YET ANOTHER INTRODUCTION TO

QUANTUM FIELD THEORY

WITH THE EMPHASIS ON

THE REAL UNDERSTANDING OF VARIOUS NOTIONS

INCLUDING, BUT NOT RESTRICTED TO:

THE LOGIC OF THE SUBJECT

CANONICAL QUANTIZATION, PATH INTEGRALS

FEYNMAN DIAGRAMS

RENORMALIZATION AND REGULARIZATION

FROM TOY MODELS UP TO THE STANDARD MODEL

FOR THE USE OF STUDENTS

AND PARTICLE PHYSICISTS IN GENERAL

(BUT PERHAPS NOT THAT MUCH FOR JOURNALISTS AND PHILOSOPHERS)

BY MARTIN MOJŽIŠ

PRICE ONE SHILLING

PRINTED, IF NECESSARY, BY THE READER HIM/HERSELF.

ENJOY.

Contents

Preface

When I started giving lectures on Quantum field theory, I had no intention to write a book on the subject. There are plenty of such books available on the market and it seemed to make a little sense to add another one. But, unfortunately, none of them was exactly to my taste. My favorite *The Quantum Theory of Fields* by Steven Weinberg did not suit that well as an introductory course. I decided to use *An Introduction to Quantum Field Theory* by Peskin and Schroeder, which was, and perhaps still is, one of the standard modern textbooks on the subject. The book, however, leaves much to be desired, so I started to write some notes to provide a set of hopefully useful comments and remarks to it. The original plan was

- to reorganize the material in a bit different way
- to offer sometimes a slightly different point of view
- to add some material

Eventually, the text became more and more self-contained, and the resemblance to the Peskin-Schroeder became weaker and weaker. At the present point, the text has very little to do with the Peskin-Schroeder, except perhaps the largely common notation.

The aim of this course is to explain not only what we are doing, but why we are doing it. I tried my best not to provide complicated answers to questions that were not asked. This applies not only to particular aspects of the subject, but also to the structure of the whole course.

In the quite extended first part almost no interesting particle physics is discussed at all. We only deal with scalar fields and spinless particles, the emphasis is on the logic of the theory (with all the necessary technicalities, of course). In this part students should learn and understand why and how do we quantize classical fields, why and how the machinery of Feynman diagrams works, why and how do we renormalize parameters of lagrangians, why and how do we utilize path integral formulation of QFT.

The physics enters only in the second part, devoted to Quantum Electrodynamics. Here the technical complications brought up by higher spins, as well as important physical results are discussed thoroughly. All this is done step by step. We start with spinless particles in classical electromagnetic field, then the QED of spinless particles is developed, and only afterwards the full (spinor) QED appears.

The third part concerns the Standard Model. Large portion of this part, however, does not deal with the SM itself, but rather with the particle physics before the SM. It is my firm belief, that students exposed directly to the SM Lagrangian, with insufficient knowledge of the prior theoretical (and experimental) development, can miss the essence of the whole business. But it is not only the historic perspective what makes the pre-SM particle physics very useful for the SM course. Virtually all the ingredients of the SM originated in pre-SM physics and so they can be introduced in a quite natural way. Only once these ingredients are grasped to a reasonable level, the SM is discussed.

Part I

Basics of QFT (Spinless Particles)

Chapter 1

Introductions

Let us state at the very beginning that Quantum field theory is

- a theory of particles (in a way Quantum mechanics is a theory of atoms)[1]
- mathematically ill-defined
- the most precise theory mankind ever had
- conceptually and technically quite demanding

Mainly because of the last feature, it seems reasonable to spend enough time with introductions. The reason for plural is that we shall try to introduce the subject in couple of different ways.

Our first introduction is in fact a summary. We shall try to show how QFT is used in practical calculations, without any attempt to understand why it is used in this way. The reason for this strange maneuver is that, surprisingly enough, it is much easier to grasp the bulk of QFT on this operational level than to really understand it. We believe that even a superficial knowledge of how QFT is usually used can be quite helpful in a subsequent, more serious, study of the subject.

The second introduction is a brief exposition of the non-relativistic many-particle quantum mechanics. This enables a natural introduction of many basic ingredients of QFT (the Fock space, creation and annihilation operators, calculation of vacuum expectation values, etc.) and simultaneously to avoid the difficult question of merging relativity and quantum theory.

It is the third introduction, which sketches that difficult question (merging relativity and quantum theory) and this is done in the spirit of the Weinberg's book. Without going into technical details we try to describe how the notion of a relativistic quantum field enters the game in a natural way. The main goal of this third introduction is to clearly formulate the question, to which the canonical quantization provides an answer.

Only after these three introductions we shall try to develop QFT systematically. Initially, the development will concern only the scalar fields (spinless particles). More realistic theories for particles with spin 1/2 and 1 are postponed to the subsequent parts of the book.

[1] According to His Envyness, The High Inquisitor of Marseille, HSP (Hrdina Statočnej Práce) and SI unit of heterosexuality, may Peroon and other slavic gods bless Him forever — QFT is a theory of many other things, like e.g. an elephant ear, trunk, tail, etc.

In spite of the fact that His Envyness is sometimes quite a liar, here he is right. QFT can be defined in such a way (a useful one) that particles are not present from the beginning as basic building blocks, but they rather emerge (not neccessarilly) as a feature of the theory. If so, the QFT is a theory of these particles.

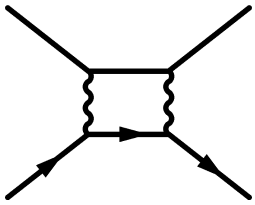
1.1 Conclusions

The machinery of QFT works like this:

- typical formulation of QFT — specification of a Lagrangian $\mathcal{L}$
- typical output of QFT — cross-sections $d\sigma/d\Omega$
- typical way of obtaining the output — Feynman diagrams

The machinery of Feynman diagrams works like this:

- For a given process (particle scattering, particle decay) there is a well defined set of pictures (graphs, diagrams). The set is infinite, but there is a simple criterion, allowing for identification of a relatively small number of the most important diagrams. Every diagram consists of several types of lines and several types of vertices. The lines either connect vertices (internal lines, propagators) or go out of the diagrams (external legs). An example:



- Every diagram has a number associated with it. The sum of these numbers is the so-called scattering amplitude. Once the amplitude is known, it is straightforward to obtain the cross-section — one just plugs the amplitude into a simple formula.
- The number associated with a diagram is the product of factors corresponding to the internal lines, external lines and the vertices of the diagram. Which factor corresponds to which element of the diagram is the content of the so-called Feynman rules. These rules are determined by the Lagrangian.
- The whole scheme is

Lagrangian
↓
Feynman rules
↓
Feynman diagrams
↓
scattering amplitude
↓
cross-section

- Derivation of the above scheme is a long and painful enterprise. Surprisingly enough, it is much easier to formulate the content of particular steps than to really derive them. And this formulation (without derivation[2]) is the theme of our introductory summary.

[2]It is perhaps worth mentioning that the direct formulation (without derivation) of the above scheme can be considered a fully sufficient formulation of the real content of QFT. This point of view is advocated in the famous Diagrammar by Nobel Prize winners 't Hooft and Veltman, where *"corresponding to any Lagrangian the rules are simply defined"*

1.1.1 Feynman rules

The role of the Lagrangian in QFT may be a sophisticated issue, but for the purposes of this summary the Lagrangian is just a set of letters containing the information about the Feynman rules. To decode this information one has to know, first of all, which letters represent fields (to know what the word *field* means is not necessary). For example, in the toy-model Lagrangian (of the so-called φ^3-theory)

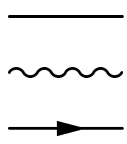
$$\mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}\mathring{m}^2\varphi^2 - \frac{1}{3!}\mathring{g}\varphi^3$$

the field is represented by the letter φ. Other symbols are whatever but not fields (as a matter of fact, they correspond to space-time derivatives, the so-called bare mass and the so-called bare coupling constant, but this is not important here). Another example is the Lagrangian of quantum electrodynamics (QED)

$$\mathcal{L}\left[\overline{\psi}, \psi, A_\mu\right] = \overline{\psi}\left(i\gamma^\mu\partial_\mu - \mathring{q}\gamma^\mu A_\mu - \mathring{m}\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and the fields are $\overline{\psi}$, ψ and A_μ (the symbol γ^μ stands for the so-called Dirac matrices, $\mathring{q}$ is the so-called bare charge and ξ is called a gauge parameter, but this information is not relevant here).

Now to the rules. Different fields are represented by different types of lines. The usual choice is a simple line for φ (called the scalar field), the wiggly line for A_μ (called in general the massless vector field, in QED the photon field) and a simple line with an arrow for $\overline{\psi}$ and ψ (called in general the spinor field, in QED usually the electron-positron field).



The arrows are commonly used for complex fields, like $\overline{\psi}$ and ψ (or φ^* and φ, if φ is complex)[3]. The arrow orientation is very important for external legs, where different orientations correspond to particles and antiparticles respectively (as we will see shortly). Every line is labelled by a momentum (and perhaps some other quantum numbers). The arrows discussed above and their orientation do not represent the momentum associated with the line!

The Feynman rules associate a specific factor with every internal line (propagator), line junction (vertex) and external line. Propagators are defined by the part of the Lagrangian quadratic in fields. Vertices are given by the rest of the Lagrangian. External line factor depends on the whole Lagrangian and usually (but not necessarily) it takes a form of the product of two terms. One of them is simple and is fully determined by the field itself, i.e. it does not depend on the details of the Lagrangian, while the other one is quite complicated and is determined by the whole Lagrangian.

[3]Actually, in practice arrows are not used for scalar field, even if it is complex. The reason is that no factors depend on the arrows in this case, so people just like to omit them (although in principle the arrows should be present).

vertices

For a theory of one field φ, the factor corresponding to the n-leg vertex is[4]

$$n\text{-leg vertex} = i\left.\frac{\partial^n \mathcal{L}}{\partial \varphi^n}\right|_{\varphi\equiv 0}$$

For a theory with more fields, like QED, the definition is analogous, e.g. the vertex with l, m and n legs corresponding to $\overline{\psi}, \psi$ and A_μ-fields respectively, is

$$(l,m,n)\text{-legs vertex} = i\left.\frac{\partial^{l+m+n}\mathcal{L}}{\partial A_\mu^l \partial\psi^m \partial\overline{\psi}^n}\right|_{\text{fields}\equiv 0}$$

Each derivative with respect to a field produces a corresponding leg entering the vertex. For terms containing space-time derivative of a field, e.g. $\partial_\mu\varphi$, the derivative with respect to φ is defined in a bit bizarre way as[5]

$$\frac{\partial}{\partial\varphi}\partial_\mu\varphi \times \text{something} = -ip_\mu \times \text{something} + \partial_\mu\varphi \times \frac{\partial}{\partial\varphi}\text{something}$$

where p_μ is the momentum (towards the vertex) of the leg produced by this derivative.

Clearly, examples are called for. In our toy-model given above (the so-called φ^3-theory) the non-quadratic part of the Lagrangian contains the third power of the field, so there will be only the 3-leg vertex

$$= \quad i\frac{\partial^3}{\partial\varphi^3}\left(-\frac{1}{3!}\mathring{g}\varphi^3\right) = -i\mathring{g}$$

In QED the non-quadratic part of the Lagrangian is $-\overline{\psi}\mathring{q}\gamma^\mu A_\mu\psi$, leading to the single vertex

$$= \quad i\frac{\partial^3\left(-\mathring{q}\overline{\psi}\gamma^\mu A_\mu\psi\right)}{\partial\overline{\psi}\partial\psi\partial A_\mu} = -i\mathring{q}\gamma^\mu$$

and for purely didactic purposes, let us calculate the vertex for the theory with the non-quadratic Lagrangian given by $-\mathring{g}\varphi^2\partial_\mu\varphi\partial^\mu\varphi$

$$\begin{aligned}
&= \quad i\frac{\partial^4}{\partial\varphi^4}\left(-\mathring{g}\varphi^2\partial_\mu\varphi\partial^\mu\varphi\right)\\
&= -i\mathring{g}\frac{\partial^3}{\partial\varphi^3}\left(2\varphi\partial_\mu\varphi\partial^\mu\varphi - 2i\varphi^2 p_1^\mu\partial_\mu\varphi\right)\\
&= -i\mathring{g}\frac{\partial^2}{\partial\varphi^2}\left(2\partial_\mu\varphi\partial^\mu\varphi - 4i\varphi p_2^\mu\partial_\mu\varphi - 4i\varphi p_1^\mu\partial_\mu\varphi - 2\varphi^2 p_1^\mu p_{2,\mu}\right)\\
&= -i4\mathring{g}\frac{\partial}{\partial\varphi}\left(-ip_3^\mu\partial_\mu\varphi - ip_2^\mu\partial_\mu\varphi - \varphi p_2 p_3 - ip_1^\mu\partial_\mu\varphi - \varphi p_1 p_3 - \varphi p_1 p_2\right)\\
&= 4i\mathring{g}\left(p_1p_2 + p_1p_3 + p_1p_4 + p_2p_3 + p_2p_4 + p_3p_4\right)
\end{aligned}$$

[4] The RHS of this definition could (should) contain a factor $(2\pi)^4\,\delta^4\left(p_1 + p_2 + \ldots + p_n\right)$ where p_i is the momentum corresponding to the i-th leg (all momenta are understood to be pointing towards the vertex). However, we prefer to include this factor elsewhere.

[5] $\frac{\partial}{\partial\varphi}\partial_\mu\varphi$ is by definition equal to $-ip_\mu$, and the Leibniz rule applies to $\frac{\partial}{\partial\varphi}$, as it should apply to anything worthy of the name derivative.

propagators

Propagators are defined by the quadratic part of the Lagrangian. They are negative inverses of the 2-leg vertices with an $i\varepsilon$ adornment (and with momenta p and $p'=-p$ pointing towards the vertex)

$$\text{propagator} = i\left(\left.\frac{\partial^2\mathcal{L}}{\partial\varphi^2}\right|_{\varphi=0,p'=-p} + i\varepsilon\right)^{-1}$$

The symbol ε stands for any positive infinitesimal quantity, therefore we will always replace $\varepsilon\times$ finite quantity by the ε itself. For complex fields one uses $\partial^2\mathcal{L}/\partial\varphi^*\partial\varphi$, definitions for other fields are similar.

The examples below are more than examples, they are universal tools to be used over and over. The point is that the quadratic parts of Lagrangians are the same in almost all theories, so once the propagators are calculated, they can be used in virtually all QFT calculations.

The quadratic part of the scalar field Lagrangian is given by $\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}\mathring{m}^2\varphi^2$, leading to $\partial^2\mathcal{L}/\partial\varphi^2|_{\varphi\equiv0,p'=-p} = -p.p' - \mathring{m}^2|_{p'=-p} = p^2 - \mathring{m}^2$, i.e.

$$= \frac{i}{p^2-\mathring{m}^2+i\varepsilon}$$

The quadratic part of the spinor field Lagrangian is given by $\overline{\psi}\left(i\gamma^\mu\partial_\mu - \mathring{m}\right)\psi$, leading to $\partial^2\mathcal{L}/\partial\overline{\psi}\partial\psi|_{\text{fields}\equiv0,p'=-p} = \gamma^\mu p_\mu - \mathring{m}$, i.e.

$$= \frac{i}{\gamma^\mu p_\mu - \mathring{m} + i\varepsilon} = \frac{i\left(\gamma^\mu p_\mu + \mathring{m}\right)}{p^2-\mathring{m}^2+i\varepsilon}$$

where we have utilized the identity $\left(\gamma^\mu p_\mu - \mathring{m}\right)\left(\gamma^\mu p_\mu + \mathring{m}\right) = p^2 - \mathring{m}^2$, which at this stage is just a God-given identity, allowing us to rewrite the propagator in the standard way with $p^2-\mathring{m}^2+i\varepsilon$ in the denominator.

Finally, for the massless vector field the quadratic Lagrangian is $-\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} - \frac{1}{2\xi}\left(\partial_\alpha A^\alpha\right)^2$ leading to[6,7,8] $\partial^2\mathcal{L}/\partial A_\mu\partial A_\nu|_{\text{fields}\equiv0,p'=-p} = (1-\frac{1}{\xi})p^\mu p^\nu - p^2\eta^{\mu\nu}$ where $\eta^{\mu\nu}$ is the metric tensor

$$= \frac{i}{\left(1-\frac{1}{\xi}\right)p^\mu p^\nu - p^2\eta^{\mu\nu}+i\varepsilon} = \frac{-i\left(\eta_{\mu\nu}-(1-\xi)\,p_\mu p_\nu/p^2\right)}{p^2+i\varepsilon}$$

Surprisingly enough, this is almost everything one would ever need as to the propagators. In the Standard Model, the spinor propagator describes quarks and leptons, the massless vector propagator describes photon and gluons, the scalar propagator describes the Higgs boson. The only missing propagator is the massive vector one, describing the $W^\pm$ and Z^0 bosons. This can be, however, worked out easily from the Lagrangian $-\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} + \frac{1}{2}\mathring{m}^2A_\alpha A^\alpha$ (the result is $-i\left(\eta^{\mu\nu} - p^\mu p^\nu/\mathring{m}^2\right)\left(p^2-\mathring{m}^2+i\varepsilon\right)^{-1}$, the derivation is left as an exercise).

