Hands-On Quantum Field Theory

Univ.-Prof. Mag. Dr. Andreas Krassnigg

Skriptum zur Spezialvorlesung im Sommersemester 2012
Institut für Physik
Karl-Franzens-Universität Graz

Contents

Chapter 0

Preliminaries

The following contains a brief summary of concepts, definitions, and notation needed in the later chapters and included here for completeness without much explanation.

0.1 Relativistic Notation

Space and time coordinates of e.g. a particle are denoted by $\vec{x} = (x_1, x_2, x_3)$ and t. In special relativity they transform under Lorentz transformations like the components of a four-vector $x^\mu := (x_0, x_1, x_2, x_3) := (c\,t, x_1, x_2, x_3)$ with c the speed of light. In general, greek indices run from 0 to 3, while latin indices run from 1 to 3. Energy and momentum E and $\vec{p}$ also transform as components of a four-vector, the *four-momentum* $p^\mu := (p_0, p_1, p_2, p_3) := (E/c, p_1, p_2, p_3)$.

The *metric tensor* in Minkowski-space special relativity is constant and given in a certain convention by

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Upper indices are called *contravariant*, lower ones *covariant*; they can be changed into one another with the help of the metric, e.g.

$$g_{\mu\nu} A^\nu{}_\rho{}^\sigma = A_{\mu\rho}{}^\sigma$$

Furthermore, for the metric one has

$$g_\mu{}^\nu = g^\mu{}_\nu = \delta^\mu{}_\nu = \mathbf{1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

which is the well-known Kronecker symbol. Another often-used expression is the Levi-Civita symbol

$$\varepsilon_{\mu\nu\rho\sigma} = \begin{cases} 0 & \text{any two indices equal} \\ 1 & \mu\nu\rho\sigma \text{ is an even permutation of (0123)} \\ -1 & \mu\nu\rho\sigma \text{ is an odd permutation of (0123)} \end{cases}$$

The analogous construction exists for ε_{ijk}.

The scalar product of two four-vectors is defined as

$$a \cdot b := \sum_{\mu=0}^{3} a_\mu b^\mu = a_0 b_0 - \vec{a} \cdot \vec{b} =: a_\mu b^\mu$$

where the Einstein summation convention is used (note that the same index cannot appear more than twice in any term). A four-vector a is called timelike, if $a \cdot a > 0$, spacelike for $a \cdot a < 0$ and lightlike for $a \cdot a = 0$, which also determines the causality implications of a.

Tensors of higher rank n have n open indices, either contravariant or covariant.

0.2 Group Theory

Consider a set $\mathcal{G}$ of elements g_i with $i = 1, 2, \ldots$ or uncountable or finite with an operation $\circ$ by which two elements can be connected. $\mathcal{G}$ is called a group, if the following axioms hold:

* $\forall\, g_1, g_2 \in \mathcal{G}: \quad g_1 \circ g_2 \in \mathcal{G}$
* $\exists\, \mathbf{1} \in \mathcal{G} : \mathbf{1} \circ g = g \circ \mathbf{1} = g \quad \forall g \in \mathcal{G}$
* $\forall g \in \mathcal{G}\ \exists\, g^{-1} \in \mathcal{G}: \quad g \circ g^{-1} = g^{-1} \circ g = \mathbf{1}$
* $\forall g_{123} \in \mathcal{G}: \quad g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$

However, $\circ$ is not necessarily commutative. If it is, the group is called *Abelian*. As an example, consider the set of 2×2 matrices with nonzero determinant together with matrix multiplication as the group operation. One can easily show that all axioms hold and that the group is not Abelian.

A *subgroup* $\mathcal{H}$ is a subset of $\mathcal{G}$ on which the same operation $\circ$ is defined, and all axioms hold for $\mathcal{H}$ alone. In the above example, a subgroup may be constructed by restricting the matrices to those with determinant $+1$.

An *invariant* or *normal subgroup* $\mathcal{I}$ is a subgroup of a group $\mathcal{G}$ such that

$$\forall\, y \in \mathcal{I}, g \in \mathcal{G}: \quad g \circ y \circ g^{-1} \in \mathcal{I}\ ; \tag{1}$$

one also writes

$$\forall\, g \in \mathcal{G}: \quad g\mathcal{I} = \mathcal{I}g\ , \tag{2}$$

In the same notation, for a subgroup $\mathcal{H}$ of a group $\mathcal{G}$, a *left coset* of $\mathcal{H}$ in $\mathcal{G}$ is defined by the set gH of all elements gh with $h \in \mathcal{H}$ generated via an element g of $\mathcal{G}$; a *right coset* is defined analogously. If $\mathcal{H}$ is actually an invariant subgroup, then left and right cosets are the same for all $g \in \mathcal{G}$. The number of left and right cosets of $\mathcal{H}$ in $\mathcal{G}$ is the same and called the *index* of $\mathcal{H}$ in $\mathcal{G}$.

0.3 Representations of Groups

A *representation* $\mathcal{D}$ of a group $\mathcal{G}$ with elements g is a homomorphism which maps $\mathcal{G}$ onto the group of automorphisms $\widetilde{\mathcal{G}}$ with elements $\tilde{g}$ on a given set with structure. The term *homomorphism* describes a map that conserves the structure of a set with respect to an operation defined among its elements, e.g. a group operation's properties are mapped onto the representation:

$$\begin{aligned} \mathcal{D}(g) &= \tilde{g} \\ \mathcal{D}(\mathbf{1}) &= \tilde{\mathbf{1}} \\ \mathcal{D}(g_2 \circ g_1) &= \tilde{g_2} \tilde{\circ} \tilde{g_1} = \mathcal{D}(g_2) \tilde{\circ} \mathcal{D}(g_1) \end{aligned}$$

A homomorphism is called *isomorphism*, if it is bijective. A homomorphism is called *endomorphism*, if it maps a set V onto itself. And finally, a bijective endomorphism is called an *automorphism*. A representation is called *faithful*, if the homomorphism $\mathcal{D}$ is injective. It is *trivial*, if all elements of the group are mapped onto the unit element of the automorphism group.

For our purposes it is interesting to talk about *linear representations*, which are obtained by mapping the group onto the group of automorphisms on a vector space. In practice this means linear transformations (matrices) in the $\mathbb{R}^n$ or $\mathbb{C}^n$. In that case one is concerned with square invertible real or complex matrices of dimension $n \times n$, which can also have more restrictive properties, like e.g. a determinant of $+1$.

Representations can be finite- or infinite-dimensional. In the latter case the representation maps the group onto the group of linear bounded operators on a complex vector (Hilbert) space. The operators' inverses must also be bounded.

Another important concept is the (ir-)reducibility of a representation. Assume a vector space with n dimensions which has a subspace $\mathcal{K}$ with $m < n$ dimensions, which is "invariant" under $\mathcal{G}$, i.e. if $k \in \mathcal{K}$ then $\mathcal{D}(g)\, k \in \mathcal{K}\ \forall\, k \in \mathcal{K}$, $g \in \mathcal{G}$. Then the matrix $\mathcal{D}(g) = \tilde{g}$ has the form

$$\tilde{g} = \begin{pmatrix} P_{\tilde{g}} & Q_{\tilde{g}} \\ 0 & S_{\tilde{g}} \end{pmatrix}$$

where $P_{\tilde{g}}$ has dimension $m \times m$, $S_{\tilde{g}}$ has $(n-m) \times (n-m)$ and $Q_{\tilde{g}}$ has $m \times (n-m)$, if one can find coordinates (unit vectors) such that an element $k \in \mathcal{K}$ has $k_i = 0$ for $i = m+1, \ldots, n$.

Such a representation is called *reducible*. If there is no such subspace, the representation is called *irreducible*. If $Q_{\tilde{g}} = 0$ the representation is *decomposable*. Every system of reducible unitary matrices is decomposable, because unitary transformations conserve orthogonality between vectors. Therefore dividing the basis (unit) vectors among e.g. $\mathcal{K}_1$ and $\mathcal{K}_2$ means that they are orthogonal to each other and $Q_{\tilde{g}} = 0$.

0.4 Quantum Theory

In quantum theory, physical states are represented by *rays* in a Hilbert space $\mathcal{H}$. A Hilbert space is a complex vector space on which an inner product is defined,

which induces a norm on the space, and the vector space is complete. If $|\Phi\rangle$ and $|\Psi\rangle$ are vectors in $\mathcal{H}$ (called *state vectors*), then so is the linear combination $\xi|\Phi\rangle + \eta|\Psi\rangle$ with the complex numbers ξ and η.

The existence of the norm on $\mathcal{H}$ means that for any two state vectors $|\Phi\rangle$ and $|\Psi\rangle$ there is a complex number defined by $\langle\Phi|\Psi\rangle = \langle\Psi|\Phi\rangle^*$, with * denoting the complex conjugate. This construction satisfies

$$
\begin{aligned}
&* \qquad \langle\Phi|\xi_1\Psi_1 + \xi_2\Psi_2\rangle = \xi_1\langle\Phi|\Psi_1\rangle + \xi_2\langle\Phi|\Psi_2\rangle \\
&* \qquad \langle\eta_1\Phi_1 + \eta_2\Phi_2|\Psi\rangle = \eta_1^*\langle\Phi_1|\Psi\rangle + \eta_2^*\langle\Phi_2|\Psi\rangle \\
&* \qquad \langle\Psi|\Psi\rangle \geq 0
\end{aligned}
$$

the latter implying that the norm only vanishes, if $|\Psi\rangle = 0$.

A *ray* is a set of normalized vectors (i.e. $\langle\Psi|\Psi\rangle = 1$) such that $|\Psi\rangle$ and $|\Psi'\rangle$ belong to the same ray if $|\Psi'\rangle = \xi|\Psi\rangle$ with the phase $\xi \in \mathbf{C}$, $|\xi| = 1$.

Observables in quantum theory are represented by hermitian operators. These are mappings $|\Psi\rangle \to A|\Psi\rangle$ of $\mathcal{H}$ onto itself, linear in the sense that

$$A|\xi\Psi + \eta\Phi\rangle = \xi A|\Psi\rangle + \eta A|\Phi\rangle \ .$$

They satisfy $A^\dagger = A$ (Hermiticity), i.e. their eigenvalues given by

$$A|\Psi\rangle = \alpha|\Psi\rangle$$

are real. The adjoint $A^\dagger$ to A is defined by

$$\langle\Phi|A^\dagger\Psi\rangle = \langle A\Phi|\Psi\rangle = \langle\Psi|A\Phi\rangle^* \ .$$

Eigenvectors corresponding to different αs are orthogonal.

As noted above, a physical state corresponds to a ray $\mathcal{R}$. Consider $\mathcal{R}_1, \mathcal{R}_2, \ldots$ to be mutually orthogonal. Then, if the state of a particular system is represented by $\mathcal{R}$, the probability of finding it in the state represented by $\mathcal{R}_n$ is

$$\mathcal{P}(\mathcal{R} \to \mathcal{R}_n) = |\langle\Psi|\Psi_n\rangle|^2$$

with

$$\sum_n \mathcal{P}(\mathcal{R} \to \mathcal{R}_n) = \mathbf{1}$$

if the set of state vectors $|\Psi_n\rangle$ is complete.

0.5 Matrices

Matrices appear in our context (and in physics in general) in many different circumstances. They can have a few interesting properties summarized here:

- The *dimension of a matrix* is written as $n \times m$ with n the number of rows and m the number of columns
- For a *square matrix* one has $n = m$

- The *determinant* of a matrix A of dimension $n \times n$ with elements $a_{ijkl...}$ carrying n indices running from 1 to n each is

$$\det(A) = \sum_{i,j,k,...,z=1}^{n} \varepsilon_{ijk...z}\, a_{1i}\, a_{2j}\, a_{3k}\, \ldots\, a_{nz}$$

- A *singular matrix* has determinant $= 0$
- A nonsingular matrix is *invertible*; the inverse of A is denoted by A^{-1}
- Matrix *operations* can be element-wise addition, subtraction, as well as matrix-multiplication $C = A \cdot B$ defined by

$$c_{ij} = \sum_k a_{ik} b_{kj}$$

  in terms of their elements. Division by a matrix can be realized via multiplication with the inverse. Caution: multiplication from the left or right in general makes a difference.
- A matrix A's transpose A^T is defined via $(A^T)_{ij} := A_{ji}$
- A matrix A with $A^T = A^{-1}$ is called *orthogonal*
- A matrix A with complex elements has an *adjoint* or *Hermitian conjugate* $A^\dagger$ defined via $(A^\dagger)_{ij} := A^*_{ji}$ with * denoting complex conjugation
- A matrix A with $A^\dagger = A$ is called *Hermitian*
- A matrix A with $A^\dagger = A^{-1}$ is called *unitary*
- Functions of matrices, e. g., the exponetial function $\exp A$, can be defined via appropriate series expansions of the desired function.

0.6 Symmetries in Quantum Theory

A symmetry transformation in physics changes our point of view (e.g. the coordinate system), but not the results of possible experimental measurements. For example, an observer $\mathcal{O}$ sees a system represented by a ray $\mathcal{R}$ or $\mathcal{R}_1$ or $\mathcal{R}_2$ while another observer $\mathcal{O}'$ has $\mathcal{R}'$ or $\mathcal{R}'_1$ or $\mathcal{R}'_2$. They must find the same probabilities

$$\mathcal{P}(\mathcal{R} \to \mathcal{R}_n) = \mathcal{P}(\mathcal{R}' \to \mathcal{R}'_n)$$

if the transformation between the observers' points of view is a symmetry transformation.

Wigner in the early 1930s showed that for any such transformation of rays $\mathcal{R} \to \mathcal{R}'$ of rays one can define an operator U on $\mathcal{H}$ such that if $|\Psi\rangle \in \mathcal{R}$ then $U|\Psi\rangle \in \mathcal{R}'$, where U is either *unitary* and *linear*

$$* \quad \langle U\Phi | U\Psi \rangle = \langle \Phi | \Psi \rangle$$
$$* \quad U|\xi\Phi + \eta\Psi\rangle = \xi U|\Phi\rangle + \eta U|\Psi\rangle$$

or *antiunitary* and *antilinear*

$$
\begin{aligned}
* \quad & \langle U\Phi|U\Psi\rangle = \langle\Phi|\Psi\rangle^* \\
* \quad & U|\xi\Phi + \eta\Psi\rangle = \xi^* U|\Phi\rangle + \eta^* U|\Psi\rangle
\end{aligned}
$$

Recall at this point that the adjoints of linear operators L and antilinear operators A are different:

$$
\begin{aligned}
* \quad & \text{linear operator } L\text{:} \quad \langle\Phi|L^\dagger\Psi\rangle = \langle L\Phi|\Psi\rangle \\
* \quad & \text{antilinear operator } A\text{:} \quad \langle\Phi|A^\dagger\Psi\rangle = \langle A\Phi|\Psi\rangle^* = \langle\Psi|A\Phi\rangle
\end{aligned}
$$

Both the unitarity and antiunitarity conditions are of the form

$$U^\dagger = U^{-1}$$

Now consider the trivial (identity) symmetry transform. It maps $\mathcal{R}$ onto $\mathcal{R}$ and is represented by $U = \mathbf{1}$, which is unitary and linear. From the latter it follows that any continuous symmetry (e.g. rotation symmetry) that can be made trivial by letting some parameter ε approach zero has to be represented by unitary and linear operators U. For infinitesimally small ε one has

$$U = \mathbf{1} + i\,\varepsilon\,\mathbf{t}$$

If U is unitary and linear, $\mathbf{t}$ must be hermitian and linear (which reminds us of observables in quantum theory). Indeed, in this way observables and symmetry transformations are connected (e.g. P^μ, $\vec{J}$, etc.).

Symmetry transformation sets can be groups:

$$
\begin{aligned}
T_1 \quad &: \quad \mathcal{R}_n \to \mathcal{R}'_n \\
T_2 \quad &: \quad \mathcal{R}'_n \to \mathcal{R}''_n \\
T_2 \circ T_1 \quad &: \quad \mathcal{R}_n \to \mathcal{R}''_n
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
T_1^{-1} \quad &: \quad \mathcal{R}'_n \to \mathcal{R}_n \\
\exists \text{ an identity } \mathbf{1} \quad &: \quad \mathcal{R}_n \to \mathcal{R}_n
\end{aligned}
$$

Unitary operators $U(T)$ mirror the group structure, but they operate on vectors, not rays. So on one hand:

$$
\begin{aligned}
|\Psi_n\rangle \quad &\in \quad \mathcal{R} \\
U(T_1)|\Psi_n\rangle \quad &\in \quad \mathcal{R}' \\
U(T_2)U(T_1)|\Psi_n\rangle \quad &\in \quad \mathcal{R}''
\end{aligned}
$$

On the other hand,

$$U(T_2 \circ T_1)|\Psi_n\rangle \quad \in \quad \mathcal{R}''$$

as well, so they can only differ by a phase:

$$U(T_2)U(T_1)|\Psi_n\rangle = \mathrm{e}^{i\varphi_n(T_2,T_1)}U(T_2 \circ T_1)|\Psi_n\rangle$$

In general, i.e. for "states with the same quantum numbers" φ_n does not depend on $_n$, i.e. the state vector $|\Psi_n\rangle$. If $\varphi = 0$ then one has a *representation* of the group by the operators U, if not one has a *projective representation.*

A special kind of group is a *Lie group*: there the transformations are characterized by a set of real continuous parameters

$$T = T(\vartheta) \; .$$

The $T(\vartheta)$ are connected continuously to $\mathbf{1}$ within the group, which means that for a unitary representation of the Lie group they are represented by unitary operators for any ϑ. The group composition law reads

$$T(\vartheta_2) \circ T(\vartheta_1) = T(f(\vartheta_2, \vartheta_1))$$

where f is some function of the two parameters from the composed transformations. Special cases are

$$f(\vartheta, 0) = f(0, \vartheta) = \vartheta$$

In the neighborhood of $\mathbf{1}$ the operator U corresponding to the transformation $T(\vartheta)$ is given by

$$U(T(\vartheta)) = \mathbf{1} + i\,\vartheta^a\,\mathbf{t}_a + \frac{1}{2}\vartheta^b\vartheta^c\,\mathbf{t}_{bc}$$

where the $\mathbf{t}_a$ are the *generators* of the group. As already mentioned above, unitary U implies hermitian $\mathbf{t}_a$, which are connected to observables.

From the group operations follow commutation relations for the generators $\mathbf{t}$ of the Lie group

$$[\mathbf{t}_b, \mathbf{t}_c] = i\,f^a_{bc}\,\mathbf{t}_a$$

The set of operators satisfying these relations is called the corresponding *Lie algebra.* The real constants f^a_{bc} are called its *structure constants.*

One can find operators C_i which commute with every generator of the Lie group, i.e. all elements of the Lie algebra. These are called *Casimir* operators and can be used to identify representations of the group via their eigenvalues. Their eigenvalues can be also used to label states together with the eigenvalues of those $\mathbf{t}_a$ that commute with the Hamiltonian.

To close this section it should be noted that any linear combination of generators is again a generator.

0.7 Exercises

This section contains a list of exercises the reader may use to intensify her/his engagement with the subject. In particular, the goal should be to make the concepts introduced above more immediate and familiar to the point, where their application feels natural.

Exercise 0.1 Consider the moments

$$d^\nu := -x_\mu T^{\mu\nu} \quad \text{and} \quad k^{\mu\nu} := (x_\sigma x^\sigma g^{\mu\varrho} - 2x^\mu x^\varrho)T_\varrho{}^\nu$$

of the (symmetric) energy-momentum tensor $T^{\mu\nu}$ of a closed system, i. e., $\partial_\mu T^{\mu\nu} = 0$ and calculate their four-divergences $\partial_\nu d^\nu$ and $\partial_\nu k^{\mu\nu}$

Exercise 0.2 Consider the tensor $F^{\mu\nu}$ to be antisymmetric $F^{\mu\nu} = -F^{\nu\mu}$.

(a) Show that F is traceless

(b) Show that the tensor defined by

$$T^{\mu\nu} := \frac{1}{4\pi}(F^{\mu\alpha}F_{\alpha}{}^{\nu} + \frac{1}{4}\, g^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta})$$

is symmetric and traceless

(c) Calculate the moments $\partial_\nu d^\nu$ and $\partial_\nu k^{\mu\nu}$ defined above for this construction of $T^{\mu\nu}$

Exercise 0.3 The tensor $\hat{F}^{\mu\nu} := \frac{1}{2}\varepsilon^{\mu\nu\varrho\sigma}F_{\varrho\sigma}$ is called *dual* to F.

(a) Show that the complex tensor $f^{\mu\nu} := \frac{1}{2}(F^{\mu\nu} + i\hat{F}^{\mu\nu})$ is dual to itself in the sense that $\hat{f}^{\mu\nu} = -if^{\mu\nu}$.

(b) How many independent components does f have?

Exercise 0.4 From everyday life, construct an example for a group, i. e., a set with an operation defined on it and explicitly prove that the group axioms hold.

Exercise 0.5 For your example from Exercise 0.4, check whether your group has any (nontrivial) invariant subgroups. If yes, find and detail them and the related cosets; determine their indices. If there is no nontrivial invariant subgroups in your example, think of a modification such that there are and perform the aforementioned tasks.

Exercise 0.6 Prove that

$$[A, B^n] = n[A, B]B^{n-1}\,, \quad \text{if} \quad [[A, B], B] = 0$$

Exercise 0.7 Construct an argument as to why it makes sense to label states in quantum theory using the eigenvalues of the Casimir operator(s) of a relevant symmetry group.

Achievement

After studying this section and performing the exercises (in fact, ideally already before that) the reader should be able and feel confident to do the following:

- Manipulate tensors in Minkowski space using covairant notation
- Work with matrices, operators, bras, and kets, and their properties
- Understand and work with the basic ideas of a group, a representation, and a symmetry in general
- Perform mathematical proofs and understand basic related techniques

Chapter 1

Relativity and Quantum Theory

We now come to the basic principles of the interconnection of quantum theory and the theory of special relativity. In the following we will use the standard particle-physics convention to set $c = 1$ and $\hbar = 1$, i.e. the units for mass, energy and momentum are given in MeV, GeV, etc., where the appropriate factors of c are implicit. Units for length and time are given in units of MeV^{-1}, etc.

