^4), f^{\mu\nu} = -f^{\nu\mu}\} \ . \tag{4.19}$$

On this space the scalar product is determined by the 2-point function,

$$\begin{aligned} \|F(f)\Omega\|^2 = \frac{1}{4}\int \mathrm{d}^8(x,y)\overline{f^{\mu\nu}(x)}f^{\sigma\tau}(y)W_{\mu\nu\sigma\tau}(x-y) = \\ c\int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|}g_{\mu\tau}p_\nu p_\sigma(2\pi)^4\overline{\hat{f}^{\mu\nu}(-p)}\hat{f}^{\sigma\tau}(-p) \ . \end{aligned} \tag{4.20}$$

We set $\hat{f}^\mu(p) = (2\pi)^2 p_\nu \hat{f}^{\mu\nu}(-p)$ and find

$$\|F(f)\Omega\|^2 = -c\int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|}\overline{\hat{f}^\mu(p)}\hat{f}_\mu(p) \ . \tag{4.21}$$

Because of the antisymmetry of $f^{\mu\nu}$, $\hat{f}^\mu$ satisfies the condition $p_\mu\hat{f}^\mu(p) = 0$. On the space $\mathcal{V}$ of these functions, equation (4.21) defines a positive semidefinite scalar product. The single particle space $\mathfrak{H}_1$ is the completion of the quotient space of $\mathcal{V}$ modulo the null space of the scalar product.

After forming the quotient, $\hat{f}^\mu$ has only two independent components. Let e.g. $p = (|p|, 0, 0, |p|)$. Then $\hat{f}^0(p) = \hat{f}^3(p)$ and

$$\overline{\hat{f}_\mu(p)}\hat{f}^\mu(p) = -|\hat{f}_1(p)|^2 - |\hat{f}_2(p)|^2 \ . \tag{4.22}$$

Because of the covariance of $F_{\mu\nu}$ and the invariance of Ω, $\mathfrak{H}_1$ is invariant under Poincaré transformations. We have

$$U(x,\Lambda)F(f)\Omega = F(f_{(x,\Lambda)})\Omega \tag{4.23}$$

mit

$$f^{\mu\nu}_{(x,\Lambda)}(y) = \Lambda^{\mu}{}_{\sigma}\Lambda^{\nu}{}_{\tau}f^{\sigma\tau}(\Lambda^{-1}(y-x)) \ . \tag{4.24}$$

On the vector space $\mathcal{V}$ one gets the representation

$$(D(x,\Lambda)\hat{f})^{\mu}(p) = e^{ipx}\Lambda^{\mu}{}_{\nu}\hat{f}^{\nu}(p\Lambda) \ . \tag{4.25}$$

Since the scalar product (4.21) is invariant under Poincaré transformations, D induces a unitary representation on the quotient space. It coincides with the representation U on $\mathfrak{H}_1$.

We now convince ourselves that this representation is a direct sum of representations with helicity ± 1. For this purpose we consider the representation d of the little group G_q of the momentum $q = (\frac{1}{2}, 0, 0, \frac{1}{2})$ on the space

$$V = \{b \in \mathbb{C}^4, b^{\mu}q_{\mu} = 0\} \ . \tag{4.26}$$

We find

$$(d(\Lambda)b)^{\mu} = \Lambda^{\mu}{}_{\nu}b^{\nu} \ , \ \Lambda \in G_q \ . \tag{4.27}$$

V has the invariant positive semidefinite scalar product

$$\langle b, b\rangle = -b_{\mu}b^{\mu} \tag{4.28}$$

with the null space

$$N = \{\lambda\hat{q}, \lambda \in \mathbb{C}\} \ , \tag{4.29}$$

where $\hat{q}$ denotes the vector corresponding to the covector q, $\hat{q}^{\mu} = q^{\mu}$. d induces a unitary representation on the 2-dimensional space V/N.

According to chapter II, G_q is isomorphic to the euclidean group in 2 dimensions. Its finite dimensional representations are trivial on the translations of the euclidean space $\mathbb{R}^2$; the rotations correspond to rotations of the 1-2 plane in the Lorentz group and act on V in a natural way by rotating the 1- and 2-component of b. Eigenvectors of these rotations are

$$e_{\pm} = \frac{1}{\sqrt{2}}(0, 1, \pm i, 0) \ . \tag{4.30}$$

The 1-dimensional subspaces $\{\lambda e_{\pm} + N, \lambda \in \mathbb{C}\}$ of V/N are invariant under G_q. This follows from

$$d(\mathbf{a}, R)(\lambda\hat{q} + \mathbf{b}) = (\lambda + \langle \mathbf{a}, R\mathbf{b}\rangle)\hat{q} + R\mathbf{b} \tag{4.31}$$

with the column vectors $\mathbf{a} \in \mathbb{R}^2$ und $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{C}^2$ and the natural scalar product in $\mathbb{C}^2$.

Let Λ_p be a Lorentz transformation with the property

$$p = q\Lambda_p \ . \tag{4.32}$$

One may for instance choose $\Lambda_p = \Lambda(B_p)$ with the notations from chapter II. But there is no choice such that Λ_p becomes a smooth function of p.

We now set $e_\pm(p) = \Lambda_p^{-1} e_\pm$. Then

$$\Lambda e_\pm(p) = \Lambda\Lambda_p^{-1} e_\pm = h(p,\Lambda)^{\pm 1} e_\pm(p\Lambda^{-1}) + \lambda\widehat{p\Lambda^{-1}} \tag{4.33}$$

with some $\lambda \in \mathbb{C}$. Here $h(p,\Lambda)$ is the Wigner phase, i.e. the value of $\Lambda_{p\Lambda^{-1}}\Lambda\Lambda^{-1}$ in the 1-dimensional representation of G_q on the subspace of V/N spanned by e_+.

We now introduce creation and annihilation operators for momentum p and helicity $\lambda = \pm 1$ with the commutation relations

$$[a_\lambda(p), a_{\lambda'}(p')] = 0 = [a^*_\lambda(p), a^*_{\lambda'}(p')] \tag{4.34}$$

and

$$[a_\lambda(p), a^*_{\lambda'}(p')] = 2|\mathbf{p}|\delta_{\lambda\lambda'}\delta(\mathbf{p}-\mathbf{p}') \tag{4.35}$$

On the corresponding Fock space we define fields by

$$F_{\mu\nu}(x) = (2\pi)^{-\frac{3}{2}} \int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|} \sum_\lambda \big(e_{\mu\nu}(p,\lambda) a_\lambda(p) e^{-ipx} + \overline{e_{\mu\nu}}(p,\lambda) a^*_\lambda(p) e^{ipx}\big) \tag{4.36}$$

with

$$e_{\mu\nu}(p,\lambda) = -i(p_\mu e_{\lambda,\nu}(p) - p_\nu e_{\lambda,\mu}(p)) \ . \tag{4.37}$$

One easily verifies that these fields satisfy the required covariance conditions and have a 2-point function of the desired form, with $c = (2\pi)^{-3}$. We convince ourselves that this value of c is determined by the principle of correspondence

Namely, the Hamiltonian for free photons is

$$H = \int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|} |\mathbf{p}| \sum_\lambda a^*_\lambda(p) a_\lambda(p) \ . \tag{4.38}$$

For the value of c given above this coincides with the spatial integral of the normal ordered energy density,

$$H = \frac{1}{2} \int \mathrm{d}^3\mathbf{x} \big(:\mathbf{E}(0,\mathbf{x})^2: + :\mathbf{B}(0,\mathbf{x})^2:\big) \ . \tag{4.39}$$

We now have found a solution of the problem of defining a quantized version of electrodynamics. It can be shown (Jost-Schroer-Pohlmeyer Theorem) that this solution is unique. Here the locality axiom plays an important role. For the 2-point function it is automatically fulfilled.

The field $F_{\mu\nu}$ satisfies the commutation relations

$$[F_{\mu\nu}(x), F_{\sigma\tau}(y)] = -\big(g_{\mu\sigma}\partial_\nu\partial_\tau - g_{\nu\sigma}\partial_\mu\partial_\tau - g_{\mu\tau}\partial_\nu\partial_\sigma + g_{\nu\tau}\partial_\mu\partial_\sigma\big) iD(x-y) \tag{4.40}$$

with the commutator function D of the massless scalar field

$$D(x) = -i(D_+(x) - D_+(-x)) = -\frac{1}{2\pi}\,\mathrm{sign}(x^0)\delta(x^2) \ . \tag{4.41}$$

As a simple example for an interaction we treat the coupling to an external conserved current j_μ. For mathematical convenience we assume that j_μ is a smooth function with compact support. The field equation for the interacting field is

$$\partial^\mu F_{\mu\nu} = j_\nu \ . \tag{4.42}$$

Let $F^0_{\mu\nu}$ denote the free Maxwell field. We consider the classical retarded solution of the inhomogeneous Maxwell equations,

$$f_{\mu\nu}(x) = -\int \mathrm{d}^4y \big(\partial_\mu D_{\mathrm{ret}}(x-y) j_\nu(y) - \partial_\nu D_{\mathrm{ret}}(x-y) j_\mu(y)\big) \ , \tag{4.43}$$

with $D_{\mathrm{ret}}(x) = \Theta(x^0)D(x)$. Then

$$F_{\mu\nu}(x) = F^0_{\mu\nu}(x) + f_{\mu\nu}(x) \tag{4.44}$$

is a solution of the field equation with the asymptotic behaviour

$$F^{\mathrm{in}}_{\mu\nu} = F^0_{\mu\nu} \ , \ F^{\mathrm{out}}_{\mu\nu} = F^0_{\mu\nu} + f^{\mathrm{out}}_{\mu\nu} \ . \tag{4.45}$$

Here $f^{\mathrm{out}}_{\mu\nu}$ is the free electromagnetic wave generated by the classical current,

$$f^{\mathrm{out}}_{\mu\nu} = -\int \mathrm{d}^4y \big(\partial_\mu D(x-y) j_\nu(y) - \partial_\nu D(x-y) j_\mu(y)\big) \ . \tag{4.46}$$

The S-matrix results, up to a phase factor, from the relation

$$F^{\mathrm{in}}_{\mu\nu}(x) S = S F^{\mathrm{out}}_{\mu\nu}(x) \tag{4.47}$$

and has the form

$$S = e^{i\alpha} e^{-ia(\hat{f})^*} e^{-ia(\hat{f})} \tag{4.48}$$

with $\hat{f}_\mu(p) = \hat{j}_\mu(p)$ and $\alpha \in \mathbb{C}$. α is given by

$$\alpha = \frac{1}{2}\int \mathrm{d}^4x \mathrm{d}^4y j_\mu(x) j_\nu(y) D_F^{\mu\nu}(x-y) \tag{4.49}$$

with the Feynman propagator $D_F^{\mu\nu} = g^{\mu\nu} D_F$, $D_F = \Delta_F(m=0)$. The massless scalar Feynman propagator is explicitly given by

$$D_F(x) = \Theta(x^0) D_+(x) + \Theta(-x^0) D_+(-x) = \lim_{\varepsilon \downarrow 0} \frac{-1}{4\pi^2(x^2 - i\varepsilon)} \ . \tag{4.50}$$

If our initial state is the vacuum, the state at late times will be a coherent state, characterized by the classical solution (4.46) or, equivalently, by the Fourier transform of the current, restricted to ∂V_+. In particular, the expected number of photons which are created is

$$\langle N \rangle = \big\langle S\Omega, NS\Omega \big\rangle = -\int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|} \overline{\hat{j}^\mu(p)} \hat{j}_\mu(p) \ . \tag{4.51}$$

As an example for a classical conserved current we choose the current which results from the movement of a classical point particle. Let

$x(\tau)$ be the world line of the particle. It is a smooth timelike curve ($\dot{x}^2 > 0$) and is future directed ($\dot{x}^0 > 0$). The associated current is

$$j^\mu(y) = e \int \mathrm{d}\tau \dot{x}^\mu \delta(y - x(\tau)) \ . \tag{4.52}$$

This is, of course, not a test function, and we have to check whether the formulas derived above remain meaningful.

First of all, it is no longer true that the interacting field coincides with the free field at early times. This is forbidden by Gauss' law since a charge is present at all times. For the same reason, if we had chosen the advanced solution in our definition of the interacting field in (4.44), the field would not coincide with the free field at late times. We may, however, ask whether the difference between the retarded and the advanced interacting field can be described by an S-matrix in Fock space. Using $D_{\mathrm{adv}}(x) = -\Theta(x^0)D(x)$ we see that formally the equation

$$F^{\mathrm{ret}}_{\mu\nu}(x)S = SF^{\mathrm{adv}}_{\mu\nu}(x) \tag{4.53}$$

has the same solution as in the previous case. It is, however, not clear whether the current defines an element of the single photon space.

The Fourier transform of the current is

$$\hat{j}^\mu(p) = (2\pi)^{-2} e \int \mathrm{d}\tau \dot{x}^\mu(\tau) e^{-ipx(\tau)} \ . \tag{4.54}$$

Let τ denote the proper time of the particle and assume that $\ddot{x} = 0$ holds outside of a finite interval $[\tau_1, \tau_2]$. Let $u_1 = \dot{x}(\tau)$ for $\tau < \tau_1$ and $u_2 = \dot{x}(\tau)$ for $\tau > \tau_2$. For momenta $p \neq 0$, $p \in \partial V_+$ we have $p\dot{x} > 0$. We now interpret the current, restricted to the boundary of the forward lightcone, as a distribution in the spatial momentum where the zero momentum is excluded. Let f be a test function in $\mathcal{D}(\mathbb{R}^3)$ with $0 \notin \operatorname{supp} f$. Then we have

$$\int d^3\mathbf{p} \hat{j}^\mu(p) f(\mathbf{p}) := \int d\tau \int d^3\mathbf{p} e^{-ipx(\tau)} f(\mathbf{p}) \ . \tag{4.55}$$

According to the principle of stationary phase, the momentum space integral is strongly decaying in τ. Physically this is connected with the fact that a massless particle in 3 spatial dimensions moves exactly with the velocity of light, hence the probability to find it in a timelike direction decreases fast with time.

We now write

$$e^{-ipx(\tau)} = \frac{i}{p\dot{x}(\tau)} \frac{d}{d\tau} e^{-ipx(\tau)} \ , \tag{4.56}$$

insert this identity into the formula (4.55) and perform a partial integration with respect to τ. The boundary terms vanish, and we obtain

$$\hat{j}^\mu(p) = -i \int d\tau e^{-ipx(\tau)} \frac{d}{d\tau} \frac{\dot{x}^\mu}{p\dot{x}(\tau)} \ . \tag{4.57}$$

By assumption the derivative of $\dot{x}^\mu$ vanishes for $\tau \notin [\tau_1, \tau_2]$. Using again the principle of the stationary phase, but now with respect to the τ-integral we observe that $\hat{j}^\mu$ is strongly decaying for large momenta on ∂V_+. In order to study the behavior for small momenta, we use the fact that the τ-integral is restricted to the interval $[\tau_1, \tau_2]$. Using again partial integration we obtain

$$\hat{j}^\mu(p) = (2\pi)^{-2} e \left(\frac{u_1^\mu}{ipu_1} e^{-ipx(\tau_1)} - \frac{u_2^\mu}{ipu_2} e^{-ipx(\tau_2)} + \int_{\tau_1}^{\tau_2} \mathrm{d}\tau \dot{x}^\mu(\tau) e^{-ipx(\tau)} \right) \tag{4.58}$$

We observe that this function is in $L^2(\mathbb{R}^3, \frac{d^3\mathbf{p}}{2|\mathbf{p}|})$ iff $u_1 = u_2$. The mean photon number in the state $S\Omega$ is

$$\langle N \rangle = \big\langle S\Omega, NS\Omega \big\rangle = - \int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|} \overline{\hat{j}^\mu(p)} \hat{j}_\mu(p) \ . \tag{4.59}$$

It holds

$$\langle N \rangle < \infty \Longleftrightarrow u_1 = u_2 \ . \tag{4.60}$$

In the case $\langle N \rangle = \infty$, S cannot be defined as an operator in Fock space, and there is no vector Φ in Fock space which plays the role of $S\Omega$, i.e. satisfies the equation

$$\big\langle \Phi, e^{iF^{\mathrm{ret}}(f)} \Phi \big\rangle = \big\langle \Omega, e^{iF^{\mathrm{adv}}(f)} \Omega \big\rangle \ . \tag{4.61}$$

The expectation values in this state, however, are well defined by the right hand side. The number of photons with energy $> \varepsilon$, e.g., is given by

$$\langle N_\varepsilon \rangle = - \int_{|\mathbf{p}|>\varepsilon} \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|} \overline{\hat{j}^\mu(p)} \hat{j}_\mu(p) < \infty \ . \tag{4.62}$$

A surprising consequence of the vacuum fluctuations of the electromagnetic field is the so-called Casimir effect. Casimir started from the assumption that also the fluctuating vacuum fields should vanish inside an ideal conductor. This would result in a change of the mean field energy and should become visible by additional forces on the conductor.

We consider the 2-point function of the electromagnetic field in the presence of a conducting plane at $x^3 = 0$. On this plane the tangential components of the electric field and the normal component of the magnetic field should vanish,

$$F_{01} = F_{02} = F_{12} = 0 \text{ for } x^3 = 0, \ . \tag{4.63}$$

Due to the Maxwell equation $\partial^\mu F_{\mu\nu} = 0$ the normal derivatives of the other components of the field have to vanish on the conducting planes,

$$\partial_3 F_{03} = \partial_3 F_{13} = \partial_3 F_{23} = 0 \text{ for } x^3 = 0 \ . \tag{4.64}$$

The associated 2-point function of the electromagnetic field can be obtained by the use of mirror charges, as in classical electrodynamics. Let S be the reflection on the plane,

$$(Sx)^\mu = S^\mu{}_\nu x^\nu \ \text{ with } S^\mu{}_\nu = (-1)^{\delta_{\mu,3}} \delta^\mu{}_\nu \ . \tag{4.65}$$

Then we obtain a 2-point function of the electromagnetic field which satisfies the boundary conditions in terms of the 2-point function without boundary conditions by

$$\langle F_{\mu\nu}(x)F_{\sigma\tau}(y)\rangle = W_{\mu\nu\sigma\tau}(x-y) - W_{\mu\nu\lambda\rho}(x-Sy)S^{\lambda}{}_{\sigma}S^{\rho}{}_{\tau} \qquad (4.66)$$

This is the 2-point function of the ground state of the electromagnetic field in the presence of the conducting plane for x and y on the same side. If x and y are separated by the plane, the 2-point function vanishes.

We may now compute the expectation values of the Wick squares of the components of the electromagnetic field. Normal ordering amounts to subtraction of the vacuum expectation values, i.e. to omitting the first term on the right hand side in the formula above. We can then set $y = x$ and find for the expectation values of the normal ordered squares of the components of the electromagnetic field at a point x with distance d from the plane

$$\langle :E_i^2(x):\rangle = -\langle :B_i^2(x):\rangle = \frac{1}{16\pi^2 d^4} \qquad (4.67)$$

We observe in particular that the energy density is not changed by the presence of the conducting plane.

One may ask how the square of the electric field can be measured. For this purpose we consider a neutral atom at the point x with vanishing permanent dipole moment and polarizability α. Under the influence of an external electric field a dipole moment

$$\mathbf{p} = 4\pi\alpha\mathbf{E}(x)$$

is induced which contributes

$$V = -2\pi\alpha\mathbf{E}(x)^2$$

to the potential energy. In first order perturbation theory, the expectation value of V determines the change in the groundstate energy due to the interaction of the electromagnetic field with the atom. The contribution of the vacuum is translation invariant and therefore does not contribute to the force acting on the atom. We subtract the vacuum expectation value and remain with the normal ordered square of the electric field. We conclude that a conducting plane attracts a neutral atom with the potential

$$V_{\mathrm{CP}}(d) = -\frac{3}{8\pi d^4} \qquad (4.68)$$

This is the so-called Casimir-Polder potential which was first derived by Casimir and Polder in 1948. It is valid for large distances d where the atom can be treated as pointlike. For short distances one finds a behavior proportional to d^{-3} as one would expect from the electrostatic interaction with the mirror image. The transition to the long distance behavior may be understood by taking into account retardation effects.

We now consider the case of two parallel conducting planes at $x^3 = 0$ and $x^3 = a$. Outside of the planes only one plane is relevant, and we obtain the results as before. But in the region between the planes we have to take into account all transformations generated by the reflections on the two planes.

Let G be the group which is generated by the reflections on both planes, and let $\mathrm{sign}(S) = (-1)^n$ if $S \in G$ is the product of n reflections. Then

$$\langle F_{01}(x)F_{01}(y)\rangle = (-\partial_0^2 + \partial_1^2)\sum_{S\in G}\mathrm{sign}(S)D_+(x - Sy)$$

$$\langle F_{02}(x)F_{02}(y)\rangle = (-\partial_0^2 + \partial_2^2)\sum_{S\in G}\mathrm{sign}(S)D_+(x - Sy)$$

$$\langle F_{12}(x)F_{12}(y)\rangle = (-\partial_1^2 - \partial_2^2)\sum_{S\in G}\mathrm{sign}(S)D_+(x - Sy)$$

$$\langle F_{03}(x)F_{03}(y)\rangle = (-\partial_0^2 + \partial_3^2)\sum_{S\in G}D_+(x - Sy)$$

$$\langle F_{13}(x)F_{13}(y)\rangle = (-\partial_1^2 - \partial_3^2)\sum_{S\in G}D_+(x - Sy)$$

$$\langle F_{23}(x)F_{23}(y)\rangle = (-\partial_2^2 - \partial_3^2)\sum_{S\in G}D_+(x - Sy)$$

The derivatives with respect to the 3rd component can be eliminated by exploiting the fact that D_+ is a solution of the wave equation. Normal ordering amounts to subtraction of the vacuum expectation value, i.e. the contribution of the unit element of G to the sums above. In the calculation of the expected value of the energy density the contributions of the odd elements of G compensate each other. The even elements of G are the translations by multiples of $2a$ in 3-direction. We obtain the following expression for the expected value of the energy density $h(x)$

$$\langle h(x)\rangle = -4\sum_{n=1}^{\infty}(\partial_0^2 D_+)(2nae_3)$$

with the unit vector e_3 in 3-direction. Inserting the expression for D_+ and computing the derivative yields

$$(\partial_0^2 D_+)(2nae_3) = \frac{1}{2^5\pi^2 a^4 n4}\ .$$

With the sum rule

$$\sum_{n=1}^{\infty}\frac{1}{n^4} = \frac{\pi^4}{90} \equiv \zeta(4)$$

(this formula can, e.g., be derived by the use of the Fourier series of x^4 on the interval $[-\pi, \pi]$) we finally obtain the result

$$\langle h(x)\rangle = -\frac{\pi^2}{720a^4}\ .$$

for a point between the two conducting planes. The force per area acting on the planes is

$$f = -\frac{\pi^2}{240a^4} \ .$$

Thus the two planes attract each other. This remarkable effect was predicted by Casimir in 1948 and is nowadays measured with increasing precision. It precise theoretical computation in less idealized situations is still a challenge for theoretical physics.

CHAPTER IV

Interacting fields: general aspects and methods

In the first 3 chapters of these lecture notes we introduced the Fock space as the space of states for a multiparticle system, the relativistic wave equations as a consequence of Poincaré symmetry and the free quantum fields. We also discussed the way interactions can be incorporated, defined the S-matrix and found the combinatorial description in terms of Feynman graphs. We did, however, not yet discuss how the various inconsistencies in this formal approach to quantum field theory can be avoided.

In the present chapter we want to analyse interacting fields from an axiomatic point of view. We first discuss the concept of the S-matrix and derive its relation to cross sections. We then discuss the so-called LSZ relations which allow to express the S-matrix directly in terms of interacting fields. The LSZ relations, however, rely on an assumption on the relation between the interacting field and asymptotic free fields. We then show that this assumption can be derived from the existence of single particle states (Haag-Ruelle scattering theory). After the clarification of these conceptual issues we describe several methods for the construction of field theories (canonical quantization, path integral) as well as basic combinatorial concepts (connected functions, vertex functions).

1. S-matrix and cross sections

In the case of a scalar field the formal expression for the S-matrix is obtained in the following way. The free scalar field is an operator valued distribution on the bosonic Fock space, given by

$$\varphi(x) = (2\pi)^{-3/2} \int \frac{d^3\mathbf{p}}{2\omega(\mathbf{p})} \left(a(p)e^{-ipx} + a^*(p)e^{ipx}\right) |_{p_0=\omega(\mathbf{p})} \qquad (1.1)$$

with the creation operators $a^*(p)$ and the annihilations operators $a(p)$ for a spinless particle with 4-momentum $p = (p^0 = \omega(\mathbf{p}), \mathbf{p})$. The annihilation and creation operators satisfy the commutation relations

$$[a(p), a^*(q)] = 2\omega(\mathbf{p})\delta^{(3)}(\mathbf{p}-\mathbf{q}) \, , [a(p), a(q)] = 0 = [a^*(p), a^*(q)] \, . \quad (1.2)$$

Polynomials of the field are the constituents of interaction Lagrangeans. In order to be well defined they have to be normal ordered which means that in the expression obtained by taking powers of the free field as

defined above the annihilation operators have to be put on the right hand side of the creation operators,

$$:\varphi(x)^n:=(2\pi)^{-\frac{3n}{2}}\int\frac{d^3\mathbf{p}_1}{2\omega(\mathbf{p}_1)}\cdots\frac{d^3\mathbf{p}_n}{2\omega(\mathbf{p}_n)}\times \sum_{I\subset\{1,\dots,n\}}\prod_{i\in I}a^*(p_i)\prod_{j\in I^c}a(p_j)e^{i(\sum_I p_i-\sum_{I^c}p_j)x}\ . \tag{1.3}$$

The interaction Lagrangean is of the form

$$\mathcal{L}_{\text{int}}(x)=:V(\varphi(x)): \tag{1.4}$$

with a polynomial V. The Dyson formula for the S-matrix is then

$$S=Te^{i\int d^4x\mathcal{L}_{\text{int}}(x)}\equiv\sum_{n=0}^{\infty}\frac{i^n}{n!}\int d^4x_1\dots d^4x_nT\mathcal{L}_{\text{int}}(x_1)\dots\mathcal{L}_{\text{int}}(x_n)\ . \tag{1.5}$$

This fundamental formula poses, unfortunately, three hard problems:

(i) the ultraviolet (UV) problem: since the interaction Lagrangean is not an operator valued function, but an operator valued distribution, the time ordered products are, in general, ill defined.

(ii) the infrared (IR) problem: in general, the integral over x will diverge.

(iii) the sum over n may not converge.

Only the UV problem is nowadays well understood by the theory of perturbative renormalization; after the first breakthrough by Tomonaga, Schwinger, Feynman and Dyson it was finally solved by Bogoliubov, Parasiuk and Hepp and further clarified by Zimmermann, Epstein and Glaser, Steinmann and others. We will discuss the solution in chapter 5 of this lecture.

The IR problem is less well understood. It amounts to the question whether the asymptotic behaviour of states can be interpreted in terms of multiparticle configurations.

Almost nothing is known about the convergence of the sum. One expects that typically the sum only describes an asymptotic expansion of the theory.