---

[6] For $\mathcal{L} = \frac{1}{2}\left[\left(\partial_\alpha A_\beta\right)\left(\partial^\beta A^\alpha\right) - \left(\partial_\alpha A_\beta\right)\left(\partial^\alpha A^\beta\right) - \frac{1}{\xi}\left(\partial_\alpha A^\alpha\right)\left(\partial_\beta A^\beta\right)\right]$ the derivatives are straightforward
$\frac{\partial\mathcal{L}}{\partial A_\mu} = -i\left(p_\alpha\eta^\mu_\beta\partial^\beta A^\alpha - p_\alpha\eta^\mu_\beta\partial^\alpha A^\beta - \frac{1}{\xi}p_\alpha\eta^{\alpha\mu}\partial_\beta A^\beta\right) = -i\left(p_\alpha\partial^\mu A^\alpha - p_\alpha\partial^\alpha A^\mu - \frac{1}{\xi}p^\mu\partial_\beta A^\beta\right)$
$\frac{\partial^2\mathcal{L}}{\partial A_\mu\partial A_\nu} = -p_\alpha p'^\mu\eta^{\alpha\nu} + p_\alpha p'^\alpha\eta^{\mu\nu} + \frac{1}{\xi}p^\mu p'_\beta\eta^{\beta\nu} = p.p'\eta^{\mu\nu} - p'^\mu p^\nu + \frac{1}{\xi}p^\mu p'^\nu$

[7] To find the matrix inverse to $M^{\lambda\mu} = (1-\xi^{-1})p^\lambda p^\mu - p^2\eta^{\lambda\mu}$ one may either make an educated guess $M^{-1}_{\mu\nu} = A\eta_{\mu\nu} + Bp_\mu p_\nu$ (there is nothing else at our disposal) and solve for A and B, or one may simply check that $\left[(1-\xi^{-1})p^\lambda p^\mu - p^2\eta^{\lambda\mu}\right]\left(-\eta_{\mu\nu} + (1-\xi)\,p_\mu p_\nu/p^2\right)/p^2 = \eta^\lambda_\nu$.

[8] Let us remark that without the term $\frac{1}{2\xi}\left(\partial_\alpha A^\alpha\right)^2$ the propagator would not exist, since the 2-leg vertex would have no inverse. Two specific choices of the parameter ξ are known as the Feynman gauge ($\xi=1$) and the Landau gauge ($\xi=0$).

external legs

The factor corresponding to an external leg is, as a rule, the product of two factors. Let us start with the simpler one. For the scalar field φ (representing a particle with zero spin) this factor is the simplest possible, it equals to 1. For other fields (representing particles with higher spins) there is a nontrivial first factor for each external leg. This factor is different for particles and antiparticles. It also distinguishes between ingoing and outgoing particles (i.e. between the initial and the final state). The factor depends on the particle momentum and spin, but we are not going to discuss this dependence in any detail here.

As to the massless vector field A_μ (e.g. for the photon, where antiparticle = particle) this factor is

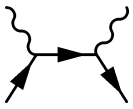
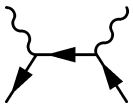
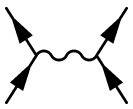
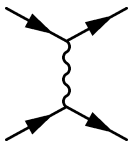
ingoing particle	ε_μ
outgoing particle	ε^*_μ

For the spinor field (e.g. for the electron and positron, which are distinguished in diagrams by the orientation of the arrow) the factor is

ingoing particle	arrow towards the diagram	u
ingoing antiparticle	arrow out of the diagram	$\overline{v}$
outgoing particle	arrow out of the diagram	$\overline{u}$
outgoing antiparticle	arrow towards the diagram	v

These rules are universal, independent of the specific form of the Lagrangian.

Examples for electrons and photons may illuminate the general rules. We will draw diagrams from the left to the right, i.e. ingoing particles (initial state) are on the left and outgoing particles (final state) on the right[9].

process	typical diagram	first external legs factors
$e^-\gamma \to e^-\gamma$		$u, \varepsilon_\mu, \overline{u}, \varepsilon^*_\nu$
$e^+\gamma \to e^+\gamma$		$\overline{v}, \varepsilon_\mu, v, \varepsilon^*_\nu$
$e^+e^- \to e^+e^-$		$\overline{v}, u, v, \overline{u}$
$e^-e^- \to e^-e^-$		$u, u, \overline{u}, \overline{u}$

[9]Note that some authors, including Peskin-Schroeder, draw the Feynman diagrams other way round, namely from the bottom to the top.

Now to the second factor corresponding to an external leg. It has a pretty simple appearance, namely it equals to $\sqrt{Z}$, where Z is a constant (the so-called wave-function renormalization constant) dependent on the field corresponding to the given leg. The definition and calculation of Z are, however, anything but simple.

Fortunately, the dominant part of vast majority of cross-sections and decay rates calculated by means of Feynman diagrams is given by the so-called tree diagrams (diagrams containing no closed loops), and at the tree level the Z constant is always equal to 1. So while staying at the tree level, one can forget about Z completely. And since our first aim is to master the tree level calculations, we can ignore the whole Z-affair until the discussion of loops and renormalization. The following sketch of the Z definition is presented only for the sake of completeness (and can be skipped safely at this moment).

Unlike all other Feynman rules, the Z constant is defined not directly via the Lagrangian, but rather via an infinite sum of Feynman diagrams[10]. The said sum, called the dressed propagator, contains all diagrams with two external legs corresponding to the field under consideration. These two external legs are treated in a specific way — the corresponding factor is not the external leg factor but rather the propagator. The dressed propagator is a function of the external leg momentum (both legs have the same momentum due to the vertex momentum δ-functions) and, as a rule, has a pole in the p^2-variable. The residuum at this pole is the wanted Z.

This definition, as it stands, applies only to the scalar fields. For higher spins the dressed propagator is a matrix and the Z constant is defined via the eigenvalues of this matrix. So one can have, in principle, several different Z constants corresponding to one field. For the electron-positron field, however, there turns out to be only one such constant and the same is true for the photon field.

In addition to this, there is yet another very important ingredient in the external leg treatment. The external leg factor stands not only for the simple (bare) external leg, but rather for the dressed external leg (with all loop corrections). In other words, when calculating a scattering amplitude, one should not include diagrams with loop corrections on external legs. These diagrams are, in a sense, taken into account via the $\sqrt{Z}$ factors[11].

Too complicated? Never mind. Simply forget everything about Z, it will be sufficient to recall it only much later, when dealing with renormalization.

Remark: *As we have indicated, in some circumstances the external leg factor may be even more complicated than the product of two terms (one of them being $\sqrt{Z}$). This happens in presence of non-vanishing sums of all diagrams with two external legs corresponding to* different *fields. This is only rarely the case and always indicates that our choice of fields was not the most appropriate one. The remedy for this trouble is quite ordinary: after a suitable re-definition (just a simple linear combination) of the fields, the trouble simply drops out.*

[10] For the defininition of Feynman diagrams see the next subsection.

[11] After being forced to calculate the loop corrections to a simple line in order to obtain Z, one does not need to calculate them again when calculating the scattering amplitude. There is at least some justice in this world.

1.1.2 Feynman diagrams

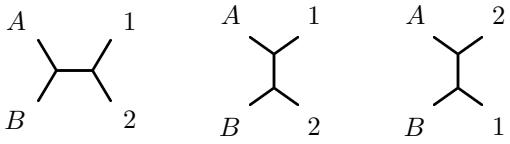
diagrams for a given process contributing at a given order

A process defines external legs, both ingoing and outgoing. A Feynman diagram corresponding to this process is any diagram (graph) with this set of external legs interconnected by the internal lines (propagators) of the theory, via the vertices of the theory, with exception of:

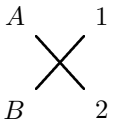
- diagrams with no vertices at all
- diagrams containing so-called "vacuum bubbles", i.e. subdiagrams not connected to any external leg
- diagrams containing so called "corrections on external legs", i.e. subdiagrams with two outgoing lines, one of which is an external leg

There is usually an infinite number of such diagrams. Still, only a finite number contribute at a given order. The order may be defined in at least three different ways, namely as **a**) the number of vertices, **b**) the power of the coupling constant or **c**) the number of (independent) loops. If there is only one type of vertex in the theory, these three definitions are equivalent[12]. If one has more types of vertices, but all characterized by the same coupling constant[13], then the first definition is not used and the other two are not equivalent.

As an example, let us consider a scattering $AB \to 12$, described by either φ^3- or φ^4-theory. At the leading order (the lowest nonzero order, tree level) one has for the φ^3-theory



while for the φ^4-theory $\left(\mathcal{L}\left[\varphi\right] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}\mathring{m}^2\varphi^2 - \frac{1}{4!}\mathring{g}\varphi^4\right)$

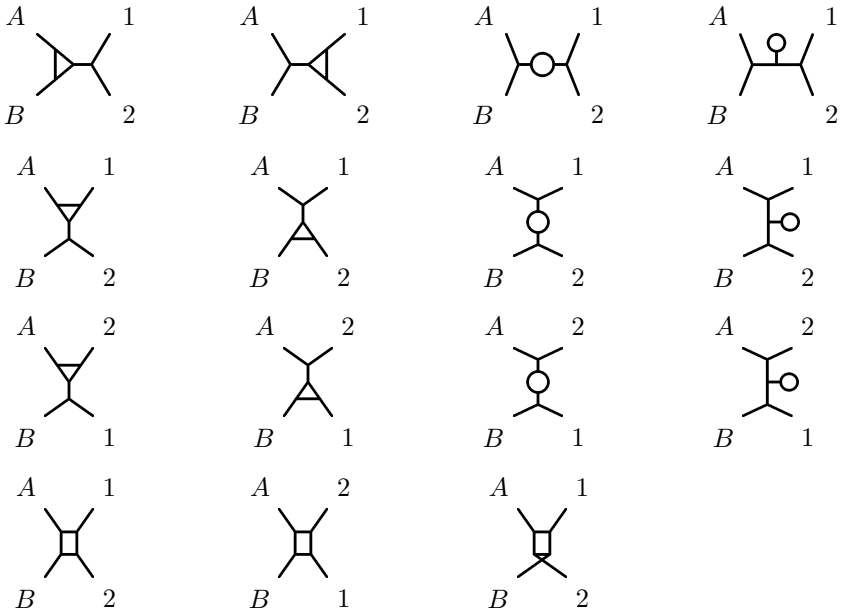


Note that the second and the third diagrams for the φ^3-theory are not equivalent, they contain different vertices (intersections of different lines).

[12]Proof: The equivalence of the first two definitions is evident (every vertex contributes by the same coupling constant). As to the equivalence of the third definition, let us denote the number of vertices, internal lines, external lines and independent loops by V, I, E and L respectively. The famous Euler's Theorem states $V = I - L + 1$. This is to be combined with $nV = 2I + E$ where n is the number of legs of the vertex of the theory. The latter equation is nothing but a simple observation that we can count the number of lines by counting vertices and multiplying their number by n, but we are double-counting internal lines in this way. When combined, the equations give $(n-2)\,V = 2L + E - 2$, i.e. for a given E the dependence of V on L is linear.

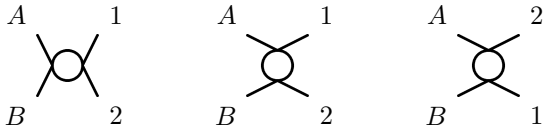
[13]A good example is the so-called scalar electrodynamics (spinless particles and photons) defined by the Lagrangian $\mathcal{L}\left[\varphi\right] = \left(D_\mu\varphi\right)^* D^\mu\varphi - \mathring{m}^2\varphi^*\varphi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2$, where $D_\mu = \partial_\mu + i\mathring{q}A_\mu$.

At the next to leading order (1-loop level) one has for the φ^3-theory

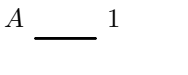


Note that in the last diagram the crossing of external legs B and 2 does not represent a vertex (we just did not manage to draw the diagram in plane without crossing). Let us remark that the 3 diagrams in the last column (so-called tadpole diagrams) are often omitted, since the factors corresponding to these diagrams usually turn out to vanish.

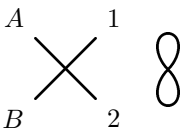
The next to leading order (1-loop level) result for the φ^4-theory is



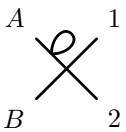
As examples of diagrams not included among Feynman diagrams corresponding to the process under consideration let us mention



no vertices at all



vacuum bubble subdiagram



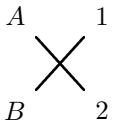
correction on external leg

the factor corresponding to a given diagram

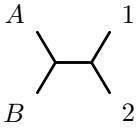
The factor corresponding to a diagram is the product[14] of factors corresponding to all external lines, internal lines and vertices of the diagram, multiplied by

- an extra factor $(2\pi)^4 \delta^4 (p_1 + p_2 + \ldots + p_n)$ for each vertex (p_i is the momentum corresponding to the i-th leg, pointing towards the vertex).
- an extra factor $\int \frac{d^4k}{(2\pi)^4}$ for each propagator (with the four-momentum k)
- an extra so-called combinatorial factor, to be discussed later
- some extra factors of (-1) related to fermionic lines[15]

Examples[16]:

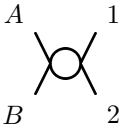


$$= -i\mathring{g}\,(2\pi)^4\,\delta^4\,(p_A + p_B - p_1 - p_2)$$



$$= -\mathring{g}^2\,(2\pi)^8 \int \tfrac{d^4k}{(2\pi)^4} \tfrac{i}{k^2-\mathring{m}^2+i\varepsilon} \delta^4\,(p_A + p_B - k)\,\delta^4\,(k - p_1 - p_2)$$

$$= -i\tfrac{\mathring{g}^2}{(p_A+p_B)^2-\mathring{m}^2+i\varepsilon}\,(2\pi)^4\,\delta^4\,(p_A + p_B - p_1 - p_2)$$

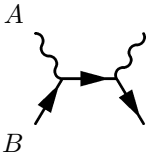


$$= -\mathring{g}^2\,(2\pi)^8\,\tfrac{1}{2}\int \tfrac{d^4k}{(2\pi)^4}\tfrac{d^4k'}{(2\pi)^4}\tfrac{i}{k^2-\mathring{m}^2+i\varepsilon}\tfrac{i}{k'^2-\mathring{m}^2+i\varepsilon}\delta^4\,(p_A + p_B - k - k') \times$$

$$\uparrow$$

$$\text{combinatorial factor} \qquad \times\,\delta^4\,(k + k' - p_1 - p_2)$$

$$= \tfrac{1}{2}\mathring{g}^2 \int d^4k \tfrac{1}{k^2-\mathring{m}^2+i\varepsilon}\tfrac{1}{(p_A+p_B-k)^2-\mathring{m}^2+i\varepsilon}\delta^4\,(p_A + p_B - p_1 - p_2)$$



$$= -\mathring{q}^2\,(2\pi)^8 \int \tfrac{d^4k}{(2\pi)^4}\overline{u}_2\gamma^\mu\delta^4\,(k - p_1 - p_2)\,\tfrac{i\left(\gamma^\lambda k_\lambda+\mathring{m}+i\varepsilon\right)}{k^2-\mathring{m}^2+i\varepsilon}\times$$

$$\times\gamma^\nu\delta^4\,(p_A + p_B - k)\,u_B\varepsilon^*_{1,\mu}\varepsilon_{A,\nu}$$

$$= -i\mathring{q}^2\tfrac{\overline{u}_2\gamma^\mu\left(\gamma^\lambda(p_A+p_B)_\lambda+\mathring{m}\right)\gamma^\nu u_B}{(p_A+p_B)^2-\mathring{m}^2+i\varepsilon}\varepsilon^*_{1,\mu}\varepsilon_{A,\nu}\,(2\pi)^4\,\delta^4\,(p_A + p_B - p_1 - p_2)$$

[14] If individual factors are simple numbers, one does not care about their ordering. In some cases, however, these factors are matrices, and then the proper ordering is necessary. The basic rule here is that for every line with an arrow, the factors are to be ordered "against the arrow", i.e. starting from the end of the arrow and going in the reverse direction.

[15] The factor (-1) for every closed fermionic loop, and the relative minus sign for the diagrams, which can be obtained from each other by an interchange of two fermionic lines. Diagrams related to each other by the omission or addition of boson lines have the same sign.

[16] All momenta in these examples are understood to flow from the left to the right. One can, of course, choose another orientation of momenta and change the signs in δ-functions correspondingly.

the factor corresponding to a given diagram — an alternative formulation

There is another, a bit more economic way, of producing the factor corresponding to a given diagram. When using this method, diagrams are drawn in such a way that all the momentum δ-functions in vertices are satisfied (e.g. in φ^3-theory instead of denoting legs going into the vertex as p_1, p_2, p_3, one denotes the third leg directly as $-p_1 - p_2$). The factor corresponding to a diagram is the product of factors corresponding to all external lines, internal lines and vertices of the diagram, multiplied by

- an overall momentum δ-function factor: $(2\pi)^4\,\delta(P_f - P_i)$
  (P_i and P_f are sums of the ingoing and outgoing momenta respectively)
- an extra factor $\int \frac{d^4k}{(2\pi)^4}$ for each independent loop
- an extra so-called combinatorial factor, to be discussed later
- some extra factors of (-1) related to fermionic lines

The new rules are obtained from the previous ones by performing all trivial integrations over the vertex momentum δ-functions. We will show now that after such integrations are performed one is indeed left with just the one non-integrated δ-function and L(= number of independent loops) integrals yet to be evaluated.

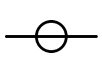
Let us start with any internal line connecting two vertices. After integration over the momentum of this line, one gets rid of one of the vertex δ-functions and the momentum of the line is fixed (in terms of the other momenta entering the vertex with the said δ-function).

The reader may check, that in all the previous examples the results are obtained more directly with this formulation. The new rules are obtained from the previous one by performing all trivial integrations over the vertex δ-functions. To see this, let us ignore all factors corresponding to a diagram except of momentum integrations for internal legs and momentum δ-functions for vertices. Each integration corresponding to an internal line connecting two different vertices can now be depicted as omission of the corresponding internal line and merging two vertices into one vertex. Repeating this procedure over and over, one eventually gets rid of all internal lines connecting two different vertices. So finally one obtains a daisy-like diagram with only one vertex and some loopy internal lines going from this vertex and returning back. The number of these loops is the same, as the number of independent loops in the original diagram (this is due to the Euler's theorem $L = I - V + 1$ and the fact that at each step the numbers I and V decrease by one). The remaining integrals are the loop-integrals mentioned in the alternative formulation. The δ-function corresponding to the last vertex is the overall δ-function. (Convince yourself about the last two statements.)