1.1 The Poincaré Group

Special relativity demands physics to be the same in inertial frames. Such frames are related by space-time translations in four dimensions

$$x'^{\mu} = x^{\mu} + a^{\mu} \quad \text{with} \quad a^{\mu} = const.$$

and Lorentz transformations, i.e. rotations around three space directions and boosts in three space directions

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

which in total amounts to ten degrees of freedom.

For inertial frames connected via Lorentz transformations one demands that the speed of light is the same in every such frame. More concretely

$$\left| \frac{d\vec{x}}{dt} \right| \Rightarrow dx^{\mu} dx_{\mu} = 0 \ ,$$

i.e. the line element for light is (obviously) light-like. With regard to a Lorentz transformation one has

$$g_{\mu\nu} dx^{\mu} dx^{\nu} \stackrel{!}{=} g_{\mu\nu} dx'^{\mu} dx'^{\nu} \quad \text{with} \quad dx'^{\mu} = \Lambda^{\mu}{}_{\nu} dx^{\nu}$$

From this it follows immediately that

$$g_{\mu\nu} \Lambda^{\mu}{}_{\rho} \Lambda^{\nu}{}_{\sigma} = g_{\rho\sigma} \tag{1.1}$$

and

$$\Lambda^{\rho}{}_{\mu}\Lambda^{\sigma}{}_{\nu}g^{\mu\nu} = g^{\rho\sigma} \ .$$

Furthermore, it is clear from Eq. (1.1) that

$$(\mathrm{Det}\Lambda)^2 = 1 \quad \text{and} \quad \mathrm{Det}\Lambda = \pm 1 \ ,$$

which means that $\Lambda^{\mu}{}_{\nu}$ has an inverse:

$$\begin{aligned} g_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma} = g_{\rho\sigma} \quad &\Rightarrow \\ g_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}g^{\sigma\tau} = \delta^{\tau}_{\rho} \quad &\Rightarrow \\ \Lambda^{\mu}{}_{\rho}g_{\mu\nu}\Lambda^{\nu}{}_{\sigma}g^{\sigma\tau} \quad &= \quad \Lambda^{\mu}{}_{\rho}\Lambda_{\mu}{}^{\tau} = \delta^{\tau}_{\rho} \\ \text{and thus } (\Lambda^{-1})^{\mu}{}_{\nu} \quad &= \quad \Lambda_{\nu}{}^{\mu} \end{aligned} \tag{1.2}$$

Space-time translations and Lorentz transformations together are the Poincaré transformations. A transformation $T(\Lambda, a)$ corresponds to

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu} + a^{\mu}$$

These transformations form a group with the group operation being the subsequent application of two transformations. One can briefly check the group axioms: the first condition is that the composition of two Poincaré transformations is again a Poincaré transformation. More concretely, one can ask what the elements Λ and a of the composed transformation

$$T(\Lambda_2, a_2) \circ T(\Lambda_1, a_1) = T(\Lambda, a)$$

are. A simple calculation yields

$$\Lambda = \Lambda_2 \circ \Lambda_1 \quad \text{and} \quad a = \Lambda_2 a_1 + a_2$$

which satisfy the criteria for a Poincaré transformation, in particular Λ satisfies Eq. (1.1) and a is a constant four-vector.

Next, there is an identity, namely

$$\mathbf{1} = T(\mathbf{1}, 0) = T(\delta_{\mu}{}^{\nu}, 0) \ ,$$

and each group element has an inverse, namely

$$T^{-1}(\Lambda, a) = T(\Lambda^{-1}, -\Lambda^{-1}a) \ .$$

Finally, the associative rule holds, since matrix multiplication is associative and the construction from the composition law preserves associativity as well.

Now let us consider the substructure of the Poincaré group. Take Eq. (1.1) and set $\rho = \sigma = 0$. Then one has

$$g_{\mu\nu}\Lambda^{\mu}{}_{0}\Lambda^{\nu}{}_{0} = g_{00} = 1$$

Writing all elements explicitly this yields

$$(\Lambda^{0}{}_{0})^2 - \sum_{i=1}^{3}(\Lambda^{i}{}_{0})^2 = 1 \quad \text{and} \quad (\Lambda^{0}{}_{0})^2 = 1 + \sum_{i=1}^{3}(\Lambda^{i}{}_{0})^2$$

where the right-hand side is a sum of one and squares and thus greater or equal than 1:

$$|\Lambda^0{}_0| \geq 1 \Rightarrow \Lambda^0{}_0 \geq 1 \quad \text{or} \quad \Lambda^0{}_0 \leq -1$$

At this point, note that the identity element **1** has $\Lambda^0{}_0 \geq 1$.

The Poincaré group is also called the *inhomogeneous Lorentz group*. Its subgroups are

- the Lorentz group, i.e. $a = 0$
- Transformations with $\mathrm{Det}\Lambda = +1$ form subgroups of both the homogeneous and inhomogeneous Lorentz groups and are named *proper* (in)homogeneous Lorentz group.
- Transformations with (additionally) $\Lambda^0{}_0 \geq 1$ form the *proper orthochronous* Lorentz group
- Only proper orthochronous Lorentz transformations are continuously connected to **1**. The jumps between $\mathrm{Det}\Lambda = \pm 1$ and $\Lambda^0{}_0 \gtrless \pm 1$ are discontinuous. Still, any Lorentz transformation can be reached from **1** by proper orthochronous transformations times the *discrete* Lorentz transformations space inversion $\mathcal{P}$, time reversal $\mathcal{T}$, and their composition $\mathcal{PT}$ given explicitly in Sec. 1.3. Together with **1** these discrete transformations are a subgroup of the Lorentz group as well.

1.2 The Poincaré Algebra

The Poincaré group is a ten-parameter Lie group. Its Lie algebra of generators can be obtained by writing a Poincaré transformation close to **1** (meaning $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu$ and $a^\mu = 0$) as

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu \quad \text{and} \quad a^\mu = \varepsilon^\mu$$

with ω and ε infinitesimal. Now consider Eq. (1.1) for such an infinitesimal Lorentz transformation and get

$$g_{\sigma\rho} = g_{\mu\nu}(\delta^\mu{}_\rho + \omega^\mu{}_\rho)(\delta^\nu{}_\sigma + \omega^\nu{}_\sigma) = g_{\sigma\rho} + \omega_{\sigma\rho} + \omega_{\rho\sigma} + \mathcal{O}(\omega^2)$$

where terms of quadratic and higher order in ω are neglected and one obtains

$$\omega_{\sigma\rho} = -\omega_{\rho\sigma}$$

which means that ω is antisymmetric. It is a tensor of rank two in four dimensions and thus the antisymmetry implies that out of the in general 16 independent components only 6 actually independent ones remain, which can be characterized via 6 independent parameters. Together with the additional four independent parameters in ε^μ one gets the ten parameters characteristic for the Poincaré group and corresponding to the ten generators.

Let us now have a look at a (projective) unitary representation of the Poincaré group, i.e. we are looking for unitary operators to represent, in general, the transformation $T(\Lambda, a)$. First, in particular, it is interesting to consider $T(\mathbf{1} + \omega, \varepsilon)$:

$$U(T(\mathbf{1} + \omega, \varepsilon)) = \mathbf{1} + \frac{1}{2} i\omega_{\rho\sigma} J^{\rho\sigma} - i\varepsilon_\rho P^\rho + \text{terms of higher order}$$

From the unitarity of U it follows that both $J^\dagger = J$ and $P^\dagger = P$ are hermitian operators. Furthermore, since ω is antisymmetric, one can also make J antisymmetric. The next question is, what the transformation properties of J and P under Poincaré transformations are. To investigate this, we look at the product

$$U(\Lambda, a) U(T(\mathbf{1} + \omega, \varepsilon)) U^{-1}(\Lambda, a)$$

and evaluate it. On the one hand, this can be written as

$$U(\Lambda, a) \left[\mathbf{1} + \frac{1}{2} i\omega_{\rho\sigma} J^{\rho\sigma} - i\varepsilon_\rho P^\rho \right] U^{-1}(\Lambda, a) \, .$$

On the other hand, one can evaluate the product completely and then compare coefficients of the infinitesimal independent terms ω and ε, respectively. More precisely, one first composes the transformations via the group composition rule to get

$$U(\Lambda, a) U(T(\mathbf{1} + \omega, \varepsilon)) U^{-1}(\Lambda, a) = U(\mathbf{1} + \Lambda\omega\Lambda^{-1}, \Lambda\varepsilon - \Lambda\omega\Lambda^{-1} a) =$$

which yields, writing everything in terms of infinitesimals again

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

Comparison of the coefficients of ε yields

$$U(\Lambda, a) P^\rho U^{-1}(\Lambda, a) = \Lambda_\mu{}^\rho P^\mu \tag{1.3}$$

and the same procedure for ω produces

$$U(\Lambda, a) J^{\rho\sigma} U^{-1}(\Lambda, a) = \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma (J^{\mu\nu} - a^\mu P^\nu + a^\nu P^\mu) \tag{1.4}$$

which are the desired transformation properties of the Poincaré generators under Poincaré transformations.

One can see that P transforms like a four vector and for a homogeneous Lorentz transformation ($a = 0$), J transforms like a tensor of rank two. For a pure translation ($\Lambda = \delta$), P is invariant, while J changes.

The next step is to determine the algebra itself, i.e. the commutation relations of the generators among each other. For this purpose, consider the transformation rules above with (Λ, a) infinitesimal as well:

$$\Lambda = \mathbf{1} + \omega \, , \quad a = \varepsilon$$

and get

$$\begin{aligned} U(\Lambda, a) &= \mathbf{1} + \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} - i\varepsilon_\mu P^\mu \\ U^{-1}(\Lambda, a) &= \mathbf{1} - \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} + i\varepsilon_\mu P^\mu \end{aligned}$$

Inserting this into Eq. (1.4) and comparing coefficients of the independent infinitesimals ω (and ε) yields

$$i\,[J^{\mu\nu}, J^{\rho\sigma}] = g^{\nu\rho}J^{\mu\sigma} - g^{\mu\rho}J^{\nu\sigma} - g^{\sigma\mu}J^{\rho\nu} + g^{\sigma\nu}J^{\rho\mu} \tag{1.5}$$

The same procedure for Eq. (1.3) yields two relations. The first comes from comparing coefficients of ω, namely

$$i\,[P^\mu, J^{\rho\sigma}] = g^{\mu\rho}P^\sigma - g^{\mu\sigma}P^\rho \tag{1.6}$$

and the second from comparing the coefficients of ε is

$$[P^\mu, P^\nu] = 0\,. \tag{1.7}$$

To better understand what the Poincaré algebra contains and implies we will rewrite it here in terms of the commonly used operators H and $\vec{P}$, the Hamiltonian and the three-momentum operator, which constitute the four-momentum operator P^μ, and the angular momentum operator $\vec{J} = (J_1, J_2, J_3)$ as well as the boost operator $\vec{K} = (K_1, K_2, K_3)$. The components of the latter two vectors transform together under Lorentz transformations as the six independent components of an antisymmetric tensor of rank two, namely $J^{\mu\nu}$ introduced above. One has

$$J^{\mu\nu} = \begin{pmatrix} 0 & -K_1 & -K_2 & -K_3 \\ K_1 & 0 & J_3 & -J_2 \\ K_2 & -J_3 & 0 & J_1 \\ K_3 & J_2 & -J_1 & 0 \end{pmatrix}$$

Inserting these definitions into Eqs. (1.5) – (1.7) one gets the commutation relations for the 10 operators $\{H, \vec{P}, \vec{J}, \vec{K}\}$ in the form

$$\begin{aligned} [J_i, J_j] &= i\varepsilon_{ijk}J_k && (1.8)\\ [J_i, K_j] &= i\varepsilon_{ijk}K_k && (1.9)\\ [K_i, K_j] &= -i\varepsilon_{ijk}J_k && (1.10)\\ [J_i, P_j] &= i\varepsilon_{ijk}P_k && (1.11)\\ [K_i, P_j] &= i\delta_{ij}H && (1.12)\\ [K_i, H] &= iP_i && (1.13)\\ [J_i, H] &= [P_i, H] = [H, H] = 0 && (1.14) \end{aligned}$$

H plays a special role in quantum theory as the generator of the time evolution of the system. Operators that commute with H are "conserved", which means that one can use their eigenvalues to label the states. Under Poincaré

transformations the situation is more complicated. The commutation relations above tell us that all components of the four-momentum commute with each other. On the other hand, one can choose any single component of the angular momentum operator, J_i, and then go ahead to diagonalize it together with the Hamiltonian simultaneously, but not together with the other momentum components. Below we will see, that a more general construction related to spin is required to achieve this. Still, the eigenvalues of neither the four-momentum operator nor a spin component are invariant under Poincaré transformations, i.e. in a general relativistic setup, states labeled by such eigenvalues (i.e. momenta and/or spin projections) will have to be properly transformed under e.g. Lorentz transformations.

1.3 Discrete Lorentz Transformations

Now let us consider the discrete Lorentz transformations space inversion $\mathcal{P}$, time reversal $\mathcal{T}$, and their product $\mathcal{PT}$:

$$\mathcal{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \mathcal{T} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathcal{PT} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

As already mentioned above, the set $\{\mathbf{1}, \mathcal{P}, \mathcal{T}, \mathcal{PT}\}$ forms a discrete subgroup of the Lorentz and Poincaré groups. While one has to note at this point, that certain interactions in physical theories can violate these symmetries, we consider them here as symmetry transformations and look for their representations in form of operators. We define, in the notation of operators representing Poincaré transformations,

$$\mathbf{P} := U(\mathcal{P}, 0) \quad \text{and} \quad \mathbf{T} := U(\mathcal{T}, 0)$$

such that for a Poincaré transformation given by (Λ, a) one has

$$\begin{aligned} \mathbf{P} U(\Lambda, a) \mathbf{P}^{-1} &= U(\mathcal{P}\Lambda\mathcal{P}^{-1}, \mathcal{P}a) \\ \mathbf{T} U(\Lambda, a) \mathbf{T}^{-1} &= U(\mathcal{T}\Lambda\mathcal{T}^{-1}, \mathcal{T}a) \end{aligned}$$

Now one can find the operations of $\mathbf{P}$ and $\mathbf{T}$ on the Poincaré generators by taking (Λ, a) infinitesimal once again and comparing the coefficients of ω and ε in analogy to the constructions above to get

$$\begin{aligned} \mathbf{P}\, i\, J^{\rho\sigma} \mathbf{P}^{-1} &= i\, \mathcal{P}_\mu{}^\rho \mathcal{P}_\nu{}^\sigma J^{\mu\nu} \\ \mathbf{P}\, i\, P^\rho \mathbf{P}^{-1} &= i\, \mathcal{P}_\mu{}^\rho P^\mu \qquad (1.15) \\ \mathbf{T}\, i\, J^{\rho\sigma} \mathbf{T}^{-1} &= i\, \mathcal{T}_\mu{}^\rho \mathcal{T}_\nu{}^\sigma J^{\mu\nu} \\ \mathbf{T}\, i\, P^\rho \mathbf{T}^{-1} &= i\, \mathcal{T}_\mu{}^\rho P^\mu \qquad (1.16) \end{aligned}$$

Factors of i have not been canceled yet, since no decision has been made on whether $\mathbf{P}$ and $\mathbf{T}$ are linear (and unitary) or antilinear (and antiunitary). Take $\rho = 0$ in Eq. (1.15), then one has

$$\mathbf{P}\, i\, H \mathbf{P}^{-1} = i\, H\ .$$

If $\mathbf{P}$ were antilinear, then the left-hand side would be

$$-i\mathbf{P}H\mathbf{P}^{-1}$$

and the equation would read

$$\mathbf{P}H\mathbf{P}^{-1} = -H \quad \text{or} \quad H\mathbf{P}^{-1} = -\mathbf{P}^{-1}H \,.$$

In this case, for each state $|\Psi\rangle$ with

$$H|\Psi\rangle = |\Psi\rangle E \,, \quad E > 0 \,,$$

there would have to also be a state $\mathbf{P}^{-1}|\Psi\rangle$ with

$$H\mathbf{P}^{-1}|\Psi\rangle = -\mathbf{P}^{-1}H|\Psi\rangle = -\mathbf{P}^{-1}|\Psi\rangle E$$

i.e. an energy smaller than that of the vacuum (zero energy). Thus, $\mathbf{P}$ cannot be antilinear, but must be linear and therefore unitary. Furthermore, $\mathbf{P}$ commutes with the Hamiltonian. For $\mathbf{T}$ the same argument works exactly the other way round; it turns out that $\mathbf{T}$ must be antilinear and therefore antiunitary.

With this knowledge, one can now evaluate the transformation properties given in covariant form above. Furthermore one can rewrite everything in the more intuitive versions using the operators $\{H, \vec{P}, \vec{J}, \vec{K}\}$ as

$$\begin{aligned} \mathbf{P}\ \vec{J}\ \mathbf{P}^{-1} &= +\vec{J} & \qquad \mathbf{T}\ \vec{J}\ \mathbf{T}^{-1} &= -\vec{J} \\ \mathbf{P}\ \vec{K}\ \mathbf{P}^{-1} &= -\vec{K} & \qquad \mathbf{T}\ \vec{K}\ \mathbf{T}^{-1} &= +\vec{K} \\ \mathbf{P}\ \vec{P}\ \mathbf{P}^{-1} &= -\vec{P} & \qquad \mathbf{T}\ \vec{P}\ \mathbf{T}^{-1} &= -\vec{P} \end{aligned} \tag{1.17}$$

In other words, the angular momentum operator transforms like an axial vector, the Lorentz-boost operator and three-momentum operator transform like polar vectors. Under time reversal, the angular momentum operator and three-momentum operator change direction, while the boost operator does not.

1.4 Casimir Operators of the Poincaré Group

There are two operators, whose eigenvalues are perfectly suited to classify states with regard to Poincaré transformations. This quest leads us to the *Casimir operators* of the Poincaré group, which per definitionem commute with all the generators. There are two Casimir operators of the Poincaré group: the first is the square of the four-momentum operator, the mass-operator squared given by

$$M^2 := P_\mu P^\mu \,. \tag{1.18}$$

If the *spectral condition* $M^2 \geq 0$ is satisfied (which it usually is in a sensible physical theory), the mass operator M can be written as the nonnegative square root of M^2 which is defined by diagonalizing M^2 and replacing all eigenvalues by their nonnegative square roots.

As a consequence, physical states labeled by the mass of a particle require no adjustment with regard to this label in different inertial frames. Another

consequence is that irreducible representations of the Poincaré group can be distinguished (labeled) by the mass of the particle under consideration (e.g. $m = 0$, $m \neq 0$).

The second Casimir operator of the Poincaré group is the square $W^\mu W_\mu$ of the *Pauli-Lubanski* vector operator, which is given by

$$W^\mu := \frac{1}{2}\varepsilon^\mu{}_{\nu\rho\sigma} J^{\nu\rho} P^\sigma \ . \tag{1.19}$$

The hermitian axial-vector operator W^μ satisfies the commutation relations

$$[W^\mu, P^\nu] = 0 \tag{1.20}$$

$$i\left[W^\mu, J^{\alpha\beta}\right] = g^{\mu\alpha} W^\beta - g^{\mu\beta} W^\alpha \tag{1.21}$$

$$[W^\mu, W^\nu] = i\varepsilon^{\mu\nu\rho\sigma} W_\rho P_\sigma \ . \tag{1.22}$$

Furthermore $P_\mu W^\mu = 0$, which follows immediately from the definition of the Pauli-Lubanski operator. Another useful defining relation is $W^\mu W_\mu =: -M^2\, \vec{j}^2$, where the vector operator $\vec{j}$ can be understood as the total intrinsic spin operator of the system (see below). At this point, one can now ask again for an as-large-as-possible set of mutually commuting operators, which can be diagonalized simultaneously. Clearly, the two Casimir operators qualify, since they commute with all generators of the Poincaré group. In addition, one has the components of the three-momentum $\vec{P}$, and one of the components of the spin vector operator $\vec{j}$ defined from W^μ can be chosen, which commutes with all other operators in the set. Thus, one arrives at $\{M, \vec{j}^2, \vec{P}, j_3\}$. When considering the transformation properties of one-particle states below, it will become clear how the eigenvalues of these operators react to Poincaré transformations of quantum states.

1.5 Spin Operators in Relativistic Quantum Theory

When discussing spin operators in relativistic quantum theory one has to also discuss boosts. In fact, different boosts have various effects such as differently constructed spin vector operators, differently constructed position operators, or different invariance considerations of objects under Poincaré transformations.

Regarding the construction of a spin vector operator in relativistic quantum theory, the choice of boost plays an important role due to the following situation: Although the spin operator is unambiguously defined in terms of the Poincaré generators, there is an infinite number of spin vector valued functions of the generators that satisfy angular-momentum commutation relations and whose square is the total spin. Essentially, one has to decide how to "perform a transformation" to the restframe (of a massive particle) by a particular boost. The most common choice is a rotationless or *canonical* boost, which is denoted as a Lorentz transformation by $\Lambda = L_c(Q)$ where Q is the velocity operator P/M and thus L_c is to be understood as an operator. The other two noteworthy boosts are called *front form* and *helicity* boosts, respectively.

To construct a spin vector operator from the Pauli-Lubanski operator, one first realizes that for a timelike four momentum and with $P_\mu W^\mu = 0$ as noted

above, W^μ is spacelike and can thus be transformed to the rest frame of the particle with the corresponding Lorentz transformation. So, one has to "perform a boost" and "divide by the mass":

$$(0, \vec{j}_c) := \frac{1}{M} L_c(Q)^\mu{}_\nu W^\nu \tag{1.23}$$

The prerequisites to this equation are that $L_c(Q)$, when applied to simultaneous eigenstates of $\vec{P}$ and M takes on the value of a Lorentz transformation that transforms P to $(M, \vec{0})$ and that all components of the four momentum operator commute with the Pauli-Lubanski operator. Analogous relations hold to define *front form spin* and *helicity spin* vectors.

From Eq. (1.23) and various commutation and transformation properties found above it follows that

$$\vec{j}_c \cdot \vec{j}_c = -\frac{1}{M^2} W^2 \tag{1.24}$$

and

$$M^2 [j_c^i, j_c^j] = i\varepsilon^{ijk} j_c^l M^2 \ . \tag{1.25}$$

If M has no vanishing eigenvalues, one can divide it out of the equation to arrive at the usual spin commutation relations.