The calculation of the time ordered products occuring in the Dyson formula can be organized by an expansion into Feynman graphs,

$$\begin{aligned}&T:V(\varphi(x_1)):\cdots:V(\varphi(x_n)):\\ &=\sum_{G\in\mathcal{G}_n}\prod_{1\le j<k\le n}\frac{(i\Delta_F(x_i-x_j))^{l_{jk}}}{l_{jk}!}:\prod_{j=1}^{n}V^{(l_j)}(\varphi(x_j)):\ .\end{aligned} \tag{1.6}$$

Here $\mathcal{G}_n$ is the set of undirected, tadpole free graphs with vertices $\{1,\dots,n\}$. l_{jk} is the number of lines connecting the vertices j and k. $l_j=\sum_{k=1}^n l_{jk}$ is the number of lines incident on the vertex j. $V^{(l_j)}$

is the l_j-th derivative of V. Note, that in contrast to standard conventions, the set of graphs is independent of the interaction, and the graphs have no external lines. For an interaction $V(\varphi) = \varphi^n$, one could instead represent the derivatives $V^{(l_j)}(\varphi) = \frac{n!}{(n-l_j)!}\varphi^{n-l_j}$ by $n - l_j$ external lines and take into account the arising combinatorial factor.

The formula for the S-matrix was originally derived from the interaction picture. There the interacting field φ is obtained from the free field φ_0 by the time evolution operators in the interaction picture

$$U(t,s) = Te^{i\int_{s\le x^0\le t} d^4x\mathcal{L}_{\rm int}(x)} \tag{1.7}$$

as

$$\varphi(t,\mathbf{x}) = U(t,0)\varphi_0(t,\mathbf{x})U(0,t)\ . \tag{1.8}$$

The S-matrix is then given by

$$S = \lim_{t\to\infty, s\to-\infty} U(t,s)\ . \tag{1.9}$$

In the interaction picture the free and the interacting field act on the same Hilbert space and coincide at time $t = 0$. Unfortunately, it can be shown, that this is not possible for a nontrivial translation invariant local interaction (Haag's Theorem). The basic reason is that the spatial translations are not changed by the interaction. But the vacuum is already fixed by the requirement that it is invariant under spatial translations, hence the interacting theory has the same vacuum as the free theory. On the other hand, local interactions as described before always contain terms which act nontrivially on the vacuum.

A way out is to assume that the interacting field behaves as a free field for very late and very early times. This may be described by the following assumption. Let $\mathfrak{H}_0$ be the Fock space where the free field acts, and let $\mathfrak{H}$ be a Hilbert space on which the interacting field acts. There exist unitary operators

$$W_\pm : \mathfrak{H}_0 \to \mathfrak{H} \tag{1.10}$$

which intertwine the representations of the Poincaré group on both Hilbert spaces. In terms of these unitaries we can define two free fields on $\mathfrak{H}$,

$$\varphi_{\rm out,in}(x) = W_\pm\varphi_0(x)W_\pm^* \tag{1.11}$$

(the outgoing and the incoming free field, respectively). These fields are supposed to satisfy the LSZ asymptotic condition

$$\varphi(t,\mathbf{x}) - \varphi_{\rm out,in}(t,\mathbf{x}) \to 0 \text{ as } t\to\pm\infty\ . \tag{1.12}$$

where convergence is understood in terms of matrix elements between suitable state vectors.

The S-matrix can then be defined by

$$S = W_+^*W_- \tag{1.13}$$

as an operator on the Fock space. Since S commutes with translations we can define the scattering amplitude $\mathcal{T}$ for the process $p, q \to p_1, \ldots, p_n$, $\{p, q\} \neq \{p_1, \ldots, p_n\}$ by

$$\left\langle a^*(p_1) \ldots a^*(p_n)\Omega, Sa^*(p)a^*(q)\Omega \right\rangle = \delta(\sum_{i=1}^{n} p_i - p - q)\mathcal{T}(p_1, \ldots, p_n; p, q) \tag{1.14}$$

In scattering experiments, one usually measures cross sections. They are related to the scattering amplitudes by the formula

$$d\sigma_{pq\to p_1 \ldots p_n} = \frac{(2\pi)^2}{4\sqrt{(pq)^2 - m^4}} \delta(\sum_i p_i - p - q)|\mathcal{T}(p_1, \ldots, p_n; p, q)|^2 \frac{d^3\mathbf{p}_1}{2\omega(\mathbf{p}_1)} \cdots \frac{d^3\mathbf{p}_n}{2\omega(\mathbf{p}_n} \ . \tag{1.15}$$

The derivation of this important formula is usually done in a somewhat handwaving way. Here we want to give a more rigorous derivation which also shows why cross sections do not give a complete information on the scattering process.

The reason for the loss of information in scattering experiments is the fact that typically the impact parameter in a two particle system cannot be prepared with sufficiently high precision. Hence the incoming state is a mixture over a whole range of impact parameters. Let $\Phi = a(f)^*a(g)^*\Omega$ be an incomimg 2-particle state with normalized single particle wave functions f and g which have disjoint support in momentum space and are concentrated around momenta p and q, respectively. Let β be the spacelike plane orthogonal to p and q. We are interested in translationally invariant positive observables A, e.g.

$$A = |a^*(p_1) \ldots a^*(p_n)\Omega\rangle\langle a^*(p_1) \ldots a^*(p_n)\Omega| \ . \tag{1.16}$$

In order to describe the uncertainty in the impact parameter we replace the pure state Φ by a mixture

$$\rho_\mu = \int_\beta d^2b\mu(b)|\Phi_b\rangle\langle\Phi_b| \tag{1.17}$$

with $\Phi_b = a(U(b)f)^*a(g)\Omega$ and a probability measure $\mu(b)d^2b$. For large values of b we expect that the particles do not influence each other. We assume that A vanishes on the initial state for all possible impact parameters

$$\left\langle \Phi_b, A\Phi_b \right\rangle = 0 \ . \tag{1.18}$$

Instead of trying to construct an initial state where the probability measure on the impact parameter plane is concentrated at a point we replace it by a distribution $\mu(b)$ which is constant in the region where scattering takes place, Let us assume that scattering can be neglected

for impact parameters $b \notin B \subset \beta$, and let area(B) denote the area of B. Then

$$\mu_B(b) = \begin{cases} \frac{1}{\text{area}(B)} & , \quad b \in B \\ 0 & , \quad \text{else} \end{cases} \tag{1.19}$$

In order to become independent of the size of B, and under the assumption that scattering outside of a sufficiently large chosen region B is arbitrarily small it is now convenient to replace the normalized measure $\mu_B(b)d^2b$ by the measure d^2b on B and to define the *cross section* as the limit as B tends to the full plane β,

$$\sigma_{f,g,\beta}(A) = \int_\beta d^2b \big\langle S\Phi_b, AS\Phi_b \big\rangle = \lim_{B\to\beta} \text{area}(B)\, \text{Tr}\, S\rho_{\mu_B} S^* A \ . \tag{1.20}$$

The cross section might diverge if the interaction does not decay sufficiently fast. This is known to happen for the total cross section for Coulomb scattering. For short range interactions, however, the scattering cross sections are finite. From the definition it is clear that cross sections have the dimension of an area (times the dimension of the observable).

We now define the T-matrix by $S = 1 + iT$. Due to (1.18) and the assumed positivity of A we have $A\Phi_b = 0$ and hence

$$\sigma_{f,g,\beta}(A) = \int_\beta \mathrm{d}^2b \big\langle \Phi_b, T^*AT\Phi_b \big\rangle \ . \tag{1.21}$$

By assumption, T and A are translation invariant. Hence the matrix elements of T^*AT in the 2-particle space have the form

$$\big\langle a^*(p)a^*(q)\Omega, T^*ATa^*(p')a^*(q')\Omega \big\rangle = \delta(p+q-p'-q')A_T(p,q,p',q') \ . \tag{1.22}$$

In the calculation of the cross section the integral over the impact parameter b yields an additional factor

$$(2\pi)^2\delta((p-p')e_1)\delta((p-p')e_2)$$

with some orthonormal basis $\{e_1, e_2\}$ of β. Together with the conservation of the total momentum this implies that the cross sections are diagonal in momenta,

$$\sigma_{f,g,\beta}(A) = (2\pi)^2 \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega(\mathbf{p})} \frac{\mathrm{d}^3\mathbf{q}}{2\omega(\mathbf{q})} |f(p)|^2 |g(q)|^2 A_T(p,q,p,q) |\det \frac{\partial\alpha}{\partial(p,q)}|^{-1} \ . \tag{1.23}$$

Here α is defined for given $p, q \in H_m^+$ by

$$\alpha(p',q') = ((p')^2 - m^2, (q')^2 - m^2, p'+q'-p-q, (p'-p)e_1, (p'-p)e_2) \tag{1.24}$$

The Jacobian of α at $(p',q') = (p,q)$ is $|\det(2p, 2q, e_1, e_2)|$. For $\beta \perp p, q$ it follows

$$|\det(2p, 2q, e_1, e_2)| = 4\sqrt{(pq)^2 - m^4} \ . \tag{1.25}$$

This expression is called the flux factor.

We now evaluate the formula (1.23) for the observable A from (1.16) under the condition that the momenta $p, q, p_1, \dots, p_n$, $p \in \operatorname{supp} f, q \in \operatorname{supp} g$, are pairwise different. A measures the probability density for n momenta,

$$\big\langle \Psi, A\Psi \big\rangle = n! |\Psi_n(p_1, \dots, p_n|^2 \tag{1.26}$$

with respect to the measure $\frac{d^3\mathbf{p}_1}{2\omega(\mathbf{p}_1)} \dots \frac{d^3\mathbf{p}_n}{2\omega(\mathbf{p}_n)}$. In the limit where f and g have sharp momenta we obtain the formula (1.15) for the cross section. Note that the loss of information appearing in this formula is due to the lack of precision in the preparation of the incoming state and might be overcome in future experiments.

As an example we consider the S-matrix for the interaction Lagrangean $\frac{g}{n!} :\varphi^n:$ in first order. One finds

$$T = \frac{g}{n!} \int \mathrm{d}^4x \, :\varphi^n(x): \tag{1.27}$$

and thus for the scattering amplitude

$$\begin{aligned} \delta(p+q-\sum p_i)\mathcal{T}(p_1, \dots, p_{n-2}; p, q) = \\ (2\pi)^{-\frac{3n}{2}} \frac{g}{n!} \int \mathrm{d}^4x \binom{n}{2} \big\langle a^*(p_1) \cdots a^*(p_{n-2})\Omega, (a^*(x))^{n-2} a(x)^2 a^*(p) a^*(q) \Omega \big\rangle \\ = g(2\pi)^{-3n/2} \int \mathrm{d}^4x e^{-i(p+q-\sum p_i)x} \end{aligned} \tag{1.28}$$

hence

$$\mathcal{T}(p_1, \dots, p_{n-2}; p, q) = (2\pi)^{4-\frac{3n}{2}} g \tag{1.29}$$

For the cross section we obtain

$$\begin{aligned} d\sigma_{p,q\to p_1,\dots,p_{n-2}} = \\ (2\pi)^{10-3n} g^2 (4\sqrt{(pq)^2 - m^4})^{-1} \delta(p+q-\sum p_i) \frac{d^3\mathbf{p}_1}{2\omega(\mathbf{p}_1)} \cdots \frac{d^3\mathbf{p}_{n-2}}{2\omega(\mathbf{p}_{n-2})} \, . \end{aligned}$$

The total cross section

$$\sigma_{\mathrm{tot}} = \frac{1}{(n-2)!} \int \frac{\mathrm{d}^3\mathbf{p}_1}{2\omega(\mathbf{p}_1)} \cdots \frac{\mathrm{d}^3\mathbf{p}_{n-2}}{2\omega(\mathbf{p}_{n-2})} \sigma_{p,q\to p_1,\dots,p_{n-2}} \, . \tag{1.30}$$

depends only on the center of mass energy $\sqrt{s}$, $s = (p+q)^2$. For $\sqrt{s} \gg (n-2)m$ much above the threshold for particle creation it behaves as s^{n-5}.

Exercise 1: *Compute the total cross section for elastic 2 particle scattering in lowest nontrivial order for the interaction Lagrangean* $\mathcal{L} = \frac{g}{4!} :\varphi^4: \ .$

2. The LSZ-relations

Due to Haag's Theorem, in a translation invariant theory the S-matrix cannot be written in the form

$$S = \lim_{t\to\infty, s\to-\infty} U(t,s) \tag{2.1}$$

with the time evolution operators $U(t,s)$ of the interaction picture (1.7). Instead one may multiply the interaction Lagrangean by a test function g with compact support and introduce the local S-matrices

$$S(g) = Te^{i\int \mathcal{L}_{\rm int} g \mathrm{d}^4x} \ . \tag{2.2}$$

Finally one may study the so-called adiabatic limit

$$S = \lim_{g\to 1} S(g) \ . \tag{2.3}$$

But even if this limit exists it is not a proiri clear in which sense the S-matrix is related to the interacting field theory.

Lehmann, Symanzik and Zimmermann (LSZ) succeeded in 1954 to derive an elegant formula for the S-matrix which is completely defined in terms of the interacting field.

Starting point is (in the simplest case) a scalar field φ which is defined as an operator valued distribution in some Hilbert space $\mathfrak{H}$, together with a unitary, strongly continuous representation U of the Poincaré group satisfying the spectrum condition. Furthermore, there exists a unique (up to a phase) Poincaré invariant unit vector (the vacuum vector Ω). The scalar field transforms covariantly under the Poincaré group,

$$U(x,\Lambda)\varphi(y)U(x,\Lambda)^{-1} = \varphi(\Lambda y + x) \ . \tag{2.4}$$

We now assume that there exists a Poincaré invariant subspace $\mathfrak{H}_1 \subset \mathfrak{H}$ which carries the irreducible representation with mass $m > 0$ and spin $s = 0$, and that there are no other states in $\mathfrak{H}$ whose mass spectrum contains m. Under these assumptions one can show (see next section) by means of the Haag-Ruelle scattering theory that there exist isometric mappings $W_\pm$ from Fock space $\mathfrak{H}_0$ to $\mathfrak{H}$ which intertwine the representations of the Poincaré group.

In addition we require that the interacting field φ has nonvanishing matrix elements between vacuum and $\mathfrak{H}_1$. Due to Poincaré covariance they are of the form

$$\langle p|\varphi(x)\Omega\rangle = (2\pi)^{-3/2}\sqrt{Z}e^{ipx} \tag{2.5}$$

with a constant $Z \neq 0$. Here the improper single particle states with sharp momentum are normalized by

$$\langle p|q\rangle = 2\omega(\mathbf{p})\delta(\mathbf{p}-\mathbf{q}) \ . \tag{2.6}$$

such that for the free field $|p\rangle = a^*(p)\Omega$. Z is called the wave function renormalization. It may be absorbed by a redefinition of the field.

As in the preceding section one now defines free fields in the Hilbert space of the interacting theory by

$$W_{\pm}\varphi_0(x) = \underset{\text{in}}{\varphi}{}^{\text{out}}(x)W_{\pm} \ , \tag{2.7}$$

the only difference being that there is no general argument why the operators $W_{\pm}$ should be invertible (problem of asymptotic completeness). The interacting field converges in the following sense towards the outgoing ($t \to \infty$) and the incoming ($t \to -\infty$) field, respectively: We have the LSZ asymptotic condition

$$\lim_{t\to\pm\infty} \Big\langle W_+\Phi, \Big(\varphi(t,\mathbf{x}) - \underset{\text{in}}{\varphi}{}^{\text{out}}(t,\mathbf{x})\Big) W_-\Psi\Big\rangle = 0 \ . \tag{2.8}$$

Here Φ and Ψ are vectors from the dense subspace with finite particle number, smooth momentum space wave functions with compact supports and non overlapping velocities. Let now f be a solution of the Klein-Gordon equation whose Cauchy data have compact support. Then

$$\underset{\text{in}}{\varphi}{}^{\text{out}}(f) = \int \mathrm{d}^3\mathbf{x}\big(\dot{\underset{\text{in}}{\varphi}}{}^{\text{out}}(t,\mathbf{x})f(t,\mathbf{x}) - \underset{\text{in}}{\varphi}{}^{\text{out}}(t,\mathbf{x})\dot{f}(t,\mathbf{x})\big) \tag{2.9}$$

is independent of t. If one replaces the free fields by the interacting field one obtains a time dependent expression which for $t \to \pm\infty$ converges towards $\underset{\text{in}}{\varphi}{}^{\text{out}}$. A slight technical complication is that we assumed that the interacting field is a distribution on Minkowski space, but did not make any assumption about the possibility to restrict it to spatial hypersurfaces. This complication can be avoided by taking a mean with a test function $h \in \mathcal{D}(\mathbb{R})$ with $\int \mathrm{d}t h(t) = 1$ over time,

$$\varphi_t(f,h) = \int \mathrm{d}\tau h(\tau) \int \mathrm{d}^3\mathbf{x}\big(\dot{\varphi}(t+\tau,\mathbf{x})f(t+\tau,\mathbf{x}) - \varphi(t+\tau,\mathbf{x})\dot{f}(t+\tau,\mathbf{x})\big) \tag{2.10}$$

and find

$$\begin{aligned} &\big\langle W_+\Phi, \big(\varphi_{\text{out}}(f) - \varphi_{\text{in}}(f)\big)W_-\Psi\big\rangle \\ &= \int \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t}\big\langle W_+\Phi, \varphi_t(f,h)W_-\Psi\big\rangle \ . \end{aligned} \tag{2.11}$$

We exploit the fact that f is a solution of the Klein-Gordon equation, and obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\varphi_t(f,h) &= \int \mathrm{d}\tau h(\tau) \int \mathrm{d}^3\mathbf{x}\big(\ddot{\varphi}(t+\tau,\mathbf{x})f(t+\tau,\mathbf{x}) - \varphi(t+\tau,\mathbf{x})\ddot{f}(t+\tau,\mathbf{x})\big) \\ &= \int \mathrm{d}\tau h(\tau) \int \mathrm{d}^3\mathbf{x}\big(\ddot{\varphi}(t+\tau,\mathbf{x})f(t+\tau,\mathbf{x}) - \varphi(t+\tau,\mathbf{x})(\Delta - m^2)f(t+\tau,\mathbf{x})\big) \\ &= \int \mathrm{d}\tau h(\tau) \int \mathrm{d}^3\mathbf{x}\big((\Box + m^2)\varphi(t+\tau,\mathbf{x})\big)f(t+\tau,\mathbf{x}) \end{aligned} \tag{2.12}$$

and thus

$$\begin{aligned}&\big\langle W_+\Phi, \big(\varphi_{\rm out}(f)-\varphi_{\rm in}(f)\big)W_-\Psi\big\rangle\\ &=\int \mathrm{d}^4x f(x)(\Box+m^2)\big\langle W_+\Phi,\varphi(x)W_-\Psi\big\rangle\ .\end{aligned}\tag{2.13}$$

We now introduce time ordered products of the field φ. They are symmetric operator valued distributions in several arguments which coincide for time ordered arguments with the operator product. They are, in general, not uniquely determined if the arguments can coincide; but this ambiguity plays no role for the following discussion. We use the following notation for two functions f, g

$$f\overset{\leftrightarrow}{\partial_t}g=(\partial_t f)g-f\partial_t g\ .\tag{2.14}$$

Let now $f_1,\dots,f_n$ be solutions of the Klein-Gordon equation with compactly supported Cauchy data. We set

$$\begin{aligned}&T(t_1,\dots,t_n)\\ &=\int \mathrm{d}^{4n}x\, T(\varphi(x_1^0+t_1,\mathbf{x}_1)\cdots\varphi(x_n^0+t_n,\mathbf{x}_n)\overset{\leftrightarrow}{\partial}_{t_1}\cdots\overset{\leftrightarrow}{\partial}_{t_n}\\ &\qquad\times f_1(x_1^0+t_1,\mathbf{x}_1)h(x_1^0)\cdots f_n(x_n^0+t_n,\mathbf{x}_n)h(x_n^0)\ .\end{aligned}\tag{2.15}$$

We have

$$\begin{aligned}&\frac{\partial^n}{\partial_{t_1}\cdots\partial_{t_n}}T(t_1,\dots,t_n)\\ &=\int \mathrm{d}^{4n}x\, h(x_1^0)f_1(x_1^0+t_1,\mathbf{x}_1)\cdots h(x_n^0)f_n(x_n^0+t_n,\mathbf{x}_n)\\ &\qquad\times(\Box_1+m^2)\cdots(\Box_n+m^2)T\varphi(x_1^0+t_1,\mathbf{x}_1)\cdots\varphi(x_n^0+t_n,\mathbf{x}_n)\ ,\end{aligned}\tag{2.16}$$

and for $t_{\pi(1)}+\operatorname{supp}h>\cdots>t_{\pi(n)}+\operatorname{supp}h$ with a permutation $\pi\in\mathrm{S}_n$ we obtain

$$T(t_1,\dots,t_n)=\varphi_{t_{\pi(1)}}(f_{\pi(1)},h)\cdots\varphi_{t_{\pi(n)}}(f_{\pi(n)},h)\ .\tag{2.17}$$

From the Haag-Ruelle theory it follows for wave functions f_i with non overlapping velocities

$$\lim_{\substack{t_1,\dots,t_k\to\infty\\ t_{k+1},\dots,t_n\to-\infty}}\big\langle\Omega,T(t_1,\dots,t_n)\Omega\big\rangle=\big\langle\prod_{i=1}^{k}\varphi_{\rm out}(f_i)^*\Omega,\prod_{j=k+1}^{n}\varphi_{\rm in}(f_j)\Omega\big\rangle\ .\tag{2.18}$$

Let $f_1,\dots,f_k$ be solutions of the Klein-Gordon equation with negative energy, and let $f_{k+1},\dots,f_n$ be solutions with positive energy. Then we

obtain the LSZ relations

$$\int \mathrm{d}^{4n}x f_1(x_1)\cdots f_n(x_n)(\Box_1+m^2)\cdots(\Box_n+m^2)\big\langle\Omega,T\varphi(x_1)\cdots\varphi(x_n)\Omega\big\rangle$$
$$= \big\langle\prod_{i=1}^{k}\varphi_{\mathrm{out}}(f_i)^*\Omega,\ \prod_{j=k+1}^{n}\varphi_{\mathrm{in}}(f_j)\Omega\big\rangle\ . \tag{2.19}$$

Fourier transformation yields

$$\prod_{i=1}^{n}(p_i^2-m^2)\hat{t}_n(-p_1,\ldots,-p_k,p_{k+1},\ldots,p_n)\restriction_{(H_m^+)^n}$$
$$= N\big\langle\prod_{i=1}^{k}a^*_{\mathrm{out}}(p_i)\Omega,\ \prod_{j=k+1}^{n}a^*_{\mathrm{in}}(p_j)\Omega\big\rangle\ . \tag{2.20}$$

with $N=i^n(2\pi)^{-\frac{n}{2}}$ and

$$\hat{t}_n(p_1,\ldots,p_n)=(2\pi)^{-2n}\int \mathrm{d}^{4n}x e^{-i\sum p_jx_j}\big\langle\Omega,T\varphi(x_1)\cdots\varphi(x_n)\Omega\big\rangle\ . \tag{2.21}$$

We observe that the Fourier transforms of the expectation values of time ordered products have for each momentum on the mass shell H_m^+ a pole of the form $(p^2-m^2)^{-1}$, and that the coefficients are given by the S-matrix elements. These are the famous LSZ relations.

Exercise 2:*Compute the vacuum expectation value of the time ordered product of 4 interacting fields in lowest nontrivial order of perturbation theory for the interaction Lagrangean* $\mathcal{L}=-\frac{g}{4!}:\varphi^4:$ *by using the interaction picture.*

3. Haag-Ruelle scattering theory

The LSZ relations deliver a connection between the time ordered products of interacting fields and the S-matrix. Their derivation is based on the LSZ asymptotic conditon which says that the interacting field converges towards a free field at asymptotic times. This assumption can be derived (in a somewhat weaker form) from a few structural assumptions. This is the result of the Haag-Ruelle scattering theory.

We consider the theory of a scalar field. We are given an operator valued distribution $\varphi(f)$ with values in the operators of some Hilbert space $\mathfrak{H}$. On the Hilbert space there exists a unitary strongly continuous representation U of the Poincaré group. We require that the field is covariant,

$$U(x,\Lambda)\varphi(y)U(x,\Lambda)^{-1}=\varphi(\Lambda y+x)\ ,$$

satisfies local commutativity for spacelike separated arguments,

$$[\varphi(x),\varphi(y)]=0 \text{ if } (x-y)^2<0\ ,$$

furthermore, we assume existence and uniqueness of the vacuum, i.e. the existence of a U-invariant unit vector $\Omega \in \mathfrak{H}$ which is unique up to a phase. The vector Ω is required to be cyclic, i.e. the application of smeared fields $\varphi(f)$ on Ω generates a dense subspace $\mathfrak{D}$ of $\mathfrak{H}$. In addition the energy should be positive independent of the Lorentz frame, i.e. $\operatorname{sp} P \subset \overline{V}_+ = \{p \in \mathbb{M}^* | p^2 \geq 0, p_0 \geq 0\}$.

The crucial assumption which implies the multiparticle structure of the theory is now the existence of single particle states whose momentum spectrum is separated by a gap from the rest of the momentum spectrum,

$$\operatorname{sp} P = \{p \in \mathbb{R}^4, p_0 \geq 0, p = 0 \text{ or } p^2 = m^2 \text{ or } p^2 > (m+\delta)^2\}$$

with the upper mass gap $\delta > 0$.

The first step for the creation of multiparticle states is the construction of almost local 1-particle generators. Let Φ_0 be a unit vector with compact momentum support on the mass shell H_m. Due to the assumed cyclicity of the vacuum there exists for every $\epsilon > 0$ a polynomial A_0 in the smeared field operators such that $||A_0\Omega - \Phi_0|| < \epsilon$. Let f be a Schwartz function whose Fourier transform is equal to 1 on the momentum spectrum of Φ_0, has compact support and intersects the momentum spectrum only on the mass shell. We set

$$A = \int d^4x \, e^{iPx} A_0 e^{-iPx} f(x)$$

and find

$$A\Omega = \tilde{f}(P) A_0 \Omega \equiv \Phi$$

with the Fourier transform

$$\tilde{f}(p) = \int d^4x \, e^{ipx} f(x) \ .$$

Φ is a vector in the single particle space which differs from Φ_0 by a vector with length less than $\epsilon \sup |\tilde{f}|$. A is a so-called almost local operator, i.e. an operator which can be fast approximated by operators A_n which are localized in double cones $\mathcal{O}_n = \{(t, \mathbf{x}), |t| + |\mathbf{x}| < n\}$, i.e. we have for all vectors $\Psi \in \mathfrak{D}$

$$\lim_{n \to \infty} n^N (A - A_n)\Psi = 0 \ \forall N \in \mathbb{N} \ .$$

We interpret A as an operation which takes place near to the origin and generates there a single particle state.

We now exploit the fact that the vector valued function $e^{iPx}\Phi$ is a positive-frequency solution of the Klein-Gordon equation. We choose a numerical negative-frequency solution of the form

$$h(t, \mathbf{x}) = (2\pi)^{-3} \int d^3\mathbf{p} \, e^{-i(\omega(\mathbf{p})t - \mathbf{p}\cdot\mathbf{x}))} h_0(\mathbf{p})$$

where h_0 is a test function with compact support which is equal to 1 on the spectrum of the spatial momentum of the vector Φ. We set

$$A(t) = \int d^3\mathbf{x}\, h(t,\mathbf{x}) e^{itP_0 - i\mathbf{P}\cdot\mathbf{x}} A e^{-itP_0 + i\mathbf{P}\cdot\mathbf{x}}$$

and observe that for all t

$$A(t)\Omega = \Phi \ .$$

The localization of the operator $A(t)$ is determined by the behaviour of the solution h at time t. The propagation of a solution of the Klein-Gordon equation of the above form depends essentially on the velocities which correspond to the momenta in the support of h_0 (the assumption that $h_0 \equiv 1$ on a certain domain plays no significant role). Let

$$V(h_0) = \{\frac{\mathbf{p}}{\omega(\mathbf{p})}, \mathbf{p} \in \operatorname{supp} h_0\} \ .$$

Then we have the following theorem:

THEOREM IV.1. *Let $\epsilon > 0$ and $N \in \mathbb{N}$. Then there are constants $c, c_N > 0$ such that*

(i) $|h(t,t\mathbf{v})| < c_N |t|^{-N} \operatorname{dist}(\mathbf{v}, V(h_0))^{-N}$ *for $t \neq 0$ and* $\operatorname{dist}(\mathbf{v}, V(h_0)) > \epsilon$.
(ii) $\int d^3\mathbf{x}\, |h(t,\mathbf{x})| < c(1 + |t|^3)$

A proof may be found in the book of Reed and Simon.