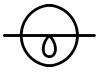
combinatorial factors

Beyond the tree level, a diagram may require the so-called combinatorial factor[17], which is usually the most cumbersome factor to evaluate. Therefore, it seems reasonable to start with some simple rules of thumb:

• If two vertices are connected by n different internal lines, the corresponding combinatorial factor is $1/n!$

 $\frac{1}{3!}$  $\frac{1}{2!2!}$

• If a line starts and ends in the same vertex, it contributes by $1/2$ to the combinatorial factor

 $\frac{1}{2}$  $\frac{1}{2}\frac{1}{2!}$

• If N permutations of n vertices do not change the diagram, the corresponding combinatorial factor is $1/N$ (note that if not all $n!$ permutations leave the diagram untouched, then $N \neq n!$)

 $\frac{1}{2}\frac{1}{3!}$  $\frac{1}{2^3}\frac{1}{3!}$

• The above list is not exhaustive, e.g. it does not allow to find the correct combinatorial factor in the following case

 $\frac{1}{8}$

A systematic, but less illustrative, prescription goes something like this: Let us assign a label to each end of every line in the diagram. The labeled diagram is characterized by sets of labels belonging to the common vertex, and by pairs of labels connected by a line. Permutation of labels would typically lead to a diagram labeled in a different way. Some permutations, however, lead to the identically labeled diagram. Find out the number N of all such permutations (leaving the labeled diagram unchanged). The combinatorial factor of the diagram is $1/N$.

Needless to say, this prescription is not that easy to follow practically. Fortunately, in simple cases (and one seldom needs to go much further) it can be reduced easily to the above rules of thumb. To provide the reader with a systematic procedure of generating all diagrams with right combinatorial factors, we will formulate one such method in the next paragraph. If found too clumsy, the paragraph can be skipped safely.

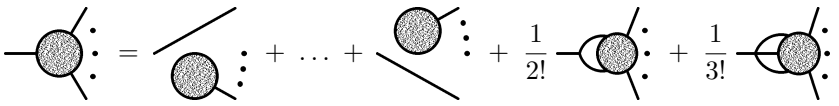
[17] Why such a name: the diagrams represent specific terms of perturbative expansion, the number of terms corresponding to a given diagram is given by some combinatorics.

a systematic way of drawing the diagrams with correct combinatorial factors

Let us consider diagrams with l external legs. We will represent the sum of all such diagrams (including diagrams with no vertices at all, diagrams containing vacuum bubbles and diagrams containing corrections on external legs) by a shaded blob with l legs.

Now let us focus on one of the external legs. What is it connected to? It can be connected either directly to one of the other external legs, or it goes to some m-leg vertex. In the former case, the sum of all diagrams equals to one leg with no vertices at all and the sum of all diagrams with l–2 external legs (a simple line disconnected to a (l–2)-legged shaded blob). In the latter case, the sum of all diagrams contains this leg going to the said vertex and the sum of all diagrams with l+m–2 external legs, m–1 coming from the said vertex and l–1 being true external legs (the vertex connected to a l+m–2 shaded blob). If the external leg can go to different vertices, all of them has to be taken into account.

The combinatorial factors come as follows: If the vertex contains n_1, n_2, ... non-external legs corresponding to the same field, then the combinatorial factor is $1/(n_1!\ n_2!\ \ldots)$. The result is the known as the Dyson-Schwinger equation. For combined φ^3- and φ^4-theories it reads

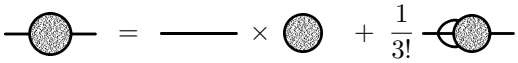


To get all diagrams up to a given order with correct combinatorial factors, the DS equation is used in an iterative way: one starts the equation itself, then one takes any leg and applies the equation to it, then the same is repeated with some other leg etc., until one reaches

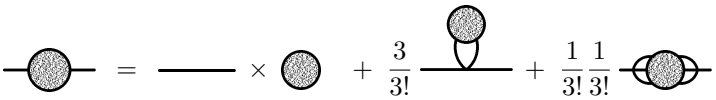
(the structure one is interested in) $\times$  $+\ldots$

with ellipsis standing for diagrams with disconnected external legs + higher orders.

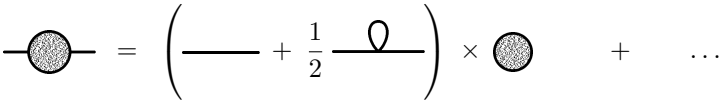
Let us illustrate the procedure by the diagram with two external legs within the φ^4-theory. The starting point is the DS equation for 2 external legs



Now the DS equation is applied to some other leg, say to the 2nd external leg

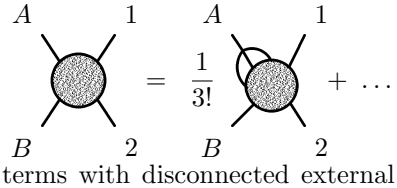


If we are interested only in the 1st order (in the number of vertices), then the last term is already of higher order, and the second term is again processed by the DS equation, to finally give

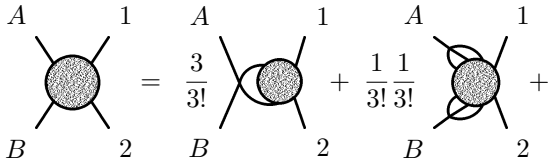


The factor in front of the second diagram in the brackets is the correct combinatorial factor for this diagram. (As an exercise the reader may try to go one order higher.)

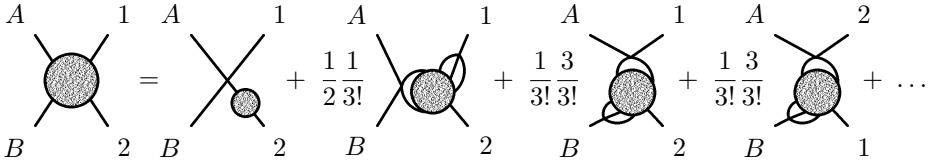
As another example let us consider the $AB \to 12$ scattering within the φ^4-theory. Again, the starting point is the DS equation



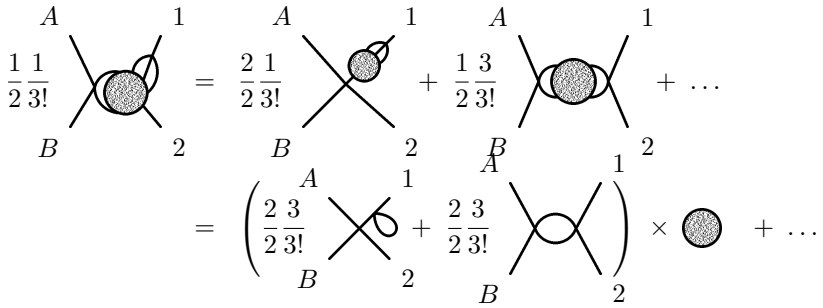
where the ellipsis stands for terms with disconnected external legs. The DS equation is now applied to some other leg, say the external leg B



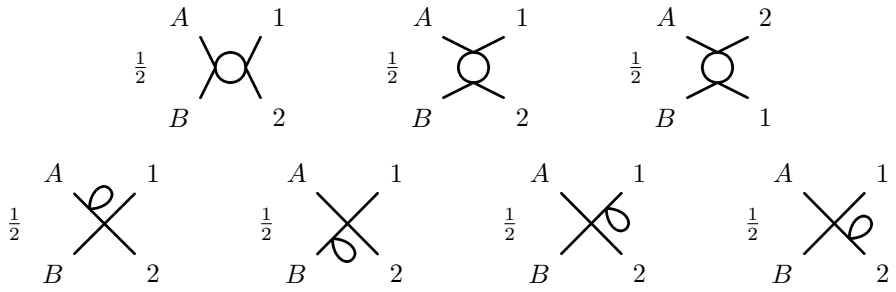
The RHS enjoys yet another DS equation, now say on the external leg 1, to give



The first two (last two) terms on the RHS come from the first (second) diagram on the RHS above, terms with more than two vertices are included into ellipsis. The first diagram is now treated with the help of the previous result for the 2-leg diagrams, the other three are treated all in the same way, which we will demonstrate only for the second diagram: we use the DS equation once more, now say for the external leg 2



Putting the pieces together, one finally obtains the one-loop diagrams with the correct combinatorial factors



1.1.3 Scattering amplitude

The definition of the scattering amplitude M_{fi} is quite simple:

$$\text{the sum of Feynman diagrams} = iM_{fi}\,(2\pi)^4\,\delta^{(4)}\,(P_f - P_i)$$

where P_i and P_f are the overall initial and final momentum respectively. By the sum of the diagrams, the sum of the corresponding factors is meant, of course.

Examples (to be checked by the reader):

- φ^3**-theory,** $AB \to 12$ **scattering**

  tree-level $M_{fi} = -\frac{\mathring{g}^2}{(p_A+p_B)^2-\mathring{m}^2+i\varepsilon} - \frac{\mathring{g}^2}{(p_A-p_1)^2-\mathring{m}^2+i\varepsilon} - \frac{\mathring{g}^2}{(p_A-p_2)^2-\mathring{m}^2+i\varepsilon}$

  1-loop-level the result is intricate and not that illuminating
  but the reader is encouraged to work out some loop diagrams

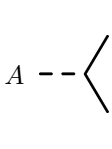
- φ^4**-theory,** $AB \to 12$ **scattering**

  tree-level $M_{fi} = -\mathring{g}$

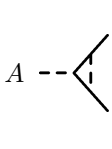
  1-loop-level $M_{fi} = -\frac{1}{2}\mathring{g}^2\left[I\,(p_A+p_B) + I\,(p_A-p_1) + I\,(p_A-p_2)\right]$

  $$I\,(p) = i\int \frac{d^4k}{(2\pi)^4}\,\frac{1}{k^2-\mathring{m}^2+i\varepsilon}\,\frac{1}{(p-k)^2-\mathring{m}^2+i\varepsilon}$$

- $\varphi^2\Phi$**-theory**[18] , $A \to 12$ **decay**

  tree-level 

  $M_{fi} = -\mathring{g}$

  1-loop-level 

  $M_{fi} = -\mathring{g}^3 J\,(p_1,p_2)$

  $$J\,(p_1,p_2) = i\int \frac{d^4k}{(2\pi)^4}\,\frac{1}{k^2-\mathring{M}^2}\,\frac{1}{(p_1+k)^2-\mathring{m}^2+i\varepsilon}\,\frac{1}{(p_2-k)^2-\mathring{m}^2+i\varepsilon}$$

[18] A theory for two different fields φ and Φ, defined by the Lagrangian

$$\mathcal{L}\,[\varphi,\Phi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}\mathring{m}^2\varphi^2 + \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi - \frac{1}{2}\mathring{M}^2\Phi^2 - \frac{\mathring{g}}{2}\varphi^2\Phi$$

1.1.4 Cross-sections and decay rates

The cross-section for a scattering $AB \to 12\ldots n$ is given by

$$d\sigma = (2\pi)^4 \delta^4 (P_f - P_i) \frac{1}{4\sqrt{(p_A.p_B)^2 - m_A^2 m_B^2}} |M_{fi}|^2 \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i}$$

while the analogous quantity for a decay $A \to 12\ldots n$ is

$$d\Gamma = (2\pi)^4 \delta^4 (P_f - P_i) \frac{1}{2E_A} |M_{fi}|^2 \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i}$$

where the so-called width Γ is related to the particle life-time τ by $\Gamma = 1/\tau$.

Because of the δ-function present in these formulae, one can relatively easily perform four integrations on the RHS. For the quite important case of $n = 2$, i.e. for $AB \to 12$ and $A \to 12$, the result after such integrations is[19] in the CMS (centre of mass system)

$$d\sigma_{\text{CMS}} = \frac{1}{64\pi^2} \frac{|\vec{p}_1|}{|\vec{p}_A|} \frac{1}{(p_A + p_B)^2} |M_{fi}|^2 \, d\Omega_1$$

$$d\Gamma_{\text{CMS}} = \frac{1}{32\pi^2} \frac{|\vec{p}_1|}{m_A^2} |M_{fi}|^2 \, d\Omega_1$$

while in the laboratory system (the rest frame of the target particle B)

$$d\sigma_{\text{lab}} = \frac{1}{64\pi^2} \frac{|\vec{p}_1|^2}{|\vec{p}_A|} \frac{1}{m_B} \frac{1}{(E_A + m_B)|\vec{p}_1| - E_1 |\vec{p}_A| \cos\vartheta} |M_{fi}|^2 \, d\Omega_1$$

and the formula for a decay is exactly the same as in the CMS (the two systems coincide in this case). All quantities (energies, momenta, angles) are, of course, understood in the corresponding frames and the non-present (integrated out) δ-functions are understood to be satisfied.

Remark: *It is perhaps worth noticing that before δ-function integrations the cross-section $d\sigma$ is a product of Lorentz scalars (indeed $\delta^4 (P_f - P_i)$, $(p_A.p_B)^2 - m_A^2 m_B^2$, $|M_{fi}|^2$ and $d^3 p_i / 2E_i$ are separately scalars), while $d\Gamma$ is a product of scalars and one non-scalar quantity $1/2E_A$. After the integrations this neat structure is completely lost and this is the reason of the sad fact that one cannot simply and directly translate cross-sections and life-times from one frame to another.*

[19] The specific formulae are obtained from the general ones by three trivial integrations over $d^3 p_2$ leading to $d\sigma = (2\pi)^4 \delta(E_1 + E_2 - E_A - E_B) \frac{1}{4\sqrt{(p_A.p_B)^2 - m_A^2 m_B^2}} |M_{fi}|^2 \frac{d^3 p_1}{(2\pi)^3 2E_1} \frac{1}{(2\pi)^3 2E_2}$ with every $\vec{p}_2$ replaced by $-\vec{p}_1$ (in the CMS). Integration over the last δ-function is usually performed in the spherical coordinates $d^3 p_1 = k^2 dk d\Omega$ where $k = |\vec{p}_1|$. The last δ-function is of the form $\delta(g(k))$ with $g(k) = \sqrt{k^2 + m_1^2} + \sqrt{k^2 + m_2^2} - E_A - E_B$. This function is monotonous and it has therefore only one zero (iff $m_1 + m_2 \leq E_A + E_B$), so one can write $\delta(g(k)) = \delta(k - k_0)/|g'(k_0)|$ where $g(k_0) = 0$. Now $g'(k) = k(E_1 + E_2)/E_1 E_2$ and after some algebra one gets in the CMS $\sqrt{(p_A.p_B)^2 - m_A^2 m_B^2} = |\vec{p}_A|(E_A + E_B)$. All this together lead to the result for $d\sigma_{CMS}$, where $|\vec{p}_1|$ is to be replaced everywhere (including M_{fi}) by the zero point k_0 of the g-function $k_0 = \frac{1}{2}\sqrt{E_{CMS}^2 - 2(m_1^2 + m_2^2) + (m_1^2 - m_2^2)^2 / E_{CMS}^2}$. The result for $d\Gamma_{CMS}$ is obtained in the same way. The calculation of the cross-section in the lab system is just slightly more cumbersome, due to $\vec{p}_2 = \vec{p}_A - \vec{p}_1$.

Examples:[20]

- φ^4-theory, $AB \to 12$ scattering, tree level

$$d\sigma_{\text{CMS}} = \frac{1}{64\pi^2}\frac{1}{s}g^2 d\Omega \qquad\qquad s = (p_A + p_B)^2$$

  In this case the differential cross-section does not depend on angles, so one can immediately write down the total cross-section $\sigma_{\text{CMS}} = g^2/16\pi s$.

- φ^3-theory, $AB \to 12$ scattering, tree level

$$d\sigma_{\text{CMS}} = \frac{g^4}{64\pi^2 s}\left(\frac{1}{s-m^2} + \frac{1}{t-m^2} + \frac{1}{u-m^2}\right)^2 d\Omega \qquad \begin{matrix} t = (p_A - p_1)^2 \\ u = (p_A - p_2)^2 \end{matrix}$$

  where s, t, u are the frequently used so-called Mandelstam variables.

Exercises:

- $AB \to 12$ scattering at the tree level, within "the φ^4-theory with derivatives", i.e. the theory of scalar fields with the non-quadratic part of the Lagrangian $\mathcal{L}_{\text{int}} = -\frac{g}{4}\varphi^2\partial_\mu\varphi\partial^\mu\varphi$.
- $\Phi \to \varphi\varphi$ decay rate, $\varphi\varphi \to \varphi\varphi$, $\varphi\Phi \to \varphi\Phi$ and $\varphi\varphi \to \Phi\Phi$ cross-sections at the tree level, for the $\varphi^2\Phi$-theory defined in the footnote on the page 17.

[20] In final results the limit $\varepsilon \to 0$ is usually understood, therefore there are no $i\varepsilon$ terms in propagators.

1.2 Many-Body Quantum Mechanics

The main characters of QFT are quantum fields or perhaps creation and annihilation operators (since the quantum fields are some specific linear combinations of the creation and annihilation operators). In most of the available textbooks on QFT, the creation and annihilation operators are introduced in the process of the so-called canonical quantization[21]. This, however, is not the most natural way. In opinion of the present author, it may be even bewildering, as it may distort the student's picture of relative importance of basic ingredients of QFT (e.g. by overemphasizing the role of the canonical quantization). The aim of this second introduction is to present a more natural definition of the creation and annihilation operators, and to demonstrate their main virtues.