Finally one can construct the spin vector operator in terms of the Poincaré generators. For canonical spin, one has

$$\vec{j}_c = \frac{1}{M} \left(\vec{W} - \frac{\vec{P} W^0}{M + H} \right) = \frac{1}{M} \left((H\vec{J} - \vec{P} \times \vec{K}) - \frac{\vec{P}(\vec{P} \cdot \vec{J})}{M + H} \right) \tag{1.26}$$

Analogous formulae can be found for front from and helicity spin.

1.6 The Position Operator in Relativistic Quantum Theory

The definition of a position operator in relativistic quantum theory is problematic in the sense that position is not an observable. To be more specific, in the discussion of the relativistically invariant localization of a particle with mass m one encounters a problem in the sense that such a particle cannot be in an invariant state *and* localized more precisely than its Compton wavelength. In a model with particle-antiparticle creation one can argue along the lines that to probe smaller and smaller distances one needs higher and higher momenta, which eventually can create particle-antiparticle pairs, where the particle created is identical to the one to be localized by the probe. In this way, uncertainty in the localization appears, also at the scale of the particle's Compton wavelength.

Nevertheless, a position operator can be constructed, which is canonically conjugate to the momentum operator and commutes with the spin operator of the system. This operator is called the *Newton-Wigner position operator*. It is constructed from the Poincaré generators; it itself is not a generator. The

construction is not unique, since it depends on the choice of spin operator via the commutation relation. As a concrete example, the Newton-Wigner operator for canonical spin is

$$\vec{X}_c = -\frac{1}{2}\left(\frac{1}{H}\vec{K} + \vec{K}\frac{1}{H}\right) - \frac{\vec{P}\times(H\vec{J} - \vec{P}\times\vec{K})}{MH(H+M)} \tag{1.27}$$

As a final note one could state that position operators are perfectly fine to use in a relativistic context, except for an interpretation as an observable.

1.7 Preview: The Poincaré Algebra in Quantum Field Theory

In a quantum field theory the central object of physics input is the Lagrangian (density). In such a formalism, symmetries are built-in via construction of that Lagrangian density and result in/are manifested by conserved currents/charges.

While one deals with Poincaré transformations from this point of view, we want to also put emphasis here on the satisfaction of the Poincaré algebra by the Poincaré generators as constructed from the Lagrangian of the theory under consideration. First, note that symmetries can be global or local (the latter depend on x). They can occur at the level of the action, the Lagrangian, or the Lagrangian density (consider, e.g., spatial translations).

In general, one can use Noether's theorem to study symmetries and the related conserved currents in a Lagrangian formalism. Regarding space-time translations one uses the energy-momentum tensor $T^{\mu\nu}$ and finds

$$\partial_\mu T^\mu{}_\nu = 0 \; .$$

In analogy, for Lorentz transformations, one gets from a Lorentz invariance of the Lagrangian density that

$$\partial_\rho \mathcal{M}^{\rho\mu\nu} = 0 \quad \text{where} \quad \mathcal{M}^{\rho\mu\nu} = -\mathcal{M}^{\rho\nu\mu}$$

(this corresponds to a conserved current for each of the elements of the familiar infinitesimal parameters $\omega_{\mu\nu}$ defined in Sec. 1.2). In both cases one can perform an integral over space, d^3x, to obtain

$$P_\nu = \int d^3x \; T^0{}_\nu \qquad \frac{d}{dt}P_\nu = 0 \tag{1.28}$$

$$\text{and} \quad J_{\mu\nu} = \int d^3x \; \mathcal{M}^0{}_{\mu\nu} \qquad \frac{d}{dt}J_{\mu\nu} = 0 \; . \tag{1.29}$$

On closer inspection, one can define a symmetric version of $T^{\mu\nu}$, the Belinfante tensor $\Theta^{\mu\nu}$ with again

$$P_\nu = \int d^3x \; \Theta^0{}_\nu$$

and subsequently write

$$\mathcal{M}^{\rho\mu\nu} := x^\mu \Theta^{\rho\nu} - x^\nu \Theta^{\rho\mu}$$

which leads to

$$\partial_\rho \mathcal{M}^{\rho\mu\nu} = \Theta^{\mu\nu} - \Theta^{\nu\mu} = 0$$

due to the symmetry of Θ mentioned above. Finally, one has

$$J^{\mu\nu} = \int d^3x \, (x^\mu \Theta^{0\nu} - x^\nu \Theta^{0\mu}) \, .$$

Starting from these constructions of T and Θ from canonical field variables as well as their commutation relations one can explicitly show the validity of the Poincaré algebra.

1.8 Exercises

Exercise 1.1 Other than usual, consider Lorentz transformations out of our context simply reduced to their matrix structure when transforming four vectors in Minkowski space. Furthermore, here we explicitly have both $c \neq 1$ and the boost velocity v.Prove Eq. (1.2) by using the explicit Lorentz transformation matrix

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & e_1\beta\gamma & e_2\beta\gamma & e_3\beta\gamma \\ e_1\beta\gamma & 1+(\gamma-1)e_1^2 & (\gamma-1)e_1e_2 & (\gamma-1)e_1e_3 \\ e_2\beta\gamma & (\gamma-1)e_1e_2 & 1+(\gamma-1)e_2^2 & (\gamma-1)e_2e_3 \\ e_3\beta\gamma & (\gamma-1)e_1e_3 & (\gamma-1)e_2e_3 & 1+(\gamma-1)e_3^2 \end{pmatrix}$$

where $\beta = \frac{v}{c}$, $\vec{e}$ is the unit vector in the direction of the velocity v and $\gamma = (1-\beta^2)^{-1/2}$

Exercise 1.2 Show that the set $\{\mathbf{1}, \mathcal{P}, \mathcal{T}, \mathcal{PT}\}$ is a subgroup of the Lorentz group.

Exercise 1.3 Prove Eqs. (1.3) and (1.4) explicitly.

Exercise 1.4 From the transformation properties of the Poincaré generators, Eqs. (1.3) and (1.4), derive the Poincaré algebra, Eqs. (1.5) – (1.7).

Exercise 1.5 In analogy to Exercise 1.3, prove the four equations under (1.15) and (1.16) explicitly.

Exercise 1.6 In analogy to the argument used to prove the unitarity of **P** in Sec. 1.3, prove explicitly that **T** is antiunitary.

Exercise 1.7 Consider the Pauli-Lubanski vector operator W^μ

(a) Prove Eq. (1.20) explicitly from the commutation relations of W^μ with the generators of the Poincaré group

(b) Prove Eq. (1.22) with the help of (1.20) and (1.21).

(c) With the help of these relations, prove explicitly that $P^\mu P_\mu$ and $W^\mu W_\mu$ are indeed Casimir operators of the Poincaré group.

(d) Calculate the general expression for $W^\mu W_\mu$ in terms of the Poincaré generators P^μ and $J^{\mu\nu}$.

Achievement

After studying this section and performing the exercises the reader should be able and feel confident to do the following:

- Understand and apply Lorentz- and Poincaré transformations at the level of group and representation theory
- Know the structure of the Poincaré group and algebra
- Be able to derive a Lie algebra's structure from the group properties and work with the generators.
- Understand the meaning and properties of the Poincaré generators
- Understand the meaning and properties of the discrete Lorentz transformations
- Know the Casimir operators of the Poincaré group and understand their properties and implications
- Have a basic understanding of the concept of the spin vector and position operators in relativistic quantum theory
- Anticipate the construction of the Poincaré generators in the formalism of a relativistic quantum field theory

Chapter 2

Representations of the Poincaré group

2.1 One-Particle States

The next goal is to classify one-particle states via their transformation properties under the Poincaré group. We will label the one-particle state to be transformed by its four momentum p and spin projection σ such that, since the four momentum and Pauli-Lubanski operators commute,

$$P^\mu |p,\sigma\rangle = |p,\sigma\rangle p^\mu$$

Note that a translation by some finite a is written as

$$U(\mathbf{1},a) = \mathrm{e}^{-iP^\mu a_\mu}$$

This can be achieved by going from $\mathbf{1}$ to the finite a via a number of N small steps so that one has a product of N transformations of the form

$$(\mathbf{1} + P^\mu a_\mu/N)^N$$

which in the limit $N \to \infty$ gives the exponential function. Analogously, for a rotation $R_{\vec{\vartheta}}$ around a certain axis by a certain angle, both defined via the vector $\vec{\vartheta}$, one has

$$U(R_{\vec{\vartheta}},0) = \mathrm{e}^{-i\vec{J}\cdot\vec{\vartheta}} \; .$$

Next, note that any Poincaré transformation can be decomposed into a pure Lorentz transformation and a pure translation afterwards:

$$T(\Lambda,a) = T(\mathbf{1},a) \circ T(\Lambda,0) \; .$$

As a result, one can study the transformation properties of a state under translations and Lorentz transformations separately.

Thus, under translations one has

$$U(\mathbf{1},a)|p,\sigma\rangle = \mathrm{e}^{-iP^\mu a_\mu}|p,\sigma\rangle = \mathrm{e}^{-ip^\mu a_\mu}|p,\sigma\rangle \; . \tag{2.1}$$

What about Lorentz transformations? To shorten the following equations, we use the notation $U(\Lambda, 0) \equiv U(\Lambda)$. First we determine the four-momentum eigenvalue of the state $U(\Lambda)|p, \sigma\rangle$ and obtain

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

Regarding the spin projection we can say at this point that $U(\Lambda)|p, \sigma\rangle$ is a linear combination of $|\Lambda p, \sigma'\rangle$ with

$$U(\Lambda)|p,\sigma\rangle = \sum_{\sigma'} C_{\sigma'\sigma}(\Lambda, p)|\Lambda p, \sigma'\rangle \tag{2.2}$$

Note that for block-diagonal C one has a reducible representation of the Lorentz group. In the following we will study the properties of the coefficients C in irreducible representations of the Lorentz group.

To study the C, note that all proper orthochronous Lorentz transformations leave p^2 invariant and that for timelike or lightlike momenta ($p^2 \leq 0$) they also leave the sign of p_0 invariant. We start by defining a *standard momentum* k, which for a massive particle is given in the particle's rest frame, and use a (canonical) boost $L_c(p/m)$ to bring it to the momentum p which the state under consideration should have

$$p^\mu = L_c{}^\mu{}_\nu(p/m)k^\nu \ . \tag{2.3}$$

Then one can define the state $|p, \sigma\rangle$ by (using p synonymously instead of p/m as the argument of the boost)

$$|p,\sigma\rangle = N(p)U(L_c(p))|k,\sigma\rangle \ ,$$

where N is some normalization factor to be determined later. For now we want to know the transformation property of $|p, \sigma\rangle$. To obtain this, apply an arbitrary Lorentz transformation to the definition just given and get

$$\begin{aligned} U(\Lambda)|p,\sigma\rangle &= U(\Lambda)N(p)U(L_c(p))|k,\sigma\rangle = N(p)U(\Lambda L_c(p))|k,\sigma\rangle \\ &= N(p)U(L_c(\Lambda p))U^{-1}(L_c(\Lambda p))U(\Lambda L_c(p))|k,\sigma\rangle \qquad (2.4) \\ &= N(p)U(L_c(\Lambda p))U(L_c^{-1}(\Lambda p)\Lambda L_c(p))|k,\sigma\rangle \qquad (2.5) \end{aligned}$$

The transformation $U(L_c^{-1}(\Lambda p)\Lambda L_c(p))$ changes the momentum of the state in steps in the following way

$$k \to p \to \Lambda p \to k$$

i.e. it belongs to a subgroup of the Lorentz group that leaves the standard momentum k invariant. This subgroup is called *little group* of transformations $\mathbf{W}$ with

$$\mathbf{W}^\mu{}_\nu k^\nu = k^\mu \ .$$

Now we can write Eq. (2.2) using the transformation $U(\mathbf{W})$ on the state $|k, \sigma\rangle$ as

$$U(\mathbf{W})|k,\sigma\rangle = \sum_{\sigma'} D_{\sigma'\sigma}(\mathbf{W})|k,\sigma'\rangle$$

where the $D_{\sigma'\sigma}(\mathbf{W})$ are a representation of the little group (one can check the group and representation axioms).

At this point let us also note that the states can be normalized according to

$$\langle k', \sigma' | k, \sigma \rangle = \delta^3(\vec{k}' - \vec{k})\delta_{\sigma'\sigma} \tag{2.6}$$

i.e. the representation of the little group must be unitary:

$$D^\dagger(\mathbf{W}) = D^{-1}(\mathbf{W}) \; .$$

Let us now take a closer look at the case of a massive particle (mass m such that $p^2 = -m^2 < 0$ and $p^0 > 0$), in which case the little group is the group of three-dimensional rotations, $SO(3)$. Unitary representations of this group can be broken up into a direct sum of irreducible unitary representations associated with a rotation R, $D^{(j)}_{\sigma'\sigma}(R)$ of dimensionality $2j+1$, $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$. Going back and inserting this into Eq. (2.2), the transformation property of a state under a Lorentz transformation Λ now reads

$$U(\Lambda)|p, \sigma\rangle = \sqrt{\frac{(\Lambda p)^0}{p^0}} \sum_{\sigma'} D^{(j)}_{\sigma'\sigma}(\mathbf{W}(\Lambda, p))|\Lambda p, \sigma'\rangle \tag{2.7}$$

where we have evaluated the normalization factor N in accord with the normalization property Eq. (2.6). The *Wigner rotation* $\mathbf{W}$ in this equation is given by

$$\mathbf{W}(\Lambda, p) = L_c^{-1}(\Lambda p)\Lambda L_c(p) \; . \tag{2.8}$$

2.2 Two-Body States and the Clebsch-Gordan Coefficients of the Poincaré Group

Irreducible representations are characterized by mass and spin. When we think about a particle in this way, it doesn't matter whether it is elementary or a bound state of other particles. Furthermore, the dynamical setup to get the bound state doesn't matter, either. We'll now begin to construct a two-body bound state via free (i.e. noninteracting) two-body states, then operators, and then introduce interaction, which will lead us to the Bakamjian-Thomas construction.

The first basic problem is the following: Consider a two-particle Hilbert space as the tensor product of the corresponding single-particle spaces. Then one has a basis, writing (e.g. for canonical spin, always denoted by the subscript $_c$)

$$|p_1\sigma_1 p_2\sigma_2\rangle_c := |p_1\sigma_1(j_1)\rangle_c \otimes |p_2\sigma_2(j_2)\rangle_c \tag{2.9}$$

normalized such that

$${}_c\langle p_1'\sigma_1' p_2'\sigma_2'|p_1\sigma_1 p_2\sigma_2\rangle_c = \delta_{\sigma_1'\sigma_1}\delta_{\sigma_2'\sigma_2}\delta^3(\vec{p_1}' - \vec{p}_1)\delta^3(\vec{p_2}' - \vec{p}_2) \tag{2.10}$$

The spin j_i has been noted explicitly in the single-particle states above in addition to the spin projection σ_i, since this plays a role for the properties of the

single-particle irreducible representations of the Poincaré group considered in each case (we'll see below, how the j_i appear in the argument).

Now the question is how this two-particle state transforms under Poincaré transformations and what the corresponding generators are. The Transformation is given by the tensor product of two single-particle representations as

$$U(\Lambda, a) := U_1(\Lambda, a) \otimes U_2(\Lambda, a) \; . \tag{2.11}$$

Using this representation, the Poincaré generators are sums of the generators for each particle

$$\begin{aligned} P^\mu &= P_1^\mu \otimes \mathbf{1}_2 + \mathbf{1}_1 \otimes P_2^\mu & (2.12)\\ \vec{K} &= \vec{K}_1 \otimes \mathbf{1}_2 + \mathbf{1}_1 \otimes \vec{K}_2 & (2.13)\\ \vec{J} &= \vec{J}_1 \otimes \mathbf{1}_2 + \mathbf{1}_1 \otimes \vec{J}_2 \quad \text{or shorter} & (2.14)\\ J^{\mu\nu} &= J_1^{\mu\nu} \otimes \mathbf{1}_2 + \mathbf{1}_1 \otimes J_2^{\mu\nu} \; . & (2.15) \end{aligned}$$

Operators like the total mass and spin for the system (without interactions) can be constructed via the definitions given earlier. However, they are in general polynomials or other nonlinear functions in the generators and thus cannot be written as the sum of their single-particle counterparts. We'll make this more explicit later.

One can/wants to use states to describe the system which are labeled by a total momentum and a total spin. This would be natural in a situation where one has no information as to whether the particle under investigation is elementary or not. Only after the assumption of inner structure one can/has to worry about constituents, their spins and their relative momentum. So the aim in this respect is to find coefficients that link two-particle states to the states with "total" quantum numbers. These are the Clebsch-Gordan coefficients of the Poincaré group. More precisely, they are the coefficients between a tensor product of two irreducible representations of the Poincaré group and a linear superposition of irreducible representations of the Poincaré group (note that this still describes noninteracting particles).

This is in analogy to the case of angular momentum in quantum mechanics, where one is dealing with irreducible representations of the rotation group. For spin s and projection m_s one has

$$|(s_1 s_2)\, s\, m_s\rangle = \sum_{m_{s_1}=-s_1}^{s_1} \sum_{m_{s_2}=-s_2}^{s_2} |s_1\, m_{s_1}\, s_2\, m_{s_2}\rangle \; C^{s m_s}_{s_1 m_{s_1} s_2 m_{s_2}} \; , \tag{2.16}$$

where the C are the corresponding Clebsch-Gordan coefficients.

So we want to express products of single-particle states as linear combinations of two-particle states that have the same Poincaré-transformation properties as a single free particle. How do we get this?

First, identify eigenstates of the four-momentum operator: consider two massive particles in a massive two-particle system. Furthermore, we will (always) use canonical spin. Then

$$P^\mu |p_1 \sigma_1 p_2 \sigma_2\rangle_c = |p_1 \sigma_1 p_2 \sigma_2\rangle_c \; (p_1^\mu + p_2^\mu) \tag{2.17}$$

from the construction of the total four-momentum operator in Eq. (2.12). Now let us construct a state in the rest frame of the system via the canonical boost

$$L_c^{-1}(P) := L_c^{-1}(\frac{P}{\sqrt{P^2}}) \tag{2.18}$$

where the two notations in Eq. (2.18) are used synonymously. We get

$$\begin{aligned} |\vec{k},\sigma_1,-\vec{k},\sigma_2\rangle_c &= [\text{Norm}]^{1/2}\, U(L_c^{-1}(P)) \sum_{\sigma_1'\sigma_2'} |p_1\sigma_1' p_2\sigma_2'\rangle_c \\ &\times D^{(j_1)}_{\sigma_1'\sigma_1}(R_c^{-1}(L_c^{-1}(P),p_1))\, D^{(j_2)}_{\sigma_2'\sigma_2}(R_c^{-1}(L_c^{-1}(P),p_2)) \end{aligned} \tag{2.19}$$

This is an eigenstate of P^μ with the eigenvalues

$$P^\mu = (M_0,\vec{0}) \quad \text{with} \quad M_0 = \sqrt{m_1^2+\vec{k}^{\,2}} + \sqrt{m_2^2+\vec{k}^{\,2}}\,. \tag{2.20}$$

Under rotations the rest-frame state transforms as

$$\begin{aligned} U(R)|\vec{k},\sigma_1,-\vec{k},\sigma_2\rangle_c &= \sum_{\sigma_1'\sigma_2'} |\overrightarrow{Rk},\sigma_1',-\overrightarrow{Rk},\sigma_2'\rangle_c \\ &\times D^{(j_1)}_{\sigma_1'\sigma_1}(R_c(R,k_1))\, D^{(j_2)}_{\sigma_2'\sigma_2}(R_c(R,k_2))\,, \end{aligned} \tag{2.21}$$

where $k_1 = (\sqrt{m_1^2+\vec{k}^{\,2}},\vec{k})$ and $k_2 = (\sqrt{m_2^2+\vec{k}^{\,2}},-\vec{k})$.

Now remember that for canonical spin

$$R_c(R,k_1) = R_c(R,k_2) = R$$

and thus

$$\begin{aligned} U(R)|\vec{k},\sigma_1,-\vec{k},\sigma_2\rangle_c &= \sum_{\sigma_1'\sigma_2'} |\overrightarrow{Rk},\sigma_1',-\overrightarrow{Rk},\sigma_2'\rangle_c \\ &\times D^{(j_1)}_{\sigma_1'\sigma_1}(R)\, D^{(j_2)}_{\sigma_2'\sigma_2}(R)\,. \end{aligned} \tag{2.22}$$

This result shows that all (three-) momenta and the spins projections in this rest-frame state transform in the same way under a rotation, which is an extremely helpful property when constructing a total angular from spins and orbital angular momenta.

With this construction we are now in the position to compute the Clebsch-Gordan coefficients of the Poincaré group. In particular, this means to evaluate the matrix elements of the rest-frame two-particle state just defined and the single-particle state for the total system in its rest frame written as

$$|(m,\vec{0}),\sigma(j)\rangle_c \tag{2.23}$$

and then use boosts to connect everything to non-rest-frame states. The "total" state transforms under rotations as

$$U(R)\,|(m,\vec{0}),\sigma(j)\rangle_c = \sum_{\sigma'} |(m,\vec{0}),\sigma'(j)\rangle_c\, D^{(j)}_{\sigma'\sigma}(R)\,. \tag{2.24}$$

The steps in the calculation are the following:

- Use an expansion in spherical harmonics to relate Eq. (2.19) to Eq. (2.23) via the Clebsch-Gordan coefficients of the rotation group
- Use a canonical boost to bring the total state to its "normal" form with total momentum P
- Use the canonical boost relation in Eq. (2.19) to relate the two-particle rest-frame state to the tensor-product momentum state for two particles

Then, one can put it all together, including proper normalization factors, to get an expression for the inner product

$$ {}_c\langle p_1\sigma_1 p_2\sigma_2|P\sigma(j)\rangle_c \ , \tag{2.25}$$

which is the desired Clebsch-Gordan coefficient of the Poincaré group.