The Theorem says that the solution is essentially localized in the kinematically allowed region $\{(t,\mathbf{x}), \frac{\mathbf{x}}{t} \in V(h_0)\}$.

We now want to show that our Hilbert space contains state vectors which can be interpreted at late or early times as outgoing or incoming multiparticle states, respectively. For the sake of transparency we restrict ourselves to the case of 2 particles.

We construct as described above 2 single particle vectors Φ_1, Φ_2 with the corresponding almost local operators A_i, $i = 1, 2$. We assume that the velocity supports of the single particle states do not overlap and choose the numerical solutions h_1, h_2 of the Klein-Gordon equation correspondingly. Using the localization properties of the wave functions h_1, h_2 as they follow from the Theorem we conclude that the operators $A_1(t)$ and $A_2(t)$ are for large $|t|$ localized far from each other, up to small corrections. We now define the Haag-Ruelle approximants

$$\Psi(t) = A_1(t) A_2(t) \Omega$$

as candidates for approximative 2-particle states. We have the following theorem:

THEOREM IV.2. (i) $\Psi(t)$ *converges for $t \to \infty$ and for $t \to -\infty$.*

(ii) *The limits depend only on the single particle vectors* Ψ_1 *and* Ψ_2*. We use the following notation:*

$$\lim_{t\to\pm\infty} \Psi(t) =: (\Phi_1 \times \Phi_2)_{out,in}$$

(iii) *The scalar product of 2 asymptotic 2-particle states is*

$$\big\langle (\Phi_1 \times \Phi_2)_{out}, (\Phi_1' \times \Phi_2')_{out} \big\rangle = \big\langle \Phi_1, \Phi_1' \big\rangle \big\langle \Phi_2, \Phi_2' \big\rangle + \big\langle \Phi_1, \Phi_2' \big\rangle \big\langle \Phi_2, \Phi_1' \big\rangle$$

hence coincides with the scalar product in the bosonic 2-particle space (an analogous statement holds for the incoming 2-particle states).

(iv) *The Poincaré group acts by*

$$U(L)(\Phi_1 \times \Phi_2)_{out,in} = (U(L)\Phi_1 \times U(L)\Phi_2)_{out,in}$$

The proof relies on the mentioned localization properties of the almost local operators $A_i(t)$ and on the t-independence of the single particle vectors $A_i(t)\Omega$. The convergence of the Haag-Ruelle approximants, e.g., follows from the fact that the time derivative vanishes fast. We have $A_i(t)\Omega = \text{const}$ and hence

$$\frac{d}{dt} A_1(t)A_2(t)\Omega = [(\frac{d}{dt} A_1(t)), A_2(t)]\Omega \ .$$

Due to the localization properties of the operators the commutator vanishes faster than every power of t. For a complete proof we refer to the literature.

The generalization of the above construction to asymptotic multiparticle states yields two isometric mappings $W_\pm$ from the bosonic Fock space $\mathfrak{F}$ built over the single particle space into the Hilbert space $\mathfrak{H}$. The ranges of these maps are the subspaces of outgoing and incoming scattering states $\mathfrak{H}_{\text{out,in}}$ The matrix elements of the S-matrix are now defined by

$$\big\langle \Phi, S\Psi \big\rangle = \big\langle W_+\Phi, W_-\Psi \big\rangle \ , \ \Phi, \Psi \in \mathfrak{F} \ .$$

In case the operators $W_\pm$ are unitary (this does not follow from the Haag-Ruelle theory), also the S-matrix is unitary. In this case the theory is called asymptotically complete. The corresponding problem in nonrelativistic quantum mechanics was solved in the eighties of the last century for quite general interactions. Within quantum field theory, not much is known about this problem.

The derivation described above yields exclusively bosons as a consequence of the assumed commutativity of observable fields at spacelike separations. Fermions would arise if one assumes anticommutativity for the fields which connect the vacuum with the single particle state. But the assumption of anticommutativity for spacelike separation has no direct physical motivation. Actually, as shown by Doplicher, Haag and Roberts in the early seventies of the last century, it suffices to start with an algebra of observables where operators which are spacelike separated commute with each other. One then can extend the algebra in

a unique way to an algebra containing also fermionic fields. Remarkably, the proof requires at least 3 spatial dimensions. If one applies quantum field theory to lower dimensional spacetimes one indeed can find particles which obey more general statistics (anyons, plektons).

4. Canonical Quantization

A standard procedure for the definition of quantum field theories is the canonical quantization of classical field theories. Here one starts from the Lagrangean formulation of classical field theory. In the same way as in classical mechanics one derives the field equations from the principle of stationary action.

Let $\mathcal{L}$ be a function which depends on the fields and their first derivatives. Let G be an open region of spacetime with compact closure, and let g be a test function with compact support which is equal to 1 on G. Then the functional

$$S_g(\varphi) = \int \mathrm{d}^4x \mathcal{L}(\varphi(x), \partial_\mu\varphi(x))g(x) \tag{4.1}$$

is required to be stationary (in typical cases minimal) for all variations of φ with support within G. For evaluating this condition we choose an arbitrary test function ψ with support within G and find

$$0 = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} S_g(\varphi + \varepsilon\psi)|_{\varepsilon=0} \ . \tag{4.2}$$

Computing the derivative yields, independently of the choice of g,

$$0 = \int \mathrm{d}^4x \big(\frac{\partial\mathcal{L}}{\partial\varphi}(x)\psi(x) + \sum_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}(x)\partial_\mu\psi(x)\big) \ . \tag{4.3}$$

Since ψ has compact support, we can, in the second term, move the differential operator from ψ to the other factor and obtain

$$0 = \int \mathrm{d}^4x \big(\frac{\partial\mathcal{L}}{\partial\varphi}(x) - \sum_\mu \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}(x)\big)\psi(x) \ . \tag{4.4}$$

If this integral has to vanish for an arbitrary choice of ψ with $\operatorname{supp}\psi \subset G$, φ has to satisfy within G the differential equation

$$\frac{\partial\mathcal{L}}{\partial\varphi} = \sum_\mu \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} \ . \tag{4.5}$$

We now require stationarity of the action for every open region G with compact closure and conclude that the field has to fulfill the field equation (*Euler-Lagrange equation*) (4.5) everywhere.

Let e.g. $\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{g}{4!}\varphi^4$, then we find

$$\frac{\partial\mathcal{L}}{\partial\varphi} = -m^2\varphi - \frac{g}{3!}\varphi^3 \ , \ \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} = \partial^\mu\varphi \tag{4.6}$$

hence the field equation

$$(\Box + m^2)\varphi = -\frac{g}{3!}\varphi^3 \ . \tag{4.7}$$

Quantization is usually done in the Hamiltonian formulation. Here one distinguishes the time coordinate and chooses some initial time (say, $t = 0$). One defines the Lagrange function as the spatial integral of the Lagrangean and the canonical momenta as the variational derivatives of the Lagrange function with respect to the time derivatives of the field. Disregarding for the moment convergence problems of the spatial integrals, we obtain for the canonical conjugated momenta

$$\pi(\mathbf{x}) = \frac{\delta L}{\delta \dot{\varphi}(\mathbf{x})} = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}}(\mathbf{x}) \ . \tag{4.8}$$

If $\dot{\varphi}$ can be written as a function of $\varphi, \vec{\partial}\varphi$ und π, we obtain the Hamiltonian as

$$H(\varphi, \pi) = \int \mathrm{d}^3\mathbf{x}\, h(\varphi(\mathbf{x}), \vec{\partial}\varphi(\mathbf{x}), \pi(\mathbf{x})) \tag{4.9}$$

with the Hamiltonian density

$$h(\varphi, \vec{\partial}\varphi, \pi) = \pi\dot{\varphi}(\varphi, \vec{\partial}\varphi, \pi) - \mathcal{L}(\varphi, \vec{\partial}\varphi, \dot{\varphi}(\varphi, \vec{\partial}\varphi, \pi)) \ . \tag{4.10}$$

The quantization prescription now consists in the replacement of φ and π by operator valued distributions in $\mathbf{x}$, such that the canonical commutation relations hold,

$$\begin{aligned} [\varphi(\mathbf{x}), \varphi(\mathbf{y})] &= 0 = [\pi(\mathbf{x}), \pi(\mathbf{y})] \\ [\varphi(\mathbf{x}), \pi(\mathbf{y})] &= i\delta(\mathbf{x} - \mathbf{y}) \end{aligned} \tag{4.11}$$

The time evolution is then determined by the Heisenberg equation,

$$\dot{\varphi}(x) = i \int \mathrm{d}^3\mathbf{y}[h(\mathbf{y}), \varphi(x)] \ . \tag{4.12}$$

This prescription suffers from some defects. Let us discuss them on the example of the φ^4-theory. In this case

$$\pi = \dot{\varphi} \ , \tag{4.13}$$

and the Hamiltonian density is

$$h = \frac{1}{2}\pi^2 + \frac{1}{2}|\vec{\partial}\varphi|^2 + \frac{m^2}{2}\varphi^2 + \frac{g}{4!}\varphi^4 \ . \tag{4.14}$$

The canonical commutation relations are realized by the free scalar field with mass m at time zero. We may also define the Hamiltonian density by replacing the products of fields by normal ordered products. Unfortunately, these products are in most cases no longer distributions on space alone but only on spacetime. But even if they would be well defined (this would be the case in a two dimensional Minkowski space), it turns out that it is not possible to define the Hamiltonian of the interacting theory as the spatial integral of the Hamiltonian density. This is the content of Haag's Theorem:

THEOREM IV.3. *Let $\mathfrak{H}$ be the Fock space of a free field φ_0, and let $U(\mathbf{x})$ be the spatial translation operator. Let φ be an operator valued distribution with the properties*

(i) $\varphi(0,\mathbf{x}) = \varphi_0(0,\mathbf{x})$, $\dot{\varphi}(0,\mathbf{x}) = \dot{\varphi}_0(0,\mathbf{x})$.
(ii) $U(\mathbf{x})\varphi(t,\mathbf{y})U(\mathbf{x})^{-1} = \varphi(t,\mathbf{y}+\mathbf{x})$.
(iii) *It exists a selfadjoint operator H which commutes with the spatial translation operators and has the property*

$$e^{itH}\varphi(0,\mathbf{x})e^{-itH} = \varphi(t,\mathbf{x}) \ . \tag{4.15}$$

Then H coincides up to an additive constant with the Hamiltonian H_0 of the free theory, and $\varphi = \varphi_0$.

PROOF. The proof relies on the fact that the vacuum vector Ω is, up to a phase, the only vector in $\mathfrak{H}$ which is invariant under spatial translations. Thus Ω is an eigenvector of H. Let λ be the corresponding eigenvalue. The fields smeared with spatial test functions $\varphi_0(0,f) = \int \mathrm{d}^3\mathbf{x} f(\mathbf{x})\varphi_0(\mathbf{x})$ generate from Ω a dense subspace $\mathfrak{D}$. On $\mathfrak{D}$ we have:

$$\begin{aligned}
&(H-\lambda)\varphi_0(0,f_1)\cdots\varphi_0(0,f_n)\Omega = (H-\lambda)\varphi(0,f_1)\cdots\varphi(0,f_n)\Omega \\
&= \sum_k \varphi(0,f_1)\cdots\dot{\varphi}(0,f_k)\cdots\varphi(0,f_n)\Omega \\
&= \sum_k \varphi_0(0,f_1)\cdots\dot{\varphi}_0(0,f_k)\cdots\varphi_0(0,f_n)\Omega \\
&= H_0\varphi_0(0,f_1)\cdots\varphi_0(0,f_n)\Omega \ .
\end{aligned} \tag{4.16}$$

Hence $H = H_0 + \lambda$ on $\mathfrak{D}$. The proof is completed by showing that H_0 is essentially selfadjoint on $\mathfrak{D}$. □

In order to avoid the consequences of Haag's Theorem one can try to define Hamiltonians with a spatial cutoff

$$H(g) = H_0 + \int \mathrm{d}^3\mathbf{x} g(\mathbf{x})h(\mathbf{x}) \tag{4.17}$$

with a test function g. One expects in view of the finite propagation speed of local perturbations that

$$\varphi(t,\mathbf{x}) = e^{itH(g)}\varphi(0,\mathbf{x})e^{-itH(g)} \tag{4.18}$$

is independent of g, provided $g \equiv 1$ within the region $\{\mathbf{y}, |\mathbf{x}-\mathbf{y}| < |t|\}$. In this way one obtains the algebra of local observables of the interacting theory, but not yet the unitary representation of the Poincaré group, under which the interacting field transforms covariantly.

In the next steps one looks for the ground state vectors $\Omega(g)$ of the locally perturbed Hamiltonians (provided they exist). The vacuum expectation values (the Wightman functions) of products of interacting fields are then obtained as the limit $g \to 1$ of

$$\big\langle \Omega(g), \varphi(x_1)\cdots\varphi(x_n)\Omega(g)\big\rangle \ . \tag{4.19}$$

Finally one applies the reconstruction theorem and constructs the vacuum Hilbert space, the interacting field and the representation of the Poincaré group. This idea was successfully used by Glimm and Jaffe for the construction of the φ^4-theory in 2 spacetime dimensions.

We now want to treat the canonical quantization of electrodynamics. The Maxwell equations can be derived from the Lagrangean

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j_\mu A^\mu \tag{4.20}$$

with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Here j_μ is a conserved current, which does not depend on A_μ. Namely, we have

$$\frac{\partial \mathcal{L}}{\partial A_\mu} = -j^\mu \ , \ \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu} \ , \tag{4.21}$$

and thus

$$j^\nu = \partial_\mu F^{\mu\nu} = \Box A^\nu - \partial^\nu \partial_\mu A^\mu \ . \tag{4.22}$$

The initial value problem for the vector potential is not well posed. Namely, if A_μ is a solution then also $A_\mu + \partial_\mu \Lambda$ with an arbitrary function Λ (gauge freedom). Therefore the transition to the Hamiltonian formalism is not possible. Formally this may be seen from the fact that the canonically conjugated momenta are not independent and that the elimination of the the time derivatives of he vector potential is not possible. The canonically conjugated momenta are

$$\pi_\mu = \frac{\partial \mathcal{L}}{\partial(\dot{A}_\mu)} = \begin{cases} 0 & , \quad \mu = 0 \\ -F^{0\mu} = E_\mu & , \quad \mu = 1,2,3 \end{cases} \ . \tag{4.23}$$

One chooses the following way out. One adds to the Lagrangean an additional term (gauge fixing) such that the modified field equation has a well posed initial value problem. Such a choice is $-\frac{\lambda}{2}(\partial^\mu A_\mu)^2$ with $\lambda \neq 0$. The field equations then are

$$\Box A_\mu + (\lambda - 1)\partial_\mu(\partial_\nu A^\nu) = j_\mu \ . \tag{4.24}$$

It follows from current conservation

$$\Box(\partial_\mu A^\mu) = 0 \ . \tag{4.25}$$

Hence $B := \partial_\mu A^\mu$ is a free massless scalar field. For $B = 0$ the modified field equations are equal to the Maxwell equations. But we cannot set B to zero since it has nontrivial commutation relations.

Exercise 3: *Show that the gauge fixed field equation* (4.24) *has a well posed initial value problem for* $\lambda \neq 0$.

The canonical conjugated momenta are

$$\pi_0 = -\lambda \partial_\mu A^\mu \ , \ \pi_k = \partial_0 A_k - \partial_k A_0 \ . \tag{4.26}$$

For the time derivatives of A_μ we find

$$\dot{A}_0 = \vec{\partial} \cdot \mathbf{A} - \lambda^{-1} \pi_0 \ , \ \dot{A}_k = \pi_k + \partial_k A_0 \ . \tag{4.27}$$

The equal time commutation relations between A_μ and $\dot{A}_\mu$ are

$$[A_\mu(0,\mathbf{x}), \dot{A}_\nu(0,\mathbf{y})] = -i(g_{\mu\nu} + (\lambda^{-1}-1)g_{\mu 0}g_{\nu 0})\delta(\mathbf{x}-\mathbf{y}) \ , \tag{4.28}$$

and those between $\dot{A}_0$ and $\dot{A}_k$

$$\begin{aligned}[\dot{A}_0(0,\mathbf{x}), \dot{A}_k(0,\mathbf{y})] &= [(\vec{\partial}\cdot\mathbf{A} - \lambda^{-1}\pi_0)(0,\mathbf{x}), (\pi_k + \partial_k A_0)(0,\mathbf{y}) \\ &= i(1-\lambda^{-1})\partial_k\delta(\mathbf{x}-\mathbf{y}) \ .\end{aligned} \tag{4.29}$$

For $\lambda = 1$ (Feynman gauge) the commutation relations are especially simple. Another often considered special case is $\lambda = \infty$ (Landau gauge).

In what follows we want to use the Feynman gauge. The field equation coincides in this case with the inhomogeneous wave equation. If j_μ commutes with A_μ, then the commutator between different components of A_μ is a solution of the homogeneous wave equation. Since the initial conditions are determined by the canonical commutation relations, we obtain

$$[A_\mu(x), A_\nu(y)] = -ig_{\mu\nu}D(x-y) \tag{4.30}$$

with the Pauli-Jordan function D. For the field B one finds

$$[B(x), A_\mu(y)] = -i\partial_\mu D(x-y) \ , \tag{4.31}$$

hence $\int \mathrm{d}^3\mathbf{x} B(t,\mathbf{x}) \overset{\leftrightarrow}{\partial_t} \Lambda(t,\mathbf{x})$, with a solution Λ of the wave equation, implements an infinitesimal gauge transformation which is compatible with the gauge fixing $\partial_\mu A^\mu = B$. This demonstrates that we cannot put B to zero.

We now consider the algebra $\mathfrak{A}_0$ of smeared fields which commute with B. This algebra is generated by $F_{\mu\nu}$ and B. In this algebra B generates a nontrivial ideal I. The quotient algebra $\mathfrak{A} = \mathfrak{A}_0/I$ then is the algebra of observables of quantum electrodynamics. It is generated by the fields $F_{\mu\nu}$. Their commutation relations and field equations coincide with those of Chapter 3.

We now want to find a representation of the fields on a Hilbert space. In a first step we try to represent the fields A_μ by hermitean operators on a Hilbert space $\mathfrak{K}$. $\mathfrak{K}$ should contain a vector Ω which represents the vacuum, and shall be generated by apllying the smeared fields to Ω. Furthermore there shall be a positive energy representation of the Poincaré group which leaves Ω invariant and under which A_μ transforms covariantly,

$$U(x,\Lambda)A_\mu(y)U(x\Lambda)^{-1} = A_\nu(\Lambda y + x)\Lambda^\nu{}_\mu \ . \tag{4.32}$$

From these conditions we obtain the 2-point function

$$\left\langle \Omega, A_\mu(x)A_\nu(y)\Omega \right\rangle = -g_{\mu\nu}D_+(x-y) \ . \tag{4.33}$$

In the single particle space

$$\mathfrak{K}_1 = \{A(f)\Omega, \ f = (f^\mu), f^\mu \in \mathcal{S}(\mathbb{R}^4)\} \tag{4.34}$$

one gets the scalar product

$$\left\langle A(f)\Omega, A(f)\Omega\right\rangle = -\int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|}\overline{\hat{f}^\mu(p)}\hat{f}_\mu(p)\ . \tag{4.35}$$

One immediately sees that the scalar product in $\mathfrak{K}$ is not positive definit if A_0 is hermitean.

The 4 components in $\mathfrak{K}_1$ can be interpreted as timelike, longitudinal and two transversal photons (related to the direction of the momentum). Only the transversal photons correspond to physical particles. To eliminate the unphysical degrees of freedom we use the field B. Let

$$\mathfrak{H}_0 = \{\Phi \in \mathfrak{K},\ B(f)\Phi = 0 \text{ falls } \operatorname{supp}\hat{f} \cap \overline{V_+} = \emptyset\}\ . \tag{4.36}$$

(Gupta-Bleuler condition). We have the following

THEOREM IV.4. $\left\langle \Phi, \Phi\right\rangle \geq 0$ *for* $\Phi \in \mathfrak{H}_0$.

PROOF. The n-particle component Φ_n of Φ is a symmetrical function $\Phi_n^{\mu_1\cdots\mu_n}(p_1,\ldots,p_n)$ with $p_i \in \partial V_+$. The Gupta-Bleuler condition says

$$\int \frac{\mathrm{d}^3\mathbf{p}}{2|\mathbf{p}|}p^\mu \hat{f}(-p)g_{\mu\mu_1}\Phi_n^{\mu_1\cdots\mu_n}(p,p_2,\ldots,p_n) = 0\ . \tag{4.37}$$

for all f with $\operatorname{supp}\hat{f} \cap \overline{V_+} = \emptyset$. Hence

$$p_{\mu_1}\Phi_n^{\mu_1\cdots\mu_n}(p,p_2,\ldots,p_n) = 0\ . \tag{4.38}$$

for $p,\ldots,p_n \in \partial\overline{V_+}$. Since $-g_{\mu\nu}$ is positive semidefinite on the orthogonal complement of a lightlike vector $p \neq 0$,

$$p^2 = 0\ ,\ qp = 0 \Longleftrightarrow p_0 = \pm|\mathbf{p}|\ ,\ q_0p_0 = \mathbf{q}\cdot\mathbf{p}\ , \tag{4.39}$$

thus due to the Cauchy-Schwarz inequality

$$|\mathbf{q}|^2 \geq \frac{(\mathbf{q}\cdot\mathbf{p})^2}{|\mathbf{p}|^2} = (q_0)^2\ , \tag{4.40}$$

we have

$$(-1)^n\overline{\Phi_{n\mu_1\ldots\mu_n}(p_1,\ldots,p_n)}\Phi_n^{\mu_1\cdots\mu_n}(p_1,p_2,\ldots,p_n) \geq 0\ . \tag{4.41}$$

But this is the integrand which occurs in the computation of the scalar product $\left\langle \Phi_n, \Phi_n\right\rangle$. Hence $\left\langle \Phi_n, \Phi_n\right\rangle \geq 0$ for all n and thus $\left\langle \Phi, \Phi\right\rangle \geq 0$. □

Let $\mathcal{N}$ be the null space of the positive semidefinite scalar product on $\mathfrak{H}_0$,

$$\mathcal{N} = \{\Phi \in \mathfrak{H}_0,\ \left\langle \Phi, \Phi\right\rangle = 0\}\ . \tag{4.42}$$

Then the quotient space $\mathfrak{H}_0/\mathcal{N}$ has a positive definite scalar product. The completion of this space is called the physical Hilbert space $\mathfrak{H}$. We now want to show that the algebra of observables $\mathfrak{A}$ can be represented by operators on $\mathfrak{H}$. According to the definition, the algebra $\mathfrak{A}_0$ consists of operators which commute with B. Therefore it leaves the space $\mathfrak{H}_0$ invariant. Also the null space is invariant under $\mathfrak{A}_0$. Namely, let

$\Phi \in \mathcal{N}$ and $C \in \mathfrak{A}_0$. Then in view of the hermiticity of B also the adjoint operator $C^* \in \mathfrak{A}_0$. The Cauchy-Schwarz inequality implies

$$\big\langle C\Phi, C\Phi\big\rangle = \big\langle C^*C\Phi, \Phi\big\rangle \leq \|C^*C\Phi\|\|\Phi\| = 0 \ . \tag{4.43}$$

Hence $\mathfrak{A}_0$ acts on the physical Hilbert space $\mathfrak{H}$. Elements of the ideal generated by B are thereby mapped to the null operator, since

$$B(f)\Phi \in \mathcal{N} \tag{4.44}$$

for all $\Phi \in \mathfrak{H}_0$ and all test functions f. We thus obtained a representation of the algebra of observables on a Hilbert space. It coincides with the representation which was constructed in Chapter 3. The advantage of the Gupta-Bleuler method is that it can be used also in the interacting case, since the field B is also in the presence of interactions a solution of the wave equation. This is no longer true in nonabelian gauge theories. For these models a generalization of this method was developed by Becchi, Rouet and Stora (BRS), and, independently, by Tsygan. It is based on a new symmetry, the BRST- symmetry. We will come back to this method in Chapter 6.

5. Path integral

Another method for the definition of quantum field theories is the method of path integrals. Based on ideas of Dirac it was developed by Feynman in his Ph.D. thesis as an alternative formulation of quantum theory.

We describe the method first on the example of a quantum mechanical particle of mass m which moves in 1-dimensional space under the influence of a potential $V(x)$. The Hamiltonian of the system is

$$H = H_0 + V \text{ with } H_0 = -\frac{1}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} \ . \tag{5.1}$$

If $H_0 + V$ is essentially selfadjoint (this is the case, if, e.g., V is continuous and bounded), then according to the Trotter product formula

$$e^{-itH}\Phi = \lim_{n\to\infty}\big(e^{-i\frac{t}{n}H_0}e^{-i\frac{t}{n}V}\big)^n\Phi \ , \ \Phi \in L^2(\mathbb{R}) \ . \tag{5.2}$$

For test functions $\Phi \in \mathcal{S}(\mathbb{R})$ the action of the free time evolution operator e^{-itH_0} can by Fourier transformation be described by the following integral

$$\begin{aligned}\big(e^{-itH_0}\Phi\big)(x) &= \frac{1}{2\pi}\int \mathrm{d}p \int \mathrm{d}y e^{ip(x-y)}e^{-it\frac{p^2}{2m}}\Phi(y) \\ &= \sqrt{\frac{m}{2\pi it}}\int \mathrm{d}y e^{i\frac{m}{2}\frac{(x-y)^2}{t}}\Phi(y)\end{aligned} \tag{5.3}$$

with

$$\sqrt{\frac{m}{2\pi it}} = \sqrt{\frac{m}{2\pi|t|}}e^{-i\frac{\pi}{4}\,\mathrm{sign}(t)} \ . \tag{5.4}$$

If V is infinitely often differentiable with polynomially bounded derivatives, one finds for the solution of the Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(t,x) = (H\psi)(t,x) \tag{5.5}$$

with initial value $\psi(0,x) = \Phi(x)$ the formula

$$\begin{aligned}\psi(t,y_0) &= (e^{-itH}\Phi)(y_0)\\ &= \lim_{n\to\infty}\big(\sqrt{\frac{mn}{2\pi it}}\big)^n \int \mathrm{d}y_1\cdots\mathrm{d}y_n e^{i\frac{t}{n}\sum_{k=1}^n\left(\frac{m}{2}\frac{(y_{k-1}-y_k)^2}{(t/n)^2}-V(y_k)\right)}\Phi(y_n)\end{aligned} \tag{5.6}$$

(Convergence in the sense of $L^2(\mathbb{R})$). The exponent in the second line is a Riemann sum which approximates the integral

$$i\int_0^t \mathrm{d}t'(\frac{m}{2}\dot{y}^2 - V(y)) = iI \tag{5.7}$$

where $y(t)$ is a path with $y(k\frac{t}{n}) = y_k$. I is the classical action of the path.

This formula for the time evolution operator suggests the following interpretation: The quantum mechanical transition amplitude

$$\langle y|e^{-itH}|x\rangle := e^{-itH}(y,x) \tag{5.8}$$

arises as a superposition of the amplitudes for all possible paths $\gamma : [0,t] \to \mathbb{R}$ with $\gamma(0) = x$ and $\gamma(1) = y$. Each of these paths contributes the complex number $e^{iI(\gamma)}$. The leading contributions come from the paths in the neighborhood of a stationary point of the action, i.e. from paths near to a classical solution.