1.2.1 Fock space, creation and annihilation operators

Fock space

1-particle system
the states constitute a Hilbert space $\mathcal{H}^1$ with an orthonormal basis $|i\rangle$, $i \in N$

2-particle system[22]
the states constitute the Hilbert space $\mathcal{H}^2$ or $\mathcal{H}^2_{\rm B}$ or $\mathcal{H}^2_{\rm F}$, with the basis $|i,j\rangle$

non-identical particles	$\mathcal{H}^2 = \mathcal{H}^1 \otimes \mathcal{H}^1$	$\|i,j\rangle = \|i\rangle \otimes \|j\rangle$
identical bosons	$\mathcal{H}^2_{\rm B} \subset \mathcal{H}^1 \otimes \mathcal{H}^1$	$\|i,j\rangle = \frac{1}{\sqrt{2}} (\|i\rangle \otimes \|j\rangle + \|j\rangle \otimes \|i\rangle)$
identical fermions	$\mathcal{H}^2_{\rm F} \subset \mathcal{H}^1 \otimes \mathcal{H}^1$	$\|i,j\rangle = \frac{1}{\sqrt{2}} (\|i\rangle \otimes \|j\rangle - \|j\rangle \otimes \|i\rangle)$

n-particle system (identical particles)
the Hilbert space is either $\mathcal{H}^n_{\rm B}$ or $\mathcal{H}^n_{\rm F} \subset \underbrace{\mathcal{H}^1 \otimes \ldots \otimes \mathcal{H}^1}_{n}$, with the basis

$$|i,j,\ldots,k\rangle = \frac{1}{\sqrt{n!}} \sum_{\text{permutations}} (\pm 1)^p \underbrace{|i\rangle \otimes |j\rangle \otimes \ldots \otimes |k\rangle}_{n}$$

where p is the parity of the permutation, the $\begin{smallmatrix}\text{upper}\\ \text{lower}\end{smallmatrix}$ sign applies to $\begin{smallmatrix}\text{bosons}\\ \text{fermions}\end{smallmatrix}$

0-particle system
1-dimensional Hilbert space $\mathcal{H}^0$ with the basis vector $|0\rangle$ (no particles, vacuum)

Fock space
direct sum of the bosonic or fermionic n-particle spaces

$$\mathcal{H}_{\rm B} = \bigoplus_{n=0}^{\infty} \mathcal{H}^n_{\rm B} \qquad\qquad \mathcal{H}_{\rm F} = \bigoplus_{n=0}^{\infty} \mathcal{H}^n_{\rm F}$$

[21] There are exceptions. In the Weinberg's book the creation and annihilation operators are introduced exactly in the spirit we are going to adopt in this section. The same philosophy is to be found in some books on many-particle quantum mechanics. On the other hand, some QFT textbooks avoid the creation and annihilation operators completely, sticking exclusively to the path integral formalism.

[22] This is the keystone of the whole structure. Once it is really understood, the rest follows smoothly. To achieve a solid grasp of the point, the reader may wish to consult the couple of remarks following the definition of the Fock space.

Remark: *Let us recall that for two linear spaces U (basis e_i, dimension m) and V (basis f_j, dimension n), the direct sum and product are linear spaces $U \oplus V$ (basis generated by both e_i and f_j, dimension $m+n$) and $U \otimes V$ (basis generated by ordered pairs (e_i, f_j), dimension $m.n$).*

Remark: *The fact that the Hilbert space of a system of two non-identical particles is the direct product of the 1-particle Hilbert spaces may come as not completely obvious. If so, it is perhaps a good idea to start from the question what exactly the 2-particle system is (provided that we know already what the 1-particle system is). The answer within the quantum physics is not that straightforward as in the classical physics, simply because we cannot count the quantum particles directly, e.g. by pointing the index finger and saying one, two. Still, the answer is not too complicated even in the quantum physics. It is natural to think of a quantum system as being 2-particle iff*
a) it contains states with sharp quantum numbers (i.e. eigenvalues of a complete system of mutually commuting operators) of both 1-particle systems, and this holds for all combinations of values of these quantum numbers
b) such states constitute a complete set of states
This, if considered carefully, is just the definition of the direct product.

Remark: *A triviality which, if not explicitly recognized, can mix up one's mind: $\mathcal{H}^1 \cap \mathcal{H}^2 = \emptyset$, i.e. the 2-particle Hilbert space contains no vectors corresponding to states with just one particle, and vice versa.*

Remark: *The fact that multiparticle states of identical particles are represented by either completely symmetric or completely antisymmetric vectors should be familiar from the basic QM course. The former case is called bosonic, the latter fermionic. In all formulae we will try, in accord with the common habit, to treat these two possibilities simultaneously, using symbols like $\pm$ and $\mp$, where the upper and lower signs apply to bosons and fermions respectively.*

Remark: *As to the basis vectors, our notation is not the only possible one. Another widely used convention (the so-called occupation number representation) denotes the basis vectors as $|n_1, n_2, \ldots\rangle$, where n_i is the number of particles in the i-th 1-particle state. So e.g. $|2,2,2,4,4\rangle \Leftrightarrow |0,3,0,2,0,0,0,\ldots\rangle$, where the LHS is in our original notation while the RHS is in the occupation number representation. The main drawback of the original notation is that it is not unique, e.g. $|1,2,3\rangle$ and $\pm|1,3,2\rangle$ denotes the same vector. One should be therefore careful when summing over all basis states. The main drawback of the occupation number representation is typographical: one cannot write any basis vector without the use of ellipsis, and even this may sometimes become unbearable (try e.g. to write $|49,87,642\rangle$ in the occupation number representation).*

Remark: *The basis vectors $|i,j,\ldots,k\rangle$ or $|n_1,n_2,\ldots\rangle$ are not all normalized to unity (they are, but only if all $i,j,\ldots,k$ are mutually different, i.e. if none of n_i exceeds 1). If some of the $i,j,\ldots,k$ are equal, i.e. if at least one $n_i > 1$, then the norm of the fermionic state is automatically zero (this is the Pauli exclusion principle), while the norm of the bosonic state is $\sqrt{n_1!\, n_2!\, \ldots}$. Prove this.*

Remark: *A triviality which, if not explicitly recognized, can mix up one's mind: the vacuum $|0\rangle$ is a unit vector which has nothing to do with the zero vector 0.*

creation and annihilation operators

Let $|i\rangle$ $(i = 1, 2, \ldots)$ be an orthonormal basis of a 1-particle Hilbert space, and $|0\rangle$, $|i\rangle$, $|i, j\rangle$, $|i, j, k\rangle$, $\ldots$ $(i \leq j \leq k \leq \ldots)$ an orthogonal basis of the Fock space. The creation and annihilation operators are defined as follows

creation operator a_i^+
is a linear operator, which maps the n-particle basis vector to the $(n+1)$-particle vector by adding one particle in the i-th state (the particle is added at the first position in the resulting vector; for bosons this rule does not matter, for fermions it determines the sign)

$$a_i^+ |0\rangle = |i\rangle \qquad a_i^+ |j\rangle = |i, j\rangle \qquad a_i^+ |j, k, \ldots\rangle = |i, j, k, \ldots\rangle$$

annihilation operator a_i
is a linear operator, which maps the n-particle basis vector to the $(n-1)$-particle vector by removing one particle in the i-th state. The particle is removed from the first position of the original vector, and if it is not there, the original vector must be reshuffled (for bosons this rule does not matter, for fermions it determines the sign). If the original vector contains more than one particle in the i-th state, the whole procedure is performed with each of them and the results are summed up. If the original vector does not contain a particle in the i-th state, the result is the zero vector.

$$a_i |0\rangle = 0 \qquad a_i |j\rangle = \delta_{ij} |0\rangle$$
$$a_i |j, k, l \ldots\rangle = \delta_{ij} |k, l, \ldots\rangle \pm \delta_{ik} |j, l, \ldots\rangle + \delta_{il} |j, k, \ldots\rangle \pm \ldots$$

Both creation and annihilation operators are linear and they are defined on the basis vectors. Consequently they are defined for any vector.

Remark: *In the occupation number representation, the definitions read*

bosons
$$a_i^+ |n_1, \ldots, n_i, \ldots\rangle = |n_1, \ldots, n_i + 1, \ldots\rangle$$
$$a_i |n_1, \ldots, n_i, \ldots\rangle = n_i |n_1, \ldots, n_i - 1, \ldots\rangle$$

fermions
$$a_i^+ |n_1, \ldots, n_i = 0, \ldots\rangle = (-1)^{p_i} |n_1, \ldots, n_i = 1, \ldots\rangle$$
$$a_i^+ |n_1, \ldots, n_i = 1, \ldots\rangle = 0$$
$$a_i |n_1, \ldots, n_i = 0, \ldots\rangle = 0$$
$$a_i |n_1, \ldots, n_i = 1, \ldots\rangle = (-1)^{p_i} |n_1, \ldots, n_i = 0, \ldots\rangle$$
$$p_i = \textstyle\sum_{k=1}^{i-1} n_k$$

Creation and annihilation operators are very useful, because
- they enable the most natural description of processes in which the number of particles is not conserved, i.e. in which particles are created and/or destroyed
- any linear operator can be expressed in terms of the creation and annihilation operators, namely as a sum of products of these operators
- there is a standard and routine method of how to calculate matrix elements of operators expressed in terms of the creation and annihilation operators.

In view of how frequent the processes of particle creation and annihilation are (decays and inelastic scatterings in the atomic, nuclear, subnuclear and solid state physics), the first point is evidently very important. And in view of how often the QM calculations are just the calculations of various matrix elements of linear operators, the other two points are clearly also very important.

key attributes of the creation and annihilation operators

Perhaps the three most important are[23]

- $a_i^+ = a_i^\dagger$ i.e. a_i^+ and a_i are Hermitian conjugated
- $a_i^+ a_i$ (no summation) is the operator of number of particles in the i-th state (and consequently $\sum_i a_i^+ a_i$ is the operator of the overall number of particles)
- $\left[a_i, a_j^+\right]_\mp = \delta_{ij} \qquad [a_i, a_j]_\mp = \left[a_i^+, a_j^+\right]_\mp = 0 \qquad$ where $[x,y]_\mp = xy \mp yx$

The proof is an easy and very useful exercise, recommended to anybody who wants to become quickly accustomed to elementary manipulations with the a_i^+, a_i operators. The following sketch of the proof is therefore intended only as a check of reader's own work (the proof is performed in the occupation number formalism, which is more convenient for this purpose).

- Hermitian conjugation

$$\begin{aligned}
\langle \ldots n_i' \ldots |\, a_i \,| \ldots n_i \ldots \rangle &= \langle \ldots n_i' .. | . n_i - 1 \ldots \rangle \, n_i = \langle \ldots | \ldots \rangle \, (n_i - 1)! \, n_i \, \delta_{n_i', n_i - 1} \\
\langle \ldots n_i \ldots |\, a_i^+ \,| \ldots n_i' \ldots \rangle &= \langle \ldots n_i .. | . n_i' + 1 \ldots \rangle \quad = \langle \ldots | \ldots \rangle \, n_i! \, \delta_{n_i, n_i' + 1}
\end{aligned}$$

where for bosons $n_i, n_i' \in N$ and for fermions $n_i, n_i' \in \{0, 1\}$

- particle number operator

bosons

$$a_i^+ a_i \,| \ldots n_i \ldots \rangle = a_i^+ n_i \,| \ldots n_i - 1 \ldots \rangle = n_i a_i^+ \,| \ldots n_i - 1 \ldots \rangle = n_i \,| \ldots n_i \ldots \rangle$$

fermions

$$\begin{aligned}
a_i^+ a_i \,| \ldots 0 \ \ldots \rangle &= 0 \\
a_i^+ a_i \,| \ldots 1 \ \ldots \rangle &= a_i^+ \, (-1)^{p_i} \,| \ldots 0 \ldots \rangle = (-1)^{2p_i} \,| \ldots 1 \ldots \rangle
\end{aligned}$$

- (anti)commutation relation

bosons

$$\begin{aligned}
\left[a_i, a_i^+\right] | \ldots n_i \ldots \rangle &= a_i \,| \ldots n_i + 1 \ldots \rangle - n_i a_i^+ \,| \ldots n_i - 1 \ldots \rangle \\
&= (n_i + 1) \,| \ldots n_i \ldots \rangle - n_i \,| \ldots n_i \ldots \rangle = | \ldots n_i \ldots \rangle \\
\left[a_i, a_j^+\right] | \ldots n_i \ldots n_j \ldots \rangle &= a_i \,| \ldots n_i \ldots n_j + 1 \ldots \rangle - a_j^+ \,| \ldots n_i - 1 \ldots n_j \ldots \rangle \\
&= | \ldots n_i - 1 \ldots n_j + 1 \ldots \rangle - | \ldots n_i - 1 \ldots n_j + 1 \ldots \rangle = 0
\end{aligned}$$

fermions

$$\begin{aligned}
\left\{a_i, a_i^+\right\} | \ldots 1 \ldots \rangle &= 0 + (-1)^{p_i} \, a_i^+ \,| \ldots 0 \ldots \rangle = (-1)^{2p_i} \,| \ldots 1 \ldots \rangle = | \ldots 1 \ldots \rangle \\
\left\{a_i, a_i^+\right\} | \ldots 0 \ldots \rangle &= (-1)^{p_i} \, a_i \,| \ldots 1 \ldots \rangle + 0 = (-1)^{2p_i} \,| \ldots 0 \ldots \rangle = | \ldots 0 \ldots \rangle \\
\left\{a_i, a_j^+\right\} | \ldots 0 \ldots 0 \ldots \rangle &= (-1)^{p_j} \, a_i \,| \ldots 0 \ldots 1 \ldots \rangle + 0 = 0 \\
\left\{a_i, a_j^+\right\} | \ldots 0 \ldots 1 \ldots \rangle &= 0 \\
\left\{a_i, a_j^+\right\} | \ldots 1 \ldots 0 \ldots \rangle &= (-1)^{p_i + p_j} \,| \ldots 0 \ldots 1 \ldots \rangle + (-1)^{p_i + p_j - 1} \,| \ldots 0 \ldots 1 \ldots \rangle = 0 \\
\left\{a_i, a_j^+\right\} | \ldots 1 \ldots 1 \ldots \rangle &= 0 + (-1)^{p_i} \, a_j^+ \,| \ldots 0 \ldots 1 \ldots \rangle = 0
\end{aligned}$$

The other (anti)commutation relations are treated in the same way.

[23] Note that a_i^+ and a_i operators could be (and often are) introduced in the reversed order. In that case, the (anti)commutation relations are postulated and the Fock space is constructed afterwards for a_i^+ and a_i to have something to live in. It is perhaps just a matter of taste, but the present author strongly prefers the "more natural logic" of this section. Later in these lectures, however, we will encounter also the reversed logic of the canonical quantization.

creation and annihilation vs. raising and lowering operators

The reader is perhaps already familiar with operators a^+ and a satisfying the above commutation relations. Such operators are introduced for the LHO (linear harmonic oscillator) in every basic QM course[24]. What is the relation between the LHO raising and lowering operators (commonly known also as ladder operators) and our creation and annihilation operators? Three important points are to be emphasized:

1. The creation and annihilation operators have (in principle) nothing to do with the LHO ladder operators. They are naturally defined (as we have seen) in the Fock space without any reference (or even presence) of harmonic oscillators.

2. The definition of raising and lowering operators (with the $m\omega$ factor replaced by an arbitrary real number λ) applies to any 1-particle system, not only to the LHO. Indeed, for a^+ and a defined as

$$a^+ = \sqrt{\tfrac{\lambda}{2\hbar}}\ \hat{x} - \tfrac{i}{\sqrt{2\hbar\lambda}}\ \hat{p} \qquad\qquad a = \sqrt{\tfrac{\lambda}{2\hbar}}\ \hat{x} + \tfrac{i}{\sqrt{2\hbar\lambda}}\ \hat{p}$$

   one can show (just like in the case of LHO) that

   (a) they are hermitian conjugate to each other
   (b) the canonical commutation relation $[\hat{x},\hat{p}] = i\hbar$ implies $[a, a^+] = 1$
   (c) the eigenvalues of the operator $\hat{N} = a^+a$ are natural numbers plus zero

   Let us stress that the definition of these operators have nothing to do with the LHO. What makes the LHO special in this respect is the Hamiltonian, which is exceptionally simple in terms of a^+, a. But ladder operators can be useful for any system. Naturally, they are most useful for systems "close to the LHO", where the difference in Hamiltonians can be treated as a small perturbation.

3. In spite of the first two points the creation and annihilation operators are usually closely related to the ladder operators. The reason is twofold.

   (a) The LHO ladder operators can be viewed as the creation and annihilation operators of some formal particles in some bizarre states. The point is that LHO is formally equivalent to the ideal gas of arbitrary number of formal particles, all of which can be, however, in just one state. This is a simple consequence of equidistant spectrum of the LHO Hamiltonian (the n-th excited state of the LHO can be treated as a state of n formal particles with equal energies)
   (b) Any ideal gas is in some sense equivalent to a system of harmonic oscillators. The point is that when expressed in terms of creation and annihilation operators, the Hamiltonian of an ideal gas is equivalent to the Hamiltonian of several formal harmonic oscillators. This will follow from the explicit form of the ideal gas Hamiltonian which we shall derive in the next section.

Due to the point 3b the LHO ladder operators play an important (even if only an auxiliary) role in one particular development of QFT, namely in the canonical quantization of classical fields (we will learn a lot about it later on). Since this is perhaps the most common development of the theory, the role of the LHO ladder operators can be easily overestimated. Let us therefore stress once more that creation and annihilation operators do not need any mention of LHO whatsoever.

[24] Recall $a^+ = \hat{x}\sqrt{m\omega/2\hbar} - i\hat{p}/\sqrt{2\hbar m\omega}$, $a = \hat{x}\sqrt{m\omega/2\hbar} + i\hat{p}/\sqrt{2\hbar m\omega}$ and they are by definition conjugated to each other. The canonical commutation relation $[\hat{x},\hat{p}] = i\hbar$ implies for a^+ and a the commutation relation of the creation and annihilation operators and the eigenvalues of the operator $N = a^+a$ are natural numbers plus zero (this follows from the commutation relation).