2.3 The Method of Induced Representations

Assume we have a representation of the little group such as the one found above. Considering massive particles we have the transformation property under a rotation R, using canonical spin and boosts,

$$ U(R)|k,\sigma\rangle_c = \sum_{\sigma'=-j}^{j} D^{(j)}_{\sigma'\sigma}(R)|k,\sigma'\rangle_c \ . \tag{2.26}$$

Now, we'll illustrate how one can get the transformation property of any state $|p,\sigma\rangle_c$ under any Poincaré transformation $U(\Lambda,a)$. First, note that $|p,\sigma\rangle_c$ is constructed via a canonical boost as

$$ U(L_c(p))|k,\sigma\rangle_c = f\times|p,\sigma\rangle_c \ .$$

The factor f can be determined from the normalization condition for the states, Eq. (2.6), and from the fact that $U(L)$ must be unitary as

$$ f = \left(\frac{\sqrt{m^2+\vec{p}^2}}{m}\right)^{1/2} \times \text{a phase}$$

Next, note that we already established the action of a translation on a state $|p,\sigma\rangle_c$ above in Eq. (2.1). These relations together with the group properties fix the action of $U(\Lambda,a)$. In particular, for a given combination p, Λ, a, one can uniquely represent $U(\Lambda,a)$ as

$$ U(\Lambda,a) = U(\mathbf{1},a)U(L_c(\Lambda p))U(R_c(\Lambda,p))U(L_c^{-1}(p)) \ . \tag{2.27}$$

In other words, the steps are

- Inverse boost of general state to its rest frame
- Wigner rotate according to (Λ,p)

- Boost to the frame with momentum Λp
- Translate

In this way one obtains a single-valued unitary representation of the Poincaré group, since all steps are unitary. The details in every step can be collected to evaluate $U(\Lambda, a)$ explicitly in terms of matrix elements between states $|p, \sigma\rangle_c$.

2.4 $ISL(2, \mathbb{C})$ and the Poincaré group

$ISL(2, \mathbb{C})$, the inhomogeneous extension of $SL(2, \mathbb{C})$ is the covering group of the Poincaré group. It is the group of ordered pairs (Λ, a) of complex 2×2 matrices Λ and a with $|\Lambda| = 1$ and $a^\dagger = a$. A Minkowski four vector x^μ can in this context be written as a complex 2×2 matrix X whose determinant is exactly equal to $x^\mu x_\mu$:

$$x^\mu \to \mathrm{X} := x^\mu \sigma_\mu = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} \quad \text{and} \quad x^\mu = \frac{1}{2}\mathrm{Tr}(\sigma^\mu \mathrm{X}) \; . \tag{2.28}$$

Here $\sigma^\mu = (\mathbf{1}, \vec{\sigma})$, with the Pauli matrices $\vec{\sigma}$. In their standard representation these are given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \; , \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \; , \tag{2.29}$$

and, together with $\mathbf{1}$ form a basis for the space of complex 2×2 matrices.

An element Λ of the Lorentz group is written as an $SL(2, \mathbb{C})$ element with determinant 1, so the determinant of the transformed matrix (which corresponds to the scalar product of the corresponding four vector in Minkowski space) is invariant under the Lorentz transformation $\mathrm{X} \to \mathrm{X}' = \Lambda \mathrm{X} \Lambda^\dagger$, as one can easily see.

Let us now consider the form of specific Lorentz transformations in this context. The first instructive example is a boost in z-direction, which turns out to simply read:

$$\Lambda(L_z) = \begin{pmatrix} e^{\alpha/2} & 0 \\ 0 & e^{-\alpha/2} \end{pmatrix} \; . \tag{2.30}$$

If we look at general rotationless boosts, we can write the element of $SL(2, \mathbb{C})$ corresponding to a boost L as a hermitian matrix H_L of the form

$$H_L = \cosh(\omega/2)\sigma^0 + \sum_{k=1}^{3} \sinh(\omega/2) n_k \sigma^k = e^{(\omega/2)\vec{n}\cdot\vec{\sigma}} \; , \tag{2.31}$$

where the unit vector $\vec{n}$ and ω define the direction and rapidity of the boost, respectively. Note that $\vec{n}$ is real and ω is positive, with $\frac{\vec{v}}{c} = \tanh(\omega)\vec{n}$.

In a similar way, one can write a rotation R around a real axis $\vec{n}$ through an angle $\varphi \in [0, 4\pi)$ by the unitary matrix U_R of the form

$$U_R = \cos(\varphi/2)\sigma^0 - i \sum_{k=1}^{3} \sin(\varphi/2) n_k \sigma^k = e^{-i(\varphi/2)\vec{n}\cdot\vec{\sigma}} \; . \tag{2.32}$$

With these two relations, any Lorentz transformation can be written as an element of $SL(2,\mathbb{C})$, because it can be decomposed into a pure boost and a pure rotation (this is often referred to in this context as the *decomposition theorem*). An analogous decomposition law is valid for the $SL(2,\mathbb{C})$ elements: The *polar decomposition theorem* states that any element of $SL(2,\mathbb{C})$ can be written as a product of a hermitian and a unitary part.

Given an element A of $SL(2,\mathbb{C})$, the corresponding Lorentz transformation $\Lambda^{\mu}{}_{\nu}$ can be constructed by

$$\Lambda^{\mu}{}_{\nu} = \frac{1}{2}\mathrm{Tr}(\bar{\sigma}^{\mu} A \sigma_{\nu} A^{\dagger}) \tag{2.33}$$

with $\bar{\sigma}^{\mu} = (\mathbf{1}, -\vec{\sigma})$.

2.5 The massless case

For the above arguments we frequently used the explicit assumption that the particle(s) under consideration have nonzero rest mass. While this is (of course) not generally true, the structure of the arguments mostly remains, which is why it was both concrete and generally instructive. Now only a few modifications and additional arguments are necessary to discuss the case of a massless particle. More precisely, we are immediately interested in the little group for this case.

The starting point is analogous to the massive case, namely analogous to Eq. (2.3)

$$p^{\mu} = L_{f\,\nu}^{\mu} k^{\nu} \,, \tag{2.34}$$

which is used to define the transformation from an arbitrary momentum p to a standard momentum k, but now for a massless particle. With the four-momentum of a massless particle being timelike, the usual choice for the standard momentum is

$$k^{\mu} = (\mathrm{p}, 0, 0, \mathrm{p}) \tag{2.35}$$

where the z-direction has been chosen without loss of generality. Now one can ask, which elements of the Poincaré group leave such a four-vector invariant. One finds that the little group in this case is $E(2)$, the Euclidean group in two dimensions. It contains two-dimensional translations as well as rotations in two dimensions (i. e. around a single real angle ϑ) as well as combinations thereof. More precisely, in general one has either a translation followed by a rotation or the other way round.

Going back to the $SL(2,\mathbb{C})$ picture, one can parameterize an element of $E(2)$ as

$$e_2(\vartheta, a) = \begin{pmatrix} e^{i\vartheta/2} & a\, e^{i\vartheta/2} \\ 0 & e^{-i\vartheta/2} \end{pmatrix} \tag{2.36}$$

with the real rotation angle ϑ and the complex constant a responsible for the translations in two dimensions.

For the following discussion it is instructive to define a special set of coordinates (*light-front coordinates*):

$$x^{+} := x^{0} + x^{3} \quad \text{and} \quad x^{-} := x^{0} - x^{3} \,, \tag{2.37}$$

the others x^1 and x^2 are unchanged and are usually denoted as $\vec{x}_\perp := (x^1, x^2)$ or $x_\perp := x^1 + ix^2$, the latter of which is used below.

In particular, writing the standard momentum of Eq. (2.35) as an element of $SL(2,\mathbb{C})$, one obtains the matrix

$$H(k) = \begin{pmatrix} k_+ & k_\perp^* \\ k_\perp & k_- \end{pmatrix} = \begin{pmatrix} 2\mathrm{p} & 0 \\ 0 & 0 \end{pmatrix} . \tag{2.38}$$

With the transformation rules defined above one obtains via the parametrization in Eq. (2.36) and using the general expression in Eq. (2.38) first, one has

$$e_2(\vartheta, a)\; H(k)\; e_2^\dagger(\vartheta, a) = \begin{pmatrix} k_+ + |a|^2 k_- + a k_\perp + a^* k_\perp^* & (k_\perp^* + a k_-) e^{i\vartheta} \\ (k_\perp + a^* k_-) e^{-i\vartheta} & k_- \end{pmatrix} , \tag{2.39}$$

from which one sees immediately that this indeed satisfies the requirements of a little-group element for the standard momentum. The next question is what the form of L in Eq. (2.35) is. The answer to this question is provided by a front-form boost, in the case of (2.35) given by

$$L_f(p) = \begin{pmatrix} \sqrt{\frac{p_+}{2\mathrm{p}}} & 0 \\ \frac{p_\perp}{\sqrt{2p_+\mathrm{p}}} & \sqrt{\frac{2\mathrm{p}}{p_+}} \end{pmatrix} \tag{2.40}$$

Starting from the right-hand side in (2.35) one can show that the transformation via L_f yields the general lightlike momentum p.

In terms of the representations of general Poincaré transformations for the massless case one finds that any element A of $SL(2,\mathbb{C})$ can be decomposed into a front form boost and an element of $E(2)$ so that one has

$$A = L_f(p)\; e_2(\vartheta, a) . \tag{2.41}$$

Concluding remarks of this section concern the form of the Lorentz generators: one obtains $J^{\mu\nu}$ using light-front components as

$$J^{\mu\nu} = \begin{pmatrix} 0 & (K_1 + J_2) & (K_2 - J_1) & -K_3 \\ -(K_1 + J_2) & 0 & J_3 & -(K_1 - J_2) \\ -(K_2 - J_1) & -J_3 & 0 & -(K_2 + J_1) \\ K_3 & (K_1 - J_2) & (K_2 + J_1) & 0 \end{pmatrix} . \tag{2.42}$$

Then we can define

$$J^{\mu\nu} =: \begin{pmatrix} 0 & E_1 & E_2 & -K_3 \\ -E_1 & 0 & J_3 & -F_1 \\ -E_2 & -J_3 & 0 & -F_2 \\ K_3 & F_1 & F_2 & 0 \end{pmatrix} . \tag{2.43}$$

This means the following replacements:

$$\begin{aligned} F_1 &= (K_1 - J_2) & E_1 &= (K_1 + J_2) \\ F_2 &= (K_2 + J_1) & E_2 &= (K_2 - J_1) \end{aligned} \tag{2.44}$$

With these definitions and the usual Poincaré commutation relations one arrives at the a set of commutation relations for the light-front generators E and F with the other generators.

2.6 Covariant Two-Fermion States

Quarks, e. g., are fermions; more precisely, they are spin-1/2 particles. Relativistically, this requires treatment of the quark spin degree of freedom via Dirac four-component spinors. In such context, e. g., a $q\bar{q}$ state is represented by an outer product of two of these four-component objects, i. e., by a 4×4 matrix. A basis for these matrices is a set of matrices constructed with the help of the Dirac γ matrices.

The Dirac matrices themselves are widely used and known. Still, their basic properties and those of the Dirac spinors are reviewed by some of the exercises at the end of this chapter.

From the point of view of the Lorentz group we are talking about representations once again. The generator of Lorentz transformations $J^{\mu\nu}$ appears in the representation $D(\Lambda)$ of an infinitesimal Lorentz transformation given by

$$D(\Lambda) = \mathbf{1} + \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} \tag{2.45}$$

where ω is infinitesimal. Furthermore, remember the commutation relations of the components of $J^{\mu\nu}$, Eq. (1.5):

$$i\,[J^{\mu\nu}, J^{\rho\sigma}] = g^{\nu\rho}J^{\mu\sigma} - g^{\mu\rho}J^{\nu\sigma} - g^{\sigma\mu}J^{\rho\nu} + g^{\sigma\nu}J^{\rho\mu} \tag{2.46}$$

Now, let γ^μ be a set of matrices with the anticommutation relations

$$\{\gamma^\mu, \gamma^\nu\} = 2\; g^{\mu\nu}\; \mathbf{1} \tag{2.47}$$

which is the well-known Clifford algebra. Then we define

$$J^{\mu\nu} = -\frac{i}{4}\,[\gamma^\mu, \gamma^\nu]\ , \tag{2.48}$$

and find that this construction satisfies Eq. (2.46). Furthermore, one gets the relation

$$[J^{\mu\nu}, \gamma^\rho] = -i\gamma^\mu g^{\nu\rho} + i\gamma^\nu g^{\mu\rho}\ . \tag{2.49}$$

This means that γ^ρ transforms like a four-vector under Lorentz transformations. Next, let's have a closer look at the Lorentz-transformation properties of products of γ matrices. We have just seen that γ^μ transforms like a vector under Lorentz transformations, i.e.

$$D(\Lambda)\gamma^\mu D^{-1}(\Lambda) = \Lambda_\nu{}^\mu \gamma^\nu\ . \tag{2.50}$$

In the same way, the 4×4 unit matrix $\mathbf{1}$ transforms like a scalar, namely

$$D(\Lambda)\mathbf{1}D^{-1}(\Lambda) = \mathbf{1}\ . \tag{2.51}$$

Furthermore, via writing $J^{\mu\nu} = -\frac{i}{4}\,[\gamma^\mu, \gamma^\nu]$ it is clear that $J^{\mu\nu}$ is an antisymmetric tensor of rank two:

$$D(\Lambda)J^{\mu\nu}D^{-1}(\Lambda) = \Lambda_\rho{}^\mu \Lambda_\sigma{}^\nu J^{\rho\sigma}\ . \tag{2.52}$$

One can now use the γ^μ to construct further antisymmetric tensors of higher rank:

$$\mathcal{A}^{\rho\sigma\tau} := \gamma^{[\rho}\gamma^{\sigma}\gamma^{\tau]} \tag{2.53}$$

which is a short-hand notation for the sum over all permutations of the indices within the brackets with a plus or minus sign for even or odd permutations, respectively. Written out explicitly, this means

$$\begin{aligned} \mathcal{A}^{\rho\sigma\tau} &= \gamma^\rho\gamma^\sigma\gamma^\tau - \gamma^\rho\gamma^\tau\gamma^\sigma \\ &+ \gamma^\tau\gamma^\rho\gamma^\sigma - \gamma^\tau\gamma^\sigma\gamma^\rho \\ &+ \gamma^\sigma\gamma^\tau\gamma^\rho - \gamma^\sigma\gamma^\rho\gamma^\tau \end{aligned} \tag{2.54}$$

The next rank is covered by

$$\mathcal{B}^{\mu\nu\rho\sigma} := \gamma^{[\mu}\gamma^{\nu}\gamma^{\rho}\gamma^{\sigma]} \; . \tag{2.55}$$

In this way one can construct tensors of any rank. This set of antisymmetric tensors forms a basis for all matrices that can be constructed from the Dirac matrices. The reason is that for any product of γs one can find a form of a sum of such antisymmetric tensors times metric tensors via the γ's anticommutation relation.

At this point, let us investigate the matrix $\beta := i\gamma^0$ and its action on the γ matrices. Since $\beta^{-1} = -i\gamma^0$, it follows that

$$\beta\gamma^i\beta^{-1} = -\gamma^i \tag{2.56}$$

$$\beta\gamma^0\beta^{-1} = \gamma^0 \tag{2.57}$$

This means that β acts on the γs like a parity transformation. On a product of γs, this yields exactly one minus sign for every spatial γ component. In particular,

$$\beta J^{ij}\beta^{-1} = J^{ij} \tag{2.58}$$

$$\beta J^{i0}\beta^{-1} = -J^{i0} \tag{2.59}$$

Let us now return to the question of finding a basis for a 2-fermion state. So far in the discussion about the γ matrices we have not made any assumption about the number of space-time dimensions or the metric. So let us consider 4 space-time dimensions (for obvious reasons). Then, one needs only the following basis of antisymmetric tensors, and the unit matrix:

$$\mathbf{1}, \quad \gamma^\mu, \quad J^{\mu\nu}, \quad \mathcal{A}^{\rho\sigma\tau}, \quad \mathcal{B}^{\mu\nu\rho\sigma} \; . \tag{2.60}$$

This set is linearly independent: they all transform differently under the Lorentz group, i.e. proper Lorentz and/or parity transformations. The scalar product for the corresponding orthogonality relation is given by matrix multiplication and subsequently taking the trace of the result. The reason for not needing more than these matrices is the number of Dirac matrices in four dimensions, namely four: γ^0, γ^1, γ^2, γ^3. This means that any product of five or more matrices of this kind has at least two matrices with the same index. Via Eq. (2.47) such a

product can be changed by anticommuting γs until those with the same index are next to each other and one has $(\gamma^\mu)^2 = \mathbf{1}$ for any value of μ, which, step by step, eventually reduces the number of γ matrices in the product to ≤ 4.

Now let's consider the above basis in more detail. The terms $\mathcal{A}$ and $\mathcal{B}$ can be written in simpler form: $\mathcal{B}^{\mu\nu\rho\sigma}$ is totally antisymmetric, i.e. it must be proportional to the Levi-Civita symbol in four dimensions and one can write

$$\mathcal{B}^{\mu\nu\rho\sigma} = i\ 4!\gamma_5\varepsilon^{\mu\nu\rho\sigma} \tag{2.61}$$

with the definition of $\gamma_5 = -i\gamma_0\gamma_1\gamma_2\gamma_3$. Similarly, for $\mathcal{A}^{\rho\sigma\tau}$ one finds

$$\mathcal{A}^{\mu\nu\rho} = i\ 3!\gamma_5\varepsilon^{\mu\nu\rho\sigma}\gamma_\sigma\ . \tag{2.62}$$

γ_5 is a *pseudoscalar*, i.e. it transforms like a scalar under Lorentz transformations and gets a minus sign from a parity transform. More precisely,

$$[J^{\mu\nu}, \gamma_5] = 0 \tag{2.63}$$

$$\beta\gamma_5\beta^{-1} = -\gamma_5 \tag{2.64}$$

In the same way, $\gamma_5\gamma^\mu$ is an axial- (or pseudo-) vector. One can easily check that $\gamma_5^2 = \mathbf{1}$ and $\{\gamma^\mu, \gamma_5\} = 0$.

Now let us count the independent components of the tensors in the basis above. A totally antisymmetric tensor with n indices in d dimensions has $\frac{d!}{n!(d-n)!}$ independent components, so we have $1+4+6+4+1 = 16$ components, which coincides with the number of independent components of an arbitrary 4×4 matrix. So we take Dirac matrices of dimension 4; this also means that they are irreducible with respect to the Lorentz group.

At this point, two more remarks are in order. First, note that the γ-matrix form for $J^{\mu\nu}$ is also written as

$$\sigma^{\mu\nu} := \frac{i}{2}\left[\gamma^\mu, \gamma^\nu\right]\ . \tag{2.65}$$

Finally, we note that the C-parity transformation in the standard representation for the γ matrices is given by

$$C := -i\gamma^2\gamma^0\ , \quad C^{-1} = -C\ ,$$

and an additional transposition of the corresponding matrix is needed so that, e.g.,

$$C\gamma^\mu C^{-1} = -(\gamma^\mu)^T \tag{2.66}$$

where the superscript ^{T} denotes matrix transposition.

2.7 Exercises

Exercise 2.1 Consider a massive particle and the Wigner rotations defined via Eq. (2.8). With the help of the definition of a canonical boost L_c, show that, if Λ is itself a rotation R, then $\mathbf{W}(R,p)$ is independent of p, more precisely, the Wigner rotation associated with a rotation is the rotation itself: $\mathbf{W}(R,p) = R$.

Exercise 2.2 Consider the following construction of a few-body state (*velocity state*) which starts from a few-particle momentum state in its rest frame:

$$|v, \vec{k}_1, \sigma_1, \vec{k}_2, \sigma_2, \ldots, \vec{k}_n, \sigma_n\rangle := U(L_c(v))|k_1, \sigma_1, k_2, \sigma_2, \ldots, k_n, \sigma_n\rangle \quad (2.67)$$

where v is the total four-velocity of the system and the k_i are rest-frame momenta of the particles defined with respect to the individual particle momenta p_i via $L_c(v)k_i = p_i$. They are considered on-shell and satisfy $\sum_{i=1}^n \vec{k}_i = 0$ (only $n-1$ of them are linearly independent).

Investigate the Lorentz-transformation properties of a velocity state under $U(\Lambda)$ and show that all internal variables ($\vec{k}_i$, σ_i) are transformed by the same rotation.

Exercise 2.3 Consider the complex 2×2 representations of four-vectors via the Pauli matrices appearing in Eq. (2.29). Without resorting to the concrete representation of the σ_i consider two four-vectors a^μ and b^μ and their matrix counterparts A and B.

(a) Show that

$$2\mathrm{Tr}[AB] = \mathrm{Tr}[A]\mathrm{Tr}[B] + \mathrm{Tr}[\vec{\sigma}A] \cdot \mathrm{Tr}[\vec{\sigma}B]$$

(b) Express $\mathrm{Det}[A]$ in terms of $\mathrm{Tr}[A]$ and $\mathrm{Tr}[A^2]$

(c) What are the components of b^μ for $B = A^{-1}$?

Exercise 2.4 Consider Eq. (2.28) and the light-front coordinates defined also above in (2.37):

$$x^+ := x^0 + x^3 \quad \text{and} \quad x^- := x^0 - x^3 \,,$$

with x^1 and x^2 unchanged.

(a) Compute the form of the scalar product of a four-vector x^μ with itself as a function of these coordinates.

(b) Show that in these coordinates, a boost of x^μ in z-direction is given by simple scale transformations of the coordinates.

(c) Consider further the $SL(2,\mathbb{C})$ matrix representation of a front-form boost as given in Eq. (2.40) Show that Λ_{L_f} maps the light front $x^+ = 0$ back onto itself.

(d) Show furthermore that transformation of the lightlike standard vector in Eq. (2.35) by a front form boost yields a general lightlike momentum p.

(e) From the $SL(2,\mathbb{C})$ matrix representation of a front-form boost, construct its $SO(1,3)$ representation.

Exercise 2.5 Properies of the Dirac matrices are essentially based on the Clifford algebra (2.47). The "Feynman-slash" short-hand notation $\not{a}$ stands for $\gamma^\mu a_\mu$.