Let now $\mathcal{W}_{x,y,t}$ be the set of continuous paths $\gamma : [0,t] \to \mathbb{R}$ with $\gamma(0) = x$ and $\gamma(1) = y$. We write

$$\langle y|e^{-itH}|x\rangle = \int_{\mathcal{W}_{x,y,t}} \mathcal{D}\gamma\, e^{iI(\gamma)}\ . \tag{5.9}$$

Here

$$\mathcal{D}\gamma = \lim_{n\to\infty}\big(\sqrt{\frac{mn}{2\pi it}}\big)^n \mathrm{d}\gamma(\frac{t}{n})\cdots\mathrm{d}\gamma(\frac{n-1}{n}t) \tag{5.10}$$

means the suitably normalized integral over all paths.

Provided one succeeds in turning this heuristic idea into a mathematical precise definition one would obtain the integrand for arbitrary potentials V by multiplication with the factor

$$e^{-i\int_0^t \mathrm{d}t' V(\gamma(t'))} \tag{5.11}$$

(Feynman-Kac-Formula). If this factor is "integrable" in the sense of path integrals one has obtained an explicit integral representation of the transition amplitude.

In the attempt to make these ideas mathematically precise, the oscillatory character of the integrals causes problems. Better treatable

are the integral kernels of the positive operators e^{-tH}, $t > 0$ (under the condition that H is bounded below). For $t > 0$ we have

$$e^{-tH_0}(x,y) = \sqrt{\frac{m}{2\pi t}} e^{-\frac{m}{2}\frac{(x-y)^2}{t}} \ . \tag{5.12}$$

The integral kernel of e^{-tH_0} has the following properties

$$\begin{aligned} e^{-tH_0}(x,y) &> 0 \ , \\ \int \mathrm{d}x e^{-tH_0}(x,y) &= 1 \ , \\ \int \mathrm{d}y e^{-tH_0}(x,y) e^{-sH_0}(y,z) &= e^{-(t+s)H_0}(x,z) \ . \end{aligned} \tag{5.13}$$

These properties can be interpreted as a probabilistic description of diffusion. The integral kernel $e^{-tH_0}(x,y)$ can be seen as the probability density for a particle driven by diffusion (with the diffusion constant $D = \frac{1}{2m}$) to arrive after the time t at the point x when it has been at the point y. The second equation is the normalization condition, saying that the particle is surely somewhere. The third property characterizes a so-called Markov process. Physically it says that the probability depends only on the starting point, but not on the previous history. In particular the initial velocity does not enter. This is a somewhat surprising aspect of diffusion and can be seen in Brownian motion.

In the theory of Brownian motion one endows the space of paths $\mathcal{W}_{x,y,t}$ with the structure of a measurable space. This means that one has to distinguish a system of measurable sets. Among them are in particular the so-called cylinder sets:

Definition IV.5. *A cylinder set $Z(t_1, \ldots, t_n; B)$ is the set of all paths γ with $(\gamma(t_1), \ldots, \gamma(t_n)) \in B$ where B is a measurable set in $\mathbb{R}^n$ and $0 < t_1 < \ldots < t_n < t$.*

The so-called Wiener integral over cylinder sets is defined by

$$\int_{Z(t_1,\ldots,t_n;B)} \mathrm{d}W^t_{xy} = \int_B \mathrm{d}y_1 \cdots \mathrm{d}y_n \, e^{-H_0(t-t_n)}(y - y_n) \cdots e^{-H_0 t_1}(y_1 - x) \ . \tag{5.14}$$

There holds the following theorem:

Theorem IV.6. *Let V be continuous and bounded from below, and let $H = H_0 + V$ be essentially selfadjoint. Then the function $e^{-\int_0^t \mathrm{d}t' V(\gamma(t'))}$ on $\mathcal{W}_{x,y,t}$ is integrable with respect to the Wiener measure, and there holds the Feynman-Kac-Formula*

$$(e^{-tH})(y,x) = \int \mathrm{d}W^t_{xy} \, e^{-\int_0^t \mathrm{d}t' V(\gamma(t'))} \ . \tag{5.15}$$

PROOF. From the Trotter product formula we have

$$(e^{-tH})(y,x) = \lim_{n\to\infty} \big(\sqrt{\frac{mn}{2\pi t}}\big)^n \int \mathrm{d}y_1 \cdots \mathrm{d}y_n e^{-\frac{t}{n}\sum_{k=1}^n \left(\frac{m}{2}\frac{(y_{k-1}-y_k)^2}{(t/n)^2}+V(y_k)\right)} \,. \tag{5.16}$$

According to the definition of the Wiener measure the right hand side is the limit $n \to \infty$ of the Wiener integrals

$$\int \mathrm{d}W^t_{xy}\, e^{-\sum_{k=1}^{n-1}\frac{t}{n}V(\gamma(\frac{kt}{n}))} \tag{5.17}$$

over the cylinder functions

$$\gamma \to e^{-\sum_{k=1}^{n-1}\frac{t}{n}\,\cdot V(\gamma(\frac{kt}{n}))} \tag{5.18}$$

These functions converge pointwise towards $e^{-\int_0^t \mathrm{d}t' V(\gamma(t'))}$. Because of the lower bound on V they are uniformly bounded by a constant. Applying the theorem on dominated convergence the limit function is integrable, and the integral coincides with the limit of the integrals over the approximating cylinder functions. □

This theorem says in particular that the integral kernel of e^{-tH} is positive. This implies that also the wave function of the ground state Ω (provided it exists) must be positive valued (up to a global phase). It can be obtained in the following way:

THEOREM IV.7. *Let $\Omega_0 \in L^2(\mathbb{R})$ be positive. Then*

$$\Omega = \lim_{t\to\infty} e^{-tH}\Omega_0 \| e^{-tH}\Omega_0\|^{-1} \,, \tag{5.19}$$

(provided the limit exists) is the unique ground state of H. If the limit does not exist, H does not have a ground state.

Exercise 4: *Prove the Theorem.*

The Theorem yields the following formula for the expectation value of a product of functions f_i of the position operator at different (imaginary) times

$$\big\langle \Omega, e^{t_1 H} f_1(x) e^{-t_1 H} \cdots e^{t_n H} f_n(x) e^{-t_n H}\Omega\big\rangle = \int \mathrm{d}\mu(\gamma) f_1(\gamma(t_1))\cdots f_n(\gamma(t_n)) \,. \tag{5.20}$$

Here μ is a probability measure on the space of all paths $\gamma : \mathbb{R} \to \mathbb{R}$. It is obtained as the limit $t \to \infty$ of the measures

$$Z(t)^{-1}\mathrm{d}x\Omega_0(x)\mathrm{d}y\Omega_0(y)\mathrm{d}W^{-t,t}_{yx} e^{-\int_{-t}^t \mathrm{d}t' V(\gamma(t'))} \tag{5.21}$$

with a normalization factor $Z(t)$. A remarkable property of this formula is that the expectation values can be computed without knowing the wave function of the ground state.

We now want to find analogous formulae for field theory. For this purpose we first construct the Schrödinger representation of the free scalar field. Here the position operators are replaced by the time zero fields $\varphi(0,\mathbf{x})$. We assume that they take values in the space of tempered

distributions $\mathcal{S}'(\mathbb{R}^3)$. We search for a measure on this space such that the Fock space can be identified with the space $L^2(\mathcal{S}'(\mathbb{R}^3), \mathrm{d}\mu)$. A Fock space vector Φ is then given by a function on $\mathcal{S}'(\mathbb{R}^3)$ with

$$\int |\Phi(T)|^2 \mathrm{d}\mu(T) = \|\Phi\|^2 \ . \tag{5.22}$$

The time zero fields smeared with test functions $f \in \mathcal{S}(\mathbb{R}^3)$,

$$(\varphi(0,f)\Phi)(T) = T(f)\Phi(T) \ . \tag{5.23}$$

act as multiplication operators on this space. An obvious difficulty consists in the fact that there is no Lebesgue measure on an infinite dimensional vector space.

We proceed in the following way. We consider the algebra generated by the operators $e^{i\varphi(0,f)}$

$$\mathcal{A} = \{ \sum_{f \in \mathcal{S}(\mathbb{R}^3)} c_f e^{i\varphi(0,f)}, c_f \in \mathbb{C}, c_f \neq 0 \text{ for finitely many } f\} \tag{5.24}$$

Each element C of this algebra defines via

$$C(T) = \sum_f c_f e^{iT(f)} \tag{5.25}$$

a continuous bounded function on $\mathcal{S}'(\mathbb{R}^3)$, and the pointwise product of these functions corresponds to the product of operators,

$$(C_1 C_2)(T) = C_1(T) C_2(T) \ . \tag{5.26}$$

A measure on $\mathcal{S}'(\mathbb{R}^3)$ can now be characterized in terms of a linear functional on this algebra of functions

$$\int \mathrm{d}\mu(T) C(T) = \mu(C) \ . \tag{5.27}$$

(*Radon measure*). In our case we can use the vacuum expectation value as the linear functional,

$$\mu(C) = \big\langle \Omega, C\Omega \big\rangle \ . \tag{5.28}$$

The corresponding measure describes the probability distribution of the configurations T of the time zero field in the vacuum.

The measure is hereby characterized by its Fourier transform,

$$\chi(f) := \hat{\mu}(f) = \mu(e^{i\varphi(0,f)}) = \int \mathrm{d}\mu(T) e^{iT(f)} \ . \tag{5.29}$$

Due to the positivity of the scalar product on Fock space (equivalent to the positivity of the measure) χ (the so-called characteristic function of the measure) satisfies the condition

$$\sum_{f,g} \chi(f-g) c_f \overline{c_g} \geq 0 \ , \tag{5.30}$$

i.e., χ is a function of *positive type*. Conversely, any continuous function on $\mathcal{S}(\mathbb{R}^3)$ which satifies the positivity property above is the Fourie transform of a measure on $\mathcal{S}'(\mathbb{R}^3)$ (Minlos' Theorem).

In our case

$$\chi(f) = e^{-\frac{1}{2}\left\langle f, \frac{1}{2\omega} f\right\rangle} \tag{5.31}$$

with

$$\left(\frac{1}{2\omega} f\right)(\mathbf{x}) = \int \mathrm{d}^3\mathbf{y} \Delta_+(0, \mathbf{x}-\mathbf{y}) f(\mathbf{y}) \ . \tag{5.32}$$

Measures whose Fourier tranforms are exponentials of a positive definite quadratic form are called Gaussian measures. The quadratic form is the covariance of the measure,

$$\int \mathrm{d}\mu(T) T(\mathbf{x}) T(\mathbf{y}) = \left\langle \Omega, \varphi(0,\mathbf{x})\varphi(0,\mathbf{y})\Omega\right\rangle = \Delta_+(0, \mathbf{x}-\mathbf{y}) \ . \tag{5.33}$$

On $\mathbb{R}^n$, Gaussian measures are characterized by a positive semidefinite $n \times n$-matrix K,

$$\int \mathrm{d}\mu_K(x) x_i x_j = K_{ij} \tag{5.34}$$

$$\int \mathrm{d}\mu_K e^{i(x,y)} = e^{-\frac{1}{2}\left\langle y, Ky\right\rangle} \ . \tag{5.35}$$

In case K is invertible, one obtains by an inverse Fourier transform

$$\mathrm{d}\mu_K(x) = (2\pi)^{-\frac{n}{2}} \det(K)^{-\frac{1}{2}} e^{-\frac{1}{2}\left\langle x, K^{-1}x\right\rangle} \mathrm{d}^n x \ . \tag{5.36}$$

In the infinite dimensional case there is no Lebesgue measure, the above factorization of the Gaussian measure thus looses its meaning. The Gaussian measure itself, however, remains well defined in the infinitely dimensional case.

Exercise 5: *Let V be an infinite dimensional real vector space. A cylinder set Z is given by a finite family of linearly independent linear functionals $l_1, \ldots, l_n$ on V and a measurable set $B \subset \mathbb{R}^n$,*

$$Z(l_1, \ldots l_n; B) = \{v \in V | (l_1(v), \ldots, l_n(v)) \in B\}$$

Let μ be a map from the set of cylinder sets to $[0, \infty]$ such that

(i) *For any countable union of pairwise disjoint cylinder sets Z_i*

$$\mu(\bigcup_{i=1}^{\infty} Z) = \sum_{i=1}^{\infty} \mu(Z_i)$$

(ii) *μ is translation invariant, i.e.*

$$\mu(Z + v) = \mu(Z)$$

for all $v \in V$.

Prove, that there is no cylinder set Z such that $0 < \mu(Z) < \infty$.

Gaussian measures can easily be defined on infinite dimensional spaces in terms of their characteristic functions. They possess, however, some somewhat surprising properties which cannot occur in finite dimensional spaces.

Let us consider a real separable infinite dimensional pre-Hilbert space $\mathfrak{D}$. The function

$$\chi(f) = e^{-\frac{1}{2}\langle f,f\rangle} \ . \tag{5.37}$$

is of positive type, i.e. satisfies (5.30). Let $\mathfrak{D}'$ be the space of (not necessarily continuous) linear functionals on $\mathfrak{D}$. We define a measure μ on $\mathfrak{D}'$ as a linear functional on the algebra of functions $l \to \sum c_f e^{il(f)}$ in terms of the characteristic function χ. We ask now on which type of functionals the measure is concentrated. The surprising statement is:

THEOREM IV.8. *The set of continuous linear functionals has the measure zero.*

PROOF. Let $\|l\| := \sup_{\|f\|=1}|l(f)|$. We define the function

$$F(l) = \begin{cases} e^{-\frac{\lambda}{2}\|l\|^2} & , \quad \|l\| < \infty \\ 0 & , \quad \|l\| = \infty \end{cases} \tag{5.38}$$

($\lambda > 0$). We want to show that

$$\int \mathrm{d}\mu(l) F(l) = 0 \ . \tag{5.39}$$

Since for $c > 0$

$$\int \mathrm{d}\mu(l) F(l) \geq \mu(\{l \in \mathfrak{D}', ||l|| < c\}) e^{-\frac{\lambda}{2}c^2} \tag{5.40}$$

this would imply that $\mu(\{l \in \mathfrak{D}', ||l|| < c\}) = 0 \forall c > 0$. But the set of bounded functionals can be written as as a disjoint union

$$\{l \in \mathfrak{D}', ||l|| < \infty\} = \bigcup_{n=0}^{\infty} \{l \in \mathfrak{D}', n < ||l|| \leq n+1\} \ . \tag{5.41}$$

Due to the σ-additivity of the measure then also the set of all bounded functionals (which coincides on a pre-Hilbert space with the space of continuous functionals) has measure zero.

In order to prove (5.39), we choose an orthonormal basis $(f_k)_{k\in\mathbb{N}}$ of $\mathfrak{D}$ and set

$$F_n(l) = e^{-\frac{\lambda}{2}\sum_{k=1}^{n} l(f_k)^2} \ . \tag{5.42}$$

We have

$$F(l) = \lim_{n\to\infty} F_n(l) \ , \ F_n(l) \leq 1 \ . \tag{5.43}$$

Hence F is the point wise limit of a uniformly bounded sequence of cylinder functions and thus integrable, and we have

$$\int \mathrm{d}\mu(l) F(l) = \lim_{n\to\infty} \int \mathrm{d}\mu(l) F_n(l) \ . \tag{5.44}$$

But the integral of F_n is the integral of the Gaussian measure

$$(2\pi)^{-n/2} \int \mathrm{d}^n x e^{-\frac{1+\lambda}{2}\sum_{k=1}^n x_k^2} = (1+\lambda)^{-n/2} \tag{5.45}$$

on $\mathbb{R}^n$, thus $\int \mathrm{d}\mu(l) F_n(l) \to 0$ as $n \to \infty$. This implies the assertion (5.39). □

A consequence of this Theorem is that in the Wiener integral the set of differentiable paths has measure zero. By a modification of the argument used in the proof above one can show that for any Hilbert-Schmidt operator A on $\mathfrak{D}$ the set of linear functionals l for which $l \circ A$ is not continuous is a set of measure zero. In this way one can show that the Wiener integral is concentrated on the continuous paths.

In the case of $\mathfrak{D} = \mathcal{S}(\mathbb{R}^3)$ and a continuous scalar product (in the sense of the topology of Schwartz space) one can show that the Gaussian measure is concentrated on the space of tempered distributions.

We realized the Fock space as the L^2 space of a Gaussian measure with covariance $\frac{1}{2\omega}$ over $\mathcal{S}'(\mathbb{R}^3)$. The time zero fields act as multiplication operators, and the vacuum vector corresponds to the function $\Omega(T) = 1$. It remains to determine the canonical conjugated momenta. They act as functional derivatives plus a term which is caused by the missing translation invariance of the Gaussian measure. One obtains for $C \in \mathcal{A}$

$$\big(\pi(0,f)C\big)(T) = \big(\pi(0,f)C\Omega\big)(T) = \big([\pi(0,f),C]\Omega\big)(T) + C\big(\pi(0,f)\Omega\big)(T) \,. \tag{5.46}$$

The term with the commutator follows from the canonical commutation relations,

$$[\pi(0,f),C] = \frac{1}{i}\int \mathrm{d}^3\mathbf{x} \frac{\delta C}{\delta\varphi(0,\mathbf{x})} f(\mathbf{x}) \,. \tag{5.47}$$

The action on the vacuum results from

$$\pi(0,f)\Omega = \dot\varphi(0,f)\Omega = iH_0\varphi(0,f)\Omega = i\varphi(0,\omega f)\Omega \,. \tag{5.48}$$

After these preparations we can compute the integral kernel of e^{-tH_0}. We consider it as a distribution in two variables,

$$\int \big(e^{-tH_0}\big)(T,T')\mathrm{d}\mu(T)\mathrm{d}\mu(T')\overline{\Phi(T)}\Psi(T') := \big\langle \Phi, e^{-tH_0}\Psi\big\rangle \,. \tag{5.49}$$

Its Fourier transform is

$$\big\langle \Omega, e^{i\varphi(0,f)}e^{-tH_0}e^{i\varphi(0,g)}\Omega\big\rangle = e^{-\frac{1}{2}\left\langle f,\frac{1}{2\omega}f\right\rangle}e^{-\frac{1}{2}\left\langle g,\frac{1}{2\omega}g\right\rangle}e^{-\left\langle f,\frac{e^{-t\omega}}{2\omega}g\right\rangle} \,. \tag{5.50}$$

This is a function of positive type and thus the characteristic function of a measure. In analogy to the theory of Brownian motion we consider this measure as the probability distribution of a diffusion process described by the operator H_0. It gives the probability for an orbit of field configurations$(T_{t'})_{t'}$ to have at time t the value T' provided it had

at time 0 the value T. Correspondingly we define also the probability distributions for field configurations at times $t_1 > \cdots > t_n$ with characteristic function

$$\exp\left(-\frac{1}{2}\sum_j \langle f_j, \frac{1}{2\omega} f_j\rangle - \sum_{j<k} \langle f_j, \frac{e^{-(t_j-t_k)\omega}}{2\omega} f_k\rangle\right) . \tag{5.51}$$

The transition to continuous times is done as usual. Let $f(x^0, \mathbf{x}) = \sum_k f_k(\mathbf{x})\delta(x^0 - t_k)$. Then the quadratic form occuring in the characteristic function of a Gaussian measure is

$$\langle f, S_2 f\rangle = \int \mathrm{d}^4x f(x) S_2(x-y) f(y) \tag{5.52}$$

with the 2-point Schwinger function

$$S_2(x) = (2\pi)^{-4} \int \mathrm{d}^4p \frac{e^{ipx}}{|p|^2 + m^2} . \tag{5.53}$$

The 2-point Schwinger function arises from the Feynman propagator by replacing p_0 by ip_0 and x_0 by ix_0 ("'*Wick rotation*"'). For $x_0 = 0$ it coincides with Δ_+. The Schwinger function is analytic for $x \neq 0$ with an analytic extension into some region of $\mathbb{C}^4$. We already know that Δ_+ is the boundary value of an analytic function H_+ on the forward tube $\mathbb{R}^4 - iV_+$; Δ_- is the boundary value of an analytic function H_- on the backward tube $\mathbb{R}^4 + iV_+$. Δ_+ and Δ_- coincide for spacelike arguments, hence, by the Edge of the Wedge Theorem there is an analytic function H with domain including a complex neighborhood of the set of spacelike points together with the forward and the backward tube which extends both H_+ and H_-. The restriction of H to the so-called euclidean points $(ix^0, \mathbf{x})$, $(x^0, \mathbf{x}) \in \mathbb{R}^4$ is the 2-point Schwinger function

$$H(ix^0, \mathbf{x}) = S_2(x^0, \mathbf{x}) . \tag{5.54}$$

The Gaussian measure with the 2-point Schwinger function as covariance defines a probability distribution μ_0 on $\mathcal{S}'(\mathbb{R}^4)$. We associate to each test function $f \in \mathcal{S}(\mathbb{R}^4)$ a random variable $\varphi(f)$ by

$$\varphi(f)(T) = T(f) \ , \ T \in \mathcal{S}'(\mathbb{R}^4) . \tag{5.55}$$

φ is called the euclidean free massive scalar field. Its correlation functions are called the Schwinger functions. S_2 is the Green's function of the differential operator $-\Delta + m^2$. We interpret μ_0 as giving a precise meaning to the formal expression

$$\mathrm{d}\mu_0 = Z^{-1} e^{-I_E(\varphi)} D\varphi \tag{5.56}$$

with the euclidean action

$$I_E(\varphi) = \int \mathrm{d}^4x \frac{1}{2}\big((\partial\varphi)^2 + m^2\varphi^2\big) \tag{5.57}$$

and the integral over all field configurations $D\varphi = \prod_x \mathrm{d}\varphi(x)$. In a lattice approximation of the euclidean field this formula can be used

directly. In the case of a continuous spacetime the formula is only heuristic, since the Lebesgue integral $D\varphi$ does not exist and, moreover, the integrand with probability 1 is equal to zero.

Exercise 6: *Prove that the Schwinger functions of the free massive scalar euclidean field satisfy the Schwinger-Dyson euqations*

$$(-\Delta_x + m^2)S_n(x, x_2, \ldots, x_n) = \sum_{i=2}^{n} \delta(x - x_i) S_{n-2}((x_j)_{j\neq i}) \ . \tag{5.58}$$

Interacting euclidean field theories are formally obtained by the Feynman-Kac formula. For the associated probability measure one sets

$$\mathrm{d}\mu(\varphi) = Z^{-1} e^{-\int \mathrm{d}^4 x V(\varphi(x))} \mathrm{d}\mu_0(\varphi) \ . \tag{5.59}$$

Unfortunately, contrary to quantum mechanics, the function $e^{-\int V}$ is not integrable, in most cases of interest. In the case of a translation invariant interaction this is a euclidean version of Haag's theorem. If one multiples V with a test function g with compact support, then in 2 dimensions, for polynomials V which are bounded from below, the function $e^{-\int gV}$ is integrable. In this way one obtains a family of probability measures μ_g on $\mathcal{S}'(\mathbb{R}^2)$. The limit points for $g \to 1$ can be considered as interacting euclidean field theories. By analytic continuation of the correlation functions one then obtains the Wightman functions of an interacting quantum field.

The euclidean action on d dimensional euclidean space is the static part of the energy of a classical scalar field in d spatial dimensions. The measure defined above can therefore be considered as the canonical ensemble for a classical statistical system with temperature $\hbar$. This surprising relation between statistical mechanics and euclidean quantum field theory allows to transfer results from one field to the other. An example is the concept of a phase transition which is known from statistical mechanics and can now be applied in quantum field theory to describe sudden changes as the parameters of the model vary.

In the attempt to define interacting euclidean theories one meets similar problems as in Minkowski space. One of these problems is the definition of powers of the fields. Since φ takes values in the space of tempered distributions, terms of the form $\varphi(x)^n$ are not well defined. In analogy to Minkowski space we try to define euclidean Wick powers $:\varphi(x)^n:$. For this purpose we consider the structure of Schwinger functions. We have

$$\begin{aligned} S_n(x_1, \ldots, x_n) :&= \int \mathrm{d}\mu_0 \varphi(x_1) \cdots \varphi(x_n) \\ &= \begin{cases} 0 & , \quad n \text{ odd} \ , \\ \sum_{\text{pairings}} \prod_{\text{pairs}} S_2(x_i, x_j) & , \quad n \text{ even} \ . \end{cases} \end{aligned} \tag{5.60}$$

This formula is completely analogous to the formulas for the time ordered functions of the free field, with S_2 replacing the Feynman propagator. It differs from the formula for the Wightman functions of the free field in the way that there the pairs have to be ordered according to the indices, since the Wightman 2- point function $\Delta_+(x-y)$ of the free field is not symmetric under permutation of x and y.

We thus can use the same combinatorial formulas as for the time ordered functions. Let $\mathcal{G}(n)$ be the set of graphs G with vertices $v \in \{1,\dots,n\}$ and undirected lines $l \in E(G)$ which connect two different vertices (external lines are not admitted), such that every vertex v is incident at exactly one line l, denoted as $v \in \partial l$. Then

$$\int \mathrm{d}\mu_0 \varphi(x_1)\cdots\varphi(x_n) = \sum_{G\in\mathcal{G}} \prod_{l\in E(G)} S_2(\{x_v, v\in\partial l\}) \ . \tag{5.61}$$

The correlation functions for Wick polynomials arise by identification of certain vertices and by omitting all lines between them (so no tadpoles occur). Let $\mathcal{G}(n_1,\dots,n_k)$ be the set of graphs G with vertices $\{1,\dots,k\}$ and undirected lines l such that the vertex i is incident at exactly n_i lines. We define the Wick polynomials first as linear forms on the space of polynomials by

$$\int \mathrm{d}\mu_0 \frac{:\varphi(x)^n:}{n!}\varphi(x_1)\cdots\varphi(x_k) = \sum_{G\in\mathcal{G}(n,\underbrace{1,\dots,1}_{k})} \prod_{l\in E(G)} S_2(\{x_v, v\in\partial l\}) \ . \tag{5.62}$$

For noncoinciding points $x_1,\dots,x_k$ also the correlation functions of Wick polynomials can be defined,

$$\int \mathrm{d}\mu_0 \prod_i \frac{:\varphi(x_i)^{n_i}:}{n_i!} = \sum_{G\in\mathcal{G}(n_1,\dots,n_k)} \prod_{i<j} \frac{S_2(x_i-x_j)^{l_{ij}}}{l_{ij}!} \ , \tag{5.63}$$

where l_{ij} is the number of lines between the vertices i and j. In the perturbative renormalization of euclidean field theories one constructs extensions of these functions to everywhere defined distributions. It should, however, be noticed that these extensions cannot, in general, be considered as the correlation function of a euclidean field, in particular $:\varphi^n:(f)$ is, in general, not a random variable.

As an example we consider the euclidean free field in 3 dimensions. In this case the Schwinger function is given by

$$S_2(x) = \frac{e^{-m|x|}}{4\pi|x|} \ . \tag{5.64}$$

The 2-point correlation function of the nth Wick power is not integrable for $n > 2$. But for $n = 2$ we find

$$\int \mathrm{d}\mu_0 |:\varphi^2:(f)|^2 = \int \mathrm{d}^6(x,y)\overline{f(x)}f(y)S_2(x-y)^2 < \infty \ , \tag{5.65}$$

$:\varphi^2:(f)$ is thus square integrable. In the case $n > 2$ the corresponding integral diverges, and for the renormalized 2-point function the result is not necessarily positive, hence cannot occur as the expectation value of a positive random variable.