Remark: *It may come as a kind of miracle that the specific complex linear combinations of $x(t)$ and $p(t)$ solve the spectrum of the LHO Hamiltonian so efficiently (as everybody knows from an elementary QM course). In quantum mechanics these linear combinations come out of thin air and they are, frankly speaking, quite mysterious. But at the classical level they are not as artificial as they may appear at the first sight. It is quite common to write down the solution of the classical equation of motion for LHO in the complex form as $x(t) = Ae^{-i\omega t} + Be^{i\omega t}$ and $p(t) = -im\omega(Ae^{-i\omega t} - Be^{i\omega t})$ (both $x(t)$ and $p(t)$ are in general complex, but if one starts with real initial conditions, then $B = A^*$, and they remain real forever). Now one sees that linear combinations $x(t) + \frac{i}{m\omega}p(t) = 2Ae^{-i\omega t}$ and $x(t) - \frac{i}{m\omega}p(t) = 2Be^{i\omega t}$ are just two independent classical solutions. It is therefore not so much surprising that the linear combinations inspired by the classical solutions (even if rescaled by the factor $\sqrt{m\omega/2\hbar}$)*[25] *solve the problem relatively quickly also at the quantum level.*

Remark: *The previous remark did not investigate any deep connection between the classical and quantum physics, it was just a rough classical inspiration of the notoriously known treatment of the quantum LHO. We will have more to say about formal connections between the classical and quantum physics,*[26] *but it is worth emphasizing that at this point our discussion is purely quantum. The same applies to the following remark on phonons, which is again purely quantum (in spite of the fact that many authors start the discussion of phonons at the classical level).*

Remark: *An ideal gas of formal particles, which arises more or less naturally in discussion of LHO, is even more appealing in case of coupled harmonic oscillators. And this is indeed a very important case, due to the famous miracle of systems in the vicinity of their stable equilibriums: any such system is well approximated by the system of coupled harmonic oscillators which, in turn, is equivalent to the system of* decoupled *harmonic oscillators.*
Stationary states of the system of the independent harmonic oscillators are characterized by the sequence $(N_1, N_2, \ldots)$ where N_n defines the energy level of the n-th oscillator. Energies of individual oscillators are $\hbar\omega_n(N_n + 1/2)$. Energy of the system is $\sum_n \hbar\omega_n(N_n + 1/2)$.
Now let us imagine a system of free particles, each of them having energy eigenstates labeled by n with eigenvalues $\hbar\omega_n$. If there are N_n particles in the n-th state, the energy of the system will be $\sum_n \hbar\omega_n N_n$. This is equivalent (up to a constant) to the Hamiltonian of independent oscillators. It is common habit to describe a system of independent harmonic oscillators in terms of the equivalent system of free formal particles. These formal particles are called phonons.
Phonons are would-be particles (often called quasiparticles or collective excitations) widely used for the formal description of real particles behaving as coupled harmonic oscillators (e.g. nuclei or positive ions in a crystal lattice). These formal particles may look like real ones, especially in systems with translational symmetry. In such systems decoupling of oscillators is provided by the Fourier transformation, which brings (quasi)momentum in the game.[27] *In that case phonons behave like having well defined energy and (quasi)momentum. Nevertheless, phonon is not a kind of particle. Strictly speaking, it is just a word.*

[25] In this section we are going to use explicit $\hbar$, just to make comparison with standard QM textbooks easier.

[26] At the end of this section we will discuss an important issue of the classical limit of a quantum theory. The whole second chapter will be, on the other hand, devoted to the reverse procedure, namely to the canonical technique of obtaining a quantum theory from the given classical one. But to appreciate the logic of this Introduction, it is important to be aware of the fact that no classical physics is involved.

[27] Many remarks in this second introduction presume a reader with a basic knowledge of the solid state theory. For those readers who are not familiar with the quoted notions, the Appendix **??** is intended to fill the gap at least to some minimal extend.

creation and annihilation operators in different bases

So far our definition of the creation and annihilation operators was based on the specific choice of basis in the 1-particle Hilbert space $\mathcal{H}^1$. Have we started from some other orthonormal basis, we would get a different set of creation and annihilation operators. What is the relation between the two sets of these operators?

Let $|\alpha\rangle$ ($\alpha \in N$) be an orthonormal basis of $\mathcal{H}^1$, different from our original basis $|i\rangle$. The new basis vectors can be expressed in terms of the old ones as $|\alpha\rangle = |i\rangle\,\langle i|\alpha\rangle$ (with the Einstein summation convention understood). Since the a^+_α operator acts by adding the particle in the α-state and the state $|\alpha\rangle$ is the specific superposition of the states $|i\rangle$, the a^+_α creation operator has to be the same linear combination of the a^+_i creation operators

$$a^+_\alpha = \sum_i \langle i|\alpha\rangle\, a^+_i \qquad\qquad a_\alpha = \sum_i \langle\alpha|i\rangle\, a_i$$

(the relation for the annihilation operators was obtained simply by hermitian conjugation). These relations become handy whenever a need arises for a switch between different bases.

Another point worth discussion is that of continuous bases. Let us recall that the generalized basis vectors $|\vec{x}\,\rangle$ or $|\vec{p}\rangle$ are not elements of the 1-particle Hilbert space, since they are not properly normalized. The generalized normalization condition for the orthonormal x-basis reads $\langle\vec{x}\,|\vec{x}\,'\rangle = \delta^3(\vec{x}-\vec{x}\,')$ and any vector $|\psi\,\rangle$ from $\mathcal{H}^1$ can be written as $|\psi\,\rangle = \int d^3x\ \ |\vec{x}\,\rangle\,\langle\vec{x}\,|\psi\,\rangle$. All the relations valid for discrete orthonormal bases hold also for continuous orthonormal bases after the Kronecker deltas are replaced by Dirac deltas and sums by integrals

$$\delta_{ij} \longrightarrow \delta^3(\vec{x}-\vec{x}\,') \qquad\qquad \sum_i \longrightarrow \int d^3x$$

The basic relations for the creation and annihilation operators in the x-representation are

- $a^+(\vec{x}) = a^\dagger(\vec{x})$ i.e. $a^+(\vec{x})$ and $a(\vec{x})$ are Hermitian conjugated
- for normalized states from the Fock space $a^+(\vec{x}\,)a(\vec{x})$ is not the operator of number of particles with the position $\vec{x}$, but rather the operator of density of particles at position $\vec{x}$ (consequently $\int d^3x\ a^+(\vec{x})\,a(\vec{x})$ is the operator of the overall number of particles)
- $[a(\vec{x}), a^+(\vec{x}\,')]_\mp = \delta^3(\vec{x}-\vec{x}\,') \qquad [a(\vec{x}), a(\vec{x}\,')]_\mp = [a^+(\vec{x}), a^+(\vec{x}\,')]_\mp = 0$

The important relations between the creation and annihilation operators in the x-representation and p-representations are just continuous versions of the discrete relations given above

$$a^+_{\vec{p}} = \int d^3x\ \ \langle\vec{x}\,|\vec{p}\rangle\, a^+_{\vec{x}} \qquad\qquad a_{\vec{p}} = \int d^3x\ \ \langle\vec{p}\,|\vec{x}\,\rangle\, a_{\vec{x}}$$

where

$$\langle\vec{x}\,|\vec{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}}\, e^{i\vec{p}.\vec{x}/\hbar}$$

for orthonormal[28] p-representation basis, i.e. for $\langle\vec{p}|\vec{p}\,'\rangle = \delta^3(\vec{p}-\vec{p}\,')$. For this basis the relations for the operators $a^+(\vec{p})$, $a(\vec{p})$ are completely analogous to the above relations for $a^+(\vec{x})$, $a(\vec{x})$.

[28] Let us remark that there is a quite common habit to use a specific unnormalized basis in the p-representation, namely the one where $\langle\vec{p}|\vec{p}\,'\rangle = (2\pi\hbar)^3\delta^3(\vec{p}-\vec{p}\,')$. For this basis $\langle\vec{x}\,|\vec{p}\rangle = e^{i\vec{p}.\vec{x}/\hbar}$, the basic (anti)commutation relation reads $[a(\vec{p}), a^+(\vec{p}\,')]_\mp = (2\pi\hbar)^3\,\delta^3(\vec{p}-\vec{p}\,')$ and the operator of density of particles with the momentum $\vec{p}$ is $(2\pi\hbar)^{-3}a^+(\vec{p})a(\vec{p})$.

Remark: *Orthonormality of the basis $|i\rangle$ in the 1-particle Hilbert space $\mathcal{H}^1$ played a crucial role in the definition of the specific non-normalized orthogonal basis in the Fock space and in the definition of the creation and annihilation operators. Important features of these operators were consequences of the specific normalization of this Fock space basis. So why should anyone use unnormalized basis in $\mathcal{H}^1$?*
The point is that, as we have mentioned several times already, the $a^+(\vec{p})$ and $a(\vec{p})$ operators are usually introduced via the so-called canonical quantization, where the starting point is some classical theory. At the classical level, the Fourier transformation is involved, bringing the function $e^{i\vec{p}.\vec{x}}$ into the game. Now the notoriously known factor of $(2\pi)^3$ can be used at different places in the definition of the Fourier transformation. And the commonly used choice leads, after quantization, to the non-unit normalization of the basis vectors in the p-representation.

Remark: *When speaking about different bases in the Fock space, yet another issue should be mentioned. When dealing with various types of particles, one needs a Fock space which contains the Fock spaces of all particle species under consideration. The obvious first choice is the direct product of the corresponding Fock spaces, e.g. $H = H_A \otimes H_B \otimes H_C$. In such a case any creation/annihilation operator for one particle type commutes with any c/a operator for any other particle type. Sometimes, however, different particle species may be viewed just as different states of the same particle (due to isospin, eightfold way, etc. symmetries). If so, it is clearly favorable to have a basis and the corresponding (anti)commutation rules which do not need a radical modification with every change of viewpoint. This is achieved by the appropriate choice of the (anti-)symmetrized subspace of the direct product of some of the Fock subspaces, i.e. by the appropriate (anti-)symmetrization of bases of these subspaces, leading to change of some commutation rules by anti-commutation ones.*

Remark: *On top of the creation and annihilation operators, one can encounter yet another – completely different – set of similar operators. The point is that the basis $|i, j, \ldots\rangle$ (or $|n_1, n_2, \ldots\rangle$ in the occupation number formalism), which arises from a particular basis $|i\rangle$ of the 1-particle Hilbert space, is perhaps the most natural, but not the only reasonable basis in the Fock space. Actually, any complete set of (physically relevant) commuting operators defines some (relevant) basis. (From this point of view, the basis $|n_1, n_2, \ldots\rangle$ is just the basis of eigenvectors of the occupation number operators.)*
If the eigenvalues of a complete system of commuting operators are discrete and bounded from below, then one can label both eigenvalues and eigenvectors by natural numbers. In such a case, the basis defined by the considered system of operators looks like $|N_1, N_2, \ldots\rangle$, and for a basis of this type we can define the so-called raising and lowering operators, just as we have defined the creation and annihilation operators: A_i^+ (A_i) raises (lowers) the quantum number N_i by one.
Of a special interest are the cases when the Hamiltonian can be written as a sum of two terms, one of which has eigenvectors $|N_1, N_2, \ldots\rangle$ with eigenvalues $E_1N_1 + E_2N_2 + \ldots$ and the second one can be understood as a small perturbation. If so, the system formally looks like an almost ideal gas (see the next section) made of a new type of particles (created by the A_i^+ operators from the state in which all the N_i vanish). These formal particles are not to be mistaken for the original particles, which the system consists of.
It may come as a kind of surprise that such formal particles do appear frequently in many-body systems. They are called elementary excitations and they come in great variety (phonons, plasmons, magnons, etc.). Their relation to the original particles is more or less known as the result of either detailed calculations, or an educated guess, or some combination of the two. The description of the system is, as a rule, much simpler in terms of the elementary excitations than in terms of the original particles. This explains the wide use of the elementary excitations language by both theorists and experimentalists.

1.2.2 Important operators expressed in terms of a_i^+, a_i

As already announced, any linear operator in the Fock space can be written as a polynomial (perhaps infinite) in creation and annihilation operators. However interesting this general statement may sound, the particular examples are even more interesting and very important in practice. We will therefore start with the examples and return to the general statement only later on.

Hamiltonian of a system of non-interacting particles

Let us consider non-interacting particles (ideal gas) in an external classical field with a potential energy $U(x)$. The most suitable choice of basis in the 1-particle Hilbert space is the set of eigenstates of the 1-particle Hamiltonian $\hat{p}^2/2m + U(x)$, i.e. the states $|i\rangle$ satisfying

$$\left(\frac{1}{2m}\hat{p}^2 + U(x)\right)|i\rangle = E_i\,|i\rangle$$

By this choice, the standard basis of the whole Fock space is determined, namely $|0\rangle$, $|i\rangle$, $|i,j\rangle$, $|i,j,k\rangle$, etc. And since the particles do not interact, each of these basis states has a sharp value of energy, namely 0, E_i, E_i+E_j, $E_i+E_j+E_k$, etc., respectively. The Hamiltonian of the system with any number of particles is the linear operator with these eigenvectors and these eigenvalues. It is very easy to guess such an operator, namely $H_0 = \sum_i E_i \hat{n}_i$, where $\hat{n}_i$ is the operator of number of particles in the i-th state. And since we know how to express $\hat{n}_i$ in terms of the creation and annihilation operators, we are done

$$H_0 = \sum_i E_i\; a_i^+ a_i$$

- If, for any reason, we would need to express H_0 in terms of another set of creation and annihilation operators $a_\alpha^+ = \sum_i \langle i|\alpha\rangle\, a_i^+$ and $a_\alpha = \sum_i \langle\alpha|i\rangle\, a_i$, it is straightforward to do so: $H_0 = \sum_{\alpha,\beta} E_{\alpha\beta} a_\alpha^+ a_\beta$ where $E_{\alpha\beta} = \sum_i E_i \langle\alpha|i\rangle\langle i|\beta\rangle$.
- If one has a continuous quantum number q, rather than the discrete index i, then (as we have already discussed) the sum $\sum_i$ is replaced by the integral $\int dq$: $H_0 = \int dq\; E(q)\, a_q^+ a_q$. Another change is that any Kronecker δ_{ij} is replaced by the Dirac delta $\delta(q-q')$.

Free particles. 1-particle states labeled by the momentum $\vec{p}$, with $E(\vec{p}) = \frac{p^2}{2m}$

$$H_0 = \int d^3p\; \frac{p^2}{2m}\; a_{\vec{p}}^+ a_{\vec{p}}$$

Periodic external field (the 0-th approximation for electrons in solids). Bloch theorem: 1-particle states are labeled by the level n, the quasi-momentum $\vec{k}$, and the spin σ. The energy $\varepsilon_n(\vec{k})$ depends on details of the periodic field

$$H_0 = \sum_{n,\sigma} \int d^3k\; \varepsilon_n(\vec{k})\; a^+_{n,\vec{k},\sigma} a_{n,\vec{k},\sigma}$$

Spherically symmetric external field (the 0-th approximation for electrons in atoms and nucleons in nuclei). 1-particle states are labeled by the quantum numbers n, l, m, σ. The energy $E_{n,l}$ depends on details of the field

$$H_0 = \sum_{n,l,m,\sigma} E_{n,l}\; a^+_{n,l,m,\sigma} a_{n,l,m,\sigma}$$

Remark: *The ideal gas approximation is very popular for electrons in atoms, molecules and solids. At the first sight, however, it looks like a rather poor approximation. The dominant (Coulomb) interaction between electrons is enormous at atomic scale and cannot be neglected in any decent approach.*
But there is no mystery involved, the ideal gas approximation used for electrons does not neglect the Coulomb interaction. The point is that the external field for the electron ideal gas contains not only the Coulomb field of positively charged nuclei, but also some kind of mean field of all negatively charged electrons. This mean field is usually given by the Hartree-Fock approximation. The corner-stone of this approximation is a restriction on electron states taken into consideration: only the direct products of single electron states are accounted for. In this restricted set of states one looks for what in some sense is the best approximation to the stationary states. This leads to the specific integro-differential equation for the 1-electron states and corresponding energies, which is then solved iteratively. [29] *The creation and annihilation operators for these Hartree-Fock states and the corresponding energies then enter the electron ideal gas Hamiltonian.*

Remark: *The ground state of a fermionic system in the Hartree-Fock approximation (the ideal gas approximation with 1-particle states and energies given by the Hartree-Fock equation) is quite simple: all the 1-particle states with energies below some boundary energy, the so-called Fermi energy ε_F, are occupied, while all the states with energies above ε_F are free. The Fermi energy depends, of course, on the number of particles in the system.*
In solids, the 1-particle Hartree-Fock states are characterized by $(n, \vec{k}, \sigma)$ (level, quasi-momentum, spin). The 1-particle n-th level Hartree-Fock energy is, as a rule, an ascending function of k^2 in any direction of $\vec{k}$. In any direction, therefore, there exists the level n and the vector $\vec{k}_F(\varphi, \vartheta)$ for which $\varepsilon_n(\vec{k}_F) = \varepsilon_F$. The endpoints of vectors $\vec{k}_F(\varphi, \vartheta)$ form a surface, called the Fermi surface. In the many-body ground state, the 1-particle states beneath (above) the Fermi surface are occupied (free).
It turns out that for a great variety of phenomena in solids, only the low excited states of the electron system are involved. They differ from the ground state by having a few 1-particle states above the Fermi surface occupied. The particular form of the Fermi surface therefore determines many macroscopic properties of the material under consideration. For this reason the knowledge of the Fermi surface is very important in the solid state physics.

Remark: *The ideal gas of fermions is frequently treated by means of a famous formal trick known as the electron-hole formalism. The ground state of the N fermion ideal gas is called the Fermi vacuum, and denoted by $|0_F\rangle$. For $i \leq N$ one defines new operators $b_i^+ = a_i$ and $b_i = a_i^+$. The original a_i^+ and a_i operators are taken into account only for $i > N$.*
Both a- and b-operators satisfy the commutation relations, and both b_i $(i \leq N)$ and a_i $(i > N)$ annihilate the Fermi vacuum (indeed, $b_i |0_F\rangle = 0$ because of anti-symmetry of fermion states, i.e. because of the Pauli exclusive principle). So, formally we have two types of particles, the holes and the new electrons, created from the Fermi vacuum by b_i^+ and a_i^+ respectively. The Hamiltonian reads $H_0 = \sum_{i \leq N} E_i b_i b_i^+ + \sum_{i>N} E_i a_i^+ a_i = \sum_{i \leq N} E_i - \sum_{i \leq N} E_i b_i^+ b_i + \sum_{i>N} E_i a_i^+ a_i$. The popular interpretation of the minus sign: the holes have negative energy.

[29] For details consult any reasonable textbook on QM or solid state physics, or for a very concise introduction perhaps the Appendix **??**.

Hamiltonian of a system of particles with the pair interaction

Perhaps the most important interaction to be added to the previous case of the ideal gas is the pair interaction, i.e. an interaction characterized by a potential energy of pairs of particles (most of applications in the solid state, atomic and nuclear physics involve such an interaction). In this case, the most suitable choice of basis in the 1-particle Hilbert space is the x-representation $|\vec{x}\rangle$, since the 2-particle states $|\vec{x},\vec{y}\rangle$ have a sharp value of the pair potential energy $V(\vec{x},\vec{y})$.