(a) Show that

- $\not{a}\not{b} = a \cdot b\, \mathbf{1} - i\sigma_{\mu\nu} a^\mu b^\nu$
- $\gamma^\nu \gamma^\mu \gamma_\nu = -2\gamma^\mu$
- $\gamma^\varrho \gamma^\mu \gamma^\nu \gamma_\varrho = 4g^{\mu\nu}\, \mathbf{1}$

(b) Prove the (Dirac-)trace identities

- Tr[any product of an odd number of γ matrices] $= 0$
- $\mathrm{Tr}[\sigma^{\mu\nu}] = 0$
- $\mathrm{Tr}[\gamma^\mu \gamma^\nu] = 4g^{\mu\nu}$

(c) In the Dirac representation the matrices are given by:

$$\gamma^0_D = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \quad \text{and} \quad \gamma^i_D = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{2.68}$$

In the chiral representation one has, on the other hand:

$$\gamma^0_\chi = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} \quad \text{and} \quad \gamma^i_\chi = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{2.69}$$

Show that the Matrix U relating these two representations via $\gamma^\mu_\chi = U \gamma^\mu_D U^\dagger$ is given by

$$U = \frac{1}{\sqrt{2}}(\mathbf{1} - \gamma_5 \gamma_0) = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{1} & -\mathbf{1} \\ \mathbf{1} & \mathbf{1} \end{pmatrix}$$

(d) Compute γ^5 in both representations

Exercise 2.6 Prove that the construction Eq. (2.48) indeed satisfies the commutation relations

(a) Eq. (2.46),

(b) Eq. (2.49), and

(c) Eq. (2.63).

Achievement

After studying this section and performing the exercises the reader should be able and feel confident to do the following:

- Understand and apply Lorentz- and Poincaré transformations at the level of one and few-body states
- Understand the meaning and context of the little group for a particular representation of the Poincaré group
- Understand the idea of the method of induced representations
- Understand the concept of a Wigner rotation
- Know and be able to use the relation between Poincaré transformations and their $SL(2,\mathbb{C})$-couterparts
- Be able to make sense of light-front coordinates
- Deal with fermion states and Pauli as well as Dirac matrices

Chapter 3

The Bethe-Salpeter Equation

The Bethe-Salpeter equation (BSE) is an integral equation to study a quantum field theoretical 2-body system with regard to possible bound states as well as to scattering solutions. It is the relativistic analog of the Lippmann-Schwinger equation in quantum mechanics. Historically it has its name from Bethe and Salpeter, who presented the equation at a meeting of the American Physical Society in 1951 and did their work in an approach using Feynman graphs. However, the equation appeared from many different directions and was investigated by various people, in particular: Nambu (1950; position-space differential equation in ladder truncation), Gell-Mann and Low (1951; field theory), Schwinger (1951; functional formalism), Hayashi and Munakata (1952, presented earlier in 1950 at meetings of the Japanese Physical Society; Feynman graphs), Kita (1952; S-matrix theory), and Mandelstam (1955; energy-plane analyticity).

While some of the above derivations have (at least partly) a perturbative character, the equation is nonperturbative. This is important, since bound states should not be accessible in perturbation theory. This can be understood in the naive picture of scattering in perturbation theory: the longer two particles interact, the more often they can exchange quanta. For a bound state, this "interaction duration" approaches infinity. Still, since Feynman graphs are both instructive and intuitive, they are used below to motivate the BSE.

3.1 The Two-Fermion Four-Point Green Function and the BSE

Assume that we have two distinct fermions interacting with each other, described by fermion operators $\psi_{(1)}(x)$, $\psi_{(2)}(x)$. We represent the two-fermion connected Green function $G(x_1, x_2; x_1', x_2')$ by the diagram in Fig. 3.1 It is given by the expression

$$G(x_1, x_2; x_1', x_2') = \frac{\langle 0|T[\psi_{(1)}(x_1)\psi_{(2)}(x_2)\bar{\psi}_{(1)}(x_1')\bar{\psi}_{(2)}(x_2')]|0\rangle}{\langle 0|U|0\rangle} \tag{3.1}$$

This is written in the interaction picture, where the Lagrangian density of the system is written as $\mathcal{L}(x) = \mathcal{L}_0(x) + \mathcal{L}_I(x)$ and $\mathcal{L}_I(x)$ is the interaction

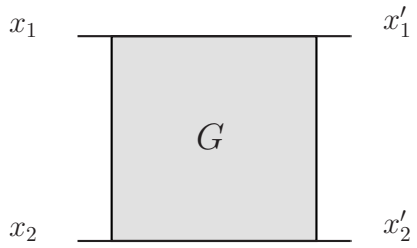

Figure 3.1: The two-fermion four-point Green function

Lagrangian density. The operator U stands for the Dyson series

$$U = T\left[\exp(i\int d^4x\ \mathcal{L}_I(x))\right]\ , \tag{3.2}$$

T is the time-ordering operator, $|0\rangle$ the vacuum state, and the denominator $\langle 0|U|0\rangle$ contains (and eliminates from G) all unconnected parts of the Green function ("vacuum bubbles"). The Green function G describes the "propagation" of the two-particle system in spacetime.

Furthermore, we need the fermion propagator $S_{(1),(2)}(x,x')$ (in pictorial notation represented as $\longleftarrow$), which is given by

$$S_{(j)}(x,x') = \frac{\langle 0|T[\psi_{(j)}(x)\bar{\psi}_{(j)}(x')]|0\rangle}{\langle 0|U|0\rangle} \tag{3.3}$$

analogous to G above.

Furthermore, we need to distinguish reducible and irreducible diagrams to build an interaction kernel that doesn't result in double-counting. A diagram in our present context is called reducible, if one can cut the diagram apart by cutting the two fermion (solid) lines once inside the diagram and none of the boson (dashed) lines. In Fig. 3.2 diagrams 2, 4, 6, 7, and 8 are reducible, the

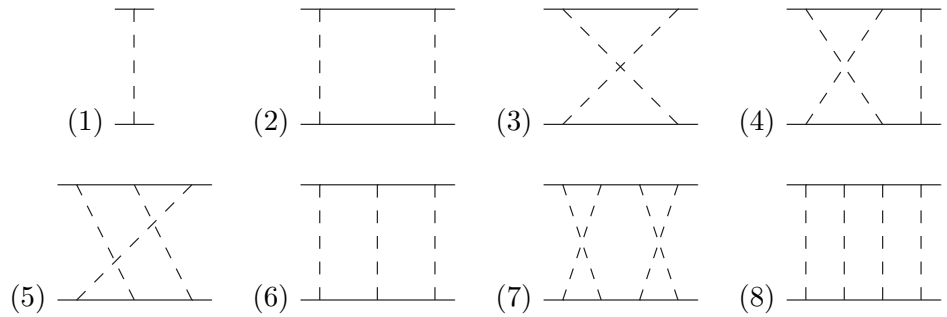

Figure 3.2: Reducible and irreducible interaction terms in the BSE for the two-fermion four-point Green function

others irreducible. The irreducible ones constitute the irreducible interaction kernel, iteration of which will reproduce all diagrams including the reducible ones.

In this way, the BSE for G becomes a sum of kernel iterations depicted in Fig. 3.3 which can be summed up completely to yield the BSE for G given in

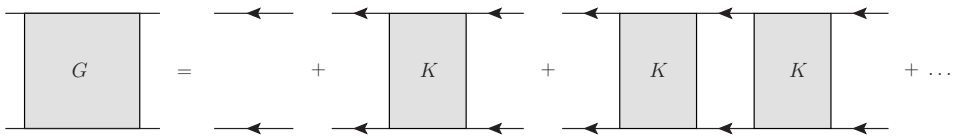

Figure 3.3: Iteration of the irreducible interaction kernel towards the BSE for the two-fermion four-point Green function

Fig. 3.4 and written as an inhomogeneous linear integral equation in Eq. (3.4)

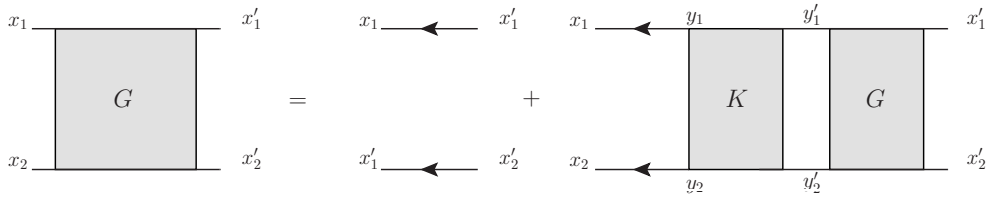

Figure 3.4: The BSE for the two-fermion four-point Green function

$$G(x_1,x_2;x_1',x_2') = -S_{(1)}(x_1,x_1')S_{(2)}(x_2,x_2') + i\int d^4y_1\; d^4y_2\; d^4y_1'\; d^4y_2'$$
$$\times S_{(1)}(x_1,y_1)S_{(2)}(x_2,y_2)\; K(y_1,y_2;y_1',y_2')\; G(y_1',y_2';x_1',x_2') \tag{3.4}$$

In the following we will write G and S in terms of matrix elements of Heisenberg operators $\psi^H_{(j)}(x)$:

$$G(x_1,x_2;x_1',x_2') = \langle 0|T[\psi^H_{(1)}(x_1)\psi^H_{(2)}(x_2)\bar{\psi}^H_{(1)}(x_1')\bar{\psi}^H_{(2)}(x_2')]|0\rangle \tag{3.5}$$
$$S_{(j)}(x,x') = \langle 0|T[\psi^H_{(j)}(x)\bar{\psi}^H_{(j)}(x')]|0\rangle \tag{3.6}$$

3.2 Momentum Space

For actual calculations, it is more convenient to work in momentum space. Therefore we now rewrite the BSE to momentum space. The first step is to define relative coordinates $x := x_1 - x_2$ and $x' := x_1' - x_2'$. This is a good choice due to the translational-invariance requirement put on the theory. The definition of a "center of mass" is arbitrary. One may choose $X := \eta x_1 + (1-\eta)x_2$ and analogously $X := \eta x_1' + (1-\eta)x_2'$; this means $x_1 = X + (1-\eta)x$ and $x_2 = X - \eta x$. The conjugate field momenta are, via Fourier transformation, translational invariance and total-momentum conservation,

$$p = (1-\eta)p_1 - \eta p_2 \tag{3.7}$$
$$P = p_1 + p_2 \quad \text{and} \tag{3.8}$$
$$p_1 = \eta P + p \tag{3.9}$$
$$p_2 = (1-\eta)P - p \tag{3.10}$$

P is the total momentum of the two fermions, p their relative momentum, and p_j their individual momenta. The Fourier transformed quantities in the BSE can be obtained via the relations

$$\begin{aligned} G(x_1,x_2;x_1',x_2') &= \frac{1}{(2\pi)^8}\int d^4P\; d^4q\; d^4q'\; e^{-iP\cdot(X-X')-iq\cdot x+iq'\cdot x'}G(P;q;q') \\ K(x_1,x_2;x_1',x_2') &= \frac{1}{(2\pi)^8}\int d^4P\; d^4q\; d^4q'\; e^{-iP\cdot(X-X')-iq\cdot x+iq'\cdot x'}K(P;q;q') \\ S_{(j)}(x_j,x_j') &= \frac{1}{(2\pi)^4}\int d^4p_j\; e^{-ip_j\cdot(x_j-x_j')}S_{(j)}(p_j) \end{aligned}$$

The BSE becomes

$$\begin{aligned} G(P;q;q') = -S_{(1)}(\eta P+q)S_{(2)}((1-\eta)P-q)\;\delta^4(q-q') \\ +iS_{(1)}(\eta P+q)S_{(2)}((1-\eta)P-q)\int d^4q''\; K(P;q;q')\; G(P;q'';q') \end{aligned} \tag{3.11}$$

This completes the discussion of the Bethe-Salpeter equation for the two-fermion four-point Green function.

3.3 The Homogeneous BSE

We now approach the problem of finding a bound state for a given system of two particles. In the following, we will start from the 4-point Green function of Eq. (3.5) and investigate what it means to find a bound state solution, which leads to the definition of an amplitude (corresponding to a wave function in quantum mechanics except for the missing probability interpretation), the Bethe-Salpeter amplitude. Once this object is defined and translated to momentum space, we'll find the homogeneous BSE that this amplitude satisfies, which is commonly used to study bound states in quantum field theory.

We start from G as given in Eq. (3.5) and let $x_1^0, x_2^0 > x_1^{0\prime}, x_2^{0\prime}$. Then one can split the time-ordered product in the definition of G and write

$$G(x_1,x_2;x_1',x_2') = \langle 0|T[\psi^H_{(1)}(x_1)\psi^H_{(2)}(x_2)]T[\bar\psi^H_{(1)}(x_1')\bar\psi^H_{(2)}(x_2')]|0\rangle \tag{3.12}$$

(we'll drop the superscript ^{H} in the following for simplicity); this assumption is reasonable, if one wants to have an intermediate state of two fermions. Now we insert a complete set of states $|P,\alpha\rangle$ with P the total momentum of the state and α all other quantum numbers. Then we get (for degenerate states, one has to also sum over the degeneracy)

$$G(x_1,x_2;x_1',x_2') = \sum_{P,\alpha}\langle 0|T[\psi^H_{(1)}(x_1)\psi^H_{(2)}(x_2)]|P,\alpha\rangle\langle P,\alpha|T[\bar\psi^H_{(1)}(x_1')\bar\psi^H_{(2)}(x_2')]|0\rangle \tag{3.13}$$

The contribution from a particular bound state B is

$$\sum_P \langle 0|T[\psi^H_{(1)}(x_1)\psi^H_{(2)}(x_2)]|P,\alpha_B\rangle\langle P,\alpha_B|T[\bar\psi^H_{(1)}(x_1')\bar\psi^H_{(2)}(x_2')]|0\rangle \tag{3.14}$$

where, since the on-mass-shell condition $P^2 = M^2$ holds for the bound state with mass M, the sum over momentum states can be written as the integral over the mass shell in momentum space in the form

$$\sum_P \ldots = \int d^4P \; \vartheta(P_0)\delta(P^2 - M^2) \ldots \tag{3.15}$$

Now define an amplitude, which is, somewhat misleadingly, called the *Bethe-Salpeter wave function* as the following matrix element

$$\chi_B(x_1, x_2) := \langle 0|T[\psi^H_{(1)}(x_1)\psi^H_{(2)}(x_2)]|P, \alpha_B\rangle \tag{3.16}$$

and its conjugate as

$$\bar{\chi}_B(x_1, x_2) := -\langle P, \alpha_B|T[\bar{\psi}^H_{(1)}(x_1')\bar{\psi}^H_{(2)}(x_2')]|0\rangle \tag{3.17}$$

With the previous definitions of X and x and with the transformation properties of the field operators $\psi(x)$ under translations this becomes

$$\chi_B(x_1, x_2) \;=:\; \chi_B(P; x)e^{-iP\cdot X} \tag{3.18}$$

$$\bar{\chi}_B(x_1^{(\prime)}, x_2^{(\prime)}) \;=:\; \bar{\chi}_B(P; x^{(\prime)})e^{iP\cdot X^{(\prime)}} \tag{3.19}$$

The corresponding momentum-space objects are obtained via

$$\chi_B(P; x) \;=\; \int d^4q \; \chi_B(P; q)e^{-iq\cdot x} \tag{3.20}$$

$$\bar{\chi}_B(P; x^{(\prime)}) \;=\; \int d^4q \; \bar{\chi}_B(P; q)e^{iq\cdot x^{(\prime)}} \tag{3.21}$$

Now putting all this together into the expression for G_B one has after Fourier transformation

$$G(P; q; q') = -i\frac{[\chi_B(P; q)\bar{\chi}_B(P; q')]|_{P \text{ on shell}}}{P^2 - M^2} + \text{terms regular in } P^2 \tag{3.22}$$

where "P on shell" means $P^2 = M^2$ and $P_0 \geq 0$. Now we insert this into the integral equation for G, Eq. (3.11), and compare the residues at the pole $P^2 = M^2$, i.e., we multiply the equation by $(P^2 - M^2)$ and afterwards take the limit $P^2 \to M^2$. This immediately yields

$$\chi_B(P; q) = iS_{(1)}(\eta P + q) \; S_{(2)}((1-\eta)P - q) \int d^4q' \; K(P; q; q') \; \chi_B(P; q') \tag{3.23}$$

The diagrammatical representation of this equation is given in Fig. 3.5. Multiplying by $S^{-1}_{(1)}(\eta P + q) \; S^{-1}_{(2)}((1-\eta)P - q)$ and setting

$$\chi_B(P; q) =: S_{(1)}(\eta P + q) \; \Gamma_B(P; q) \; S_{(2)}((1-\eta)P - q) \tag{3.24}$$

one obtains the homogeneous BSE for the *Bethe-Salpeter amplitude* $\Gamma_B(P; q)$

$$\Gamma_B(P; q) = i \int d^4q' \; K(P; q; q') \; S_{(1)}(\eta P + q')\Gamma_B(P; q')S_{(2)}((1-\eta)P - q') \tag{3.25}$$

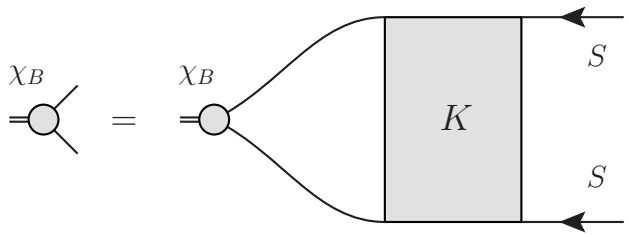

Figure 3.5: The homogeneous BSE for the Bethe-Salpeter wave function χ_B

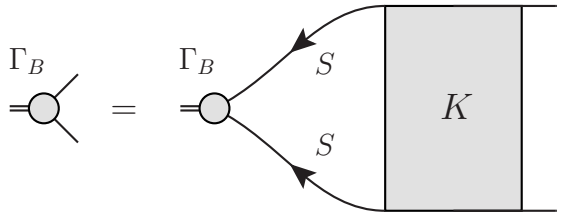

Figure 3.6: The homogeneous BSE for the Bethe-Salpeter amplitude Γ_B

In diagrammatical representation, this equation is given in Fig. 3.6

At this point, a comment on fermion-(anti)fermion systems is in order: in this case the propagators S and the amplitudes χ and Γ are represented by 4×4-matrices in Dirac space. This can be made explicit by adding Dirac indices (we'll use $a, b, c, \ldots$ in the following) to the equations, e. g., the homogeneous BSE for Γ in (3.25):

$$[\Gamma(P;q)]_{ab} = i\int d^4q' \; [K(P;q;q')]^{cd}_{ab} \; [\chi(P;q')]_{cd} \tag{3.26}$$

and for (3.24)

$$[\chi(P;q)]_{ab} =: [S_{(1)}(\eta P+q)]_{ac}[\Gamma(P;q)]_{cd}[S_{(2)}((1-\eta)P-q)]_{db}\,. \tag{3.27}$$

In general, these indices can be omitted if one takes care to put the terms in an expression in the correct order as one would need them for the implied (Dirac) matrix multiplications. While we omit explicit Dirac indices in the derivation of the canonical normalization condition below for more clarity, it should be noted that the order is not kept and only given explicitly at the very end of the next section.

3.4 The Canonical Normalization Condition

Equations (3.23) and (3.25) are homogeneous, linear integral equations. The solution in each case is only determined up to a constant normalization factor. The corresponding normalization condition is derived from the inhomogeneous

BSE for G; the associated norm is referred to as the canonical norm of the amplitude. Such a condition can be derived for both χ_B and Γ_B. We'll concentrate on the former.

First, rewrite the integral equation for G using the notation

$$D(P;q;q') := \delta^4(q-q')\left[S_{(1)}(\eta P+q)\; S_{(2)}((1-\eta)P-q)\right]^{-1} \tag{3.28}$$

and without writing arguments. Then one has

$$DG = \mathbf{1} + iKG\;. \tag{3.29}$$

Formally, the solution of this equation is

$$(D-iK)G = \mathbf{1} \quad \Rightarrow \quad G = (D-iK)^{-1}\;. \tag{3.30}$$

With the help of the relation

$$\frac{\partial}{\partial\lambda}A^{-1} = -A^{-1}\left(\frac{\partial}{\partial\lambda}A\right)A^{-1}$$

which is valid for any invertible operator A, one can get the following relation by differentiating the above expression for G by λ, which we'll later replace by P^μ:

$$\begin{aligned}\frac{\partial}{\partial\lambda}G &= -(D-iK)^{-1}\left[\frac{\partial}{\partial\lambda}D - i\frac{\partial}{\partial\lambda}K\right](D-iK)^{-1} \\ &= -G\left[\frac{\partial}{\partial\lambda}D - i\frac{\partial}{\partial\lambda}K\right]G\;.\end{aligned} \tag{3.31}$$

Now again, we use pole terms for G and equate the corresponding residues. This yields

$$i\bar{\chi}_B\;\left[\frac{\partial}{\partial\lambda}D - i\frac{\partial}{\partial\lambda}K\right]\;\chi_B = \frac{\partial(M^2)}{\partial\lambda} \tag{3.32}$$

Now we replace $\lambda \to P^\mu$ and write everything explicitly with momentum-space arguments as the canonical normalization condition for χ_B:

$$\begin{aligned}2P^\mu &= i\int d^4q\;\bar{\chi}_B(P;q)\left[\frac{\partial}{\partial P^\mu}\left(S_{(1)}(\eta P+q)\; S_{(2)}((1-\eta)P-q)\right)^{-1}\right]\chi_B(P;q) \\ &+ \int d^4q\;d^4q'\;\bar{\chi}_B(P;q)\left[\frac{\partial}{\partial P^\mu}K(P;q;q')\right]\chi_B(P;q)\end{aligned} \tag{3.33}$$

An analogous condition exists for $\Gamma_B(P;q)$, which can be easily derived from Eq. (3.33). One can see that there are two terms; one involving a derivative of (the inverse of) the propagators $S_{(1)}$ and $S_{(2)}$ and another involving a derivative of the kernel K with respect to the total momentum. In a simple enough interaction, it is possible that the kernel doesn't depend on the total momentum and only the first term remains nonzero.