Exercise 7: *The path integral for Fermi fields is the so-called Berezin integral. Let $\mathcal{A}$ denote the unital associative algebra which is generated by the elements $a_1, \ldots, a_n$ with the relations*

$$a_i a_j + a_j a_i = 0 \ , \ i, j = 1, \ldots, n$$

(Grassmann algebra). The Berezin integral is a linear functional $\int da_1 \ldots da_n$ on $\mathcal{A}$ with the properties

$$\int da_1 \ldots da_n \, a_1 \ldots a_n = 1,$$

$$\int da_1 \ldots da_n \prod_{j \in I} a_j = 0 \text{ for } I \overset{\neq}{\subset} \{1, \ldots, n\}.$$

Compute the Gaussian integral

$$\int da_1 \ldots da_n \, e^{\sum_{i<j} a_i a_j A_{ij}}, A_{ij} \in \mathbb{C} \ .$$

6. Connected functions (truncated functions)

In the graphical expansion of the terms of perturbation theory it is convenient to decompose the graphs into connected components. The distribution corresponding to the graph is the tensor product of the distributions corresponding to its connected components of the graph, hence it is sufficient to construct the distributions for connected graphs. Actually, the decomposition of correlation functions into connected pieces can be defined independent of perturbation theory and can in particular also performed for interacting theories.

Let ω be a linear functional on a (not necessarily commutative) unital algebra $\mathcal{A}$ with the normalization condition $\omega(1) = 1$. We think hereby, e.g., on the Wightman functional on the tensor algebra of test functions,

$$\omega(f_1 \otimes \cdots \otimes f_n) = \left\langle \Omega, \varphi(f_1) \cdots \varphi(f_n) \Omega \right\rangle \ , \tag{6.1}$$

or on the system of time ordered functions as a functional on the symmetrical tensor algebra of test functions,

$$\omega(f_1 \otimes \cdots \otimes f_n) = \left\langle \Omega, T\varphi(f_1) \cdots \varphi(f_n) \Omega \right\rangle \ . \tag{6.2}$$

A further possibility are probability measures, considered as linear functionals on the algebra of random variables which associate to each random variable its expected value.

Let us first assume that ω has an expansion labeled by graphs,

$$\omega(A_1 \cdots A_n) = \sum_{G \in \mathcal{G}} \omega_G(A_1, \ldots, A_n) \ . \tag{6.3}$$

Here ω_G is for each graph G a multilinear functional on $\mathcal{A}$ which factorizes if the graphs can be decomposed into disconnected subgraphs. We decompose every graph into its connected components. Thereby the set of vertices is decomposed into pairwise disjoint nonempty subsets,

$$\{1,\dots,n\} = I_1 \cup \dots \cup I_k \ , \ I_j \cap I_l = \emptyset \text{ für } j \neq l \ . \tag{6.4}$$

Such a decomposition is called a partition, and we denote the set of partitions of $\{1,\dots,n\}$ by $\mathrm{Part}(\{1,\dots,n\})$. We now, in a first step, fix some partition and sum only over those graphs whose decomposition into connected components induces the given partition, and sum afterwards over the set of all partitions. Let

$$\omega_c(A_1,\dots,A_n) = \sum_{G\in\mathcal{G}_c} \omega_G(A_1,\dots,A_n) \ , \tag{6.5}$$

where $\mathcal{G}_c \subset \mathcal{G}$ denotes the subset of connected graphs. Then

$$\omega(A_1 \dots A_n) = \sum_{P\in\mathrm{Part}(\{1,\dots,n\})} \prod_{I\in P} \omega_c(A_i, i\in I) \ . \tag{6.6}$$

We now can use the same formula also in the case where no expansion labeled by graphs is given, and consider it as an implicit definition of connected functions ω_c as multilinear functionals on $\mathcal{A}$. Indeed it is possible to solve the equation above and to give an explicit formula for the connected functions, e.g., we have $\omega_c(A) = \omega(A)$, $\omega_c(A_1,A_2) = \omega(A_1A_2) - \omega(A_1)\omega(A_2)$, and, more generally, the recursion relation

$$\omega_c(A_1,\dots,A_n) = \omega(A_1\cdots A_n) - \sum_{\sharp(P)>1} \prod_{I\in P} \omega_c(A_i, i\in I) \ , \tag{6.7}$$

where $\sharp(P)$ denotes the number of elements of the partition P. The connected functions in this abstract meaning are often called truncated functions.

There also exists a closed formula for the connected functions. Multilinear functionals on vector spaces can always be considered as linear functional on the tensor product of the vector spaces. In our case we consider the algebra $\mathcal{A}$ as a vector space, i.e., we ignore for the moment the multiplicative structure. The connected functions form a system of multilinear maps and can formally combined to a linear functional on the tensor algebra

$$T\mathcal{A} = \bigoplus_{n=0}^{\infty} \mathcal{A}^{\otimes n} \tag{6.8}$$

(with $\omega_c(1) = 0$). On the set of linear functionals on $T\mathcal{A}$ we introduce the following associative product,

$$(FG)(A_1\otimes\dots\otimes A_n) = \sum_{I\subset\{1,\dots,n\}} F(\bigotimes_{i\in I} A_i) G(\bigotimes_{j\in I^c} A_j) \ , \tag{6.9}$$

where I^c denotes the complement of I in $\{1,\dots,n\}$. The unit for this product is the linear functional

$$1(A_1 \otimes \cdots \otimes A_n) = \delta_{n0} \ . \tag{6.10}$$

The defining equation (6.6) for the connected functions can be written in terms of this product in the following form,

$$\omega \circ m = e^{\omega_c} = \sum_{n=0}^{\infty} \frac{\omega_c^n}{n!} \ . \tag{6.11}$$

Here we used the multiplication in the algebra $\mathcal{A}$ for the definition of a linear map

$$m : \left\{ \begin{array}{ccc} T\mathcal{A} & \to & \mathcal{A} \\ A_1 \otimes \cdots \otimes A_n & \mapsto & A_1 \cdots A_n \end{array} \right. \tag{6.12}$$

Namely, we have

$$\omega_c^k(A_1 \otimes \cdots \otimes A_n) = \sum_{I_1,\dots,I_k \subset \{1,\dots,n\}} \prod_j \omega_c(\bigotimes_{i \in I_j} A_i) \tag{6.13}$$

where the index sets are pairwise disjoint with union $\{1,\dots,n\}$. The contributions of the empty sets vanish because of $\omega_c(1) = 0$, hence we can restrict the sum to the partitions $P \in \mathrm{Part}(\{1,\dots,n\}$ with k elements. Each partition occurs $k!$ times, corresponding to the number of possibilities to number the index sets. After division by $k!$ and summation over k we obtain equation (6.6).

From (6.11) we obtain, by inversion of the power series of the exponential function, the wanted formula for the connected functions,

$$\omega_c = \log \omega \circ m = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} (\omega \circ m - 1)^k \ . \tag{6.14}$$

The series above converges because of $(\omega \circ m - 1)(1) = 0$. In n-th order we find

$$\omega_c(A_1,\dots,A_n) = \sum_{P \in \mathrm{Part}(\{1,\dots,n\})} (-1)^{\sharp(P)} (\sharp(P)-1)! \prod_{I \in P} \omega(\prod_{i \in I} A_i) \ . \tag{6.15}$$

We want to apply the formula for elements of the form

$$\exp_\otimes A = \sum_{k=0}^{\infty} \frac{1}{k!} A^{\otimes k} \ , \ A \in \mathcal{A} \ . \tag{6.16}$$

If one evaluates linear functionals on $T\mathcal{A}$ on these elements, then the product of functionals yields the product of values,

$$(FG)(\exp_\otimes A) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} F(A^{\otimes k}) G(A^{\otimes (n-k)}) = F(\exp_\otimes A) G(\exp_\otimes A) \ . \tag{6.17}$$

Thus we find

$$\omega_c(\exp_\otimes A) = \log(\omega(e^A)) \ . \tag{6.18}$$

(Here we used that $m(\exp_\otimes A) = e^A$.)

In these formulas nothing is assumed on the convergence of the series. Instead they are interpreted in the sense of formal power series.

For the characteristic function of the probability measure of the interacting theory we obtain

$$\chi(f) = \frac{\mu_0(e^{i\varphi(f)}e^{-\int V})}{\mu_0(e^{-\int V})} = \exp\,(\mu_0)_c((\exp_\otimes i\varphi(f) - 1) \otimes \exp_\otimes(-\int V)) \ . \tag{6.19}$$

In the case $V = (g/4!) : \varphi^4 :$ we get the following graphical expansion for the connected correlation functions,

$$\mu_c(\varphi(x_1), \dots, \varphi(x_n)) =$$

$$\sum_{k=0}^{\infty} \frac{(-g)^k}{k!} \int \mathrm{d}^{4k}(x_{n+1}, \dots, x_{n+k}) \sum_{G \in \mathcal{G}_c(n\times 1, k\times 4)} \prod_{1\le i<j\le n+k} \frac{S_2(x_i - x_j)^{l_{ij}}}{l_{ij}!} \ . \tag{6.20}$$

In this expansion the first n vertices which are not integrated out are called external vertices, and the other inner vertices.

A completely analogous formula holds for the connected time ordered functions of the φ^4 theory. One only replaces $-g$ by ig and S_2 by $i\Delta_F$. By the use of the LSZ relations one then can determine the connected S-matrix elements.

There exists a further formula for the computation of the connected n-point function. For the 2-point function we have

$$\omega_c(A_1, A_2) = \frac{1}{2}(\omega \otimes \omega)(\tilde{A}_1 \tilde{A}_2) \tag{6.21}$$

with $\tilde{A} = A \otimes 1 - 1 \otimes A$ where $\mathcal{A} \otimes \mathcal{A}$ is considered as an algebra with the product

$$(A_1 \otimes A_2)(B_1 \otimes B_2) = A_1 B_1 \otimes A_2 B_2 \ . \tag{6.22}$$

A corresponding formula exists for the n-point function

$$\omega_c(A_1, \dots, A_n) = \frac{1}{n}\omega^{\otimes n}(\tilde{A}_1 \cdots \tilde{A}_n) \tag{6.23}$$

with

$$\tilde{A} = \sum_{k=1}^{n} e^{2\pi i(k-1)/n} 1 \otimes \cdots \underset{k\text{-th position}}{A} \otimes \cdots 1 \ . \tag{6.24}$$

This formula is especially useful when one wants to exploit positivity properties of ω (ω_c is i.g. not a positive functional). If, e.g., ω is a probability measure μ, then $(\omega_c)_n$ can be computed in terms of the product measure $\mu \times \cdots \times \mu$.

Exercise 8: *Prove equation* (6.23).

7. One-particle-irreducible functions (vertex functions)

The correlation functions of a translation invariant theory depend only on the relative coordinates. This leads to further factorization property. Let G be some graph which consists of two subgraphs G_1 and G_2 which are connected by a line l_0. Let S_{G_1} and S_{G_2} be the associated contributions to the Schwinger function. Then

$$S_G(x_i, i \in V(G)) = S_{G_1}(x_i, i \in V(G_1)) S_2(x_j, j \in \partial l_0) S_{G_2}(x_i, i \in V(G_2)) \,. \tag{7.1}$$

This formula makes sense even in the case that the factors are distributions. Namely, let v_1, v_2 be the end points of the line l_0 in the graphs G_1 and G_2, respectively. Because of translation invariance, S_{G_1} depends only on the relative coordinates $y_i = x_i - x_{v_1}$, $i \in V(G_1) \setminus \{v_1\}$ and S_{G_2} only on the relative coordinates $y_i = x_i - x_{v_2}$, $i \in V(G_2) \setminus \{v_2\}$. Together with $y_{l_0} = x_{v_1} - x_{v_2}$ one obtains a system of independent relative coordinates, and the product above is a tensor product.

We may therefore decompose connected graphs into so-called *one particle irreducible* (**1PI**) subgraphs (a better name would be *one line irreducible graphs*). A connected graph is called 1PI if it remains connected after removal of an arbitrary line.

For performing this decomposition we introduce an equivalence relation on the set of vertices. Two vertices i and j of a graph are called strongly connected if they are connected by a path in any graph which arises from G by erasing one line. Equivalence classes of strongly connected vertices form together with their internal lines maximal 1PI subgraphs. The external lines which connect a 1PI subgraph with the rest of the graph are called amputated legs. We associate to a 1PI graph G with vertex set $V(G)$, inner lines $l \in E_i(G)$ and amputated legs $l \in E_a(G)$ (where ∂l denotes the vertex at which the leg is attached) a function

$$\Gamma_G(y_l, l \in E_a(G)) = \int \prod_{v \in V(G)} d^4 y_v \prod_{l \in E_i(G)} S_2(y_v, v \in \partial l) \prod_{l \in E_a(G)} \delta(y_l - y_{\partial l}) \,. \tag{7.2}$$

After contractions of the 1PI subgraphs of a connected graph to a single vertex there remains a tree graph (i.e. a graph without loops) which characterizes the way the full graph is built from its 1PI subgraphs, together with the information to which amputated leg of the 1PI subgraph the lines of the tree graph are attached.

For a connected graph with two external vertices (these are the graphs which contribute to the 2-point function), the arising tree graph has no bifurcations, and on its vertices of order 2 there are 1PI subgraphs. We now sum the functions Γ_G for all 1PI graphs with two

amputated legs and obtain a function $\Sigma(x,y)$. The perturbative formula for the connected 2-point function now reads

$$G_2(x,y) = S_2(x-y) + \sum_{n=1}^{\infty} \int \mathrm{d}^{4n} z S_2(x-z_1)\Sigma(z_1,z_2)S_2(z_2-z_3)\cdots\Sigma(z_{n-1},z_n)S_2(z_n-y) \ . \tag{7.3}$$

We interpret the functions G_2, Σ and S_2 as integral kernels of operators in $L^2(\mathbb{R}^4)$. Then the expansion in equation (7.3) assumes the form

$$G_2 = \sum_{n=0}^{\infty} S_2(\Sigma S_2)^n \ . \tag{7.4}$$

This is a geometric series with the sum

$$G_2 = (-\Delta + m^2 - \Sigma)^{-1} \ . \tag{7.5}$$

Σ is called the self energy (i.e. m^2 is replaced by $m^2 - \Sigma$).

In an analogous way we sum Γ_G over all 1PI graphs with n amputated legs, $n \neq 2$, and obtain a function $\Gamma_n(x_1,\ldots,x_n)$. Γ_n is called the n-point vertex function.

We now can write the connected n-point functions G_n (also called Green's functions) in terms of the vertex functions and the connected 2-point-function G_2.

In the case of $n = 1$ we find

$$G_1(x) = \int d^4y \, G_2(x,y)\Gamma_1(y) \tag{7.6}$$

Let $\mathcal{T}_n$ be the set of tree graphs T with n external vertices $v \in V_e(T)$ and an arbitrary finite number of inner vertices $v \in V_i(T)$ with $n_v > 2$ incident lines. The number of lines $l \in E(T)$ of a tree graph is equal to the number of vertices minus 1. Each line connects two vertices. Thus we have

$$2\sharp(E(T)) = n + \sum_{v\in V_i(T)} n_v = 2(n + \sharp(V_i(T)) - 1) \ , \tag{7.7}$$

hence

$$\sum_v (n_v - 2) = n - 2 \ . \tag{7.8}$$

We now associate to every pair (v,l) with $v \in \partial l$ a variable y_{vl}. We obtain

$$G_n = \sum_{T\in\mathcal{T}_n} G_T \tag{7.9}$$

with

$$G_T(y_{vl}|v \in V_e(T)) = \int \prod_{(v,l),v\in V_i(T)} \mathrm{d}^4 y_{vl} \prod_{v\in V_i(T)} \Gamma_{n_v}(y_{vl}, \partial l \ni v) \prod_{l\in E(T)} G_2(y_v, v \in \partial l) \ . \tag{7.10}$$

Also the vertex functions can be introduced independently of a graphical expansion. We set $\Gamma_2 = -G_2^{-1}$ (in the sense of operators), $\Gamma_1 = 0$ (we restrict ourselves to the case $G_1 = 0$) and $\Gamma_0 = 0$.

First we consider an apparently completely different problem. Let $\phi(x;j)$ be the expected value of the field $\varphi(x)$ under the influence of an additional interaction term $-\varphi(j) = -\int \mathrm{d}^4x\varphi(x)j(x)$ with a test function j,

$$\phi(x;j) = \mu_j(\varphi(x)) = \frac{\mu(\varphi(x)e^{\varphi(j)})}{\mu(e^{\varphi(j)})} = \frac{\delta}{\delta j(x)}G(j) \tag{7.11}$$

with the generating function of connected functions $G(j) = \log \mu(e^{\varphi(j)})$,

$$G(j) = \sum_{n=0}^{\infty} \frac{1}{n!} G_n(j^{\otimes n}) \ . \tag{7.12}$$

We look now for the so-called effective action, i.e. a functional Γ on the space of classical field configurations with ϕ as stationary point. We set

$$\Gamma(\phi) = \sup_j \big(\phi(j) - G(j)\big) \ . \tag{7.13}$$

If the supremum is reached for some j and if G is differentiable there, j satisfies the equation $\phi = \frac{\delta G}{\delta j}$. Γ is the Legendre transform of G. Inversely,

$$G(j) = \sup_\phi \big(\phi(j) - \Gamma(\phi)\big) \ . \tag{7.14}$$

is the Legendre transform of Γ (provided G is convex). The supremum is reached at $j = \frac{\delta \Gamma}{\delta \phi}$. We observe that $\frac{\delta \Gamma}{\delta \phi} : \phi \mapsto j$ is the inverse of the mapping $\frac{\delta G}{\delta j} : j \mapsto \phi$.

The inverse of a power series in one variable is the so-called Bürmann-Lagrange series. Here we have to find an infinite dimensional version of this formula. We shall show that the vertex functions Γ_n are, up to a sign, the coefficients of the power series expansion of Γ.

Let

$$\Gamma^{(n)}(x_1, \dots, x_n; \phi) = \frac{\delta^n \Gamma(\phi)}{\delta\phi(x_1) \cdots \delta\phi(x_n)} \tag{7.15}$$

and

$$G_n(x_1, \dots, x_n; j) = \frac{\delta^n G(j)}{\delta j(x_1) \cdots \delta j(x_n)} \ . \tag{7.16}$$

$\phi(x) = G_1(x, j)$ implies

$$\frac{\delta\phi(x)}{\delta j(y)} = G_2(x, y; j) \ . \tag{7.17}$$

Thus by the chain rule

$$\frac{\delta\Gamma^{(n)}(x_1, \dots, x_n; \phi(j))}{\delta j(y)} = \int \mathrm{d}x_{n+1} \Gamma^{(n+1)}(x_1, \dots, x_{n+1}; \phi(j)) G_2(x_{n+1}, y; j) \ . \tag{7.18}$$

Hence we find

$$\delta(x-y) = \frac{\delta j(x)}{\delta j(y)} = \int \mathrm{d}z \Gamma^{(2)}(x,z;\phi(j)) G_2(z,y;j) \ . \tag{7.19}$$

As an integral kernel, $\Gamma^{(2)}$ is the inverse of G_2.

We differentiate this equation with respect to j and obtain

$$\begin{aligned} 0 = \int \mathrm{d}z \int \mathrm{d}z' \big(& \Gamma^{(3)}(x,z,z';\phi(j)) G_2(z,y;j) G_2(z',y';j) \big) \\ & + \Gamma^{(2)}(x,z;\phi(j)) G_3(x,y,y';j) \ . \end{aligned} \tag{7.20}$$

Solving for G_3 yields

$$\begin{aligned} & G_3(y,y',y'';j) \\ = & - \int \mathrm{d}z \mathrm{d}z' \mathrm{d}z'' G_2(y,z;j) G_2(y',z';j) G_2(y'',z'',j) \Gamma^{(3)}(z,z',z'';\phi(j)) \ . \end{aligned} \tag{7.21}$$

But this is identical to the formula for the expansion of G_3 into 1PI functions after insertion of $\Gamma_3 = -\Gamma^{(3)}$.

We now assume that the formula (7.10) with $\Gamma_n = -\Gamma^{(n)}$ holds up to order n, and differentiate with respect to j.

We obtain

$$\begin{aligned} & G_{n+1}(x_1,\dots,x_n,x;j) = \sum_T \int \prod_{(v,l),v\in V_i(T)} \mathrm{d}y_{vl} \\ & \int \big(\mathrm{d}z \sum_v \Gamma_{n_v+1}(y_{vl}, \partial l \ni v, z; \phi(j)) G_2(z,x;j) \\ & \qquad \times \prod_{v' \neq v} \Gamma_{n_{v'}}(y_{n_{v'}};\phi(j)) \prod_l G_2(y_{vl}, v \in \partial l; j) \big) \\ & + \prod_v \Gamma_{n_v}(y_{vl}, \partial l \ni v; \phi(j)) \sum_l G_3(y_{vl}, v \in \partial l, x; j) \prod_{l' \neq l} G_2(y_{vl'}, v \in \partial l'; j) \end{aligned} \tag{7.22}$$

We then insert equation (7.21). The claim follows from the fact that every tree graph with boundary points $1,\dots,n+1$ and inner vertices of order larger than 2 can be obtained from one with boundary points $1,\dots,n$ by either connecting the point $n+1$ with a line (thereby a new vertex of order 3 is generated, this corresponds to the first sum in the formula above) or connecting with an inner vertex (2nd part of the sum).

A byproduct of the derived relation is that one may also expand the vertex functions in terms of connected functions One only has to replace G_n by $-\Gamma_n$ and Γ_n by $-G_n$.

The relations above hold independently of the nature of the connected functions. In case they are the Green's functions of a translation

invariant theory and depend only on the relative coordinates it is useful to go to momentum space. The Fourier transforms of the n-point functions are of the form

$$\hat{G}_n(p_1, \ldots, p_n) = \delta(\sum p_i) g_n(p_1, \ldots, p_n) \tag{7.23}$$

and

$$\hat{\Gamma}_n(p_1, \ldots, p_n) = \delta(\sum p_i) \gamma_n(p_1, \ldots, p_n) \ , \tag{7.24}$$

where g_n and γ_n are defined only on the hypersurface $\sum p_i = 0$. One now inserts these expressions into equation (7.10). All integrals over the internal vertices can be easily performed because of the δ-functions, and there remain only the momenta at the boundary points with the condition $\sum p_i = 0$. All momenta at the inner vertices are uniquely determined by the external momenta. We obtain

$$g_n(p_1, \ldots, p_n) = \sum_T \prod_l g_2(p_{\partial l}) \prod_v \gamma_{n_v}(p_v) \ . \tag{7.25}$$

We may now use the LSZ relations for calculating the S-matrix. We have to take into account that the self energy $\hat{\Sigma}(p, q) = \delta(p+q)\sigma(p)$ will shift the mass of the particle. In case the Green's functions are Lorentz invariant, σ is a function of p^2. If $p^2 = M^2$ is a simple zero of the function $p^2 - m^2 - \sigma(p)$, we interpret M as the mass of the interacting particle. According to the LSZ relations we obtain the S-matrix elements by multiplying the Fourier transformed Green's functions with $p_i^2 - M^2$ for every external vertex i and putting afterwards the external momenta onto the mass shell $p_i^2 = M^2$. On the mass shell we have

$$(p^2 - M^2) g_2(p, -p) = \text{const} \ . \tag{7.26}$$

The contribution of the connected function to the S-matrix thus arises (up to a normalization factor) from the expansion (7.25) in terms of vertex functions by omitting all external lines ("amputation of external legs").

CHAPTER V

Renormalization

1. Mass and wave function renormalization

The perturbative construction of interacting quantum field theories yields expressions which are ill defined due to the singularities of the Feynman propagator. As an example let us look at a scalar field theory with interaction $\frac{g}{4!}\varphi^4$. The first contribution to the self energy is

$$\Sigma(x,y) = \frac{g^2}{3!}(i\Delta_F(x-y))^3 \ .$$

Δ_F is singular for lightlike arguments. Nevertheless, the powers of $\Delta_F(x)$ are well defined for $x \neq 0$. Namely, outside of the origin $i\Delta_F$ coincides either with Δ_+ (for $x \notin \overline{V_-}$) or with Δ_- (for $x \notin \overline{V_+}$). Both distributions have well defined powers, e.g. for $f \in \mathcal{D}(\mathbb{R}^4)$ we have

$$\int \mathrm{d}^4x \Delta_+(x)^3 f(x) = (2\pi)^{-7} \int \frac{\mathrm{d}^3\mathbf{p}_1}{2\omega_1}\frac{\mathrm{d}^3\mathbf{p}_2}{2\omega_2}\frac{\mathrm{d}^3\mathbf{p}_3}{2\omega_3}\hat{f}(\omega_1+\omega_2+\omega_3, \mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3)$$

with the Fourier transform $\hat{f}$ of f. Since $\hat{f}$ decays fast at infinity, the integral on the right hand side converges and defines a distribution.

There remains the singularity at $x = 0$. In a first attempt we try, as for the powers of Δ_+, a representation by the Fourier transforms and obtain

$$\int \mathrm{d}^4x (i\Delta_F(x)^3 f(x) =$$
$$(2\pi)^{-10} \int \mathrm{d}^4p \hat{f}(p) \int\int \frac{\mathrm{d}^4k \mathrm{d}^4q}{(k^2-m^2+i\varepsilon)(q^2-m^2+i\varepsilon)((p-k-q)^2-m^2+i\varepsilon)} \ .$$

The problem is that the integral over k and q does not converge. We now expand the integrand in a Taylor series up to 2nd order in p around

$p = 0$,

$$\begin{aligned}
&\frac{1}{(k^2-m^2+i\varepsilon)(q^2-m^2+i\varepsilon)((p-k-q)^2-m^2+i\varepsilon)} = \\
&\frac{1}{(k^2-m^2+i\varepsilon)(q^2-m^2+i\varepsilon)((k+q)^2-m^2+i\varepsilon)} \\
+&\frac{2p(k+q)}{(k^2-m^2+i\varepsilon)(q^2-m^2+i\varepsilon)((k+q)^2-m^2+i\varepsilon)^2} \\
+&\frac{4(p(k+q))^2+p^2((k+q)^2-m^2)}{(k^2-m^2+i\varepsilon)(q^2-m^2+i\varepsilon)((k+q)^2-m^2+i\varepsilon)^3} \\
&+R(p,k,q)\ .
\end{aligned}$$

The remainder R decays for large momenta with the 9th power and should not raise problems for the integration. For the first term we expect a quadratic divergence; this term does not depend on the external momentum p. The second term diverges linearly and is a linear function of p. The 3rd term finally seems to diverge logarithmically and depends quadratically on p.

For making these considerations precise, we regularize the integrand by multiplying it with a test function which vanishes for large momenta. Let $w \in \mathcal{D}(\mathbb{R}^8)$ with $w(0) = 1$. We multipy the integrand with $w(\frac{k}{\Lambda}, \frac{q}{\Lambda})$ with the cutoff momentum $\Lambda > 0$. In the limit $\Lambda \to \infty$ this factor converges towards 1. From the terms of the Taylor expansion we obtain after integration over k and q the terms

$$a(\Lambda) + b_\mu(\Lambda)p^\mu + c_{\mu\nu}(\Lambda)p^\mu p^\nu + R(p,\Lambda)\ .$$

For large Λ, $a(\Lambda)$ behaves as Λ^2, $b_\mu(\Lambda)$ as Λ and $c_{\mu\nu}(\Lambda)$ as $\ln(\Lambda)$. $R(p,\Lambda)$ converges for $\Lambda \to \infty$ against a distribution $R(p)$. The question is how to interpret the divergent quantities..