Due to the fact that we are dealing with the pair interaction, the 3-particle state $|\vec{x}_1,\vec{x}_2,\vec{x}_3\rangle$ does also have the sharp value of the potential energy, namely $V(\vec{x}_1,\vec{x}_2)+V(\vec{x}_1,\vec{x}_3)+V(\vec{x}_2,\vec{x}_3)$, and the same holds for other multiparticle states (this, in fact, is the definition of the pair interaction).

What is the potential energy of the state with $n(\vec{x}_i)$ particles at the position $\vec{x}_i$, where $i=1,2,\ldots$? The number of pairs contributing by $V(\vec{x}_i,\vec{x}_j)$ is $\frac{1}{2}n_{\vec{x}_i}n_{\vec{x}_j}$ for $i\neq j$, by which we understand also $\vec{x}_i\neq\vec{x}_j$ (the $\frac{1}{2}$ is there to avoid double-counting). For $i=j$ there is a subtlety involved. One has to include the potential energy of a particle with all other particles sharing the same position, but not with itself (a particle with itself does not constitute a pair). The number of pairs contributing by $V(\vec{x}_i,\vec{x}_i)$ is therefore $\frac{1}{2}n_{\vec{x}_i}(n_{\vec{x}_i}-1)$. This makes the total potential energy in the state under consideration equal to $\frac{1}{2}\sum_{i,j}V(\vec{x}_i,\vec{x}_j)n_{\vec{x}_i}n_{\vec{x}_j}-\frac{1}{2}\sum_i V(\vec{x}_i,\vec{x}_i)n_{\vec{x}_i}$.

Using the same logic as in the case of the ideal gas, it is now easy to write down the operator of the total potential energy in terms of operators $\hat{n}_{\vec{x}}=a^+_{\vec{x}}a_{\vec{x}}$. Using the commutation relations for the creation and annihilation operators the resulting expression can be simplified to the form[30]

$$H_{\text{pair}}=\frac{1}{2}\int d^3x\; d^3y\; V(\vec{x},\vec{y})\;\; a^+_{\vec{x}}a^+_{\vec{y}}a_{\vec{y}}a_{\vec{x}}$$

Note the order of the creation and annihilation operators, which is mandatory. It embodies the above mentioned subtlety.

As we have seen before, the x-representation is usually not the most suitable for the ideal gas Hamiltonian. To have the complete Hamiltonian of a system with the pair interaction presented in a single representation, it is useful to rewrite the potential energy operator H_{pair} in the other representation.

Free particles. All one needs is $a_{\vec{x}}=\int d^3p\;\langle\vec{x}|\vec{p}\rangle\,a_{\vec{p}}=\int\frac{d^3p}{\sqrt{(2\pi\hbar)^3}}e^{i\vec{p}.\vec{x}/\hbar}a_{\vec{p}}$

$$H_{\text{pair}}=\frac{1}{2}\int d^3p_1d^3p_2d^3p_3d^3p_4\;V(\vec{p}_1,\vec{p}_2,\vec{p}_3,\vec{p}_4)\;\; a^+_{\vec{p}_1}a^+_{\vec{p}_2}a_{\vec{p}_3}a_{\vec{p}_4}$$

$$V(\vec{p}_1,\vec{p}_2,\vec{p}_3,\vec{p}_4)=\int\frac{d^3x}{(2\pi\hbar)^3}\frac{d^3y}{(2\pi\hbar)^3}V(\vec{x},\vec{y})\exp\frac{i\,(\vec{p}_4-\vec{p}_1)\,.\vec{x}}{\hbar}\exp\frac{i\,(\vec{p}_3-\vec{p}_2)\,.\vec{y}}{\hbar}$$

Periodic external field. Replace the plane waves $\langle\vec{x}|\vec{p}\rangle$ by the Bloch functions $\langle\vec{x}|n,\vec{k}\rangle=u_n(\vec{k})e^{i\vec{k}.\vec{x}}$.

Spherically symmetric external field. Replace the plane waves $\langle\vec{x}|\vec{p}\rangle$ by the product of the radial Schrödinger equation solutions and spherical harmonics $\langle\vec{x}|n,l,m\rangle=R_{nl}(r)\,Y_{lm}(\varphi,\vartheta)$.

[30] $U=\frac{1}{2}\int d^3x\;d^3y\;V(\vec{x},\vec{y})\;a^+_{\vec{x}}a_{\vec{x}}a^+_{\vec{y}}a_{\vec{y}}-\frac{1}{2}\int d^3x\;V(\vec{x},\vec{x})\;a^+_{\vec{x}}a_{\vec{x}}$
$=\frac{1}{2}\int d^3x\;d^3y\;V(\vec{x},\vec{y})\;a^+_{\vec{x}}\left(\delta(x-y)\pm a^+_{\vec{y}}a_{\vec{x}}\right)a_{\vec{y}}-\frac{1}{2}\int d^3x\;V(\vec{x},\vec{x})\;a^+_{\vec{x}}a_{\vec{x}}$
$=\pm\frac{1}{2}\int d^3x\;d^3y\;V(\vec{x},\vec{y})\;a^+_{\vec{x}}a^+_{\vec{y}}a_{\vec{x}}a_{\vec{y}}=\frac{1}{2}\int d^3x\;d^3y\;V(\vec{x},\vec{y})\;a^+_{\vec{x}}a^+_{\vec{y}}a_{\vec{y}}a_{\vec{x}}$

Remark: *Even if it is not necessary for our purposes, it is hard to resist the temptation to discuss perhaps the most important pair interaction in the non-relativistic quantum physics, namely the Coulomb interaction. This is of general interest not only because of the vast amount of applications, but also due to the fact that the standard way of dealing with the Coulomb potential in the p-representation involves a mathematical inconsistency. The way in which this inconsistency is treated is in a sense generic, and therefore quite instructive.*
In the x-representation $V_{\mathrm{Coulomb}}(\vec{x},\vec{y}) = \frac{e^2}{4\pi}\frac{1}{|\vec{x}-\vec{y}|}$*, i.e. in the p-representation (which is relevant in the case with no external field)*

$$V_{\mathrm{Coulomb}}(\vec{p}_1,\vec{p}_2,\vec{p}_3,\vec{p}_4) = \frac{e^2}{4\pi}\int \frac{d^3x}{(2\pi)^3}\frac{d^3y}{(2\pi)^3}\frac{1}{|\vec{x}-\vec{y}|}e^{i(\vec{p}_4-\vec{p}_1).\vec{x}}e^{i(\vec{p}_3-\vec{p}_2).\vec{y}}$$

(for the sake of brevity, we use the Heaviside-Lorentz convention in electrodynamics and $\hbar = 1$ *units in QM). This integral, however, is badly divergent. The integrand simply does not drop out fast enough for* $|\vec{x}| \to \infty$ *and* $|\vec{y}| \to \infty$*.*
Instead of giving up the use of the p-representation for the Coulomb potential energy, it is a common habit to use a dirty trick. It starts by considering the Yukawa (or Debey) potential energy $V_{\mathrm{Debey}}(\vec{x},\vec{y}) = \frac{e^2}{4\pi}\frac{1}{|\vec{x}-\vec{y}|}e^{-\mu|\vec{x}-\vec{y}|}$*, for which the p-representation is well defined and can be evaluated readily*[31]

$$V_{\mathrm{Debey}}(\vec{p}_1,\vec{p}_2,\vec{p}_3,\vec{p}_4) = \frac{e^2}{4\pi}\frac{1}{2\pi^2}\frac{1}{\mu^2+(\vec{p}_4-\vec{p}_1)^2}\delta(\vec{p}_1+\vec{p}_2-\vec{p}_3-\vec{p}_4)$$

Now comes the dirty part. It is based on two simple (almost trivial) observations: the first is that $V_{\mathrm{Coulomb}}(\vec{x},\vec{y}) = \lim_{\mu\to 0} V_{\mathrm{Debey}}(\vec{x},\vec{y})$*, and the second is that the limit* $\lim_{\mu\to 0} V_{\mathrm{Debey}}(\vec{p}_1,\vec{p}_2,\vec{p}_3,\vec{p}_4)$ *is well defined. From this, a brave heart can easily conclude that* $V_{\mathrm{Coulomb}}(\vec{p}_1,\vec{p}_2,\vec{p}_3,\vec{p}_4) = \frac{e^2}{8\pi^3}\frac{1}{(\vec{p}_4-\vec{p}_1)^2}\delta(\vec{p}_1+\vec{p}_2-\vec{p}_3-\vec{p}_4)$*. And, believe it or not, this is indeed what is commonly used as the Coulomb potential energy in the p-representation.*
Needless to say, from the mathematical point of view, this is an awful heresy (called illegal change of order of a limit and an integral). How does it come about that physicists are working happily with something so unlawful?
The most popular answer is that the Debey is nothing else but a screened Coulomb, and that in most systems this is more realistic than the pure Coulomb. This is a reasonable answer, with a slight off-taste of a cheat (the limit $\mu \to 0$ *convicts us that we are nevertheless interested in the pure Coulomb).*
Perhaps a bit more fair answer is this one: For μ *small enough, one cannot say experimentally the difference between Debey and Coulomb. And the common plausible belief is that measurable outputs should not depend on immeasurable inputs (if this was not typically true, whole science would hardly be possible). If mathematics nevertheless reveals inconsistencies for some values of an immeasurable parameter, one should feel free to choose another value, which allows for mathematically sound treatment.*

---

[31] For $\vec{r} = \vec{x}-\vec{y}$, $V_{\mathrm{Yukawa}}(\vec{p}_1,\vec{p}_2,\vec{p}_3,\vec{p}_4) = \frac{e^2}{4\pi}\int \frac{d^3r}{(2\pi)^3}\frac{d^3y}{(2\pi)^3}\frac{1}{r}e^{-\mu r}e^{i(\vec{p}_4-\vec{p}_1).\vec{r}}e^{i(\vec{p}_3-\vec{p}_2+\vec{p}_4-\vec{p}_1).\vec{y}}$
$= \frac{e^2}{4\pi}\delta(\vec{p}_1+\vec{p}_2-\vec{p}_3-\vec{p}_4)\int \frac{d^3r}{(2\pi)^3}\frac{e^{-\mu r}}{r}e^{i\vec{q}.\vec{r}}$, where $\vec{q} = \vec{p}_4-\vec{p}_1$. The remaining integral in the spherical coordinates: $\frac{1}{(2\pi)^2}\int_0^\infty dr\ r\ e^{-\mu r}\int_{-1}^1 d\cos\vartheta\ e^{iqr\cos\vartheta} = \frac{1}{2\pi^2}\frac{1}{q}\int_0^\infty dr\ e^{-\mu r}\sin qr$. For the integral $I = \int_0^\infty dr\ e^{-\mu r}\sin qr$ one obtains by double partial integration the relation $I = \frac{q}{\mu^2} - \frac{q^2}{\mu^2}I$, and putting everything together, one comes to the quoted result.

Hamiltonian of a system of unstable particles

Let us consider a system of particles of three types A, B and C, one of which decays into other two, say $C \to AB$. The decay can be understood as the annihilation of the decaying particle and the simultaneous creation of the products. The corresponding interaction Hamiltonian, i.e. the part of the time-evolution generator responsible for this kind of change, is almost self-evident. It should contain the combination $a_i^+ b_j^+ c_k$, where the lowercase letters for creation and annihilation operators correspond to the uppercase letters denoting the type of particle, and subscripts specify the states of the particles involved.

The decay is usually allowed for various combinations of the quantum numbers i, j, k, so the interaction Hamiltonian should assume the form of the sum of all legal alternatives. This is usually written as the sum of all alternatives, each of them multiplied by some factor, which vanishes for the forbidden combinations of quantum numbers: $\sum_{i,j,k} g_{ijk} a_i^+ b_j^+ c_k$.

There is still one problem with this candidate for the interaction Hamiltonian: it is not Hermitian. But this is quite easy to take care of, one just adds the Hermitian conjugate operator $g_{ijk}^* c_k^+ b_j a_i$. So the Hermiticity of the Hamiltonian requires, for any decay, the existence of the reverse process $AB \to C$. All in all

$$H_{\text{int}} = \sum_{i,j,k} g_{ijk} a_i^+ b_j^+ c_k + g_{ijk}^* c_k^+ b_j a_i$$

Generalizations (decays with 3 or more products, or even some more bizarre processes, with more than one particle annihilated) are straightforward.

For a Hamiltonian of this type the C particles need not to be necessary unstable. If the mass of the C particle is smaller than the masses of A na B particles then decay may be kinematically forbidden (due to the momentum δ-function in the final formula for the decay rate). Many of the realistic Hamiltonians of this type describe stable particles, in spite of the fact that we have introduce it naturally for unstable particles.

The factor g_{ijk} is usually called the coupling constant, although it depends on the quantum numbers i, j, k. The reason for the name constant is that most of decays are local and translation invariant (they do not vary with changes in position). In the x-representation the locality means that $g_{\vec{x},\vec{y},\vec{z}} = g_{\vec{x}} \delta(\vec{x}-\vec{y}) \delta(\vec{x}-\vec{z})$ and translational invariance requires that g does not depend on $\vec{x}$

$$H_{\text{int}} = \int d^3x \; g \, a_{\vec{x}}^+ b_{\vec{x}}^+ c_{\vec{x}} + g^* c_{\vec{x}}^+ b_{\vec{x}} a_{\vec{x}}$$

Remark: *electron-photon interaction*

Remark: *electron-phonon interaction*

any linear operator expressed in terms of a^+, a

The only purpose of this paragraph is to satisfy a curious reader (if there is any). It will be of no practical importance to us.

First of all, it is very easy to convince oneself that any linear operator can be expressed in terms of creation operators a_i^+, annihilation operators a_i and the vacuum projection operator $|0\rangle\langle 0|$. Indeed, if $\hat{A}$ is a linear operator, then

$$\hat{A} = \sum_{\substack{i,j,\dots \\ k,l,\dots}} A_{ij\dots,kl\dots} a_i^+ a_j^+ \dots |0\rangle\langle 0| a_k a_l \dots$$

where $A_{ij\dots,kl\dots} = \frac{\langle i,j,\dots|\hat{A}|k,l,\dots\rangle}{\langle i,j,\dots|i,j,\dots\rangle\langle k,l,\dots|k,l,\dots\rangle}$. Proof: both LHS and RHS have the same matrix elements for all combinations of basis vectors (check this).

The only question therefore is how to get rid of $|0\rangle\langle 0|$. This is done by induction. First, one expresses the $\hat{A}$ operator only within the 0-particle subspace $\mathcal{H}^0$of the Fock space, where it is nothing else but the multiplication by the constant

$$\hat{A}_0 = \tilde{A}_{0,0} \equiv \langle 0|\hat{A}|0\rangle$$

Then, one expresses the $\hat{A}_1 = \hat{A} - \hat{A}_0$ operator within the 0- and 1-particle subspace $\mathcal{H}^0 \oplus \mathcal{H}^1$. Here one gets (check it)

$$\hat{A}_1 = \tilde{A}_{i,j} a_i^+ a_j + \tilde{A}_{i,0} a_i^+ + \tilde{A}_{0,j} a_j$$

where $\tilde{A}_{ij} = \langle i|\hat{A} - \hat{A}_0|j\rangle$, $\tilde{A}_{i,0} = \langle i|\hat{A} - \hat{A}_0|0\rangle$ and $\tilde{A}_{0,j} = \langle 0|\hat{A} - \hat{A}_0|j\rangle$. If one restricts oneself to $\mathcal{H}^0 \oplus \mathcal{H}^1$, then $\hat{A} = \hat{A}_0 + \hat{A}_1$ (why?). So we have succeeded in writing the operator $\hat{A}$ in terms of a_i^+, a_i, even if only in the subspace of the Fock space. This subspace is now expanded to $\mathcal{H}^0 \oplus \mathcal{H}^1 \oplus \mathcal{H}^2$, etc.

It may be instructive now to work out the operator $\hat{A}_2 = \hat{A} - \hat{A}_0 - \hat{A}_1$within $\mathcal{H}^0 \oplus \mathcal{H}^1 \oplus \mathcal{H}^2$ in terms of a_i^+, a_i (try it). We will, however, proceed directly to the general case of $\hat{A}_n = \hat{A} - \sum_{m=0}^{n-1} \hat{A}_m$ within $\bigoplus_{m=0}^{n} \mathcal{H}^m$

$$\hat{A}_n = \sum_{\substack{\text{allowed} \\ \text{combinations}}} \tilde{A}_{ij\dots,kl\dots} a_i^+ a_j^+ \dots a_k a_l \dots$$

$$\tilde{A}_{ij\dots,kl\dots} = \frac{\langle i,j,\dots|\hat{A} - \sum_{m=0}^{n-1} \hat{A}_m |k,l,\dots\rangle}{\langle i,j,\dots|i,j,\dots\rangle\langle k,l,\dots|k,l\dots\rangle}$$

and the "allowed combinations" are either $\underbrace{ij\dots}_{n}, \underbrace{kl\dots}_{m\le n}$ or $\underbrace{ij\dots}_{m\le n}, \underbrace{kl\dots}_{n}$.

If restricted to $\bigoplus_{m=0}^{n} \mathcal{H}^m$, then $\hat{A} = \sum_{m=0}^{n} \hat{A}_m$, i.e. we have $\hat{A}$ expressed in terms of a_i^+, a_i. To get an expression valid not only in subspaces, one takes

$$\hat{A} = \sum_{m=0}^{\infty} \hat{A}_m$$

1.2.3 Calculation of matrix elements — the main trick

One of the most valuable benefits of the use of the creation and annihilation operator formalism is the completely automatous way of matrix elements calculation. The starting point is twofold:
• any ket (bra) vector can be written as a superposition of the basis vectors, which in turn can be obtained by a_i^+ (a_i) operators acting on $|0\rangle$
• any linear operator can be written as a linear combination of products of the a_i^+ and a_i operators
Consequently, any matrix element of a linear operator is equal to some linear combination of the vacuum expectation values (VEVs) of the products of the creation and annihilation operators.