To make the case of two fermion constituents clearer once again, we repeat the normalization condition here and put all terms in the correct order with respect to matrix multiplication in Dirac space. In this case, the calculation

also involves the Dirac trace over the expressions on the right hand side of Eq. (3.33) and one obtains, writing $q_+ := \eta P + q$ and $q_- := (1-\eta)P - q$

$$\begin{aligned} 2P^\mu &= i\mathrm{Tr}\left[\int d^4q\ \bar{\chi}(P;q)\frac{\partial S^{-1}_{(1)}(q_+)}{\partial P^\mu}\ \chi_B(P;q)S^{-1}_{(2)}(q_-)\right] \\ &+ i\mathrm{Tr}\left[\int d^4q\ \bar{\chi}(P;q)\ S^{-1}_{(1)}(q_+)\chi_B(P;q)\frac{\partial S^{-1}_{(2)}(q_-)}{\partial P^\mu}\right] \\ &+ \mathrm{Tr}\left[\int d^4q\ d^4q'\ \bar{\chi}_B(P;q)\frac{\partial K(P;q;q')}{\partial P^\mu}\chi_B(P;q)\right] \end{aligned} \tag{3.34}$$

3.5 The Inhomogeneous Vertex BSE

In a quantum field theory one can also investigate the properties of general vertices (three-point Green functions). As an example we write a vertex $\Gamma_x(P;q)$ where the subscript x is there to remind us that this object is restricted to certain properties (quantum numbers). These vertex functions in an interacting theory satisfy an inhomogeneous BSE analogous to the homogeneous BSE for Γ. It reads, e.g.,

$$\Gamma_x(P;q) = i\Gamma^0_x + \int d^4q'\ K(P;q;q')\ S_{(1)}(\eta P + q')\Gamma_x(P;q')S_{(2)}((1-\eta)P - q') \tag{3.35}$$

In diagrammatical representation, this equation is given in Fig. 3.7. The in-

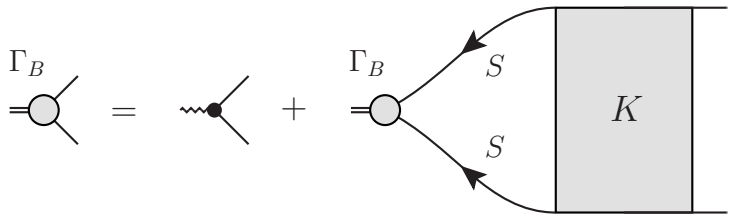

Figure 3.7: The inhomogeneous BSE for the vertex function Γ_x

homogeneous term here essentially represents the corresponding current *before* the interactions of the theory. We will discuss possible solution strategies for this and the homogeneous BSE for Γ later. At this point it is interesting to note that, similar to the 4-point function G, the pseudoscalar vertex has a pole at a bound state's mass of the system, but only if this bound state has the quantum numbers specified by x. G on the other hand contains poles of all kinds and quantum numbers. The pole structure in the case here can be written as

$$\Gamma_x(P;q)|_{P^2\cong M^2_{ps}} = \frac{r_x}{P^2 - M^2_x}\Gamma^B_x(P;q) + \text{ regular terms at this pole} \tag{3.36}$$

where M_x is the pseudoscalar bound state's mass and $\Gamma^B_x(P;q)$ canonically normalized. Inserting this into the Eq. (3.35) and once more equating residues yields the corresponding homogeneous BSE.

3.6 Solution Strategies for the Homogeneous and the Inhomogeneous BSEs

Both (vertex) BSE types, Eqs. (3.25) and (3.35) are integral equations of the second Fredholm kind. More precisely, this means

- linear integral equation
- constant integration boundaries (Fredholm)
- solution appears both under the integral and outside (second kind)

In principle, analytic solutions are possible. In practice, e.g. in QCD, we can essentially forget about that, which means that solutions will in general be found numerically. While these notes do not deal with numerical methods at all (please refer to the literature to learn about the numerical details of iterative, inversion, or other methods), the basic ideas and strategies for a solution are discussed in the following.

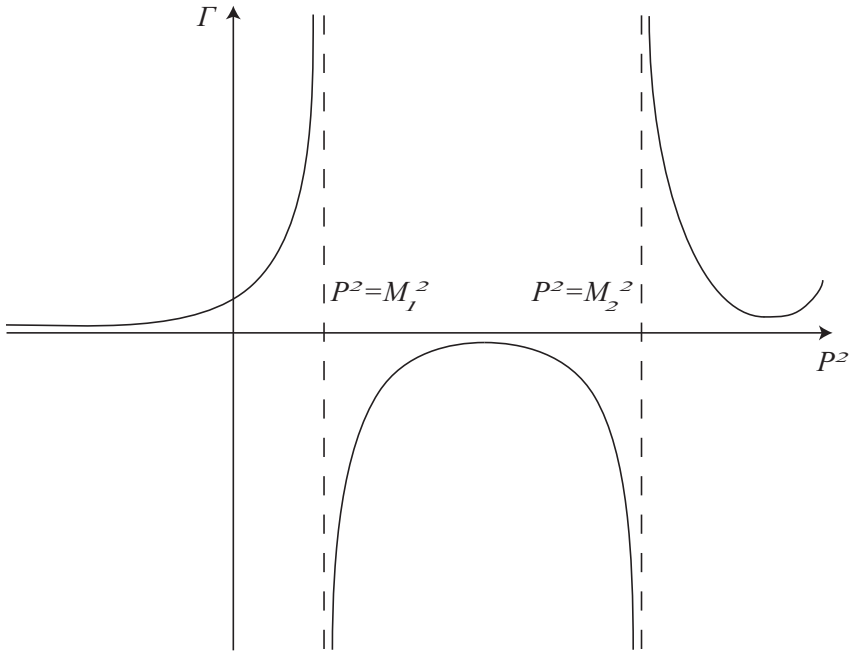

Figure 3.8: The form of the solution of the inhomogeneous BSE for the pseudoscalar vertex function Γ_5

Let's start with the inhomogeneous BSE for Γ_x and analyze the solution in terms of its dependence on the momentum variables involved. With two independent four-momenta available, the total momentum P and the relative momentum q, one has three Lorentz-invariant variables, which can be constructed as the Lorentz-scalar products obtainable from P and q, namely P^2, q^2, and $q \cdot P$. In terms of finding a bound state's mass M, it is interesting to study the solution primarily as a function of P^2, since the solution will show

a pole behavior at some $P^2 = M^2$, as seen in Eq. (3.36). For this purpose, one sets the other two variables to some particular values, e.g. $q^2 = q \cdot P = 0$ and obtains a function of P^2, whose pole structures can be easily spotted upon inspection, see Fig. 3.8. To actually numerically find the solutions precisely enough, it is not ideal to extract pole positions of $\Gamma_x(P^2)$, but it is advisable to investigate $1/\Gamma_x(P^2)$ and find zeros of this function. This procedure is illustrated in Fig. 3.9. The pole with the smallest M corresponds to the ground

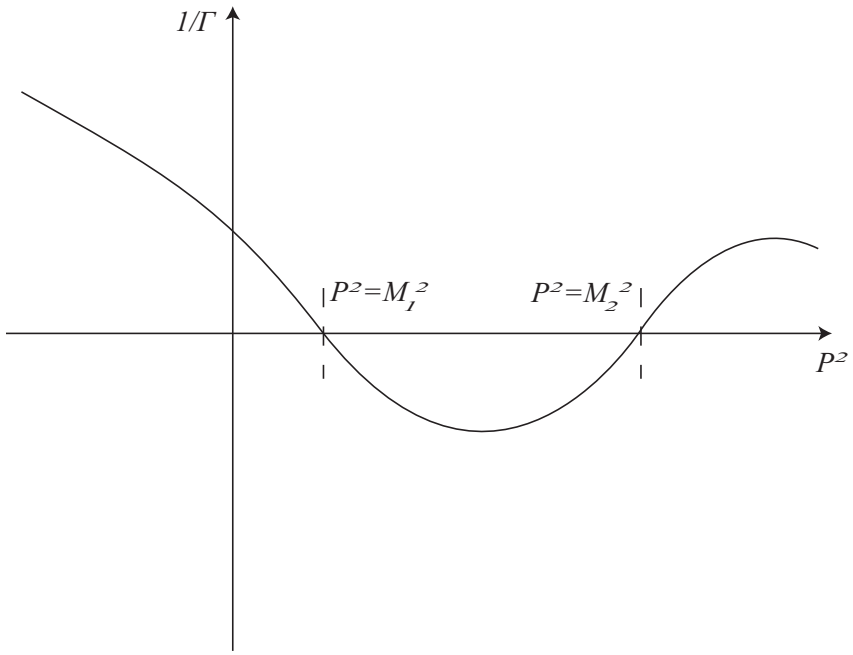

Figure 3.9: Illustration of the solution strategy for the inhomogeneous BSE for the pseudoscalar vertex function Γ_5

state M_1, the next to the first excited state with M_2, etc. Note that both figures depict the situation in Minkowski space; in Euclidean space, where a solution is found at $P^2 = -M^2$, the situation is mirrored with respect to the vertical axis. Note furthermore that for a pole

$$\Gamma(P^2) = \frac{r}{P^2 - M^2}$$

one has

$$\frac{1}{\Gamma(P^2)} = \frac{P^2 - M^2}{r}$$

and $\frac{\partial}{\partial(P^2)}$ yields

$$\left(\frac{1}{\Gamma}\right)'\Big|_{(P^2=M^2)} = \frac{1}{r}.$$

In this way, one can extract both the pole position and residue.

The equation as such can be solved by iterative methods. However, due to the structure of the equation and due to the failure of the iterative method

close to and beyond the first pole structure, a better way is to rewrite the equation in the following way. Dropping all arguments, defining the formal kernel $\tilde{K} := KS_{(1)}S_{(2)}$, and writing the driving term as D, Eq. (3.35) becomes

$$\Gamma = D + \tilde{K}\,\Gamma\,. \tag{3.37}$$

The formal solution of this equation is

$$\Gamma = (\mathbf{1} - \tilde{K})^{-1} D \tag{3.38}$$

which can be implemented numerically via matrix-inversion methods and provides a stable solution for all values of P^2.

Now, consider the homogeneous BSE, Eq. (3.25) and rewrite it formally using the formal kernel $\tilde{K}$ once again as

$$\Gamma = \tilde{K}\,\Gamma\,. \tag{3.39}$$

The problem with this equation is that, in contrast to the inhomogeneous BSE, where all three Lorentz-invariant variables are actually independent, the homogeneous equation is only valid on-shell i.e. $P^2 = M^2$ is fixed a priori, but actually an unknown parameter. In other words, to numerically find the solution at $P^2 = M^2$, one has to already know the value of M^2. How can one bypass this problem?

Upon closer inspection, Eq. (3.39) is an eigenvalue equation with an eigenvalue equal to one. As a result, the solution method can make use of the eigenvalue concept via the introduction of the artificial and P^2-dependent eigenvalue $\lambda(P^2)$ in the BSE to obtain

$$\lambda(P^2)\,\Gamma = \tilde{K}\,\Gamma\,. \tag{3.40}$$

Now, for any given value of P^2, one can solve this eigenvalue equation and obtain a spectrum of eigenvalues λ_1, λ_2, ... Doing this for several values of P^2 yields each λ_i as a function of P^2. Then the only step left is to find those values of P^2, where $\lambda_i(P^2) = 1$ for each i and thereby obtain the bound-state masses M_i of the system, which of course coincide with those found from the corresponding inhomogeneous BSE. This procedure is illustrated in Fig. 3.10

Together with the mass M one also obtains the homogeneous BSA Γ, which can be canonically normalized and subsequently used to compute further properties of the bound state, such as decay properties, various bound-state form factors or couplings in various scattering processes.

In the following we will look at a concrete application of this formalism and start with a discussion of mesons in the context of quantum chromodynamics (QCD).

3.7 General Remarks on mesons

Mesons are not baryons. Nor are they leptons. Apart from the historical "in-between" situation regarding baryon, meson, and lepton mass, this definition

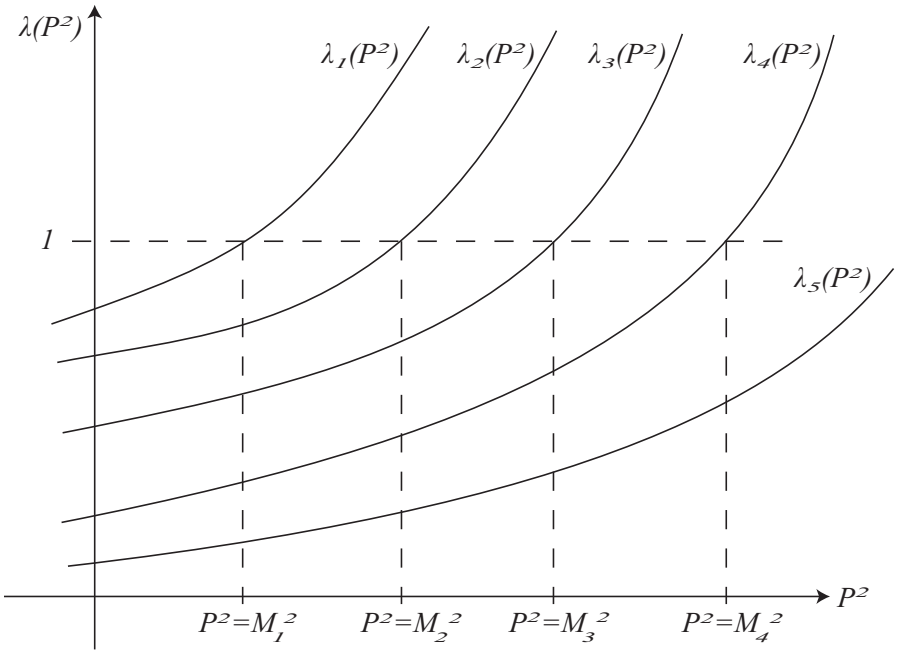

Figure 3.10: Illustration of the solution strategy for the homogeneous BSE

is still precise. Mesons have neither baryon nor lepton number. A conserved quantity such as a "meson number" doesn't seem to exist, i.e. neutral mesons can be produced in high-energy experiments in large numbers. This high abundance makes especially the very light pion a very interesting and important object to study.

Mesons are *hadrons*, i.e. they are subject to the strong interaction. This interaction is responsible for both the inner structure of hadrons as bound states of quarks and gluons in the context of QCD as well as the interaction of hadrons among each other (e.g. nucleon-nucleon interaction via pion exchange).

Mesons come in various kinds. With respect to the Casimir operators of the Poincaré group we note that mesons have mass $\neq 0$ and integer spin $j = 0, 1, 2, \ldots$. Other than that, quark *flavor* is a key to the systematics of the meson spectrum, which is implemented in the *quark model*. In its most simple form, the quark model explains the quantum numbers of mesons (and analogously baryons) via constructing a meson state out of a quark and an antiquark, while a baryon is made out of three quarks. In this successful concept, (anti)quarks carry charges $\pm 2/3$ or $\pm 1/3$ times the elementary charge, which yields integer charges for both mesons and baryons. Furthermore, quarks carry baryon number $+1/3$, antiquarks $-1/3$. For example, the up-quark u has charge $+2/3$, the down-quark d $-1/3$, such that a proton p^+ is given by the quark content uud, a neutron n^0 by udd, and a positively charged pion π^+ by $u\bar{d}$. For more information on flavors, a "complete" but ever-changing list of quark- and meson- as well as baryon-properties, see the website and publications of the Particle-Data Group at `http://pdg.lbl.gov`.

On further notes, quarks (and gluons) cannot be observed individually, i.e. no single quark or gluon has ever reached a detector. They occur only inside hadrons, a property which is called *confinement* and is a result of the underlying quantum field theory of the strong interaction, quantum chromodynamics (QCD), which is a non-Abelian gauge theory. While the confinement picture is modified under extreme conditions like high temperatures and/or densities, for our purposes it makes sense to consider mesons as quark-antiquark bound states in QCD via the Bethe-Salpeter equation. To do this, we have to construct the Bethe-Salpeter amplitude as the bound-state amplitude of a fermion-antifermion system. This is accomplished in the following.

3.8 Structure of the Meson BSA

Let us recall that a meson's BSA $\Gamma(P;q)$ depends on two independent momenta, namely the total momentum P and the relative momentum q of the meson's constituents. Together with the γ matrices discussed above one has three four-vectors to build the amplitude: γ^μ, q^μ, and P^μ. To understand the details of the construction, we must characterize each meson by a set of quantum numbers. We have J^{PC} with the total angular momentum J, the parity P, and the C-parity C, which is a parity quantum number for a state under charge conjugation. Note that C is only a good quantum number, if the meson is a bound state of equal-mass and flavor-antiflavor constituents (i.e. it is its own antiparticle). In the general case of unequal constituent masses one therefore has simply J^P.

Now look at all possible scalar products of the three four-vectors just mentioned, but with the idea in mind that we would like to use them as basis elements in Dirac space. We find that q^2, P^2, $\gamma_\mu\gamma^\mu$, and $q \cdot P$ are all proportional to $\mathbf{1}$ in Dirac space, i.e., as soon as we have $\mathbf{1}$ as a basis element, we can ignore these. The remaining possibilities are $q \cdot \gamma =: \not{q}$ and $P \cdot \gamma =: \not{P}$. All of these products are, of course, Lorentz scalar or pseudoscalar by construction. In addition, we know that all vectors involved are actually polar vectors, so all of these products are *scalar* and not pseudoscalar: they have parity $+$.

The next interesting quantity is C-parity. To investigate this, we have to keep in mind that a charge conjugation inverts all internal quantum numbers, including internal momenta (it corresponds to making an antiparticle out of a particle). More precisely, e.g. charge and flavor are inverted, but also the relative (internal) momentum $q \to -q$. However, the total momentum P remains unchanged. The charge conjugation of an amplitude Γ is denoted by a bar over the Γ and given by the following prescription

$$\Gamma(P;q) \quad \longrightarrow \quad \bar{\Gamma}(P;q) = \left(C\,\Gamma(P;-q)\,C^{-1}\right)^T \tag{3.41}$$

In this way and with Eq. (2.66) we find that $\mathbf{1}$ and $\not{q}$ have positive C-parity, while $\not{P}$ has negative C-parity. With this information we can now investigate and analyze the properties of any amplitude, best by example.

For the sake of simplicity let us start with a meson with spin zero, i.e. $J = 0$. There are two possibilities for parity, which make the meson either

a *scalar* (positive parity) or a *pseudoscalar* (negative parity). We begin with scalars and ask ourselves the question, what possibilities there are to construct an amplitude Γ which corresponds to the quantum numbers $J^P = 0^+$ (we'll worry about positive C-parity later, as soon as we have a construction).

To know what we are doing, let us first discuss the dependence on the variables once more, but in more detail. In terms of the four-vectors γ^μ, q^μ, and P^μ we realize that only γ^μ introduces a Dirac structure different from $\mathbf{1}$. Generally, we need to find all possible and linearly independent Dirac structures obtainable from the vectors given here. Once these structures, called *Dirac covariants*, are found we know that any function multiplying them can only depend on scalar products of the four vectors q^μ, and P^μ, since the Dirac part is already taken care of. In other words, for $J = 0$ one can write the BSA as a sum over Dirac covariants T_i multiplied by Lorentz-invariant functions F_i in the way

$$\Gamma(P;q) = \sum_{i=1}^{N_J} T_i(\gamma;P;q)\ F_i(P;q) = \sum_{i=1}^{N_J} T_i(\gamma;P;q)\ F_i(P^2,q^2,q\cdot P)\ , \quad (3.42)$$

where the number of covariants N_J depends on the spin J of the meson and $N_J = 4$ for $J = 0$ ($N_J = 8$ for all other values of J – we'll see below where that comes from).

Now we can attack the question what the covariants for $J^P = 0^+$ are. Since we want scalars, we can use scalar products of the four-vectors at hand. In addition we'll be able to use products of those scalar products. The goal is to arrive at a linearly independent and complete set of covariants. We will demonstrate this construction explicitly for the scalar case and do more, namely construct an orthogonal basis in Dirac space. The starting point are, as already mentioned, the four-vectors at hand, γ^μ, q^μ, and P^μ. In particular, we have already shown that a number of the scalar products constructed from this set are proportional to $\mathbf{1}$, so this is our first basis element. To see linear independence or orthogonality, we actually need a scalar product, which in this case is defined for two matrix-valued structures A and B as the trace of the matrix product of A and B:

$$\langle A|B\rangle := \mathrm{Tr}(A\cdot B) \quad (3.43)$$

The remaining two scalar products from our set of four-vectors are $\not{q}$ and $\not{P}$. Now let us check orthogonality in these cases.

$$\mathrm{Tr}(\mathbf{1}\cdot\not{q}) = \mathrm{Tr}(\not{q}) = 0 = \mathrm{Tr}(\not{P}) = \mathrm{Tr}(\mathbf{1}\cdot\not{P})$$

so both $\not{q}$ and $\not{P}$ are orthogonal to $\mathbf{1}$, but are they also orthogonal to each other? Obviously not, since

$$\mathrm{Tr}(\not{q}\cdot\not{P}) = 4q\cdot P\ , \quad (3.44)$$

which can be proven using Eq. (2.47) and the fact that cyclic permutation of all factors inside the trace doesn't change the value of the trace. However, these two terms can be made orthogonal by a small modification of one of them, namely

$$\not{q} \longrightarrow \not{q} - \not{P}\frac{q\cdot P}{P^2}\ ,$$

such that

$$\mathrm{Tr}\left((\not{q} - \not{P}\frac{q\cdot P}{P^2})\cdot\not{P}\right) = 0\,. \tag{3.45}$$

It is interesting to note at this point that the construction just introduced is an example for a *transversal projection* with respect to the total momentum P. In general one writes, denoting transversality by the superscript ^{T}, for an arbitrary four-vector V^μ

$$V_\mu^T := V_\mu - P_\mu\frac{V\cdot P}{P^2}\,. \tag{3.46}$$

This will be useful below when we discuss BSAs of mesons with higher spin.