In a first step we exploit the freedom in the choice of w to achieve at $b_\mu = 0$ and $c_{\mu\nu} = cg_{\mu\nu}$. For finite values of Λ one obtains as quadratic term in the effective action

$$\Gamma(\phi) = \int \mathrm{d}^4x\big(\frac{1}{2}(1+c)\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}(m^2+a)\phi^2\big) + \int \mathrm{d}^4x \int \mathrm{d}^4y \phi(x)\hat{R}(x-y;\Lambda)\phi(y)\ .$$

We set $Z = 1 + c$ (wave function renormalization) and $m^2_{\mathrm{ren}} = (m^2 + a)Z^{-1}$ (mass renormalization). Then we redefine the field as $\hat{\phi} = \sqrt{Z}\phi$ and the effective action as a function of the redefined field, $\hat{\Gamma}(\hat{\phi}) = \Gamma(\phi)$. The local contributions to $\hat{\Gamma}$ look like the classical action with the renormalized mass. We now fix $\hat{\phi}$ and m_{ren} and calculate ϕ and m (the "bar" mass) as functions of Λ. In the limit $\Lambda \to \infty$ the bar quantities diverge, but all observables are expected to depend only on the renormalized quanatities. This method is called multiplicative renormalization.

Another (equivalent) possibility for eliminating the divergences consists in adding Λ-dependent counter to the Lagrangean which compensate the divergent terms in the effective action. In our case suitable counter terms are

$$-\frac{c}{2}\partial_\mu\phi\partial^\mu\phi + \frac{a}{2}\varphi^2 \ .$$

This procedure is called additive renormalization.

2. Coupling constant renormalization

The 4-point vertex function in 2nd order is

$$\Gamma_4(x_1,x_2,x_3,x_4) = g\delta(x_1-x_2)\delta(x_2-x_3)\delta(x_3-x_4)$$
$$+\frac{g^2}{2}\int \mathrm{d}^4y\mathrm{d}^4z(i\Delta_F(y-z))^2\delta(x_1-y)\big(\delta(x_2-y)\delta(x_3-z)\delta(x_4-z)$$
$$+\delta(x_3-y)\delta(x_2-z)\delta(x_4-z)+\delta(x_4-y)\delta(x_2-z)\delta(x_3-z)\big) \ .$$

The calculation of Δ_F^2 leads to a logarithmically divergent integral in momentum space,

$$\widehat{(i\Delta_F)^2}(p) = (2\pi)^{-6}\int \mathrm{d}^4k \frac{1}{(k^2-m^2+i\varepsilon)((p-k)^2-m^2+i\varepsilon)} =$$
$$(2\pi)^{-6}\int \mathrm{d}^4p\left(\frac{1}{(k^2-m^2+i\varepsilon)^2}+\frac{p^2-2kp}{(k^2-m^2+i\varepsilon)^2((p-k)^2-m^2+i\varepsilon)}\right) \ .$$

The second term is finite, the first term diverges logarithmically. We regularize the integral as in the preceding section and obtain a contribution of the form $d(\Lambda)$. Therefore, the coefficient of ϕ^4 in the effective action is (up to nonlocal terms which remain finite in the limit $\Lambda\to\infty$)

$$\frac{g}{4!}+\frac{g^2}{2\cdot 4!}d(\Lambda) \ .$$

As before we have the two alternatives: we could introduce the renormalized coupling constant

$$g_{\mathrm{ren}} = (g+\frac{g^2}{2}d)Z^{-2}$$

and express the bar coupling constant g as a function of Λ (multiplicative renormalization) or we add a counterterm $-\frac{g^2}{2\cdot 4!}d\phi^4$ to the Lagrangean.

One can show that in the ϕ^4-theory all divergent terms in the effective action are of the form $\frac{a}{2}(\partial\phi)^2$, $\frac{c}{2}\phi^2$ and $\frac{d}{4!}\phi^4$. Thus the additive as well as the multiplicative renormalization can be performed in every order of perturbation theory.

In the proof of this claim divergences of subdiagrams create problems which could be solved only around 20 years after the work of Tomonaga, Schwinger, Feynman and Dyson. In the momentum representation these problems already arise in the treatment of the so-called "setting sun"-graph from Section 1, since it contains the "fish"-graph

from this section three times as divergent subgraph. Thus the integral over the remainder R does not converge absolutely, and the result depends on the choice of the test function w.

Fortunately, in the position space representation the problem does not occur in this order. This may be seen in the following way.

We already know that the powers of Δ_+ are well defined distributions. We may write them in the following way

$$\Delta_+(x)^n = \int_0^\infty \mathrm{d}M^2 \rho_n(M^2) \Delta_+(x, M) \ .$$

with the so-called phase space density

$$\rho_n(M^2) = (2\pi)^{-3(n-1)} \int \prod_j \frac{\mathrm{d}^3\mathbf{p}_j}{2\omega_j} \delta(M - \sum \omega_j) \delta(\sum \mathbf{p}_j) \ .$$

This expression is a special case of the Källan-Lehmann representation of the 2-point function of a scalar field, applied to the 2-point function of $:\varphi^n:$. One easily verifies that $\rho_n(M^2)$ grows for $M \to \infty$ as M^{2n-4} and vanishes for $M < nm$. For $n = 2$, e.g., we find after one momentum space integration

$$\rho_2(M^2) = (2\pi)^{-3} \int \frac{\mathrm{d}^3\mathbf{p}}{4\omega^2} \delta(M - 2\omega) \ .$$

and arrive finally at

$$\rho_2(M^2) = \frac{1}{16\pi^2} \sqrt{1 - \frac{4m^2}{M^2}} \ .$$

A corresponding expression with the same function ρ holds for Δ_-.

It is suggestive to use the same representation also for the powers of the Feynman-propagators. But this would lead to a divergent intergral over M^2. But one can exploit the fact that $\Delta_\pm(x; M)$ are solutions of the Klein-Gordon equation for mass M. We set

$$(i\Delta_F(x))^n_{\mathrm{ren}} = (-\Box + a)^n \int_{n^2m^2}^\infty \mathrm{d}M^2 \frac{\rho_n(M^2)}{(M^2 + a)^{n-1}} i\Delta_F(x, M) \ .$$

This expression coincides for $x \neq 0$ with the unrenormalized expression. It depends nontrivially on the choice of the constant a.

As an example we compute the difference of the renormalized squares of the Feynman propagator with constants a and 0. We obtain

$$\begin{aligned} &(i\Delta_F(x)^2_{\mathrm{ren},a=0} - (i\Delta_F(x)^2_{\mathrm{ren},a} \\ &= (-\Box) \int \mathrm{d}M^2 \rho_2(M^2) (\frac{1}{M^2} - \frac{1}{M^2 + a}) i\Delta_F(x, M) \\ &\quad -a \int \mathrm{d}M^2 \frac{\rho_2(M^2)}{M^2 + a} i\Delta_F(x, M) \ . \end{aligned} \tag{2.1}$$

In the first term, the d'Alembertian $\Box$ can be interchanged with the integration, because of the faster decay in M^2. $i\Delta_F$ is a Green's function for the Klein-Gordon operator for mass M, thus

$$-\Box i\Delta_F(x,M) = M^2 i\Delta_F(x,M) + \delta(x) \ .$$

The first term on the right hand side compensates precisely the second integral in (2.1). We observe that the renormalized expressions differ by a multiple of the δ-function.

3. Regularization and Renormalization Methods

3.1. Pauli-Villars Regularization. A simple method for regularizing the Feynman propagator goes back to Pauli and Villars. In this method one subtracts from the Feynman propagator suitable linear combinations of Feynman propagators to other masses,

$$\Delta_F^{\rm reg}(x) := \sum_{i=1}^{n} c_i \Delta_F(x, M_i) \tag{3.1}$$

with $M_1 = m$, $c_1 = 1$, $\sum c_i M_i^{2k} = 0$, $k = 0, \ldots, n-2$. In the limit $M_i \to \infty$, $i > 1$ one gets back the original Feynman propagator. In the simplest case ($n = 2$) one chooses

$$\Delta_F^{\rm reg}(x) = \Delta_F(x,m) - \Delta_F(x,M) \ . \tag{3.2}$$

The behavior of the Feynman propagator at high energies is in leading order independent of M, hence the Fourier transform of the regularized propagator behaves as

$$\widehat{\Delta_F^{\rm reg}} \sim \frac{1}{(k^2)^2} \ , \ k^2 \to \infty \ (n=2) \ .$$

In terms of the regularized propagator (3.2) the fish graph becomes finite. For stronger divergent graphs one has to choose higher values of n.

The Pauli-Villars regularization can be understood as follows: we choose n independent free fields φ_{M_i} with masses M_i, $i = 1, \ldots, n$, form the tensor product of the associated Fock spaces and define a field

$$\phi = \sum \sqrt{c_i}\varphi_{M_i} \ . \tag{3.3}$$

The Feynman propagator of ϕ is the regularized propagator (3.1). Since not all of the coefficients c_i can be positive, the field ϕ is not hermitean. One may change the scalar product in the corresponding Fock spaces, such that ϕ is again hermitean, but then the scalar product is no longer positive definite.

A nice feature of the Pauli-Villars method is that it is Lorentz invariant. The Pauli-Villars method can, however, not directly be applied to nonabelian gauge theories.

Another version of Pauli-Villars regularization is to subtract the contributions of higher masses for a given graph, e.g. for the setting-sun graph

$$(\Delta_F(x)^3)^{\mathrm{reg}} = \sum_{i=1}^{3} c_i \Delta_F(x, M_i)^3$$

with $c_1 = 1$, $M_1 = m$ and $c_2 = \frac{M_3^2 - M_1^2}{M_2^2 - M_3^2}$, $c_3 = \frac{M_1^2 - M_2^2}{M_2^2 - M_3^2}$.

3.2. Feynman Parameters. For the calculation of Feynman graphs often the following method is useful. Namely, one can rewrite the momentum space integral over a product of n Feynman propagators as an integral over the $(-n)$th power of a quadratic form in the momenta, followed by an integral over the parameters determining the quadratic form. Then the momentum space integration can be performed, and it remains an integral over the parameters.

The method is based on the following lemma:

LEMMA V.1. *Let $a_1, \dots, a_n > 0$. Then*

$$\frac{1}{a_1 \cdots a_n} = (n-1)! \int_{x_1,\dots,x_n>0} \mathrm{d}x_1 \cdots \mathrm{d}x_n \frac{\delta(\sum x_i - 1)}{(\sum x_i a_i)^n} \ .$$

PROOF. For $n = 1$ the formula is obviously correct. We assume that it is correct for some $n \geq 1$. For $n+1$ the right hand side is

$$n! \int \mathrm{d}x_1 \cdots \mathrm{d}x_{n+1} \frac{\delta(\sum^{n+1} x_i - 1)}{(\sum^{n+1} x_i a_i)^{n+1}} \ .$$

We may evaluate the δ-function over x_{n+1} and obtain the equivalent expression

$$n! \int_{\sum x_i < 1} \mathrm{d}x_1 \cdots \mathrm{d}x_n \frac{1}{\left(\sum^n x_i(a_i - a_{n+1}) + a_{n+1}\right)^{n+1}} \ .$$

We now can perform the x_n-integration. The result is

$$(n-1)! \frac{1}{a_n - a_{n+1}} \int_{\sum x_i < 1} \mathrm{d}x_1 \cdots \mathrm{d}x_{n-1} \Big(\sum^{n-1} x_i(a_i - b) + b)\Big)^{-n} \Big|_{b=a_n}^{b=a_{n+1}} \ .$$

According to our induction hypothesis this coincides with

$$\frac{1}{a_n - a_{n+1}} \left(\frac{1}{a_1 \cdots a_{n-1} a_{n+1}} - \frac{1}{a_1 \cdots a_{n-1} a_n} \right) = \frac{1}{a_1 \cdots a_{n+1}} \ .$$

□

We use this method for calculating the fish graph. In order to avoid problems with varying signs we consider the corresponding problem for the euclidean theory.

The square of the 2-point Schwinger function is formally given by

$$S_2(x)^2 = (2\pi)^{-8} \int \mathrm{d}^4 p \, e^{ipx} \int \mathrm{d}^4 k \frac{1}{(k^2 + m^2)((p-k)^2 + m^2)} \ .$$

According to the lemma above we may replace the integrand of the k-integral by

$$\int_0^1 \mathrm{d}x \frac{1}{\big(x((p-k)^2+m^2)+(1-x)(k^2+m^2)\big)^2} \ .$$

The quadratic form in the denominator is

$$k^2+m^2+xp^2-2xpk \ .$$

We set $q=k-xp$ and $a=m^2+x(1-x)p^2$. There remains the integral

$$\int \mathrm{d}^4q(q^2+a)^{-2} \ .$$

This integral diverges logarithmically. We cut it off at $|q|=\Lambda$ and obtain

$$\int_{|q|\leq\Lambda} \mathrm{d}^4q(q^2+a)^{-2} = 2\pi^2 \int_0^\Lambda \mathrm{d}r\, r^3(r^2+a)^{-2} = -\pi^2 \int_0^\Lambda \mathrm{d}r\, r^2 \frac{\mathrm{d}}{\mathrm{d}r}\frac{1}{r^2+a} \ .$$

Here we used in the first step that the (3-dimensional) surface area of a ball in $\mathbb{R}^4$ with radius r is $2\pi^2r^3$. We now perform partial integration and find

$$-\pi^2\frac{\Lambda^2}{\Lambda^2+a} + \pi^2 \int_0^\Lambda \mathrm{d}r \frac{2r}{r^2+a} \ .$$

The antiderivative of the remaining integrand is $\ln(r^2+a)$. Therefore we obtain

$$\int_{|q|<\Lambda} \mathrm{d}^4q(q^2+a)^{-2} = -\pi^2\frac{\Lambda^2}{\Lambda^2+a} + \pi^2 \ln\frac{\Lambda^2+a}{a} \ .$$

The first term converges for $\Lambda\to\infty$ towards $-\pi^2$. The second term diverges logarithmically. It holds for $b>0$

$$\ln\frac{\Lambda^2+a}{a} = \ln\frac{\Lambda^2+a}{\Lambda^2+b} + \ln\frac{\Lambda^2+b}{b} + \ln\frac{b}{a} \ .$$

We observe that the divergent term can be chosen to be independent of a. Thus the divergence does not depend on the external momentum. We now subtract the value of the integral at $p=0$ and get in the limit $\Lambda\to\infty$

$$\int \mathrm{d}^4q \left((q^2+m^2+x(1-x)p^2)^{-2} - (q^2+m^2)^{-2}\right) = -\pi^2 \ln(1+x(1-x)\frac{p^2}{m^2}) \ .$$

It remains the integral over the Feynman parameter x. For this purpose we factorize the argument of the logarithm

$$1+x(1-x)\frac{p^2}{m^2} = -\frac{p^2}{m^2}(x-\frac{1}{2}-\sqrt{\frac{1}{4}+\frac{m^2}{p^2}})(x-\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{m^2}{p^2}}) \ .$$

The antiderivative of the logarithm is $x(\ln x - 1)$. Hence

$$\int_0^1 \mathrm{d}x \ln(1 + x(1-x)\frac{p^2}{m^2}) = \ln \frac{p^2}{m^2} + 2x(\ln x - 1)\Big|_{-\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{m^2}{p^2}}}^{\frac{1}{2}+\sqrt{\frac{1}{4}+\frac{m^2}{p^2}}} .$$

After a few manipulations we obtain the result

$$2(\theta \coth \theta - 1)$$

with $\sinh \theta = \frac{p^2}{4m^2}$.

The contribution of the fish graph, subtracted at momentum $p = 0$, is therefore

$$\lim_{\Lambda\to\infty} \left(\hat{S}_2^2(p)_\Lambda - \hat{S}_2^2(0)_\Lambda\right) = \frac{1}{32\pi^4}\left(1 - \sqrt{1+\frac{4m^2}{p^2}}\mathrm{Arsinh}\sqrt{\frac{p^2}{4m^2}}\right) =: f(p^2) .$$

For the vertex function $\Gamma_4(p_1,p_2,p_3,p_4) = \gamma_4(p_1,p_2,p_3,p_4)\delta(\sum p_i)$ the fish graph gives the contribution

$$\gamma_4^{\text{fish}}(p_1,p_2,p_3,p_4) = f(s) + f(t) + f(u)$$

with the so-called Mandelstam variables $s = (p_1+p_2)^2$, $t = (p_1+p_3)^2$ und $u = (p_1+p_4)^2$ (under the condition $\sum p_i = 0$).

3.3. α-Parameter or Schwinger's Proper Time Formalism. We use the integral representation

$$\frac{1}{p^2+m^2} = \int_0^\infty \mathrm{d}\alpha\, e^{-\alpha(p^2+m^2)}$$

of the propagator. The contribution of the fish graph is

$$\begin{aligned} I(p) &:= \int \frac{\mathrm{d}^4 q}{(q^2+m^2)((p-q)^2+m^2)} \\ &= \int_0^\infty \mathrm{d}\alpha \int_0^\infty \mathrm{d}\beta \int \mathrm{d}^4 q\, e^{-(\alpha+\beta)m^2 - \alpha q^2 - \beta(p-q)^2} . \end{aligned}$$

The integral over q is of the Gaussian form. We compute

$$\int \mathrm{d}^4 q\, e^{-\alpha q^2 - \beta(p-q)^2} = \int \mathrm{d}^4 q\, e^{-(\alpha+\beta)(q - \frac{\beta}{\alpha+\beta}p)^2 - \frac{\alpha\beta}{\alpha+\beta}p^2} = \frac{\pi^2}{(\alpha+\beta)^2} e^{-\frac{\alpha\beta}{\alpha+\beta}p^2} .$$

We set $\alpha+\beta = \gamma$ und $\frac{\alpha}{\alpha+\beta} = x$. Then

$$I(p) = \pi^2 \int_0^1 \mathrm{d}x \int_0^\infty \mathrm{d}\gamma\, \gamma^{-1} e^{-\gamma(m^2 + x(1-x)p^2)} .$$

The integral over γ diverges at $\gamma = 0$ logarithmically.

We now proceed as in the preceding section and subtract the value at $p = 0$ before integrating over γ. We obtain an expression of the form

$$F(a,b) = \int_0^\infty \frac{\mathrm{d}\gamma}{\gamma}(f(\gamma a) - f(\gamma b)) .$$

For an integral of this form one finds

$$\begin{aligned} F(a,b) &= \lim_{\varepsilon\downarrow 0} \int_\varepsilon^\infty \frac{\mathrm{d}\gamma}{\gamma}(f(\gamma a) - f(\gamma b)) \\ &= \lim_{\varepsilon\downarrow 0} \int_{\varepsilon a}^{\varepsilon b} \frac{\mathrm{d}\gamma}{\gamma} f(\gamma) = f(0) \ln \frac{b}{a} \ . \end{aligned}$$

This yields the renormalized contribution of the fish graph

$$I(p)_{\mathrm{ren}} = -\pi^2 \int_0^1 \mathrm{d}x \ln(1 + x(1-x)\frac{p^2}{m^2}) \ ,$$

in agreement with the calculation in terms of Feynman parameters.

3.4. Dimensional Regularization. The ultraviolet divergences become weaker in lower spacetime dimensions. In the α-parameter-representation the spacetime dimension occurs as an exponent in an homogeneous polynomial of the α-parameter. In the example of the preceding section we obtain for the unrenormalized integral as a function of the dimension d the expression

$$I_d(p) = \mu^{4-d}\pi^{d/2} \int \mathrm{d}x \int_0^\infty \mathrm{d}\gamma\, \gamma^{-\frac{d-2}{2}} e^{-\gamma(m^2 + x(1-x)p^2)} \ . \tag{3.4}$$

Here μ is an arbitrarily chosen parameter with the dimension of a mass such that the integral remains dimensionless for $d \neq 4$. The integrand is well defined for complex d, and the integral is absolutely convergent for $\mathrm{Re}\, d < 4$. For $a > 0$ we have

$$\int \mathrm{d}\gamma\, \gamma^{-\frac{d-2}{2}} e^{-\gamma a} = a^{\frac{d-4}{2}} \int \mathrm{d}\gamma\, \gamma^{-\frac{d-2}{2}} e^{-\gamma} = a^{\frac{d-4}{2}} \Gamma(\frac{4-d}{2}) \ .$$

The Γ-function is meromorphic with poles at the points $0, -1, -2, \ldots$. It satisfies the functional equation

$$x\Gamma(x) = \Gamma(x+1) \ .$$

Thus

$$\Gamma(\frac{4-d}{2}) = \frac{\Gamma(\frac{6-d}{2})}{\frac{4-d}{2}} \ .$$

We observe that $I_d(p)$ has an analytic extension with a simple pole at $d = 4$. The coefficient of the pole term is π^2. After subtraction of the pole term we obtain a function which is analytic in a neighborhood of $d = 4$,

$$I_d(p)_{\mathrm{ren}} = I_d(p) - \frac{\pi^2}{\frac{4-d}{2}} = \frac{2}{d-4} \left(\int_0^1 \mathrm{d}x \pi^{d/2} \left(\frac{m^2 + x(1-x)p^2}{\mu^2} \right)^{\frac{d-4}{2}} \Gamma(\frac{6-d}{2}) - \pi^2 \right)$$

Let $\varepsilon = \frac{4-d}{2}$. The limit $d \to 4$ is then the derivative of the function

$$\int_0^1 \mathrm{d}x \pi^{2-\varepsilon} \left(\frac{\mu^2}{m^2 + x(1-x)p^2} \right)^{\varepsilon} \Gamma(1+\varepsilon)$$

with respect to ε at $\varepsilon = 0$. The result is

$$I(p)_{\text{ren}} = \pi^2 \int_0^1 \mathrm{d}x \left(-\ln \pi - \ln \frac{m^2 + x(1-x)p^2}{\mu^2} + \Gamma'(1) \right) .$$

It depends on the choice of the parameter μ. The obtained expression is called minimally subtracted (MS-scheme). The constant $\gamma = -\Gamma'(1) = 0.5772$ is the so-called Euler-Mascheroni constant. Often one uses the so-called modified minimal subtraction ($\overline{\text{MS}}$, speak *MS bar*). Here in addition the constant $-\pi^2(\gamma + \ln 4\pi)$ is subtracted.

These conventions have no physical meaning. Physical meaningful quantities are, e.g., the differences $I(p)_{\text{ren}} - I(0)_{\text{ren}}$. They coincide with the previously derived expressions.

3.5. Extension of Distributions. The methods discussed up to now rely essentially on the form of the Feynman propagator in momentum space. But we already observed that in the case of graphs with two vertices the arising powers of the Feynman propagators in position space can be immediately given for $x \neq 0$. The remaining problem is the extension of a distribution which is defined on test functions which vanish in a neighborhood of the origin to a distribution which is defined everywhere.

The mathematical question of existence and uniqueness of extensions can be answered according to Steinmann in terms of the so-called scaling degree of the distribution [**?**].

Definition V.2. *Let T be a distribution on an open subset $U \subset \mathbb{R}^d$ with $\lambda U \subset U$ for $\lambda < 1$. The scaling degree* $\mathrm{sd}(T)$ *is the smallest element of $\mathbb{R} \cup \{\pm\infty\}$ with the property*

$$\lim_{\lambda \downarrow 0} \lambda^\delta T(\lambda x) = 0 \ \forall \delta > \mathrm{sd}(T)$$

(convergence in the sense of distributions).

Examples 1. (i) *Let f be a continuous function of one real variable with $f(0) \neq 0$. Then*

$$\lim_{\lambda \downarrow 0} \lambda^\delta \int \mathrm{d}x\, f(\lambda x)\varphi(x) \to \begin{cases} 0 & , \ \delta > 0 \\ f(0) \int \varphi & , \ \delta = 0 \\ \infty \int \varphi & , \ \delta < 0 \end{cases}$$

hence $\mathrm{sd}(f) = 0$.

(ii) *The δ-function on $\mathbb{R}^d$ satisfies*

$$\delta(\lambda x) = \lambda^{-d} \delta(x) ,$$

hence $\mathrm{sd}(\delta) = d$.

(iii) *The function $f(x) = e^{\frac{1}{x}}$ defines a distribution on $\mathbb{R} \setminus \{0\}$. Since $\lambda^\delta e^{\frac{1}{\lambda x}}$ diverges for all $\delta \in \mathbb{R}$, we have* $\mathrm{sd}(f) = \infty$.

(iv) *The function* $f(x) = e^{-\frac{1}{x^2}}$ *on* $\mathbb{R} \setminus \{0\}$ *has the property*

$$\lambda^\delta f(\lambda x) \to 0 \ \forall \delta \in \mathbb{R} \ ,$$

thus the scaling degree of the corresponding distribution is $\mathrm{sd}(f) = -\infty$.

(v) *The Feynman propagator in 4 dimensios has the scaling behavior*

$$\Delta_F(\lambda x, m) = \lambda^{-2} \Delta_F(x, \lambda m) \ .$$

In the limit $m \to 0$ *the Feynman propagator converges towards the Feynman propagator of the massless theory. Therefore the scaling degree of the Feynman propagator is*

$$\mathrm{sd}(\Delta_F) = 2 \ . \tag{3.5}$$

(vi) *Let* T *be a tempered distribution on* $\mathbb{R}^d$. *Its Fourier transform is again a tempered distribution* $\hat{T}$ *which formally satisfies*

$$T(\varphi) = \int \mathrm{d}^n p \, \hat{T}(p) \hat{\varphi}(-p)$$

Now assume that $\lambda^s \hat{T}(\lambda^{-1} p) \to 0$ *for* $\lambda \to 0$ *for some* $s \in \mathbb{R}$. *Then*

$$\int \mathrm{d}^d x \, T(\lambda x) \varphi(x) = \lambda^{-d} \int \mathrm{d}^n x \, T(x) \varphi(\lambda^{-1} x) \ .$$

With

$$\varphi(\lambda^{-1} x) = (2\pi)^{-d/2} \lambda^d \int \mathrm{d}^d p \, e^{ipx} \hat{\varphi}(\lambda p)$$

we find

$$\begin{aligned} \lambda^{s+d} \int \mathrm{d}^d x \, T(\lambda x) \varphi(x) &= \int \mathrm{d}^d p \, \hat{T}(p) \hat{\varphi}(-\lambda p) \\ &= \lambda^s \int \mathrm{d}^d p \, \hat{T}(\lambda^{-1} p) \hat{\varphi}(-p) \\ &\to 0 \ \textit{for} \ \lambda \to 0 \ . \end{aligned}$$

Thus the scaling degree of T *satisfies* $\mathrm{sd}(T) \leq d + s$.

The scaling degree has the following properties:

LEMMA V.3. (i) $\mathrm{sd}(\partial^\alpha T) \leq \mathrm{sd}(T) + |\alpha|$ *(Here and in what follows* α *denotes a multi-Index* $\alpha = (\alpha_1, \ldots, \alpha_d)$ *mit* $\alpha_j \in \mathbb{N}_0$. $|\alpha| = \sum_j \alpha_j$ *is the length of the multi-Index.)*

(ii) $\mathrm{sd}(x^\alpha T) \leq \mathrm{sd}(T) - |\alpha|$

(iii) $\mathrm{sd}(S + T) \leq \max(\mathrm{sd}(S), \mathrm{sd}(T))$

(iv) $\mathrm{sd}(fT) \leq \mathrm{sd}\, T$ *for* $f \in \mathcal{C}^\infty(\mathbb{R}^d)$.

PROOF. The last property needs a proof. Let $\delta > \mathrm{sd}(T)$. Then

$$\lambda^\delta \int \mathrm{d}^d x\, T(\lambda x)\varphi(x) \to 0 \ , \ \forall \varphi \in \mathcal{D}(\mathcal{U})$$

with $\mathcal{U} = \mathbb{R}^d$ or $\mathbb{R}^d \setminus \{0\}$. Thus the family of distributions $T_\lambda(x) = \lambda^\delta T(\lambda x)$, $\lambda \leq 1$ is weakly bounded. Due to the principle of uniform boundedness (Banach-Steinhaus) it is then also uniformly bounded, i.e. for every compact region $K \subset \mathcal{U}$ there is a polynomial P, such that for $\varphi \in \mathcal{D}(K)$

$$|T_\lambda(\varphi)| \leq \sup_x |P(\partial)\varphi(x)| \ .$$

For the scaling degree of fT follows

$$|\lambda^\delta \int \mathrm{d}^d x f(\lambda x) T(\lambda x)\varphi(x)| \leq \sup_x |P(\partial_x) f(\lambda x)\varphi(x)| \ .$$

The right hand side of the inequality is bounded in λ, $\lambda \leq 1$, hence the scaling degree satisfies $\mathrm{sd}(fT) \leq \delta$. This inequality holds for all $\delta > \mathrm{sd}(T)$. This proves the claim. □

We now arrive at the fundamental existence and uniqueness theorem.