Some of the VEVs are very easy to calculate, namely those in which the last (first) operator in the product is the annihilation (creation) one. Indeed, due to $a_i|0\rangle = 0$ and $\langle 0| a_i^+ = 0$, any such VEV vanishes. Other VEVs are easily brought to the previous form, one just has to use the (anti)commutation relations $\left[a_i, a_j^+\right]_\mp = \delta_{ij}$ to push the creation (annihilation) operators to the right (left). By repeated use of the (anti)commutation relations, the original VEV is brought to the sum of scalar products $\langle 0|0\rangle$ multiplied by pure numbers, and the VEVs vanishing because of $\langle 0| a_i^+ = 0$ or $a_i|0\rangle = 0$. An example is perhaps more instructive than a general exposition.

Example: *Let us consider a decay-like Hamiltonian for just one type of particles* $H_{\mathrm{decay}} = \sum_{i,j,k} g_{ijk}(a_i^+ a_j^+ a_k + a_k^+ a_j a_i)$ *(e.g. phonons, as well as gluons, enjoy this kind of interaction). Note that the coupling "constant" is real* $g_{ijk} = g_{ijk}^*$. *And let us say we want to calculate* $\langle l| H_{\mathrm{decay}} |m,n\rangle$. *First, one writes*

$$\langle l| H_{\mathrm{decay}} |m,n\rangle = \sum_{i,j,k} g_{ijk} \left(\langle 0| a_l a_i^+ a_j^+ a_k a_m^+ a_n^+ |0\rangle + \langle 0| a_l a_k^+ a_j a_i a_m^+ a_n^+ |0\rangle \right)$$

Then one starts to reshuffle the operators. Take, e.g., the first two and use $a_l a_i^+ = \delta_{li} \pm a_i^+ a_l$ *(or with i replaced by k in the second term), to obtain*

$$\begin{aligned}\langle l| H_{\mathrm{decay}} |m,n\rangle = \sum_{i,j,k} g_{ijk} \big(& \delta_{li} \langle 0| a_j^+ a_k a_m^+ a_n^+ |0\rangle \pm \langle 0| a_i^+ a_l a_j^+ a_k a_m^+ a_n^+ |0\rangle \\ & + \delta_{lk} \langle 0| a_j a_i a_m^+ a_n^+ |0\rangle \pm \langle 0| a_k^+ a_l a_j a_i a_m^+ a_n^+ |0\rangle \big)\end{aligned}$$

Three of the four terms have a^+ *next to* $\langle 0|$, *and consequently they vanish. In the remaining term (the third one) one continues with reshuffling*

$$\begin{aligned}\langle l| H_{\mathrm{decay}} |m,n\rangle &= \sum_{i,j,k} g_{ijk}\delta_{lk} \left(\delta_{im} \langle 0| a_j a_n^+ |0\rangle \pm \langle 0| a_j a_m^+ a_i a_n^+ |0\rangle \right) \\ &= \sum_{i,j,k} g_{ijk}\delta_{lk} \left(\delta_{im}\delta_{jn} \pm \delta_{jm} \langle 0| a_i a_n^+ |0\rangle + 0 \right) \\ &= \sum_{i,j,k} g_{ijk}\delta_{lk} \left(\delta_{im}\delta_{jn} \pm \delta_{jm}\delta_{in} \right) = g_{mnl} \pm g_{nml}\end{aligned}$$

The result could be seen directly from the beginning. The point was not to obtain this particular result, but rather to illustrate the general procedure. It should be clear from the example, that however complicated the operator and the states (between which it is sandwiched) are, the calculation proceeds in the same way: one just reshuffles the creation and annihilation operators.

The example has demonstrated an important common feature of this type of calculations: after all the rubbish vanishes, what remains are just products of deltas (Kronecker or Dirac for discrete or continuous cases respectively), each delta originating from some pair of $a\,a^+$. This enables a short-cut in the whole procedure, namely to take into account only those terms in which all a and a^+operators can be organized (without leftovers) in the pairs $a\,a^+$(in this order!), and to assign the appropriate δ to every such pair.

This is usually done within the "clip" notation, like e.g. $\ldots \overbrace{a_i \ldots a_j^+} \ldots$, which tells us that a_i is paired with a_j^+. The factor corresponding to this particular clip is therefore δ_{ij}. In the case of fermions one has to keep track of signs. The reader may try to convince him/herself that the rule is: every pair of clips biting into each other, generates the minus sign for fermions.

Example: *The same example as above, now using the clip short-cut.*

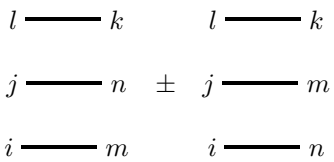
$$\langle l|\, H_{\rm decay}\, |m,n\rangle = \sum_{i,j,k} g_{ijk}\left(\langle 0|\, a_l a_i^+ a_j^+ a_k a_m^+ a_n^+\, |0\rangle + \langle 0|\, a_l a_k^+ a_j a_i a_m^+ a_n^+\, |0\rangle\right)$$

The first term cannot be organized (without leftovers) in pairs of aa^+, so it does not contribute. As to the rest, one has

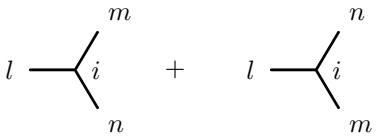
$$\langle 0|\, a_l a_k^+ a_j a_i a_m^+ a_n^+\, |0\rangle = \langle 0|\, \overbrace{a_l a_k^+}\, \overbrace{a_j \underbrace{a_i a_m^+}\, a_n^+}\, |0\rangle + \langle 0|\, \overbrace{a_l a_k^+}\, \overbrace{a_j \underbrace{a_i a_m^+}\, a_n^+}\, |0\rangle = \delta_{lk}\delta_{jn}\delta_{im} \pm \delta_{lk}\delta_{jm}\delta_{in}$$

leading immediately to the same result as before ($g_{mnl} \pm g_{nml}$).

The common habit is to make the short-cut even shorter. Instead of writing the clips above or under the corresponding operators, one draws just the clips and then assigns the corresponding factor to the whole picture. In this comics-like formalism our matrix element would become

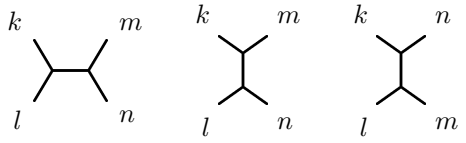


If the interaction is local, the picture is changed slightly. Local interactions contain products of operators in the same position, e.g. $a_{\vec{x}}^+ b_{\vec{x}}^+ c_{\vec{x}}$, which in our discrete case would correspond to $H_{\rm decay} = \sum_i g(a_i^+ a_i^+ a_i + a_i^+ a_i a_i)$. For fermions this Hamiltonian vanishes (why?), for bosons result becomes $g\sum_i(\delta_{li}\delta_{in}\delta_{im} + \delta_{li}\delta_{im}\delta_{in})$ and the above picture is changed to



Exercise: ***(a useful one)*** *For bosonic particles with $H_{\rm decay} = \sum_i g(a_i^+ a_i^+ a_i + a_i^+ a_i a_i)$ calculate $\langle k,l|\, H_{\rm decay} H_{\rm decay}\, |m,n\rangle$ and draw the corresponding pictures.*

Answer: $12g^2$ for $k=l=m=n$ (and 0 otherwise). The pictures corresponding to the matrix element are (contributing 4 times each).



Feynman diagrams – a comics version of a perturbation theory

The pictures from the exercise at the end of the previous paragraph may suggest some relation to the Feynman diagrams and indeed they are closely related. In this subsection (which can be skipped safely) we will demonstrate that in a theory with the interaction Hamiltonian written in terms of creation and annihilation operators various terms of a perturbation series can be represented by such diagrams. We will not derive the Feynman rules presented in Conclusions yet (this will be achieved only later on), but the basic idea should become clear already now.

One can illustrate the whole procedure on a simple example of time-independent perturbation theory, where eigenvalues E_n and eigenstates $|\psi_n\rangle$ of a complete Hamiltonian $H = H_0 + \alpha H'$ are expressed in terms of eigenvalues $\mathcal{E}_n$ and eigenstates $|\varphi_n\rangle$ of the unperturbed H_0. Let us recall that this is achieved by expanding the E_n and $|\psi_n\rangle$ in powers of α: $E_n = \mathcal{E}_n + \sum_{k=1}^{\infty} \alpha^k E_n^{(k)}$ and $|\psi_n\rangle = |\varphi_n\rangle + \sum_{k=1}^{\infty} \alpha^k |\psi_n^{(k)}\rangle$, where $H_0 |\varphi_n\rangle = \mathcal{E}_n |\varphi_n\rangle$ and $\langle\varphi_n|\psi_n^{(k)}\rangle = 0$. Comparing coefficients of various powers of α in the equation $(H_0 + \alpha H') |\psi_n\rangle = E_n |\psi_n\rangle$ one obtains explicit formulae for the expansion coefficients. The lowest order results are well known from any textbook on quantum mechanics, in non-degenerate case one gets

$$E_n^{(1)} = \langle\varphi_n| H' |\varphi_n\rangle \qquad E_n^{(2)} = \sum_{m\neq n} \frac{\langle\varphi_n| H' |\varphi_m\rangle \langle\varphi_m| H' |\varphi_n\rangle}{\mathcal{E}_n - \mathcal{E}_m}$$

The higher orders are more involved, e.g.

$$E_n^{(3)} = \sum_{m,m'\neq n} \frac{\langle\varphi_n| H' |\varphi_m\rangle \langle\varphi_m| H' |\varphi_{m'}\rangle \langle\varphi_{m'}| H' |\varphi_n\rangle}{(\mathcal{E}_n - \mathcal{E}_m)(\mathcal{E}_n - \mathcal{E}_{m'})} + \dots$$

where ellipses stand for additional terms, which have a similar structure (product of the interaction Hamiltonians sandwiched between various states in the numerator and product of energy differences in the denominator) and their complexity increases with increasing order.

Let us consider, as an example, a system of free non-relativistic bosons with a local decay-like interaction (with real g). The unperturbed and the perturbation Hamiltonians are

$$H_0 = \int d^3p \, \frac{p^2}{2m} a_{\vec{p}}^+ a_{\vec{p}} \qquad\qquad H' = g \int d^3x \, (a_{\vec{x}}^+ a_{\vec{x}}^+ a_{\vec{x}} + a_{\vec{x}}^+ a_{\vec{x}} a_{\vec{x}})$$

respectively. Since the eigenstates of the H_0 are the vectors $|\vec{p}\rangle$, it is convenient to rewrite H' into the p-representation, using the relations $a_{\vec{x}} = \int \frac{d^3k}{(2\pi\hbar)^{3/2}} e^{i\vec{k}.\vec{x}/\hbar} a_{\vec{k}}$ and $a_{\vec{x}}^+ = \int \frac{d^3k}{(2\pi\hbar)^{3/2}} e^{-i\vec{k}.\vec{x}/\hbar} a_{\vec{k}}^+$. After plugging this into H' and integrating over d^3x (recall $\int dx \, e^{-ikx} = 2\pi\delta(k)$) one obtains

$$H' = g \int \frac{d^3k_1}{(2\pi\hbar)^{3/2}} \frac{d^3k_2}{(2\pi\hbar)^{3/2}} \frac{d^3k_3}{(2\pi\hbar)^{3/2}} \, (2\pi\hbar)^{3/2} \, \delta(\vec{k}_1 + \vec{k}_2 - \vec{k}_3) \, a_{\vec{k}_1}^+ a_{\vec{k}_2}^+ a_{\vec{k}_3} + h.c.$$

For such a H' the only non-vanishing matrix elements are those between states in which number of particles differ precisely by one. The first order correction to energy of one-particle state $E_{\vec{p}}^{(1)} = \langle\vec{p}| H' |\vec{p}\rangle$ is therefore zero. In the second order correction one has non-vanishing matrix elements $\langle\vec{p}| H' |\varphi_m\rangle$ for two-particle states $|\varphi_m\rangle = |\vec{p}', \vec{p}''\rangle$, i.e.

$$E_{\vec{p}}^{(2)} = \int d^3p' \, d^3p'' \, \frac{\langle\vec{p}| H' |\vec{p}', \vec{p}''\rangle \langle\vec{p}', \vec{p}''| H' |\vec{p}\rangle}{(p^2 - p'^2 - p''^2)/2m}.$$

Now using the technique of the previous subsection one can represent the matrix elements in the numerator of $E^{(2)}_{\vec{p}}$ by two pictures, namely  and  . And since the RHS two-particle state of the first picture is the same as the LHS two-particle state of the second picture, one can draw their product as  . Now just like in the previous subsection one gets

$$\langle\vec{p}\,|\, a^+_{\vec{k}_1}\, a^+_{\vec{k}_2}\, a_{\vec{k}_3} + h.c.\,|\vec{p}\,',\vec{p}\,''\rangle = \delta(\vec{p}-\vec{k}_3)\,\delta(\vec{p}\,'-\vec{k}_2)\,\delta(\vec{p}\,''-\vec{k}_1) \pm \delta(\vec{p}-\vec{k}_3)\,\delta(\vec{p}\,'-\vec{k}_1)\,\delta(\vec{p}\,''-\vec{k}_2)$$

and so for bosons

$$\langle\vec{p}\,|\, H'\,|\vec{p}\,',\vec{p}\,''\rangle = \frac{2g}{(2\pi\hbar)^3}\;\delta(\vec{p}\,'+\vec{p}\,''-\vec{p})$$

while for fermions this matrix element vanishes. Our expression for the second order correction to the energy of bosons (not only the numerator, but rather the whole second order correction) can be therefore written as the diagram

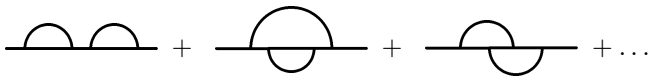
$E^{(2)}_n =$ 

with the following rules[32]

- To every internal line assign the factor $\int d^3p$
- To every vertex assign the factor $\frac{2g}{(2\pi\hbar)^3}\;\delta(\vec{p}\,'+\vec{p}\,''-\vec{p})$
- To every intermediate state $|\varphi_m\rangle$ assign the factor $\frac{1}{\mathcal{E}_n-\mathcal{E}_m}$

The intermediate states are the states of several free particles (with specific momenta) represented by internal lines of the diagram. When going from left to right, every vertex changes the number of particles, i.e. the intermediate state is changed and the corresponding energy denominator should be added.

The point of the diagrams and the rules is that they work in the same way also for higher order corrections, e.g. with these rules one can easily write (draw)

$E^{(4)}_n =$ 

We are not going to elaborate further and to derive the rules in their full complexity (e.g. we are not going to formulate the rules for terms hidden in ellipses). The purpose of this paragraph was just to give a basic taste of how do the Feynman diagrams and rules emerge in a quantum theory. All the details are to be discussed only when the specific perturbation theory within the relativistic quantum theory is developed in the next chapter.

[32] Note that these diagrams are not exactly the ones described in Conclusions. They represent different quantities (the energy eigenvalues, not the scattering amplitude) and the rules are also different (the factors for intermediate states, not for internal lines). But what they do have in common is that they are just pictorial representations of various terms in perturbative expansion of some quantity in some version of the perturbation theory.

Remark: *It is probably worth mentioning that there exists an equivalent, but slightly different version of the perturbation theory for energy eigenvalues, in which complexity of terms does not increase with the increasing order. It is known as the Brillouin–Wigner perturbation theory (while the standard version notoriously known from textbooks is called the Rayleigh–Schrödinger perturbation theory). In the BW version one obtains equally simple expressions in all orders (for derivation see the next remark)*

$$E_n^{(k)} = \sum_{m_1,\ldots,m_{k-1}\neq n} \frac{\langle\varphi_n| H' |\varphi_{m_1}\rangle \ldots \langle\varphi_{m_{k-1}}| H' |\varphi_n\rangle}{(E_n - \mathcal{E}_{m_1}) \ldots (E_n - \mathcal{E}_{m_{k-1}})}$$

Note the presence of E_n (instead of $\mathcal{E}_n$) on the RHS of the last relation. Because of this, the relation is only an implicit one and has to be solved (either iteratively or by expansion of E_n in powers of α) to get the explicit result for $E_n^{(k)}$ (which is, as it should be, equal to the standard version of the perturbation theory). Due to its simpler structure the BW version of the perturbation theory is better suited for the Feynman diagrammatical representation than the RS version. Reason: no ellipses in the BW version (the terms corresponding to the ellipses in the RS version can be also represented by Feynman diagrams, but with more complicated rules). Moral: for some versions of the perturbation theory the Feynman diagrams are more suitable than for the others.

Remark: *For the sake of completeness we will sketch here the derivation of the Brillouin-Wigner perturbation theory (see the previous remark). One possible way is to start with another version of a perturbation theory, namely the one for the particle scattering. This is of interest by itself, because it provides us with yet another example of Feynman diagrams in Quantum mechanics.*

In QM textbooks the potential scattering is usually treated by solving the time-independent Schrödinger equation in the x-representation.[33] *This, however, is not best suited for the formalism of creation and annihilation operators. Fortunately, the whole idea can be generalised in a relatively easy way. One writes the Schrödinger equation in the form $(E - H_0)|\psi_l\rangle = \alpha H' |\psi_l\rangle$, where l stays for a label (which is continuous in the case of scattering). This equation is solved formally using the inverse operator* [34] *$(E - H_0)^{-1}$. The inverse operator for $E - H_0$, however, does not exist, since the operator $E - H_0$ has the zero eigenvalue. It is therefore a common habit to consider the operator $E - H_0 + i\epsilon$ instead (this can be viewed as a generalisation of the $i\epsilon$ prescription in the complex plane integration for the Green function). The implicit solution of the Schrödinger equation is written as the so-called Lippman-Schwinger equation*

$$|\psi\rangle = |\varphi\rangle + (E - H_0 + i\epsilon)^{-1} \alpha H' |\psi\rangle$$

where $|\varphi\rangle$ is a solution of the homogeneous equation $(E - H_0)|\varphi\rangle = 0$.