So far, so good. But have we yet found all linearly independent basis elements? Since we have already used all possible scalar products that our four-vectors permit us to construct, the only possibility to find more basis elements is to investigate products of these scalar products that are linearly independent with respect to all the basis elements already known. Obviously, multiplying $\mathbf{1}$ with either $\not{q}$ or $\not{P}$ doesn't bring anything new. What about $\not{q}\not{q}$? A short calculation shows that

$$\not{q}\not{q} = q^2\,\mathbf{1}\,, \tag{3.47}$$

the analogous result holds for $\not{P}\not{P}$. So those don't bring anything new, either. The next possibility is $\not{q}\not{P}$, which *is* new, i.e., it is linearly independent with respect to the other three. We'll prove this by showing that, actually, the construction $\not{q}\not{P} - \not{P}\not{q} = [\not{q},\not{P}]$ is orthogonal to the three elements that we already have. Indeed, we find

$$\mathrm{Tr}\,(\mathbf{1}\cdot[\not{q},\not{P}]) = 0 \tag{3.48}$$

$$\mathrm{Tr}\left((\not{q} - \not{P}\frac{q\cdot P}{P^2})\cdot[\not{q},\not{P}]\right) = 0 \tag{3.49}$$

$$\mathrm{Tr}\,(\not{P}\cdot[\not{q},\not{P}]) = 0\,. \tag{3.50}$$

The remaining question is whether or not any further linearly independent covariants can be found. To clarify that the answer is no, it is enough to argue that any products of $[\not{q},\not{P}]$ with $\not{q}$ or $\not{P}$ or itself can be reduced to already known covariants via the Clifford algebra, Eq. (2.47). As a result, the basis in Dirac space for a scalar meson with $J^P = 0^+$ can be given by

$$\begin{array}{ll} T_1 = \mathbf{1} & T_2 = i\not{P} \\ T_3 = i\not{q} & T_4 = [\not{q},\not{P}] \end{array} \tag{3.51}$$

Before we determine the C-parity of these covariants and discuss the C-parity of the corresponding amplitude Γ, we note that a pseudoscalar-meson set of covariants can be constructed from Eq. (3.51) by multiplication of each covariant with a factor γ_5, since γ_5 has negative parity and in this way one exactly gets pseudoscalar transformation properties of each covariant. The result is

$$\begin{array}{ll} T_1 = i\,\gamma_5 & T_2 = \gamma_5\not{P} \\ T_3 = \gamma_5\not{q} & T_4 = i\,\gamma_5\,[\not{q},\not{P}] \end{array} \tag{3.52}$$

Now let us investigate the C-parity of each of the covariants in Eqs. (3.51) and (3.52). Trivially, the C-parity of $\mathbf{1}$ is positive. For $\not{q}$ we have a combination of

$$q^\mu \to -q^\mu \quad \text{and} \quad (C\gamma^\mu C^{-1})^T = -\gamma^\mu$$

under C such that the C-parity of their product is positive. Since P^μ doesn't change its sign under charge conjugation, $\not{P}$ has negative C-parity. To see that $[\not{q},\not{P}]$ has positive C-parity, we explicitly evaluate

$$\begin{aligned}(C[\not{q},\not{P}]C^{-1})^T &= (C(\not{q}\not{P}-\not{P}\not{q})C^{-1})^T = (C\not{q}\not{P}C^{-1})^T - (C\not{P}\not{q}C^{-1})^T \\ &= (C\not{q}C^{-1}C\not{P}C^{-1})^T - (C\not{P}C^{-1}C\not{q}C^{-1})^T \\ &= (-\not{P}\not{q}) - (-\not{q}\not{P}) = [\not{q},\not{P}] \end{aligned} \tag{3.53}$$

Now that we know all C-parities for the covariants in the scalar case, Eq. (3.51), we can put everything together and go back to Eq. (3.42): Assuming that the total amplitude for a scalar meson with $J^{PC} = 0^{++}$ is demanded to have positive C-parity, with the C-parities of T_1, T_2, T_3, and T_4 being $+$, $-$, $+$, and $+$, one must have C-parities $+$, $-$, $+$, and $+$ as well for F_1, F_2, F_3, and F_4. But how can we guarantee this? Remember that the variables that the F_i depend on are P^2 (which has positive C-parity), q^2 (which has positive C-parity), and $q \cdot P$ (which has negative C-parity). So to get a function F_i with even (odd) C-parity one needs to ensure that F_i is a function symmetric (antisymmetric) in $q \cdot P$. In this way one can construct amplitudes with arbitrary parity and C-parity.

Let us quickly do the same analysis for the pseudoscalar covariants in Eq. (3.52). There γ_5 has positive C-parity, which is also important to know for the following analysis. Next, one has

$$(C\gamma_5\not{P}C^{-1})^T = (C\gamma_5 C^{-1}C\not{P}C^{-1})^T = -\not{P}\gamma_5 = \gamma_5\not{P} \tag{3.54}$$

and, analogously

$$(C\gamma_5\not{q}C^{-1})^T = (C\gamma_5 C^{-1}C\not{q}C^{-1})^T = \not{q}\gamma_5 = -\gamma_5\not{P}\,, \tag{3.55}$$

since γ_5 anticommutes with γ^μ. Finally,

$$(C\gamma_5[\not{q},\not{P}]C^{-1})^T = \gamma_5[\not{q},\not{P}]\,. \tag{3.56}$$

Thus, in order to have a pseudoscalar meson with $J^{PC} = 0^{-+}$ one needs C-parities $+$, $+$, $-$, and $+$ as well for F_1, F_2, F_3, and F_4.

The next task is to investigate the possibilities to construct a Dirac basis for the BSA of a meson with spin $J = 1$. In order to achieve this, we must create an object that transforms like a four-vector (or axialvector) under Lorentz transformations. Since we already have constructed a Lorentz-scalar amplitude with four structures, we can construct a Lorentz-vector basis by multiplying those four structures with the three four-vectors that are at hand. More precisely one thereby finds the 12 covariants

$$\begin{array}{llll} T_1^\mu = \gamma^\mu & T_2^\mu = \gamma^\mu\,\not{q} & T_3^\mu = \gamma^\mu\,\not{P} & T_4^\mu = \gamma^\mu\,[\not{q},\not{P}] \\ T_5^\mu = q^\mu\mathbf{1} & T_6^\mu = q^\mu\,\not{q} & T_7^\mu = q^\mu\,\not{P} & T_8^\mu = q^\mu\,[\not{q},\not{P}] \\ T_9^\mu = P^\mu\mathbf{1} & T_{10}^\mu = P^\mu\,\not{q} & T_{11}^\mu = P^\mu\,\not{P} & T_{12}^\mu = P^\mu\,[\not{q},\not{P}] \end{array} \tag{3.57}$$

At this point, we have to implement a property of massive spin-1 particles, namely that they are transverse, i.e. there is a set of polarization vectors $\varepsilon^\mu(\lambda, P)$, $\lambda = 1, 2, 3$ such that $P^\mu \varepsilon_\mu(\lambda, P) = 0$. In other words, the 12 covariants just defined have to be transversely projected via Eq. (3.46) with respect to P^μ. In addition, for some cases, the resulting covariant doesn't have a well-defined C-parity. $\gamma^{T\mu}\,[\not{q},\not{P}]$ is such an example, in which case one obtains a covariant with a well-defined C-parity by adding a corresponding term, in this example, $-q^{T\mu}\not{P}$. One also sees immediately that all covariants in Eq. (3.57) that are proportional to P^μ vanish by this projection. So one remains with eight transverse covariants, e.g.

$$\begin{array}{ll} T_1^\mu = \gamma^{T\mu} & T_2^\mu = q^{T\mu}\not{q} \\ T_3^\mu = q^{T\mu}\not{P} & T_4^\mu = \gamma^{T\mu}\,[\not{q},\not{P}] - q^{T\mu}\not{P} \\ T_5^\mu = i\,q^{T\mu}\,\mathbf{1} & T_6^\mu = i\,(\gamma^{T\mu}\not{q} - \not{q}\,\gamma^{T\mu}) \\ T_7^\mu = i\,\gamma^{T\mu}\not{P} & T_8^\mu = i\,q^{T\mu}\,[\not{q},\not{P}] \end{array} \tag{3.58}$$

Here, T_3 and T_6 have positive C-parity, the others negative. To obtain a vector state with $J^{PC} = 1^{--}$, F_3 and F_6 must be odd functions of $q \cdot P$, the others even. A basis for an axialvector state can be constructed analogously to the construction of the pseudoscalar from the scalar state in the $J = 0$ case, namely by multiplication of each covariant with a factor γ_5.

A general prescription for the construction of a BSA for any spin J is given in the following. First, note that the BSA for a meson with spin J is given by

$$\Gamma^{\mu\nu\cdots}(q; P; \gamma) = \sum_{i=1}^{N_J} T_i^{\mu\nu\cdots}(q; P; \gamma)\; F_i(q^2, q \cdot P, P^2)\;, \tag{3.59}$$

where the generalized scalar product for the covariants $T_i^{\mu\nu\cdots}$ is defined via the Dirac trace

$$\sum_{\mu\nu\ldots} \mathrm{Tr}[T_i^{\mu\nu\cdots} T_j^{\mu\nu\cdots}] = t_{ij} f(i, j)\;. \tag{3.60}$$

One may also choose the basis elements orthogonal such that $t_{ij} = \delta_{ij}$, with the $f(i, j)$ functions of q^2, P^2, and $q \cdot P$, or orthonormal such that in addition $f(i, j) = 1$ for all i, j. The sum is carried out over the J indices $\mu, \nu, \ldots$.

Next, define the projections

$$q_\mu^T := q_\mu - P_\mu \frac{q \cdot P}{P^2}\;, \tag{3.61}$$

$$\gamma_\mu^T := \gamma_\mu - P_\mu \frac{\not{P}}{P^2}\;, \tag{3.62}$$

$$\gamma_\mu^{TT} := \gamma_\mu - P_\mu \frac{\not{P}}{P^2} - q_\mu^T \frac{\not{q}^T}{(q^T)^2}\;, \tag{3.63}$$

$$g_{\mu\nu}^T := g_{\mu\nu} - \frac{P_\mu P_\nu}{P^2} \tag{3.64}$$

as well as

$$\sigma^{q,P} := i/2\,[\not{q},\not{P}] \tag{3.65}$$

and keep in mind the 4 orthogonal scalar covariants

$$\mathbf{1}, \quad \not{P}, \quad \not{q}^T, \quad i\sigma^{q,P} \; . \tag{3.66}$$

Then, to consider states of any particular spin J, one has to construct Lorentz-tensors of rank J which are totally symmetric, transverse in all open indices and Lorentz-traceless: such an object has the $2J+1$ spin degrees of freedom as demanded in quantum mechanics of a massive particle. These restrictions, together with the algebra properties of the Dirac matrices, lead to eight covariant structures for all $J \geq 1$. In particular, one can define two tensor structures $M_{\mu\nu}$ and $N_{\mu\nu}$ such that $N_{\mu\nu\ldots\tau}$ is the traceless part of

$$q^T_\mu q^T_\nu \; \ldots \; q^T_\tau \tag{3.67}$$

and $M_{\mu\nu\ldots\tau}$ is the traceless part of the totally symmetric sum constructed from

$$\gamma^{TT}_\mu q^T_\nu \; \ldots \; q^T_\tau \; . \tag{3.68}$$

Each of these multiplied by the four terms in (3.66) defines four rank-J tensor covariants, in total eight, orthogonal in the sense of Eq. (3.60).

To make this a bit more concrete, we now generate orthogonalized tensor covariants, i. e., $J^P = 2^+$ and obtain as a first step the symmetric and transverse expressions

$$\tilde{M}_{\mu\nu} = \gamma^{TT}_\mu q^T_\nu + q^T_\mu \gamma^{TT}_\nu \;\; \text{and} \tag{3.69}$$

$$\tilde{N}_{\mu\nu} = q^T_\mu q^T_\nu \; , \tag{3.70}$$

which automatically satisfy Eq. (3.60). The next step is to implement the tracelessness, which is equivalent to orthogonality with respect to $g^T_{\mu\nu}$. This yields

$$M_{\mu\nu} = \tilde{M}_{\mu\nu} - g^T_{\mu\nu} \frac{\tilde{M}_{\rho\sigma} g^T_{\rho\sigma}}{(g^T)^2} \tag{3.71}$$

$$N_{\mu\nu} = \tilde{N}_{\mu\nu} - g^T_{\mu\nu} \frac{\tilde{N}_{\rho\sigma} g^T_{\rho\sigma}}{(g^T)^2} \; , \tag{3.72}$$

which is the desired structure, and by multiplication with the four scalar covariants in (3.66) gives the eight desired orthogonal tensor covariants. Subsequently, normalization can be achieved via

$$\hat{T}^{\mu\nu\ldots}_i = \frac{T^{\mu\nu\ldots}_i}{\sqrt{\sum_{\alpha\beta\ldots} \mathrm{Tr}[T^{\alpha\beta\ldots}_i \cdot T^{\alpha\beta\ldots}_i]}} \; .$$

3.9 Exercises

Exercise 3.1 Derive the canonical normalization condition for the BSA Γ from the condition for χ, Eq. (3.34).

Exercise 3.2 Based on Eq. (3.40), consider the following simple iteration procedure: Insert an initial guess for the amplitude vector Γ on the right-hand side of (3.40), say Γ_0. Then, compute the result, i.e., the left-hand side, which is proportional to Γ_1, the "improved guess", reinsert Γ_1 on the r.h.s. and repeat the procedure until convergence is achieved.

In this procedure, argue

(a) how one can extract λ

(b) why this procedure yields the largest eigenvalue λ_0 of $\tilde{K}$

(c) how one can obtain the next-to-largest eigenvalue λ_1, and further eigenvalues

(d) how a convergence criterion could be formulated

Exercise 3.3 Prove

(a) Eqs. (3.48), (3.49), and (3.50)

(b) Eq. (3.56)

Exercise 3.4 Analyze the C-parities of the covariants Nr. 4, 6, and 8 in Eq. (3.58) and show that they are indeed as indicated after the equation. In addition, starting from the more general T_4 in Eq. (3.57), show how one can obtain an expression of definite C-parity, namely the one given in Eq. (3.58).

Exercise 3.5 Prove that a totally symmetric tensor of rank J in n space-time dimensions has

$$\binom{J+n-1}{n-1}$$

independent components. Prove furthermore, starting from the actual case $n = 4$, that demanding Lorentz-tracelessness and transversality reduces this number to the desired $2J + 1$.

Achievement

After studying this section and performing the exercises the reader should be able and feel confident to do the following:

- Understand the differences between the homogeneous and inhomogeneous versions of the BSE
- Sketch and compare the solution strategies for these two versions of the BSE
- Construct Dirac covariants for 2-fermion systems and determine their basic properties for various values of the total spin of the 2-particle state

Chapter 4

Example Model Calculations

In the last section we have introduced the BSE and specified how to construct the BSA, e. g., for a fermion-antifermion bound state with spin J, but without actually going into any details with respect to what S and K concretely are or where they come from. Still this enabled us to discuss general solution strategies like inserting the appropriate structure into the BSE and solving the equation by projecting on each covariant via the corresponding trace relations to obtain coupled integral equations for the amplitudes F_i. With the general notion of dealing with matrices and vectors, we arrived at viable approaches to finding bound-state masses or rather values of P^2 for which the homogeneous BSE is solved.

Now it is time to detail the interaction kernel K and the dressed propagators of the constituents of the bound states and obtain the corresponding concrete solutions. The context for this is provided to us by the Dyson-Schwinger equations of QCD, which we'll introduce and make use of in the hands-on spirit.

4.1 Dyson-Schwinger Equations in a Nutshell

Although there aren't many in-depth treatments of Dyson-Schwinger Equations (DSEs) in textbooks, the subject is reasonably covered in several excellent recent review articles which can be found in the bibliography at the end of these notes. For our purposes and context, the most important information about DSEs is summarized here:

- The name comes from the works of Dyson (1949) and Schwinger (1951), who showed that, from the field equations of a quantum field theory, one can derive a set of coupled integral equations, which relate the Green functions of a theory to each other.
- This set of equations is called the set of Dyson-Schwinger equations (equivalently Schwinger-Dyson equations in part of the literature).
- The set of DSEs is an infinite and coupled set of nonlinear integral equations, satisfied by the Green functions; there is one particular equation for each Green function.

- In general, the DSE for an n-point function involves other Green functions of both lower and higher order
- The question arises how one can get actual solutions for this infinite tower of coupled equations, which is possible via an expansion or truncation:
- One possibility (usually mentioned in textbooks, but inappropriate for our purposes) is perturbation theory: although the DSEs are non-perturbative *a priori*, they can be used to generate every diagram in perturbation theory via a weak-coupling expansion. We will not use such an expansion.
- The other possibility is that of a truncation scheme: a *truncation* of the set of DSEs means that one first selects a subset of equations to be solved explicitly, and then makes Ansätze for all Green functions, whose equations are *not* solved explicitly (this will become clear in the example below). The truncation scheme then guides the practitioner from a simpler truncation to a more complicated one in a way that, e. g., respects symmetries of the theory.
- Symmetries are represented in this context by Ward-(Green)-Takahashi or Slavnov-Taylor identities. These are relations (additional to the DSEs themselves) between Green functions which can be used as constraints on several occasions to, e. g., determine the quark-antiquark interaction kernel, or properties of vertices.

Below we will examine in some detail the DSE for the quark propagator, which provides part of the input for our example study of the meson BSE further below.

4.2 Euclidean Space

It is of computational advantage to work (calculate/solve equations) in Euclidean rather than Minkowski space. For example, it is possible to simply use spherical variables in four dimensions in the integration. In terms of interpretation one can choose to define the particular quantum field theory under consideration in Euclidean space from the very beginning and calculate everything from there.

Since we are dealing with integral equations in the following, we'll shortly summarize all necessary ingredients and their properties here. In quantities and equations which can be expressed as functions of the metric one replaces the Minkowski $g_{\mu\nu}$ by the Euclidean $g_{\mu\nu} = \delta_{\mu\nu}$. As a result of this simplification, co- and contravariant indices are no longer distinguishable and can be written either up or low on expressions throughout any calculation. To avoid confusion, it is also common to transform a vector's Minkowski-space time-component a_0 into a fourth component a_4 in Euclidean space. The scalar product of two four vectors a and b then becomes $a \cdot b = a_4 b_4 + \vec{a} \cdot \vec{b}$ and one has $a^2 = a \cdot a < 0$ for timelike and $a^2 > 0$ for a spacelike vector a, respectively (which is exactly opposite to the Minkowski-space situation).

Regarding integration we consider a four momentum vector q and parameterize its components via the radial variable $q^2 \in [0,\infty)$ and three angles $\beta, \vartheta \in [0,\pi]$, and $\varphi \in [0,2\pi)$. One has

$$q^\mu = \sqrt{q^2} \begin{pmatrix} \sin\beta \sin\vartheta \cos\varphi \\ \sin\beta \sin\vartheta \sin\varphi \\ \sin\beta \cos\vartheta \\ \cos\beta \end{pmatrix} \tag{4.1}$$

The corresponding momentum-space integration measure is given by

$$\begin{aligned} \int d^4q &= \int_0^\infty \frac{q^2}{2} d(q^2) \int_0^\beta \sin^2\beta d\beta \int_0^\pi \sin\vartheta d\vartheta \int_0^{2\pi} d\varphi = \\ &= \int_0^\infty \frac{x}{2} dx \int_{-1}^1 \sqrt{1-z^2} dz \int_{-1}^1 dy \int_0^{2\pi} d\varphi \end{aligned} \tag{4.2}$$

where one defines $x := q^2$, $z := \cos\beta$, and $y := \cos\vartheta$. We'll get back to this below when it is time to parameterize the momentum variables in the quark DSE and the BSE.

4.3 The Quark Propagator and its DSE

The free propagator $S(q)$ for a fermion of mass m and four-momentum q in Euclidean space (with a certain sign convention) is given by

$$S(q) = \frac{1}{i\not{q} + m\mathbf{1}} \,. \tag{4.3}$$

Dealing with quarks, one has to ask how quarks, which are actually not observed as free particles in our detectors, can be described in this context. The answer is the dressed quark propagator, in which the two scalar dressing functions $A(q^2)$ and $B(q^2)$ modify S in Eq. (4.3) to its most general form,

$$S(q) = \frac{1}{i\not{q} A(q^2) + \mathbf{1} B(q^2)} \,. \tag{4.4}$$

Bringing the Dirac structure of the propagator from the denominator to the numerator, we get

$$S(q) = \frac{-i\not{q} A(q^2) + \mathbf{1} B(q^2)}{q^2 A^2(q^2) + B^2(q^2)} \,, \tag{4.5}$$

which is sometimes also written as

$$S(q) = -i\not{q}\sigma_v(q^2) + \mathbf{1}\sigma_s \tag{4.6}$$

with

$$\sigma_v(q^2) = \frac{A(q^2)}{q^2 A^2(q^2) + B^2(q^2)} \quad \text{and} \quad \sigma_s(q^2) = \frac{B(q^2)}{q^2 A^2(q^2) + B^2(q^2)} \,. \tag{4.7}$$

The subscripts $_v$ and $_s$ refer to the *vector* and *scalar* parts of the propagator, respectively. The dressing functions σ_v, σ_s, A, and B are functions of the

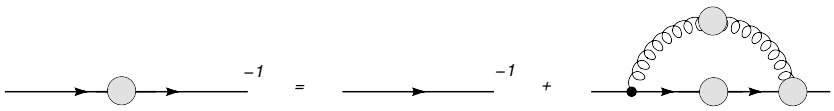

Figure 4.1: The DSE for the quark propagator

four-momentum squared q^2 only, i. e., all the Dirac structure is already given explicitly.

Going back to the form in Eq. (4.4), one can define the *quark wave-function renormalization* $Z(q^2) := 1/A(q^2)$ and the *quark mass function* $M(q^2) := B(q^2)/A(q^2)$ to write

$$S(q) = \frac{Z(q^2)}{i\not{q} + \mathbf{1}M(q^2)} \tag{4.8}$$

and also

$$S(q) = Z(q^2)\frac{-i\not{q} + \mathbf{1}M(q^2)}{q^2 + M^2(q^2)} \,, \tag{4.9}$$

which makes the analogy to a free propagator and the motivation for the term "mass function" more evident.

Now that we have clarified the structure of the quark propagator, we'll have a look at its DSE. The quark DSE in QCD essentially describes how the various interactions of quarks and gluons among each other influences quark propagation. The inverse of the dressed quark propagator $S^{-1}(p)$ is obtained from the inverse of the bare (free) propagator $S_0^{-1}(p)$ via adding the quark self-energy $\Sigma(p)$. The self-energy involves integration over an internal (loop) momentum q and depends on the dressed quark propagator $S(q)$ itself, but also on the (so far unknown) dressed gluon propagator $D_{\mu\nu}(p-q)$ and the bare and dressed quark-gluon vertices $\gamma^\mu \frac{\lambda^a}{2}$ and $\Gamma^a_\nu(q;p)$, respectively. In addition, the coupling constant g appears. Putting all this together, one obtains the quark DSE or *QCD gap equation*

$$\begin{aligned} S^{-1}(p) &= S_0^{-1}(p) + \Sigma(p) & (4.10)\\ &= i\not{p} + \mathbf{1}m_q + \int \frac{d^4q}{(2\pi)^4} g^2 D_{\mu\nu}(p-q)\frac{\lambda^a}{2}\gamma^\mu S(q)\Gamma^a_\nu(q;p) & (4.11) \end{aligned}$$

which is depicted in Fig. 4.1.