THEOREM V.4. *Let $T_0 \in \mathcal{D}(\mathbb{R}^n \setminus \{0\})$. Then:*

(i) *If $\mathrm{sd}(T_0) < n$, then there exists a unique distribution $T \in \mathcal{D}(\mathbb{R}^n)$ with the properties*

$$T(\varphi) = T_0(\varphi) \ , \ \varphi \in \mathcal{D}(\mathbb{R}^n \setminus \{0\})$$

and $\mathrm{sd}(T) = \mathrm{sd}(T_0)$.

(ii) *In the case $n \leq \mathrm{sd}(T_0) < \infty$ there are distributions $T \in \mathcal{D}(\mathbb{R}^n)$ with the properties*

$$T(\varphi) = T_0(\varphi) \ , \ \varphi \in \mathcal{D}(\mathbb{R}^n \setminus \{0\})$$

and $\mathrm{sd}(T) = \mathrm{sd}(T_0)$. Two such distributions T_1 und T_2 differ by a derivative of the δ-function,

$$T_1 - T_2 = P(\partial)\delta \ ,$$

with a polynomial P with degree $\deg(P) \leq \mathrm{sd}(T) - n$.

(iii) *If $\mathrm{sd}(T_0) = \infty$, there is no distribution $T \in \mathcal{D}(\mathbb{R}^n)$ with the property*

$$T(\varphi) = T_0(\varphi) \ , \ \varphi \in \mathcal{D}(\mathbb{R}^n \setminus \{0\}) \ .$$

PROOF. The proof of this theorem is somewhat lengthy. We start with item (iii). Since every distribution T can be locally represented as a (distributional) derivative of a continuous function, i.e., for every compact set $K \subset \mathbb{R}^n$ there exists a polynomial P such that

$$T(\varphi) = P(\partial) f(\varphi) \text{ for } \varphi \in \mathcal{D}(K) \tag{3.6}$$

the scaling degree of T is finite. Since T_0 is a restriction of T to a smaller test function space, also the scaling degree of T_0 must be finite.

We now prove the uniqueness statement of item (i). Two extensions T_1 and T_2 differ by a distribution R with $\operatorname{supp} R = \{0\}$. Such a distribution is a derivative of the δ-function and has scaling degree $\geq n$ if it does not vanish. Since $\operatorname{sd}(R) \leq \max\{\operatorname{sd}(T_1), \operatorname{sd}(T_2)$ we have $R = 0$, i.e. $T_1 = T_2$.

Next we prove the existence of an extension of T_0 in the case $\operatorname{sd}(T_0) < n$. For this purpose we choose a test function $\chi \in \mathcal{D}(\mathbb{R}^n)$ with $\chi(x) = 1$ in some neighborhood of the origin. For all $\mu > 0$ and all $\varphi \in \mathcal{D}(\mathbb{R}^n)$, $(1-\chi(\mu x))\varphi(x) = 0$ in some neighborhood of the origin, hence

$$T_\mu(x) := T_0(x)(1 - \chi(\mu x)), \mu > 0$$

is a family of distributions which are defined on all of $\mathbb{R}^n$. For each $\varphi \in \mathcal{D}(\mathbb{R}^n \setminus \{0\})$ there is some $\mu_0 > 0$ such that $\chi(\mu x)\varphi(x) = 0$ for all x and for all $\mu \geq \mu_0$, hence

$$\lim_{\mu\to\infty} T_\mu(\varphi) = T_0(\varphi) \ , \ \varphi \in \mathcal{D}(\mathbb{R}^n \setminus \{0\}) \ .$$

We now prove that the limit $(T_\mu)_{\mu\to\infty}$ exists for all test functions. The space of distributions is sequentially complete, hence the limit defines a distribution T.

We write

$$T = T_1 + \int_1^\infty d\mu \frac{d}{d\mu} T_\mu \tag{3.7}$$

and compute

$$\begin{aligned} \langle \frac{d}{d\mu} T_\mu, f\rangle &= -\int d^n x T_0(x) \sum_i^n x_i (\partial_i \chi)(\mu x) f(x) \\ &= -\sum_i^n \mu^{-(n+1)} \int d^n x f(\mu^{-1}x) T_0(\mu^{-1}x) x_i \partial_i \chi(x) \end{aligned} \tag{3.8}$$

Now let $\operatorname{sd}(T_0) < \delta < n$. We have $\operatorname{sd}(fT_0) \leq \operatorname{sd}(T_0)$ by Lemma V.3 , hence

$$\mu^{-\delta} \int d^n x f(\mu^{-1}x) T_0(\mu^{-1}x) x_i \partial_i \chi(x) \to 0 \text{ for } \mu \to \infty \tag{3.9}$$

according to the definition of the scaling degree, and thus

$$|\langle \frac{d}{d\mu} T_\mu, f\rangle| \leq \text{const}\, \mu^{\delta-n-1} \tag{3.10}$$

We conclude that the integral in (3.7) exists.

In case $\operatorname{sd}(T_0) \geq n$ but finite we can draw the same conclusion if we restrict the distribution to test functions f which vanish at the origin at order $\operatorname{sd}(T_0) - n$. Let

$$\mathcal{D}_\omega(\mathbb{R}^n) = \{f \in \mathcal{D}(\mathbb{R}^n), \partial^\alpha f(0) = 0, |\alpha| \leq \omega\} \ .$$

We use the fact that every $f \in \mathcal{D}_\omega(\mathbb{R}^n)$ can be represented in the form

$$f(x) = \sum_{|\alpha|=[\omega]+1} x^\alpha g_\alpha(x)$$

with $g_\alpha \in \mathcal{D}(\mathbb{R}^n)$. This follows from the Taylor expansion of $f(\lambda x)$ at $\lambda = 0$; the terms up to order $[\omega]$ vanish, and the remainder

$$f(x) = \int_0^1 \mathrm{d}\lambda \frac{(1-\lambda)^{[\omega]}}{[\omega]!} \frac{\mathrm{d}^{[\omega]+1}}{\mathrm{d}\lambda^{[\omega]+1}} f(\lambda x)$$

is of the form above but with $g_\alpha \in \mathcal{C}^\infty(\mathbb{R}^n)$. We then multiply the equation by a test function which is equal to 1 on the support of f and obtain the desired representation. We conclude that $\mathrm{sd}(fT_0) \leq \mathrm{sd}(T_0) - [\omega]$.

We may now repeat the argument for the convergence of the integral in (3.7) and get a unique distribution T_ω on $\mathcal{D}_\omega(\mathbb{R}^n)$.

In the last step we extend the distribution T_ω to all of $\mathcal{D}(\mathbb{R}^n)$. For this purpose we choose within $\mathcal{D}(\mathbb{R}^n)$ a complementary subspace $\mathcal{W}$ to $\mathcal{D}_\omega(\mathbb{R}^n)$ and set

$$T = T_\omega \oplus l$$

with an arbitrary linear functional l on $\mathcal{W}$. In this way we obtain all possible extensions of T_ω. The dual space of $\mathcal{W}$ can be identified with the set of distributions which vanish on $\mathcal{D}_\omega(\mathbb{R}^n)$,

$$l \in \mathcal{D}_\omega(\mathbb{R}^n)^\perp = \{P(\partial)\delta, \mathrm{deg} P \leq \omega\} \ .$$

The distributions $\{\partial^\alpha \delta, |\alpha| \leq \omega\}$ form a basis of $\mathcal{D}_\omega(\mathbb{R}^n)^\perp$. The dual basis in $\mathcal{W}$ consists of functions w_α with $\partial^\alpha w_\beta(0) = (-1)^{|\alpha|}\delta_{\alpha\beta}, |\alpha|, |\beta| \leq \omega$, and any family of test functions w_α with this property generates a complementary space of $\mathcal{D}_\omega(\mathbb{R}^n)$. An example are the functions $w_\alpha = \frac{(-x)^\alpha}{\alpha!} w$ with a function w which is equal to 1 in some neighborhood of the origin.

We may now construct a projection W from $\mathcal{D}(\mathbb{R}^n)$ onto $\mathcal{D}_\omega(\mathbb{R}^n)$,

$$W = 1 - \sum_{|\alpha| \leq \omega} |w_\alpha\rangle\langle\partial^\alpha\delta| \ .$$

W may be considered as a modified Taylor subtraction.

If W is fixed, the extensions of T_ω are given by

$$T = T_\omega \circ W + l$$

with $l \in \mathcal{D}_\omega(\mathbb{R}^n)^\perp$.

Actually one can absorb the linear functional l by modifying $\mathcal{W}$. Namely let $g \in \mathcal{D}_\omega(\mathbb{R}^n)$ with $T_\omega(g) = 1$. Then also

$$W' = W + |g\rangle\langle l|$$

is a projection onto $\mathcal{D}_\omega(\mathbb{R}^n)$, and we have

$$T_\omega \circ W' = T_\omega \circ W + l \ .$$

The corresponding dual basis is $\{w'_\alpha = w_\alpha - c_\alpha\psi\}$ for $l = \sum c_\alpha \partial^\alpha \delta$.

It remains to show that the obtained extensions have the correct scaling degree. For this we refer to the literature. □

EXAMPLES 2. (i) *Let* $T_0(\varphi) = \int \mathrm{d}x \, \frac{\varphi(x)}{|x|}$ *for* $\varphi \in \mathcal{D}(\mathbb{R} \setminus \{0\})$. *We have* $\operatorname{sd} T_0 = 1$, *hence* T_0 *can be uniquely extended to* $\mathcal{D}_0(\mathbb{R}) = \{\varphi \in \mathcal{D}(\mathbb{R}), \varphi(0) = 0\}$.

Let $w \in \mathcal{D}(\mathbb{R})$ *with* $w(0) = 1$. *Then*

$$\mathcal{D}(\mathbb{R}) = \mathcal{D}_0(\mathbb{R}) \oplus \{\lambda w, \lambda \in \mathbb{C}\} \ .$$

This holds since for each $\varphi \in \mathcal{D}(\mathbb{R})$

$$\varphi_0(x) = \varphi(x) - w(x)\varphi(0) \in \mathcal{D}_0(\mathbb{R})$$

and thus

$$\varphi = \varphi_0 + \varphi(0)w \ .$$

We define an extension of T_0 *by*

$$\begin{aligned} T(\varphi) = T_0(\varphi_0) &= \int \mathrm{d}x \frac{\varphi(x) - \varphi(0)w(x)}{|x|} \\ &= \lim_{\varepsilon \downarrow 0} \int_{|x|>\varepsilon} \mathrm{d}x \frac{\varphi(x) - \varphi(0)w(x)}{|x|} \\ &= \lim_{\varepsilon \downarrow 0} (-\operatorname{sign} x \ln|x|) \left(\varphi(x) - \varphi(0)w(x)\right) |_{-\varepsilon}^{\varepsilon} \\ &\quad - \int \mathrm{d}x \ \operatorname{sign} x \ln|x| \frac{\mathrm{d}}{\mathrm{d}x} \left(\varphi(x) - \varphi(0)w(x)\right) \\ &= \int \mathrm{d}x \left(\frac{\mathrm{d}}{\mathrm{d}x} \operatorname{sign} x \ln|x|\right) \varphi(x) - \varphi(0) \int \mathrm{d}x \left(\frac{\mathrm{d}}{\mathrm{d}x} \operatorname{sign} x \ln|x|\right) w(x) \ , \end{aligned}$$

hence

$$T = (\operatorname{sign} x \ln|x|)' + c\delta \ \textit{with} \ c = -(\operatorname{sign} x \ln|x|)'(w) \ .$$

(ii) *Let* $T_0(x) = (i\Delta_F(x)^n$, $x \neq 0$. *With the spectral function* ρ_n *we have the representation*

$$T_0(x) = \int \mathrm{d}M^2 \rho_n(M) i\Delta_F(x, M)$$

The scaling degree of T_0 *is* $2n$. *We choose test functions* w_α *with* $\partial^\beta w_\alpha(0) = (-1)^{|\alpha|}\delta_\alpha^\beta$, $|\alpha|, |\beta| \leq 2n - 4$. *Then*

$$W\varphi = \varphi - \sum_\alpha w_\alpha \partial^\alpha \varphi(0)(-1)^{|\alpha|} \in \mathcal{D}_{2n-4}(\mathbb{R}^4) \ .$$

Since the δ*-function and its derivatives vanish up to order* $2n - 4$ *on* $W\varphi$, *we find*

$$\int \mathrm{d}^4x \, \Delta_F(x, \mathcal{M})(W\varphi)(x) = \frac{(-1)^{n-1}}{M^{2(n-1)}} \int \mathrm{d}^4x \, \Delta_F(x, M) \Box^{n-1}(W\varphi)(x)$$

Thus T_ω, $\omega = 2n-4$ is given by

$$T_\omega = (-1)^{n-1}\Box^{n-1}\int \mathrm{d}M^2 \frac{\rho_n(M)}{M^{2(n-1)}} i\Delta_F(\cdot, M) \ .$$

$T = T_\omega \circ W$ is then

$$T = (-1)^{n-1}\Box^{n-1}\int \mathrm{d}M^2 \frac{\rho_n(M)}{M^{2(n-1)}} i\Delta_F(\cdot, M) + \sum_\alpha c_\alpha \partial^\alpha \delta$$

with

$$c_\alpha = -(-1)^{n-1}\Box^{n-1}\int \mathrm{d}M^2 \frac{\rho_n(M)}{M^{2(n-1)}} \int \mathrm{d}^4x i\Delta_F(x, M) w_\alpha(x) \ .$$

In the extension procedure described above no divergences occur. The connection with the previously discussed methods is the following. Let (T^Λ) be a sequence of (everywhere defined) distributions which converges on $\mathcal{D}_\omega(\mathbb{R}^n)$ towards T_ω. Then

$$\begin{aligned} T = T_\omega \circ W &= \lim_\Lambda T^\Lambda \circ W \\ &= \lim_\Lambda T^\Lambda - \sum_\alpha \langle T^\Lambda, w_\alpha\rangle \partial^\alpha \delta \ . \end{aligned}$$

The divergent counter terms are thus given by $\langle T^\Lambda, w_\alpha\rangle \partial^\alpha \delta$. If, on the other hand, $C^\Lambda = \sum_\alpha c^\Lambda_\alpha \partial^\alpha \delta$ are counter terms such that $T^\Lambda - C^\Lambda$ converges, then the limits

$$\lim c^\Lambda_\alpha - \langle T^\Lambda, w_\alpha\rangle =: c_\alpha \tag{3.11}$$

exist, and one may modify the projection W such that the limit is of the form $T_\omega \circ W$.

In case T_ω is tempered (i.e. defined on $\mathcal{S}_\omega(\mathbb{R}^n) = \{\varphi \in \mathcal{S}(\mathbb{R}^n), \partial^\alpha\varphi(0) = 0, |\alpha| \leq \omega\}$), and the sequence $T^\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ converges on $\mathcal{S}_\omega(\mathbb{R}^n)$ towards T_ω, the extension may also be discussed in momentum space.

The counter terms are

$$C^\Lambda(p) = \sum_\alpha \langle T^\Lambda, w_\alpha\rangle i^{|\alpha|} p^\alpha. \tag{3.12}$$

Since $w_\alpha(x) = \frac{1}{\alpha!} x^\alpha v_\alpha$ with $v_\alpha(0) = 1$ and $\partial^{\beta-\alpha} v_\alpha(0) = 0$ for $\alpha < \beta$, $|\beta| \leq \omega$ we find

$$\langle T^\Lambda, w_\alpha\rangle = \frac{1}{\alpha!}\langle x^\alpha T^\Lambda, v_\alpha\rangle \tag{3.13}$$

Since the Fourier transform of $x^\alpha T^\Lambda$ is $(-i)^{|\alpha|}\partial^\alpha \hat{T}^\Lambda$ we conclude that the counter term in momentum space is a modified Taylor expansion around zero where the derivative at the origin is replaced by a mean over the derivative. In case $\hat{T^\Lambda}$ is, in the limit, sufficiently nicely behaved at $p = 0$, one can choose an extension with $v_\beta \equiv 1$. This corresponds to the BPHZ (Bogoliubov-Parasiuk-Hepp-Zimmermann) method of subtracting the Taylor series at $p = 0$ up to the order ω. In the Epstein Glaser framework this case is called the central solution.

The Taylor expansion at zero momentum works for Feynman integrals in massive theories. In theories with massless particles one subtracts instead the Taylor expansion at some spacelike momentum q. In the framework described above this would mean to choose as a basis in $\mathcal{S}_\omega(\mathbb{R}^n)^\perp$ the distributions $(e^{iqx}\partial^\alpha\delta)$. The dual base in $\mathcal{W}$ is then characterized by the condition $\partial^\beta(e^{-iqx}w_\alpha)_{x=0} = \delta^\beta_\alpha$. w_β is of the form $w_\beta = x^\beta e^{iqx}v_\beta$ where v_β satisfies the same condition as before. If we again set $v_\beta \equiv 1$ we obtain the Taylor expansion at q.

We now want to treat the question whether symmetries of T_0 can be preserved during the extension.

This is not always the case as may be seen in the example of $\frac{1}{|x|}$. $\frac{1}{|x|}$ is a homogeneous distribution on $\mathcal{D}(\mathbb{R}\setminus\{0\})$, but none of its extensions $\frac{\mathrm{d}}{\mathrm{d}x}\left(\operatorname{sign} x \ln\frac{|x|}{a}\right)$, $a > 0$ to $\mathcal{D}(\mathbb{R})$ is homogeneous. In field theory this phenomenon occurs, e.g., in the renormalization of the fish graph in the massless φ^4-theory. The renormalized theory is no longer scale invariant ("dimensional transmutation").

Let V be a representation of a group G on $\mathcal{D}(\mathbb{R}^n)$ such that $\mathcal{D}_\omega(\mathbb{R}^n)$ and T_ω are invariant under V,

$$\langle T_\omega, V(g)\varphi\rangle = \langle T_\omega, \varphi\rangle \ , \ \varphi \in \mathcal{D}_\omega(\mathbb{R}^n) \ .$$

Let V^t be the transposed representation of G on the space of distributions $S \in \mathcal{D}'(\mathbb{R}^n)$,

$$\langle V^t(g)S, \varphi\rangle = \langle S, V(g^{-1})\varphi\rangle \ .$$

Then we have for the extensions T of T_ω

$$V^t(g)T = T + l(g) \ , \ l(g) \in \mathcal{D}_\omega(\mathbb{R}^n)^\perp \ .$$

$V^t(g)$, restricted to $\{\lambda T, \lambda \in \mathbb{C}\}\oplus\mathcal{D}_\omega(\mathbb{R}^n)^\perp$, is a finite dimensional representation of G. It contains the restriction $D(g)$ of $V^t(g)$ on $\mathcal{D}_\omega(\mathbb{R}^n)^\perp$ as a subrepresentation.

We now search for an $l_0 \in \mathcal{D}_\omega(\mathbb{R}^n)^\perp$, such that $T + l_0$ is invariant. Then

$$V^t(g)(T + l_0) = T + l(g) + V^t(g)l_0 = T + l_0 \ ,$$

i.e., l_0 has to fulfill the equation

$$l(g) = l_0 - D(g)l_0 \ .$$

We know that $l(g)$ satisfies the equation

$$l(gh) = V^t(gh)T - T = V^t(g)\left(V^t(h)T - T\right) + V^t(g)T - T = D(g)l(h) + l(g) \ .$$

The solution of such an equation is called a cocycle. If $l(g)$ can be expressed as above in terms of l_0 the cocycle condition is automatically fulfilled; such an $l(g)$ is called a coboundary. The space of cocycles modulo the space of coboundaries is called the cohomology of G with respect to the representation D.

In case the finite dimensional representations of the group are completely reducible the cohomology of a finite dimensional representation is always trivial . Namely, we may look at the matrix representation

$$\begin{pmatrix} 1 & 0 \\ l(g) & D(g) \end{pmatrix}$$

Due to the complete reducibility of the representation there exists a 1-dimensional invariant subspace which is complementary to the representation space of D. Such a subspace is of the form

$$\{\lambda \begin{pmatrix} 1 \\ l_0 \end{pmatrix}, \lambda \in \mathbb{C}\} \tag{3.14}$$

On this subspace the representation is trivial. This is equivalent to the above condition on l_0.

For the Lorentz group every finite dimensional representation is completely reducible. Therefore Lorentz-invariant distributions always possess Lorentz-invariant extensions. For the scaling transformations, on the other hand, one has to study the representations of $\mathbb{R}_+$ as multiplicative group. But they are not always completely reducible, as may be seen from the example

$$\mathbb{R}_+ \ni \lambda \mapsto \begin{pmatrix} 1 & 0 \\ \ln \lambda & 1 \end{pmatrix} .$$

The existence of these representations is responsible for the breaking of scaling invariance in the massless φ^4 theory, and probably also for the occurence of a mass scale for light hadrons in QCD.

4. Renormalization in all orders

We discussed divergences and their removal in some examples. We now want to address the problem of performing renormalization systematically in all orders of perturbation theory. A very transparent method was developed by Epstein und Glaser on the basis of older ideas of Stückelberg and Bogoliubov. It consists in an inductive construction of time ordered products.

Time ordering does not commute with time derivatives. Therefore it is preferable to perform time ordering before the fields are represented as operator valued distributions on Fock space. We thus consider the arguments of time ordered products as classical fields, as they occur in the Lagrangean and in the path integral, and without imposing the field equation ("off-shell"-formalism). The space of functionals on smooth classical field configurations ϕ has the structure of a commutative associative algebra. The possible interactions are of the form

$$\mathcal{L}(g) = \int d^4x \sum_{i=1}^{n} g_i(x) A_i(\phi(x), \partial\phi(x)) \tag{4.1}$$

with n-tuples of test functions $g = (g_i)$ and classical fields $A_i = A_i(\phi, \partial\phi)$ (local functionals). We endow the space of functionals of classical field configurations with a noncommutative product $\star$ which takes after identification of polynomial functionals of ϕ with normal ordered products of quantum fields φ the form of Wick's theorem.

Let F, G be functionals of ϕ. Then

$$(F \star G)(\phi) = \exp \langle \hbar\Delta_+, \frac{\delta^2}{\delta\phi_1\delta\phi_2}\rangle F(\phi_1)G(\phi_2)|_{\phi_1=\phi_2=\phi} \ . \tag{4.2}$$

The exponential function is here defined as a formal power series in $\hbar$. The terms of finite order are well defined provided the functional derivatives of F and G are sufficiently well behaved. This is true for iterated $\star$ products of local functionals and relies as the operator product of Wick polynomials on the positive energy property of Δ_+. The S-matrix for an interaction of the form (4.1) is the generating functional of time ordered products of the fields A_i,

$$S(\mathcal{L}(g)) = \sum_n \frac{i^n}{n!} \sum_{i_1,\dots,i_n} \langle T(A_{i_1}, \dots, A_{i_n}), g_{i_1} \otimes \cdots \otimes g_{i_n}\rangle$$

The time ordered products are here considered as distributions with values in the space of functionals of classical field configurations. We often use the somewhat sloppy notation $TA_{i_1}(x_1)\cdots A_{i_n}(x_n)$. The properties of the time ordered products can be encoded in properties of the S-matrix:

Unitarity:

$$S(\mathcal{L}(g))^* = S(\overline{\mathcal{L}(g)})^{-1} \tag{4.3}$$

where the inverse refers to the $\star$-product.

Causality: If there is a Cauchy surface with $\operatorname{supp} g$ in its future and $\operatorname{supp} h$ in its past, the S-matrix factorizes

$$S(\mathcal{L}(g+h)) = S(\mathcal{L}(g)) \star S(\mathcal{L}(h)) \tag{4.4}$$

Covariance:

$$\alpha_{(a,\Lambda)} S(\mathcal{L}(g)) = S(\alpha_{(a,\Lambda)}\mathcal{L}(g)) \tag{4.5}$$

where the Poincaré transform (a, Λ) acts on functionals F on $\mathcal{C}^\infty(\mathbb{M})$ by

$$(\alpha_{(a,\Lambda)}F)(\varphi) = F(\varphi_{(a,\Lambda)}) \tag{4.6}$$

with $\varphi_{(a,\Lambda)}(x) = \varphi(\Lambda x + a)$.

For the time ordered products one obtains the following requirements

Symmetry: As expansion coefficients of a power series the time ordered products are symmetrical in their indices.

Causality:

$$TA_1(x_1)\cdots A_n(x_n) = TA_1(x_1)\cdots A_n(x_k) \star TA_1(x_{k+1})\cdots A_n(x_n)$$

if none of the points $x_1, \ldots, x_k$ lies in the past of some of the points $x_{k+1}, \ldots, x_n$. In case the two sets of points are spacelike to each other it follows in particular that the time ordered products on the right hand side commute with each other.

Unitarity:

$$(TA_1(x_1)\cdots A_n(x_n))^* = \sum_{I_1+\ldots+I_k=\{1,\ldots,n\}} (-1)^k T(A_i^*(x_i), i\in I_1)\star\cdots\star T(A_i^*(x_i), i\in I_k) \tag{4.7}$$

Covariance:

$$\alpha_{a,\Lambda} TA_1(x_1)\cdots A_n(x_n) = T\Lambda A_1(\Lambda x_1 + a)\cdots \Lambda A_n(\Lambda x_1 + a) \tag{4.8}$$

These general rules are amended by the initial condition

$$TA(x) = A(x).$$

The initial condition together with the axioms of symmetry and causality implies that for pairwise different arguments the the time ordered product can be written in form of an operator product. Namely, let $x_i \neq x_j$ for all $i \neq j$. Then there is a permutation π such that $x_{\pi(i)}$ is not contained in the past of $x_{\pi(j)}$ for $i < j$. Therefore

$$TA_1(x_1)\cdots A_n(x_n) = A_{\pi(1)}(x_{\pi(1)}) \star \cdots \star A_{\pi(n)}(x_{\pi(n)}) \ . \tag{4.9}$$

In case two arguments are spacelike separated the permutation is not unique. Thus the definition of time ordered products is consistent only for fields which $\star$-commute for spacelike separations.

The time ordered product of 2 fields is uniquely determined for noncoinciding arguments. Renormalization amounts to extending this distribution to all of $\mathbb{R}^4 \times \mathbb{R}^4$. We proceed by induction and assume that time ordered products with less than n factors are everywhere defined and satisfy the conditions above. Here the causality condition in the case of n factors reduces to the requirements that the $\star$ product on the right hand side is commutative for spacelike localization.

Now the time ordered product of n factors is uniquely determined outside of the thin diagonal $x_1 = \ldots = x_n$. Namely, assume that not all points x_i are equal. Then the set of points can be separated by a spacelike hyperplane such that a nonempty subset lies in the future and another nonempty subset in the past of the hypersurface Due to the causality axiom the time ordered product of n factors can be written as a $\star$-product of time ordered products with less than n factors. The result is independent of the choice of the hyperplane because of the commutativity of the $\star$-product of spacelike localized time ordered products.