For our present purposes the following remarks are important: First of all, we notice that the equation is nonlinear (in contrast to the BSE) and we'll discuss solution strategies below. Then, this integral equation for the quark two-point function involves one other two-point function (the gluon propagator) and a three-point function, the quark-gluon vertex. If we knew these two additional functions, we could solve the equation for S "immediately" (at least in principle). However, $D_{\mu\nu}$ and Γ^a_ν satisfy their own DSEs, involving S, themselves, and other Green functions like, e. g., the three-gluon vertex or the four-gluon vertex. These, in turn, satisfy their own DSEs, involving other (higher) Green functions, etc. This is characteristic of the infinite coupled

tower of DSEs. In this context, a truncation helps by, e. g., making Ansätze for both these unknown functions from general or particular properties of the theory under consideration and/or phenomenological input, and then solving the equation under the assumption that the Ansätze give a reasonable account of what one would have obtained by actually and self-consistently solving the entire tower of DSEs. Before we actually give an example of such a truncation and explore its consequences, we sketch simple solution strategies for the quark DSE in the following.

The first step towards a solution is turning Eq. (4.11) into two coupled integral equations for the dressing functions A and B (or, alternatively, for Z and M) via suitable Dirac projections involving different projectors and taking the Dirac-trace of the resulting equations. The next step is finding a solution strategy for solving the coupled integral equations for, say, A and B. As mentioned above, the quark DSE is nonlinear, which doesn't a priori preclude success of a simple iterative approach. However, more sophisticated techniques can be helpful, one of which we'll now introduce. This particular technique is analogous to the Newton method for finding the zero of a function $F(x)$, where one starts at an initial value x_0 and obtains the next guess x_1 for the zero by constructing a tangent to $F(x)$ at the point $(x_0, F(x_0))$ and intersecting it with the abscissa. This yields the procedure

$$x_i = x_{i-1} - \left(\frac{dF(x)}{dx}\right)^{-1}\Bigg|_{x=x_{i-1}} F(x_{i-1}) \; . \tag{4.12}$$

Now consider a matrix equation:

$$\vec{x} = \vec{a} + \vec{K}(\vec{x}) \tag{4.13}$$

with the constant vector $\vec{a}$ and the vector-valued kernel function $\vec{K}(\vec{x})$. Define the vector-valued function

$$\vec{F}(\vec{x}) := \vec{x} - \vec{a} - \vec{K}(\vec{x}) \; . \tag{4.14}$$

To find solutions to the original equation (4.13), one must now find zeros of $\vec{F}(\vec{x})$, which we achieve by a method along the lines of (4.12) and get

$$\vec{x}_i = \vec{x}_{i-1} - \left(\mathbf{1} - \frac{\partial \vec{K}(\vec{x})}{\partial \vec{x}}\right)^{-1}\Bigg|_{\vec{x}=\vec{x}_{i-1}} \left(\vec{x}_{i-1} - \vec{a} - \vec{K}(\vec{x}_{i-1})\right) \; . \tag{4.15}$$

Remember that we used a special case of this method for the inhomogeneous BSE above, namely that $\vec{K}(\vec{x}) = K \cdot \vec{x}$ with the matrix K, which (corresponding to Eq. (3.38)) yields

$$\vec{x}_i = (\mathbf{1} - K)^{-1} \cdot \vec{a} \; . \tag{4.16}$$

Obviously, this requires "only one iteration step" and the subscript $_i$ on the left-hand side of the equation can be dropped to immediately give the solution for $\vec{x}$.

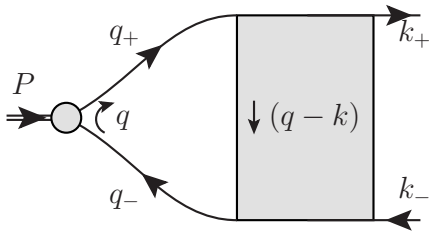

Figure 4.2: The momentum flow in the BSE (adapted from Ref. [10])

4.4 The BSE Kernel: Concerning S

A priori, it is not clear what the BSE kernel for a quark-(anti)quark system is. The first early applications of the BSE were in nuclear physics, where people studied N-N systems (scattering). The interaction for such a system was proposed to be π-exchange. In this setting the simplest (leading) kernel in the BSE is a *ladder* kernel proportional to the square of the πNN coupling. Naturally, the full N-N interaction is more complicated than this, but as a first step, this particular interaction provided some treatability of the problem.

In meson studies in QCD, the situation is similar: we are dealing with quarks and antiquarks, interacting via gluons. But first, a few general statements are in order: formally, part of the BSE kernel is the quark-(anti)quark interaction, the other part is the product of the two constituent dressed propagators. We'll deal with them first, after a short outline of the overall kinematical setup of the BSE and its components. As a reminder, the BSE (now in Euclidean space and omitting all sorts of indices, but paying attention to the variables) reads

$$\Gamma(k;P) = \int \frac{d^4q}{(2\pi)^4} K(P;k;q)\; S(q_+)\; \Gamma(q;P)\; S(q_-)\;; \tag{4.17}$$

the corresponding momentum flow on the right-hand side of this equation is illustrated in Fig. 4.2. The Euclidean four-momenta are: the total momentum of the $q\bar{q}$-system P, and the corresponding relative momenta k and q as well as the quark momentum $q-+ = q+\eta_+ P$ and the antiquark momentum $q_- = q-\eta_- P$ (as already stated above in a slightly different manner, $\eta_\pm$ can take values $\in [0,1]$ and $\eta_+ + \eta_- = 1$). To fully appreciate the situation when we work everything out concretely, let us start by parameterizing the four-momenta P, k, and q in terms of the general setup given above in Eq. (4.1), but with the possibility to freely choose the (4D)-direction of P, then almost freely k. One has:

$$P^\mu = iM \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \;, \quad k^\mu = \sqrt{k^2} \begin{pmatrix} 0 \\ 0 \\ \sin\beta_k \\ \cos\beta_k \end{pmatrix} \;, \quad q^\mu = \sqrt{q^2} \begin{pmatrix} \sin\beta_q \sin\vartheta \cos\varphi \\ \sin\beta_q \sin\vartheta \sin\varphi \\ \sin\beta_q \cos\vartheta \\ \cos\beta_q \end{pmatrix} . \tag{4.18}$$

To understand the form of P and in particular the appearance of iM, let us recall that in Euclidean space we have the homogeneous BSE under the assump-

tion of the existence of a bound state solution at $P^2 = -M^2$ with M being the bound state's mass. Using the notation x, y, z defined above after Eq. (4.2), we get

$$k^\mu = \sqrt{k^2} \begin{pmatrix} 0 \\ 0 \\ \sqrt{1-z_k^2} \\ z_k \end{pmatrix} \quad , \quad q^\mu = \sqrt{q^2} \begin{pmatrix} \cos\varphi\sqrt{1-z_q^2}\sqrt{1-y_q^2} \\ \sin\varphi\sqrt{1-z_q^2}\sqrt{1-y_q^2} \\ y_q\sqrt{1-z_q^2} \\ z_q \end{pmatrix} . \tag{4.19}$$

Next, we investigate the Lorentz-invariant variables in the problem. We have, aside from the trivial P^2, q^2, and k^2,

$$q \cdot P = iM\sqrt{q^2} z_q \tag{4.20}$$

$$k \cdot P = iM\sqrt{k^2} z_k \tag{4.21}$$

$$q \cdot k = \sqrt{q^2}\sqrt{k^2}(z_q z_k + y_q\sqrt{1-z_q^2}\sqrt{1-z_k^2}) \, . \tag{4.22}$$

One should note already at this point that in general it is more advantageous to keep the complex variable $\sqrt{P^2}$ in favor of iM, since the latter is only reasonable for negative values of P^2 as encountered in the solution of the homogeneous BSE. When solving the inhomogeneous vertex-BSE, on the other hand, one has no restriction like this and P^2 is better set up as an input, which could, in general, be a complex number. Still, for the following demonstration focussed on the homogeneous BSE in particular, explicit factors of iM are written for instructional purposes.

The corresponding problem arises from the quark propagator arguments under the integral on the right-hand side of Eq. (4.17), which are q_+ and q_-. Written out explicitly, one sees that with the all-real components of the integration momentum q and real M and $\eta_\pm$, both q_+ and q_- are actually complex four-momenta. That said, it is not surprising to see also their squares (which appear as the arguments of the dressing functions in the quark propagators in Eq. (4.17)) turn out to be complex. More details can be easily obtained by writing q_+^2 and q_-^2 in terms of the Lorentz-invariants given above as

$$q_+^2 = q^2 - \eta_+^2 M^2 + 2i\eta_+ M\sqrt{q^2} z_q \tag{4.23}$$

$$q_-^2 = q^2 - \eta_-^2 M^2 - 2i\eta_- M\sqrt{q^2} z_q \, . \tag{4.24}$$

As an immediate consequence, the domain in the complex q_+^2, sampled via a numerical solution of the BSE in Eq. (4.17)), on which the quark propagator $S(q_+)$ needs to be known is surrounded by a parabolic shape depicted in Fig. 4.3.

Let us now summarize the situation just described as well as some of its consequences:

- The quark propagator is needed for complex arguments as input in the BSE, if $P^2 < 0$

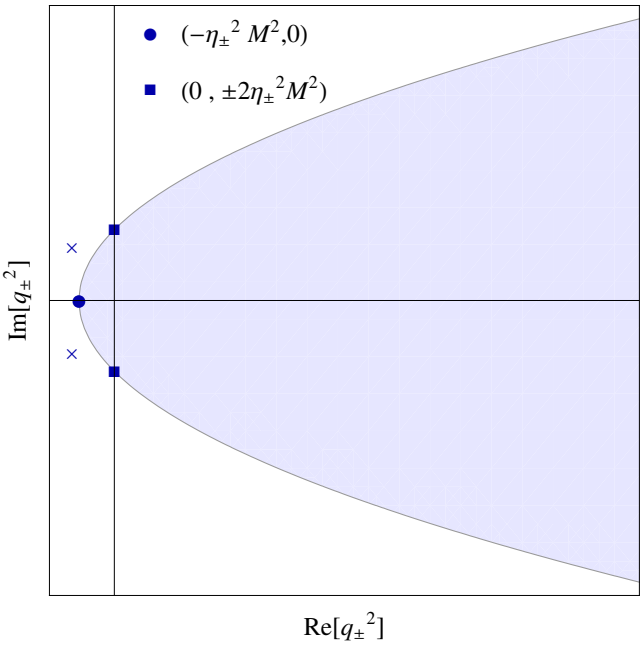

Figure 4.3: The complex domain of the argument of the quark propagator in the BSE as sampled via the momentum configuration in Euclidean space. The cross inside the parabola is the origin, the coordinates of the intersection points with the real and imaginary axes are given explicitly in the figure (adapted from Ref. [11])

- The sampling region is a parabolic domain in the $q_\pm^2$-planes, whose size is characterized by $\eta_\pm$ and M
- For large M (heavier states) a large region in such a complex plane must be sampled, which is numerically nontrivial
- $\eta_\pm$ can be used to manipulate the situation by shifting parabola size from one constituent propagator to the other, which is particularly useful for unequal-mass constituents
- It is possible (i. e., it happens) that there are singularities of the propagator functions σ_v and σ_s at finite $q_\pm^2$ in the complex plane (sketched in Fig. 4.3 by two small x outside the parabolic region). Then, for simple/standard numerical methods there is a limiting bound-state mass, for which the integration domain starts to overlap these singularities. Larger bound-state masses can be studied only using more appropriate numerical methods or sophisticated extrapolation techniques.

To finish this section it is instructive to mention the sources for the data/values needed for S in the BSE during a solution. The simplest possibility would simply be to use free propagators, i. e., $Z(q^2) = 1$ and $M(q^2) = m$. While this may be reasonable in general, it is not for our considerations of mesons as bound

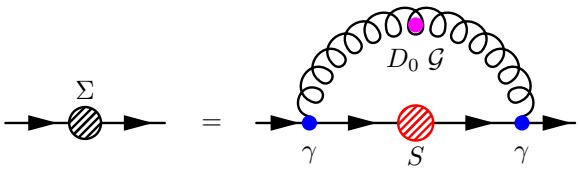

Figure 4.4: The rainbow-truncated quark self-energy.

states of dressed (anti)quarks. The next and still simple possibility is to use Ansätze for the propagator dressing functions which yield properties desired of quarks. This is good enough for, e. g., purely phenomenological approaches to hadron properties, but cannot reveal much information about the dynamics of the system as it would be obtained from an underlying quantum field theory. Thus, the method of choice in the following is to solve the quark DSE as introduced above and analytically (in practice numerically) continue its solutions to complex values as needed in the BSE setup. To have a consistent dynamical setup of both the DSE and the BSE kernels (as described below) is one of the keys to phenomenological success.

4.5 The BSE Kernel: Concerning K

What remains to be found is the form of K as it appears in Eq. (4.17). Since we only know that the kernel should contain all possible interactions resulting from the theory under consideration, we need some starting point for the construction of a tractable problem. One possibility already mentioned above is that of a truncation; however, such a step should be performed in a reasonable manner. On a general note, one can find from the functional formalism that $K = -\delta\Sigma/\delta S$, i. e., the two-particle interaction kernel is the functional derivative of the one-particle self-energy with respect to the dressed particle propagator, which one can attempt to satisfy. More concretely, as a guiding principle for truncating the DSEs and at the same time finding a consistent kernel for the BSE one can make use of the symmetries of the underlying theory. These are provided by Ward-Takahashi or Slavnov-Taylor identities, relations which constrain the Green functions beyond their satisfaction of the DSEs themselves. In the context of meson studies with the background of QCD one has, e. g., the vector and axial vector (and color singlet) Ward-Takahashi identities as possibilities to constrain the integral equation kernels (these two particular identities are related to charge conservation and chiral symmetry, respectively). It was found some time ago that a very simple truncation of the set of quark DSE + quark-antiquark BSE, the so-called *rainbow-ladder truncation* can actually be employed to satisfy both identities, which is a promising setup for a phenomenological model of hadron properties.

In more detail, employing the rainbow-ladder truncation in our system of coupled integral equations means that the quark DSE, Eq. (4.11), neglecting renormalization/regularization issues completely here, in the *rainbow truncation*

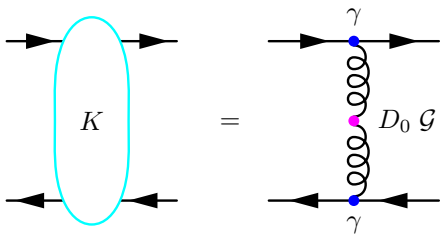

Figure 4.5: The ladder-truncated quark-antiquark BSE interaction kernel.

becomes

$$S^{-1}(p) = i\not{p} + \mathbf{1}m_q + \frac{4}{3}\int \frac{d^4q}{(2\pi)^4}\mathcal{G}((p-q)^2)D^f_{\mu\nu}(p-q)\gamma_\mu S(q)\gamma_\nu \ , \qquad (4.25)$$

where the factor 4/3 comes from the color trace, $D^f_{\mu\nu}(p-q)$ is the free gluon propagator, and $\mathcal{G}((p-q)^2)$ is an effective interaction that depends on the gluon momentum-squared. For a graphical representation of the rainbow quark self-energy, see Fig. 4.4.

Correspondingly, the BSE in *ladder truncation* reads (see also Fig. 4.5 for a graphical representation)

$$\Gamma(p;P) = -\frac{4}{3}\int \frac{d^4q}{(2\pi)^4}\mathcal{G}((p-q)^2)D^f_{\mu\nu}(p-q)\gamma_\mu S(q_+)\Gamma(q;P)S(q_-)\gamma_\nu \ . \qquad (4.26)$$

With these consistent kernels one can show that general results of the theory are preserved in the truncation. For example, in the chiral limit (i. e., for vanishing current quark mass) one obtains a zero-mass pion solution from the homogeneous pseudoscalar BSE without fine-tuning the effective interaction.

4.6 Equations for Components and Chebyshev Expansion

Starting from, e. g., the ladder BSE in Eq. (4.26) one has to deal with the Dirac structure of the equation and BSA to arrive at a set of coupled integral equations for scalar components. This is achieved by projecting onto one of the BSA's Dirac covariants by multiplying both sides of the equation with the particular covariant T_i (Lorentz indices are omitted here for brevity) and taking the Dirac trace. In the most straight-forward case of a set of orthogonal covariants, the left-hand side of the equation,

$$\Gamma(p;P) = \sum_{j=1}^{N_J} T_j(p;P;\gamma)F_j P^2, p^2, p\cdot P \ , \qquad (4.27)$$

is reduced to the i-covariant's normalization times the component F_i, i. e., a single term. On the right-hand side the kernel is projected as well, and via $\Gamma(q;P)$

under the integral and its expansion in covariants times components a matrix structure emerges, which couples the integral equations for the F_i together. A typical equation of this kind is satisfied by $F_j P^2, p^2, p \cdot P$, where P^2 is actually a parameter with respect to the solution process. This reduces the variables to the radial p^2 and the cosine z_p. At this point one can apply numerical methods for the solution of the emerging matrix equation by discretizing the momentum and cosine variables and, e. g., proceed with standard iteration techniques in two variables, which makes the kernel comparatively large.

An alternative, advantageous and widely used in the calculations of hadron spectra is to apply the momentum discretization only to $\int d(q^2)$ and to deal with the z-dependence by an expansion in Chebyshev polynomials of the second kind. The components are then written as

$$F^i(q^2, z_q) = \sum_{j=1}^{N_z} {}^j F_i(q^2) U_j(z_q) \ , \tag{4.28}$$

with ${}^j F_i(q^2)$ the so-called Chebyshev *moments*, which retain only the functional dependence on q^2 (and in principle also P^2). The number of terms N_z taken into account is finite in practice, but infinite in principle. The Chebyshev polynomials of the second kind $U_j(z)$ satisfy the orthogonality relation

$$\frac{2}{\pi} \int dz \sqrt{1-z^2}\, U_i(z)\, U_j(z) = \delta_{ij} \ . \tag{4.29}$$

To obtain a matrix structure not only in the covariants, but also in the Chebyshev moments, the above expansion is inserted into the Dirac-projected equation resulting from Eq. (4.26), and is then projected further on one moment by use of Eq. (4.29).

4.7 Exercises

Exercise 4.1 Consider the quark DSE in the form (4.25), ignore regularization/renormalization issues, and from there derive two coupled and non-linear integral equations for the functions $A(p^2)$ and $B(p^2)$.

Exercise 4.2 Consider the effective interaction in Eq. (4.25) to be

$$\mathcal{G}(k^2) D^f_{\mu\nu}(k) = (2\pi)^4 D\ \delta^4(k)(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}) = (2\pi)^4 D\ \delta^4(k) \frac{3}{4} \delta_{\mu\nu} \ , \tag{4.30}$$

where $k = p - q$, and D is a constant scale (a form like this was prominently used by Munczek and Nemirovsky). In this simple model,

(a) solve the quark DSE explicitly in the chiral limit, i. e., give expressions for $A(p^2)$, $B(p^2)$, $M(p^2)$, $\sigma_v(p^2)$, and $\sigma_s(p^2)$ under the assumption that $m = 0$. Note here that the equations should be solved under the assumptions that $A(p^2)$, $B(p^2)$ be real and non-negative for real p^2 and that they have the asymptotic behavior $A(p^2) \to 1$ and $B(p^2) \to m$ for $p^2 \to \infty$.

(b) consider the pseudoscalar homogeneous BSE (4.26) in the chiral limit. Observe that, since the BSE kernel contains the delta-function in the difference of relative momenta q and we are considering bound states, we set $q = 0$ and terms proportional to q don't contribute in any BSA, which leaves only two terms for the pseudoscalar case. Construct the BSE as a 2×2 matrix eigenvalue equation.

(c) show that this eigenvalue equation is solved for $P^2 = 0$ (the pion is massless in the chiral limit!), if both covariants are retained. Investigate, whether this situation changes if one uses only the γ_5 covariant.

(d) consider next the homogeneous vector-meson BSE and compute the vector-meson mass, again in the chiral limit. In this case, use the vectors P^μ and the polarization vector ε^μ together with γ^μ to construct the non-vanishing covariants of the BSA.

(e) argue, why one cannot study any state with $J > 1$ in this model.

(f) show that the homogeneous axial-vector and scalar BSEs in this model (again in the chiral limit) don't have any solutions.

(g) argue, why one doesn't obtain radial meson excitations in this model.

Achievement

After studying this section and performing the exercises the reader should feel confident to do the following:

- Brag about having solved a BSE by hand.

Acknowledgement

I would like to thank the students of this course, given during the summer term 2012 at the University of Graz, Austria, for their continued enthusiasm and feedback as well as for their "test-driving" the exercises, and for pointing out a number of errors in the manuscript.

Bibliography

[1] S. Weinberg. *The Quantum Theory of Fields, Volume I.* Cambridge University Press, Cambridge, England, 1995.

[2] C. Itzykson and J.-B. Zuber. *Quantum Field Theory.* Dover, 2005.

[3] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory.* Westview Press, 1995.

[4] B. D. Keister and W. N. Polyzou. *Adv. Nucl. Phys.*, 20:225–479, 1991.

[5] N. Nakanishi. *Prog. Theor. Phys. Suppl.*, 43:1–81, 1969.

[6] M. T. Noda, N. Nakanishi, and N. Seto. *Prog. Theor. Phys. Suppl.*, 95:78–111, 1988.

[7] C. H. Llewellyn-Smith. *Ann. Phys.*, 53:521–558, 1969.

[8] C. D. Roberts, M. S. Bhagwat, A. Holl, and S. V. Wright. *Eur. Phys. J. Special Topics*, 140:53–116, 2007.

[9] K. Nakamura et al. *J. Phys. G*, 37:075021, 2010.

[10] M. Blank and A. Krassnigg. *Comput. Phys. Commun.*, 182:1391, 2011.

[11] A. Krassnigg and M. Blank. *Phys. Rev. D*, 83:096006, 2011.