The time ordered product is a formal series of functional differential operators. It has the structure

$$T_n = m_n \circ \sum_{G \in \mathcal{G}_n} \langle t_G, \delta^G \rangle \tag{4.10}$$

with

$$\delta^G = \frac{\delta^{2|E(G)|}}{\prod_{v \in V(G)} \delta \varphi_v^{\alpha_v}} \tag{4.11}$$

where α_v is the number of lines incident at the vertex v and where t_G is an extension of

$$t_G^0(x_{v,e}, v \in \partial e, e \in E(G)) = \prod_{e \in E(G)} i\Delta_F(x_{v,e}, v \in \partial e) \tag{4.12}$$

m_n denotes the pointwise product of n functionals of φ,

$$m_n(F_1 \otimes \cdots \otimes F_n)(\varphi) = F_1(\varphi) \dots F_n(\varphi) \tag{4.13}$$

After application of the functional differential δ^G to an n-tuple of local functionals of φ one obtains, for every field configuration φ, a distribution with support on the partial diagonals $x_{v,e} = x_{v,e'}$ if $v \in \partial e \cap \partial e'$. We can write it as a test function of compact support in the center of mass coordinates $x_v = \sum_{e: v \in \partial e} \frac{1}{\alpha_v} x_{v,e}$ with values in the space

$$Y^G = \bigotimes_{v \in V(G)} Y^{\alpha_v} \tag{4.14}$$

where Y^{α_v} is the space of distributions in the relative coordinates α_v independent coordinates $x_{v,e} - x_v$, $v \in \partial e$ with support at the origin. Application of a distribution from Y^G to t_G^0 yields the unrenormalized Feynman amplitude. Its extension to all test functions proceeds as described in the previous section. The ambiguity in the extension process amounts to the addition of a local functional in nth order.

The result can be summarized in the Main Theorem of Renormalization:

THEOREM V.5. (i) *There exist formal series of multilinear maps*

$$S(\mathcal{L}(g)) = \sum_{n=0}^{\infty} \frac{i^n}{n!} \langle T_n(\mathcal{L}(g), \dots, \mathcal{L}(g))$$

which satisfy the axioms.

(ii) *Two solutions S and $\hat{S}$ differ by an analytic map $Z : \mathcal{L} \to \hat{\mathcal{L}}$ with $Z'(0) = \mathrm{id}$ in the space of admissible Lagrangeans such that*

$$S(\hat{\mathcal{L}}(g)) = \hat{S}(\mathcal{L}(g)) \ .$$

The maps Z characterize the ambiguity of the construction of the theory from a Lagrangean. They form a group which was called renormalization group by Stückelberg und Petermann.

One may impose additional constraints on the renormalization procedure. These constraints can be fulfilled if the associated cocycle in the renormalization group is a coboundary.

Also in cases the constraint cannot be fulfilled the associated cocyle is of interest. An important example is the behavior under scaling. Let us start from a free theory which is scale invariant, as e.g. the free massless scalar field. In order to fix a renormalization procedure we have to introduce a scale μ with the dimension of a mass. Changing the scale from μ to μ' amounts to a renormalization group transformation $Z(\mu, \mu')$ with the composition law

$$Z(\mu, \mu') \circ Z(\mu', \mu'') = Z(\mu, \mu'') \tag{4.15}$$

The parameter μ characterizes different renormalization procedures. One may now, for a fixed value of μ, investigate the behavior of the theory at scaled coordinates. But this is equivalent to keeping the coordinates fixed and to scaling the renormalization parameter. Therefore the renormalization group transformations $Z(\mu, \mu')$ characterize also, for a given theory, the behavior at different scales. As a result one obtains the so-called running coupling constants.

CHAPTER VI

Nonabelian Gauge Theories

1. Introduction

In chapter III we constructed fields which generate massive particles with given spin s. The associated Feynman propagator arises as a $2s$-fold derivative of the scalar Feynman propagators and thus has the scaling degree $2 + 2s$. For $s \geq 1$ there is therefore no local field in the Fock space of the theory which can be used as the interaction Lagrangean of a (power counting) renormalizable theory. The W- and Z-bosons as quanta of the weak interaction, however, have spin 1, their interaction (provided it is power counting renormalizable) must therefore be of another type.

A similar problem already occurred in quantum electrodynamics. The Feynman propagator for the electromagnetic field has scaling degree 4, for the Dirac field scaling degree 3; also there is no suitable interaction term in the Fock space of the free theory.

In QED the problem is solved by extending the theory such that also the vector potential becomes a local field. The time ordered 2-point function in Feynman gauge is

$$\langle \Omega, TA_\mu(x)A_\nu(y)\Omega\rangle = -g_{\mu\nu}D_F(x-y)\ . \tag{1.1}$$

Formally one can construct a Fock space for the extended theory, but with a scalar product which is not positive.

In the extended theory the Feynman propagator of the vector potential has scaling degree 2, and the interaction term of QED

$$\mathcal{L} = A_\mu j^\mu\ , \tag{1.2}$$

with the current $j^\mu = \overline{\psi}\gamma^\mu\psi$ of the Dirac field, has dimension 4. One obtains thus a renormalizable theory. The renormalization ambiguity can be used to fulfill the so-called Ward identity which allows to go back to the original theory.

Crucial for this procedure is the fact that the Nakanishi-Lautrup field $B = \partial_\mu A^\mu$ is a free field also in the presence of interactions. It implements the remaining gauge transformations after gauge fixing (see (IV,4.31))

$$Q(\Lambda) = \int d^3\mathbf{x}\, B(t,\mathbf{x})\overset{\leftrightarrow}{\partial}_t\Lambda(t,\mathbf{x})$$

with solutions Λ of the wave equation, and can be used for the construction of the algebra of observables and the physical Hilbert space as described in Chapter IV.

In the presence of several vector fields A^a_μ, $a = 1, \ldots, n$ which interact with each other, the B-fields are no longer free, and one has to use another method.

The best known method is the BRST method. In this method the c-number function Λ is replaced by fields c^a (called the ghost fields). c^a is a scalar Fermi field and fulfills a modified wave equation involving the vector potential. In terms of c one now can form an operator Q which implements the remaining gauge transformations, but with the gauge parameters replaced by the ghost fields. A further family of fields, the so-called anti ghost fields $\bar{c}_a$, satisfy canonical anticommutation relations with the ghost fields. Q is chosen to have vanishing square, and the quotient of the kernel of Q by the range of Q (the cohomology of Q) is identified with a dense subset of the physical Hilbert space (actually the positivity of the scalar product on the quotient has to be checked directly; it does not follow automatically from the construction).

The observables are constructed in a similar way. The algebra generated by the vector fields, the ghost fields and the anti ghost fields has a natural grading (ghost number). The graded commutator with Q defines a graded derivation on this algebra (the BRST transformation s). $Q^2 = 0$ implies $s^2 = 0$, and the algebra of observables is defined as the cohomology of s, and it is straightforward to show that it has a natural representation on the physical Hilbert space.

The remarkable fact is now that this construction is possible only under very special conditions on the interaction between the vector fields. The most important class of interactions for which the method works is that of nonabelian gauge theories with Lagrangean

$$\mathcal{L} = \frac{1}{4} F^a_{\mu\nu} F^{\mu\nu,b} \kappa_{ab}$$

with the field strength tensor

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + f^a_{bc} A^b_\mu A^b_\nu \ ,$$

the Killing form κ_{ab} and the structure constants f^a_{bc} of a Lie algebra.

The present theory of elementary particles, the so-called standard model, is a gauge theory, based on the Lie group $\mathrm{U}(1) \times \mathrm{SU}(2) \times \mathrm{SU}(3)$.

2. Classical Gauge Theories

The idea (and the name) of gauge theory goes back to an attempt of Hermann Weyl. He intended to extend Einstein's theory of gravity in order to include also classical electromagnetism. His ansatz was that the theory should not only be generally covariant as general relativity, but should allow also the freedom to choose at every point independently the unit of length (which is in Einstein's theory fixed

by the parallel transport in terms of the Levi Civita connection). This additional freedom should correspond to the charge.

In the original form, the idea did not work, but after the invention of quantum mechanics with its complex valued wave functions the idea worked in a modified way. Namely, one can vary the phase of the wave function independently at every point. Taking the derivative amounts to

$$\partial_\mu e^{i\Lambda}\psi = e^{i\Lambda}(\partial_\mu + i\partial_\mu\Lambda)\psi$$

The additional term is nothing else as the gauge freedom of a vector potential, hence the covariant derivative

$$D_\mu = \partial_\mu - iA_\mu$$

satifies

$$D_\mu e^{i\Lambda}\psi = e^{i\Lambda}D_\mu\psi$$

if both, ψ and A_μ are transformed.

The arising structure was generalized to other groups. In particular the concept of isospin suggested to look at SU(2) as a possible gauge group, and, after the observation of higher symmetry groups, also to SU(3).

In the mean time mathematicians had developed the theory of fiber bundles by which general properties of gauge theories could be analyzed.

Let us first define the notion of a vector bundle. Let M be a smooth manifold, and let V be a finite dimensional vector space. We associate to each point x a copy V_x of V and define the disjoint union

$$\bigcup_{x\in M} V_x =: E$$

as the total space of the bundle. The map $\pi : E \to M$ with $\pi(v) = x$ if $v \in V_x$ is assumed to be smooth. Maps $s : M \to E$ with $\pi \circ s = \mathrm{id}$ are called sections of the bundle.

The idea of gauge theory is that there is no a priori given identification between different fibers. Instead the relation between different fibers is given by an additional field, the gauge field. In the case of a trivial bundle (actual all bundles over Minkowski space are trivial) sections can be identified with functions with values in V. The crucial fact is that even then there is no distinguished identification. The gauge field (in mathematics: a linear connection) now can be characterized by associating to each path on M a linear isomorphism between the fibers on the end points of the graph (the parallel transport), such that this association is compatible with the concatenation of paths and such that the reversed path leads to the inverse isomorphism. Given these isomorphisms, one can introduce covariant directional derivatives along curves in M. They have on trivial bundles the form

$$D_\mu = \partial_\mu + A_\mu \tag{2.1}$$

where $A_\mu(x)$ is a linear operator on V. If one transforms the fiber at x with the automorphism $g(x)$ (gauge transformation in physics, bundle isomorphism in mathematics) one obtains the gauge transformed gauge field

$$A'_\mu = g^{-1}\partial_\mu g + g^{-1}A_\mu g \tag{2.2}$$

The gauge field is nontrivial if the field strength (the curvature in mathematics) is nontrivial,

$$F_{\mu,\nu} = [D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu . A_\nu] \tag{2.3}$$

A gauge invariant Lagrangean can be be defined as

$$\mathcal{L}(A, \partial A) = \frac{1}{4g^2} \operatorname{Tr} F_{\mu\nu} F^{\mu\nu} \tag{2.4}$$

(Yang-Mills action). This interaction induces the Yang-Mills equations

$$[D^\mu, F_{\mu\nu}] = 0 \tag{2.5}$$

A class of solutions is given by the so-called (anti-)self dual gauge fields (instantons) which satisfy

$$F = \pm * F \tag{2.6}$$

They exist only in the euclidean theory. They are thought to give important contributions to the path integral.

3. Perturbative construction of nonabelian gauge theories; BRST symmetry

The Yang-Mills equation does not possess a well posed initial value problem. The reason is, as for the Maxwell eqaution for the vector potential, the gauge freedom which allows gauge transformations at different times. To fix the gauge one can introduce a condition

$$\partial^\mu A_\mu = B \tag{3.1}$$

with an arbitrary function B with values in the Lie algebra of the gauge group. Then the arising nonlinear hyperbolic equation

$$\Box A_\nu - \partial_\nu B - [A_\nu, B] + [A_\mu, \partial^\mu A_\nu] + [A^\mu, F_{\mu\nu}] = 0 \tag{3.2}$$

has (at least locally) a well posed initial value problem provided the initial values satisfy the gauge fixing condition (3.1).

We now follow the historical development and describe the heuristic solution to the problem of quantizing Yang-Mills theories. We use the path integral formulation.

Let $S(A)$ the Yang-Mills action and $F(A)$ a gauge invariant functional of the gauge field A. We want to determine the expectation value

$$\langle F \rangle = \frac{1}{Z} \int \mathcal{D}A \, e^{iS(A)} F(A) \ . \tag{3.3}$$

We consider the integral of the gauge fixing condition over the gauge group

$$\int dg\,\delta(\partial^\mu A^g_\mu - B) \equiv \int \prod_x dg(x)\delta(\partial^\mu A^g_\mu(x) - B(x)) =: \Delta(A,B)^{-1}\ . \tag{3.4}$$

and obtain

$$\langle F\rangle = \frac{1}{Z}\int \mathcal{D}A\, e^{iS(A)}F(A)\Delta(A,B)\int dg\,\delta(\partial^\mu A^g_\mu - B)\ . \tag{3.5}$$

independently of B. We now interchange the order of integration over g and A, use the facts that S, F, and the measure $\mathcal{D}A$ are gauge invariant and arrive at

$$\langle F\rangle = \frac{1}{Z}\int \mathcal{D}A\, e^{iS(A)}F(A)\Delta(A,B)\delta(\partial^\mu A_\mu - B) \tag{3.6}$$

We now compute the functional $\Delta(A,B)$. Under the condition that there is exactly one gauge transformation g_0 such that $\partial^\mu A^{g_0}_\mu = B$ the transformation law for the δ-function yields

$$\Delta(A,B)^{-1} = \left|\det \frac{\delta\partial^\mu A^g_\mu}{\delta g}(g_0)\right|^{-1}\ . \tag{3.7}$$

Actually, the assumption is in general not fulfilled, there can be several gauge equivalent gauge fields which satisfy the gauge fixing condition (so-called Gribov copies). In perturbation theory, however, one can restrict oneself to an infinitesimal neighborhood of a gauge field which satisfies the gauge fixing condition. There the solution is unique.

The functional derivative in the determinant is for a gauge field A which satisfies the gauge fixing condition the linear operator on the Lie algebra of the gauge group

$$(M(A)\xi)_a(x) = \left((\Box + \partial^\mu \mathrm{ad}A_\mu)\,\xi\right)_a(x) \tag{3.8}$$

where we use a gauge field which satisfies the gauge fixing condition. We set $\Delta(A) = \Delta(A,\partial^\mu A_\mu) = \det M(A)$ and find for a gauge invariant functional F

$$\int \mathcal{D}A\, e^{iS(A)}F(A) = \int \mathcal{D}A\,\delta(\partial A - B)\det M(A) e^{iS(A)}f(A)\ . \tag{3.9}$$

In order to eliminate the δ-function we integrate over B with a Gaussian measure, formally given by $\mathcal{D}B\, e^{i\frac{\lambda}{2}\int dx\,\mathrm{tr}\,B(x)^2}$ and obtain the additional term $S_{\mathrm{gf}}(A) = \frac{\lambda}{2}(\partial A)^2$ (gauge fixing term) in the Lagrangean.

The determinant can, according to an idea of Fadeev and Popov, be considered as a fermionic Gaussian integral. We introduce fermionic scalar fields (Fadeev-Popov ghosts c^a and antighosts $\bar c_a$) and add to the action the term

$$S_{\mathrm{gh}}(A,c,\bar c) = \int dx\big\langle \bar c, M(A)c\big\rangle = \int dx\big\langle \partial^\mu \bar c(x), \partial_\mu c(x) + [A_\mu, c(x)]\big\rangle \tag{3.10}$$

Here we assumed that the determinant has a constant phase such that it can be absorbed in the normalization factor. Altogether we get an action

$$S(A,c,\bar{c}) = S(A) + S_{\rm gf}(A) + S_{\rm gh}(A,c,\bar{c}) \ . \tag{3.11}$$

It is convenient to add an auxiliary field (called B) by writing the exponential of the gauge fixing term as the Fourier transform of a Gaussian measure over B,

$$e^{i\int dx \frac{\lambda}{2}(\partial A)^2} = \int \mathcal{D}B\, e^{i\int dx\, {\rm tr}\, B\partial A + \frac{2}{\lambda}\,{\rm tr}\, B^2} \ . \tag{3.12}$$

B satisfies the field equation $B = \frac{1}{\lambda}\partial A$ and can therefore be eliminated. It plays however an important role in the so-called BRST symmetry of the action

$$S(A,B,c,\bar{c}) = S(A) + S_{\rm gf}(A,B) + S_{\rm gh}(c,\bar{c}) \ . \tag{3.13}$$

with $S_{\rm gf}(A,B) = \int dx\, {\rm tr}\, B\partial A + \frac{2}{\lambda}\,{\rm tr}\, B^2$. The augmented action is no longer gauge invariant. In particular, the gauge fixing term changes under an infinitesimal gauge transformation to

$$\langle \frac{\delta S_{\rm gf}(A^g,B)}{\delta g}|_{g=e}, \xi\rangle = \int dx\, {\rm tr}\, BM(A)\xi \tag{3.14}$$

We now identify the ghost field with the Maurer-Cartan form of the gauge group at the identity, i.e. $c(\xi) = \xi$. The equation above can then be written as

$$d_g S_{\rm gf}(A^g,B)|_{g=e} = \int dx\, {\rm tr}\, BM(A)c \tag{3.15}$$

in terms of the exterior differential d_g of the gauge group. This term coincides up to a sign with the ghost action when $\bar{c}$ is replaced by B. One now introduces the BRST transformation s as a graded (with respect to ghost number) translation invariant derivation,

$$s(A) = \partial c + [A,c] \tag{3.16}$$

$$s(c) = -[c,c] \tag{3.17}$$

$$s(\bar{c}) = -B \tag{3.18}$$

$$s(B) = 0 \tag{3.19}$$

On A and c, the BRST transformation coincides with the differential d_g. In particular, (3.17) is the Maurer-Cartan equation. Moreover, s itself is a differential (i.e. it satisfies $s^2 = 0$).

We now verify that the extended action is BRST invariant. But this follows from $M(A)c = s(\partial A)$, hence $s(M(A)c) = 0$ and

$$s(S_{\rm gh}(c,\bar{c})) = -\int dx\, {\rm tr}\, BM(A)c = -s(S_{\rm gf}(A,B)) \ . \tag{3.20}$$

Renormalization of gauge theories is now performed such that BRST invariance is preserved in every step. In terms of the effective action

(the vertex functional) the procedure is the following. One constructs the effective action $\Gamma(A, B, c, \bar{c})$ as a formal power series in $\hbar$ with ghost number zero, $\Gamma = \sum_{n=0}^{\infty} \Gamma_n$. In zeroth order $\Gamma_0 = S$. We now assume $s(\Gamma_k) = 0$ for $k < n$. One then shows that $s(\Gamma_n)$ is a local functional of the fields. The crucial question is now whether there exists another local functional S_n with ghost number zero and $s(S_n) + s(\Gamma_n) = 0$. If yes, one adds $\hbar^n S_n$ to the classical action S and obtains an effective action which has ghost number zero and is BRST invariant up to nth order.

The possibility of preserving BRST invariance is guaranteed if the cohomology of the BRST transformation on local functionals with ghost number 1 is trivial. This is generally true for semisimple gauge groups.

4. The triangle anomaly

An instructive and extremely important example of an anomaly (i.e. a nontrivial cohomology class) is the so-called triangle anomaly. It occurs already in QED in connection with the axial current.

Let ψ be a Dirac field and $j_\mu = \overline{\psi}\gamma_\mu\psi$ the vector current and $j^5_\mu = \overline{\psi}\gamma^5\gamma_\mu\psi$ the axial current.

The Dirac equation for ψ yields

$$\partial^\mu j_\mu = 0 \ , \ \partial^\mu j^5_\mu = 2im\overline{\psi}\gamma^5\psi \ . \tag{4.1}$$

We want to define the vacuum expectation value of the product of 3 currents,

$$\left\langle \Omega, T j_\mu(x) j_\nu(y) j^5_\lambda(z)\Omega \right\rangle \equiv t_{\mu\nu\lambda}(u, v) \quad , \ u = x - z, \ v = y - z \ . \tag{4.2}$$

These distributions have scaling degree 9; they are symmetric under permutation of the two vector currents, i.e. under the interchange of $(u, \mu) \leftrightarrow (v, \nu)$ and transform under orthochronous Lorentz transformations according to

$$t_{\mu\nu\lambda}(\Lambda u, \Lambda v)\Lambda^\mu_\alpha\Lambda^\nu_\beta\Lambda^\lambda_\gamma \det\Lambda = t_{\alpha\beta\gamma}(u, v) \tag{4.3}$$

They are unique up to a derivative of the δ-function of first order with the same symmetry. The possible finite counter terms are multiples of

$$C_{\mu\nu\lambda}(u, v) = \varepsilon_{\mu\nu\lambda\rho}(\partial^\rho_u - \partial^\rho_v)\delta(u)\delta(v) \ . \tag{4.4}$$

We now want to impose the conservation of the vector current (this is actually needed for the renormalizability of QED). The distribution

$$\partial^\mu_u t_{\mu\nu\lambda}(u, v) \equiv V_{\nu\lambda}(u, v) \tag{4.5}$$

has scaling degree ≤ 10 and vanishes outside of the origin. The symmetry properties of $t_{\mu\nu\lambda}$ imply that $V_{\nu\lambda}$ has the structure

$$V_{\nu\lambda}(u, v) = a\varepsilon_{\nu\lambda\alpha\beta}\partial^\alpha_u\partial^\beta_v\delta(u)\delta(v) \tag{4.6}$$

with some constant a. If we replace $t_{\mu\nu\lambda}$ by $t_{\mu\nu\lambda} + bC_{\mu\nu\lambda}(u,v)$, the divergence is

$$\partial_u^\mu(t_{\mu\nu\lambda}(u,v) + bC_{\mu\nu\lambda}(u,v)) = V_{\nu\lambda}(u,v) - b\varepsilon_{\mu\nu\lambda\rho}\partial_u^\mu\partial_v^\rho\delta(u)\delta(v) \quad (4.7)$$

We observe that we have to choose $a = b$. Due to the symmetry $(u,\mu) \leftrightarrow (v,\nu)$ then also the divergence with respect to the other vector current vanishes.

But now $t_{\mu\nu\lambda}$ is uniquely determined, and there is no freedom left to achieve the conservation of the axial current (in the sense of (4.1)). In fact, this relation turns out to be violated after renormalization.

We want to compute

$$\begin{aligned} W_{\mu\nu}(u,v) &= -(\partial_u^\lambda + \partial_v^\lambda)t_{\mu\nu\lambda}(u,v) - 2im\big\langle\Omega, Tj_\mu(x)j_\nu(y)\overline{\psi}\gamma_5\psi(x)\Omega\big\rangle \\ &= c\varepsilon_{\mu\nu\alpha\beta}\partial_u^\alpha\partial_v^\beta\delta(u)\delta(v)\ . \end{aligned} \quad (4.8)$$

For this purpose we use the Pauli-Villars method and subtract the corresponding expression for another mass M. $r^{(M)}_{\mu\nu\lambda} = t_{\mu\nu\lambda} - t^{(M)}_{\mu\nu\lambda}$ has scaling degree 8. For $(u,v) \neq 0$ it satisfies the equation

$$\begin{aligned} &-(\partial_u^\lambda + \partial_v^\lambda)r^{(M)}_{\mu\nu\lambda}(u,v) - 2im\big\langle\Omega, Tj_\mu(x)j_\nu(y)\overline{\psi}\gamma_5\psi(x)\Omega\big\rangle \\ &= -2iM\big\langle\Omega, Tj_\mu(x)j_\nu(y)\overline{\psi}\gamma_5\psi(x)\Omega\big\rangle^{(M)}\ . \end{aligned} \quad (4.9)$$

Both sides have scaling degree 9. Therefore they are uniquely determined since there is no counter term with the correct symmetry, hence equation (4.9) holds everywhere. Moreover, for the same reason, $r^{(M)}_{\mu\nu\lambda}$ converges towards $t_{\mu\nu\lambda}$ for $M \to \infty$. We thus conclude that

$$W_{\mu\nu}(u,v) = \lim_{M\to\infty}(-2iM)\big\langle\Omega, Tj_\mu(x)j_\nu(y)\overline{\psi}\gamma_5\psi(x)\Omega\big\rangle^{(M)}\ . \quad (4.10)$$

It remains to compute the right hand side of (4.10). In terms of the Feynman propagator for Dirac fields with mass M,

$$S_F(x) = (i\partial\!\!\!/ + M)\Delta_F(x) = (2\pi)^{-4}\int \mathrm{d}^4p\, e^{-ipx}(p\!\!\!/ - M + i\varepsilon)^{-1}\ , \quad (4.11)$$

we have

$$\begin{aligned} &\big\langle\Omega, Tj_\mu(x)j_\nu(y)\overline{\psi}\gamma_5\psi(x)\Omega\big\rangle^{(M)} = \\ &\qquad \operatorname{tr}\gamma_\mu S_F(x-y)\gamma_\nu S_F(y-z)\gamma^5 S_F(z-x) + ((x,\mu)\leftrightarrow(y,\nu))\ . \end{aligned} \quad (4.12)$$

We perform the trace over the Dirac matrices. Using

$$\frac{1}{4}\operatorname{tr}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\gamma^5 = \varepsilon_{\mu\nu\rho\sigma} \quad (4.13)$$

and the fact that the trace of an odd number of γ-matrices as well as that of $\gamma_\mu\gamma_\nu\gamma^5$ vanish, we find

$$\big\langle\Omega, Tj_\mu(x)j_\nu(y)\overline{\psi}\gamma_5\psi(x)\Omega\big\rangle^{(M)} = 8iM\varepsilon_{\mu\nu\rho\sigma}\partial_u^\rho\partial_v^\sigma\Delta_F(u-v)\Delta_F(-u)\Delta_F(v)\ . \quad (4.14)$$

It remains to compute the limit $M^2\Delta_F(u-v)\Delta_F(-u)\Delta_F(v)$ for $M \to \infty$. Smearing with a test function $f \in \mathcal{S}(\mathbb{R}^8)$ with a Fourier transform with compact support yields

$$\int d^4u d^4v M^2 \Delta_F(u-v)\Delta_F(-u)\Delta_F(v) f(u,v) = \tag{4.15}$$

$$(2\pi)^{-8} \int \frac{M^2 \hat{f}(p,q) d^4p d^4q d^4k}{((p+k)^2 - M^2 + i\epsilon)(k^2 - M^2 + i\epsilon)((p+q+k)^2 - M^2 + i\epsilon)} \tag{4.16}$$

Here the k integral is not absolutely convergent. We define it as

$$\int dk = \lim_{\Lambda\to\infty} \int_{|\mathbf{k}|<\Lambda} d^3\mathbf{k} \int_{-\infty}^{\infty} dk_0 \tag{4.17}$$

For finite Λ the integral is absolutely convergent. Moreover, the integrand is a meromorphic function of k_0 with poles at

$$\begin{aligned} &\pm \sqrt{|\mathbf{k}|^2 + M^2 - i\epsilon}, \\ &-p_0 \pm \sqrt{|\mathbf{p}+\mathbf{k}|^2 + M^2 - i\epsilon}, \\ &-p_0 - q_0 \pm \sqrt{|\mathbf{p}+\mathbf{q}+\mathbf{k}|^2 + M^2 - i\epsilon}\ . \end{aligned} \tag{4.18}$$

For M sufficiently large all poles with positive real part have negative imaginary part and vice versa. We may now deform the integration contour for k_0 and integrate instead over the imaginary axis. The arising integral over k is then absolutely convergent, moreover, for all $(p,q) \in \operatorname{supp} \hat{f}$ it is uniformly bounded by an integrable function. We then change variables $k \to Mk$, perform the limit $M \to \infty$ pointwise in p and q and obtain

$$\lim_{M\to\infty} \int d^4u d^4v M^2 \Delta_F(u-v)\Delta_F(-u)\Delta_F(v) f(u,v) = \tag{4.19}$$

$$f(0)(2\pi)^{-4} \int d^4k (|k|^2+1)^{-3} = \frac{1}{32\pi^2} f(0)\ .$$

We finally arrive at the numerical value of the constant c in the axial anomaly

$$c = -\frac{1}{2\pi^2}\ . \tag{4.20}$$