[33] A brief reminder: The Schrödinger equation for potential scattering is rewritten as a Helmholtz-like equation $(E - H_0)\psi(\vec{r}) = \alpha H'\psi(\vec{r})$ with $H_0 = -\frac{\hbar^2}{2m}\Delta$ and $H' = U(\vec{r})$. The solution of this equation is expressed via the Green function $G_+(\vec{r}, \vec{r}')$ satisfying $\left(\frac{\hbar^2}{2m}\Delta + E\right) G_+(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}')$ and the "outgoing wave" boundary condition $G_+(\vec{r}, \vec{0})_{r\to\infty} \to \frac{f(\varphi,\vartheta)}{r} e^{ikr}$ where $E = \frac{k^2}{2m}$. The Green function is usually found by Fourier transforming the equation, solving the resulting algebraic equation and backward Fourier transforming the solution with integration in the complex plane, the result is $G_+(\vec{r}, \vec{r}') = -\frac{1}{4\pi|\vec{r}-\vec{r}'|} e^{i\vec{k}.(\vec{r}-\vec{r}')}$. The implicit solution of the Schrödinger equation is given by $\psi(\vec{r}) = \varphi(\vec{r}) + \int d^3r' G_+(\vec{r}, \vec{r}')U(\vec{r}')\psi(\vec{r}')$, where $\varphi(\vec{r})$ is a solution of the homogeneous Helmholtz equation (usually one takes $\varphi(\vec{r}) = e^{ikz}$, which is a general solution of the homogeneous equation with free parameters fixed by initial conditions)). As a final step the implicit equation is solved iteratively.

[34] An inverse operator is, so to speak, a generalisation of the Green function. Indeed, the inverse operator is defined by $A\,A^{-1} = \mathbb{1}$, the Green function by $A\,G(\vec{r} - \vec{r}') = \delta(\vec{r} - \vec{r}')$ and the delta-function can be understood as a matrix-like representation of the unit operator $\mathbb{1}$.

It is quite straightforward to find the operator $(E-H_0+i\epsilon)^{-1}$. *One just has to realise that in the basis* $|\varphi_l\rangle$ *in which the operator* H_0 *is diagonal, one has* $E-H_0+i\epsilon = \sum_l (E-\mathcal{E}_l+i\epsilon)\,|\varphi_l\rangle\,\langle\varphi_l|$, *where sum, of course, stands for integrals over the continuous label* l. *As a direct consequence* $(E-H_0+i\epsilon)^{-1} = \sum_l \frac{|\varphi_l\rangle\langle\varphi_l|}{E-\mathcal{E}_l+i\epsilon}$ *and therefore*

$$|\psi\rangle = |\varphi\rangle + \sum_l \frac{|\varphi_l\rangle\,\langle\varphi_l|}{E-\mathcal{E}_l+i\epsilon}\alpha H'\,|\psi\rangle$$

This implicit relation is solved iteratively, leading to the same structure as in the previous remark (matrix elements of H' *in the numerator, energy differences in the denominator). The diagrams are as natural and useful as in the previous case. This version of perturbation theory with energy denominators is called the old-fashioned perturbation theory. To get the modern version presented in Conclusions, one should work with the time-dependent perturbation theory. We will derive this version later on, in the framework of the relativistic quantum theory.*

As the last point let us focus not on the continuous, but rather of the discrete part of the spectrum. The analogue of the Lippmann-Schwinger equation becomes (in the $\epsilon \to 0$ *limit)*

$$|\psi_n\rangle = |\varphi_n\rangle + \sum_{m\neq n} \frac{|\varphi_m\rangle\,\langle\varphi_m|}{E_n-\mathcal{E}_m}\alpha H'\,|\psi_n\rangle$$

This is now solved iteratively, leading to

$$|\psi_n\rangle = |\varphi_n\rangle + \alpha\sum_{m\neq n}\frac{|\varphi_m\rangle\langle\varphi_m|H'|\varphi_n\rangle}{E_n-\mathcal{E}_m} + \alpha^2\sum_{m\neq n}\sum_{m'\neq n}\frac{|\varphi_m\rangle\langle\varphi_m|H'|\varphi_{m'}\rangle\langle\varphi_{m'}|H'|\varphi_n\rangle}{(E_n-\mathcal{E}_m)(E_n-\mathcal{E}_{m'})} + \begin{matrix}\textit{higher}\\ \textit{terms}\end{matrix}$$

where the higher terms have the same structure as the explicitly presented ones (note that the iterative solution provided the result in the form of power expansion in the parameter α*). The final step is to multiply the iterative solution by* $\langle\varphi_n|\,H$ *and to use* $\langle\varphi_n|H|\psi_n\rangle = E_n\langle\varphi_n|\psi_n\rangle$ *and* $\langle\varphi_n|H|\varphi_n\rangle = \mathcal{E}_n\langle\varphi_n|\varphi_n\rangle + \alpha\langle\varphi_n|H'|\varphi_n\rangle$. *With the standard normalization (the same as in the Rayleigh–Schrödinger perturbation theory), i.e. for* $\langle\varphi_n|\varphi_n\rangle = \langle\varphi_n|\psi_n\rangle = 1$ *one obtains the Brillouin–Wigner expansion*

$$\begin{aligned} E_n = \;& \mathcal{E}_n + \alpha\langle\varphi_n|H'|\varphi_n\rangle + \alpha^2\sum_{m\neq n}\frac{\langle\varphi_n|H'|\varphi_m\rangle\langle\varphi_m|H'|\varphi_n\rangle}{E_n-\mathcal{E}_m} \\ & + \alpha^3\sum_{m\neq n}\sum_{m'\neq n}\frac{\langle\varphi_n|H'|\varphi_m\rangle\langle\varphi_m|H'|\varphi_{m'}\rangle\langle\varphi_{m'}|H'|\varphi_n\rangle}{(E_n-\mathcal{E}_m)(E_n-\mathcal{E}_{m'})} + \begin{matrix}\textit{higher}\\ \textit{terms}\end{matrix} \end{aligned}$$

1.2.4 Quantum particles and classical fields

We are going to finish this second introduction by discussing a remarkable classical limit (of quantum theory expressed in terms of the creation and annihilation operators) which differs significantly from the notoriously known classical limit provided by the Ehrenfest theorems. In their simplest version these theorems state that the mean values of particle position and momentum operators in one dimension fulfill the following equations

$$\frac{d\bar{p}}{dt} = -\overline{V'(x)} \qquad \frac{d\bar{x}}{dt} = \frac{\bar{p}}{m}$$

where $\bar{A} = \langle\psi|\,\hat{A}\,|\psi\rangle$ for some state $|\psi\rangle$ (not denoted explicitly, to avoid overloaded notation). Generalization to three dimensions and to several particles is straightforward.

Let us emphasize that since in general $\overline{V'(x)} \neq V'(\bar{x})$, the Ehrenfest equation $\frac{d\bar{p}}{dt} = -\overline{V'(x)}$ is not a classical one. The difference between the Ehrenfest and Newton equations is given by the Taylor expansion around the mean value $\bar{x}$: $f(x) = f(\bar{x}) + f'(\bar{x})\Delta x + \frac{1}{2}f''(\bar{x})(\Delta x)^2 + \cdots$ (where $\Delta x = x - \bar{x}$) leading to $\overline{f(x)} = f(\bar{x}) + \frac{1}{2}f''(\bar{x})\overline{(\Delta x)^2} + \cdots$ (the term linear in $\overline{\Delta x}$ is missing, due to $\overline{\Delta x} \equiv 0$). The Ehrenfest equation

$$\frac{d\bar{p}}{dt} = -V'(\bar{x}) - \frac{1}{2}V'''(\bar{x})\overline{(\Delta x)^2} + \cdots$$

is a quantum equation for mean values, which looks like the Newton equation $\frac{d\bar{p}}{dt} = -V'(\bar{x})$ with the corrections proportional to the variance and higher central momenta of the position operator in the state $|\psi\rangle$. If these corrections are small enough, the Ehrenfest equation becomes quasi-classical. If the corrections are negligible (or even vanishing), the Ehrenfest equation becomes truly classical Newton equation.

The Ehrenfest equations demonstrate how the classical physics can emerge from the quantum mechanics. One possibility is the (at most) quadratic potential $V(x) = a + bx + cx^2$, for which the third and higher derivatives vanish. For such a potential the Ehrenfest equations are truly classical for any state $|\psi\rangle$.

Another possibility is a potential which contains a quadratic part and a "small correction". In this case the Ehrenfest equations are only quasi-classical (since the terms with higher derivatives are present, even if small), again for any state $|\psi\rangle$.

Yet another example of classical or quasi-classical behavior of quantum systems is provided not by the properties of the potential, but rather by the smallness of the higher central momenta $\overline{(\Delta x)^n}$ $(n \geq 2)$ for some specific states $|\psi\rangle$.

Remark: *1*

2

3

4

5

6

7

8

9

10

11

The Ehrenfest theorems are special cases of general time evolution of mean values in QM. In the Schrödinger picture operators are time independent and states evolve according to the Schrödinger equation $i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle$ (for the bra-vectors $-i\hbar\frac{d}{dt}\langle\psi(t)| = \langle\psi(t)|\hat{H}$). As a consequence, the evolution of the mean value of an operator $\hat{A}$ is described by

$$i\hbar\tfrac{d}{dt}\bar{A} = \overline{[\hat{A},\hat{H}]}$$

The same result is, of course, obtained in the Heisenberg picture, where states are time independent and operators evolve according to the Heisenberg equation $i\hbar\frac{\partial}{\partial t}\hat{A}(t) = [\hat{A}(t),\hat{H}]$. This equation is quite similar to the classical Hamilton equation $\frac{\partial}{\partial t}A(t) = \{A(t),H\}$ and relation between the two is much closer than mere similarity.

The Poisson brackets in the Hamilton equation are defined as $\{B,C\} = \sum_i \frac{\partial B}{\partial q_i}\frac{\partial C}{\partial p_i} - \frac{\partial B}{\partial p_i}\frac{\partial C}{\partial q_i}$ (beware of different overall sign used by different authors) and they are usually evaluated by explicit differentiation. This, however, is not the only possibility. An alternative version of their evaluation is provided by two properties of the brackets which follow directly from the definition:

- $\{B,CD\} = \{B,C\}D + C\{B,D\} \qquad \{BC,D\} = B\{C,D\} + \{B,D\}C$
- $\{q_i,q_j\} = \{p_i,p_j\} = 0 \qquad \{q_i,p_j\} = \delta_{ij}$

If the functions $A(q,p)$ and $H(q,p)$ are polynomials in the variables q_i, p_j (as they usually are) then the brackets $\{A,B\}$ can be calculated by repeated use of the first property. In each step one q_i or p_j is brought outside a bracket, so that at the end only the brackets of q_i and p_j are present. Once these are evaluated according to the second property, the result becomes a sum of polynomials in q_i, p_j. We will denote this specific sum of polynomials as $S(q,p)$, so that

$$\{A(q,p),H(q,p)\} = S(q,p)$$

Now the same technique can be used for evaluation of the commutator $[\hat{A},\hat{H}]$ if the operators $\hat{A}$ and $\hat{H}$ are polynomials in operators $\hat{q}_i, \hat{p}_j$ satisfying the canonical commutation relations. The point is that commutators also have the property analogous to the first property of Poisson brackets and canonical commutation relations are (up to a constant) analogous to the second property of Poisson brackets

- $[\hat{B},\hat{C}\hat{D}] = [\hat{B},\hat{C}]\hat{D} + \hat{C}[\hat{B},\hat{D}] \qquad [\hat{B}\hat{C},\hat{D}] = \hat{B}[\hat{C},\hat{D}] + [\hat{B},\hat{D}]\hat{C}$
- $[\hat{q}_i,\hat{q}_j] = [\hat{p}_i,\hat{p}_j] = 0 \qquad [\hat{q}_i,\hat{p}_j] = i\hbar\delta_{ij}$

So the same line of evaluation leads us to the same sum of polynomials as it did in the case of Poisson brackets

$$[\hat{A}(\hat{q},\hat{p}),\hat{H}(\hat{q},\hat{p})] = i\hbar\hat{S}(\hat{q},\hat{p})$$

where $\hat{S}$ is (up to the ordering of the operators in products) the same polynomial of $\hat{q},\hat{p}$ as S was of q,p.

So at the quantum level we obtained $i\hbar\frac{d}{dt}\bar{A} = i\hbar\overline{\hat{S}(\hat{q},\hat{p})}$. Taylor expanding the function $\hat{S}(\hat{q},\hat{p})$ around the mean values $\bar{q},\bar{p}$ one obtains (just like in the case of the Ehrenfest theorem) $i\hbar\frac{d}{dt}\bar{A} = i\hbar S(\bar{q},\bar{p}) + \cdots$ where ellipsis stand for corrections proportional to variances and higher central momenta of $\hat{q}$ and $\hat{p}$. Taking everything together one gets

$$\tfrac{d}{dt}\bar{A} = \{A(\bar{q},\bar{p}),H(\bar{q},\bar{p})\} + \cdots$$

which is the generalization of the Ehrenfest theorems. Indeed, without ellipsis this is nothing else but the truly classical equation for time evolution in the Hamiltonian formalism of the classical mechanics (written in terms of Poisson brackets).

What does the Hamiltonian formalism in the classical mechanics have to do with the formalism of creation and annihilation operators in the quantum mechanics? The bridge between the two is spanned by a simple fact that for every pair of bosonic creation and annihilation operators one can define another two operators satisfying the canonical commutation relations. Indeed, for any creation and anihilation operators satisfying $[a_i, a_j^+] = \delta_{ij}$ and $[a_i, a_j] = [a_i^+, a_j^+] = 0$ one can define specific hermitian linear combinations[35]

$$\hat{\Phi}_i = \sqrt{\tfrac{\hbar}{2}}\,(a_i^+ + a_i) \qquad\qquad \hat{\Pi}_i = i\sqrt{\tfrac{\hbar}{2}}\,(a_i^+ - a_i)$$

These operators do not have a physical meaning of position and momentum operators, nevertheless they do satisfy the canonical commutation relations (check it)

$$[\hat{\Phi}_i, \hat{\Pi}_j] = i\hbar\delta_{ij} \qquad\qquad [\hat{\Phi}_i, \hat{\Phi}_j] = [\hat{\Pi}_i, \hat{\Pi}_j] = 0$$

As a consequence, their mean values satisfy the generalized Ehrenfest theorems

$$\dot{\bar{\Phi}}_i = \{\bar{\Phi}_i, H(\bar{\Phi}, \bar{\Pi})\} + \cdots \qquad\qquad \dot{\bar{\Pi}}_i = \{\bar{\Pi}_i, H(\bar{\Phi}, \bar{\Pi})\} + \cdots$$

for any Hamiltonian $\hat{H}(\hat{\Phi}, \hat{\Pi})$ and in any state. If the terms hidden in the ellipses are negligible in some states then the mean values $\bar{\Phi}_i$ and $\bar{\Pi}_i$ in these states are classical quantities emerging naturally from the many-particle quantum mechanics.

The question now is, if there are states in many-particle systems in which the said ellipses are negligible and if these states are somehow typical for these systems. The answer is affirmative. The point is that the Hamiltonian for non-interacting particles (ideal gas) turns out to be quadratic in canonical operators. Indeed, since $a_i^+ = (\hat{\Phi}_i - i\hat{\Pi}_i)/\sqrt{2\hbar}$ and $a_i = (\hat{\Phi}_i + i\hat{\Pi}_i)/\sqrt{2\hbar}$

$$H = \sum_i E_i a_i^+ a_i = \tfrac{1}{2\hbar}\sum_i E_i(\hat{\Phi}_i^2 + \hat{\Pi}_i^2 - \hbar)$$

For such a Hamiltonian the generalized Ehrenfest equations for the mean values of the canonical operators are truly classical, with no ellipses present (check it)[36]

$$\dot{\bar{\varphi}}_i(t) = E_i\bar{\pi}_i(t) \qquad\qquad \dot{\bar{\pi}}_i = -E_i\bar{\varphi}_i(t)$$

So for an ideal gas the mean values $\bar{\Phi}_i, \bar{\Pi}_i$ in any state are truly classical quantities. For a slightly non-ideal gas (with the interaction Hamiltonian being just a small correction to the ideal gas Hamiltonian), the ellipses in the generalized Ehrenfest equations are present again, but only as a small correction to the ideal gas case.

We have achieved quite a remarkable result. For an ideal gas of bosons (with no bound states at all, not to mention macroscopic ones) canonical operators with mean values (in any state) obeying classical equations of motion are always present. These operators do not have (so far) an obvious physical interpretation, but they are hermitian and therefore they should correspond to some observables of the quantum theory of this gas. The mean values of these observables, on the other hand, should represent some classical quantities. What are these classical quantities and how do we measure them?

[35] Similar linear combinations should be familiar from the notoriously known treatment of the LHO exploiting the raising and lowering operators. Let us stress, however, that we are not discussing raising and lowering operators for harmonic oscillators here, but creation and annihilation operators of some real particles (which may have nothing to do with harmonic oscillators).

[36] Up to the irrelevant (infinite) constant this is the Hamiltonian of a system of harmonic oscillators. So even if the harmonic oscillators were of no importance for the definition of the operators $\hat{\Phi}_i, \hat{\Pi}_i$, they are quite important in other respect: the ideal gas is equivalent to a system of harmonic oscillators and for harmonic oscillators the Ehrenfest theorems are not only quasi-classical, but truly classical for any state.

To reveal the physical nature of the new classical quantity emerging from quantum mechanics let us consider an ideal gas of free spinless particles. For such a gas the discrete quantum number i is to be replaced by continuous momentum $\vec{p}$ and the Ehrenfest equations become

$$\dot{\bar{\varphi}}(\vec{p},t) = E(\vec{p})\ \bar{\pi}(\vec{p},t) \qquad\qquad \dot{\bar{\pi}}(\vec{p},t) = -E(\vec{p})\ \bar{\varphi}(\vec{p},t)$$

with $E(\vec{p}) = \frac{1}{2m}p^2$ and $E(\vec{p}) = \sqrt{m^2+p^2}$ for non-relativistic and relativistic gases respectively.

It is now instructive to obtain more combine the two equations into one

$$\ddot{\bar{\varphi}}(\vec{p},t) = -E^2(\vec{p})\ \bar{\varphi}(\vec{p},t)$$

and then to trade the powers of $\vec{p}$ for partial derivatives by means of the Fourier transformation $f(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}.\vec{x}} f(\vec{p},t)$. The resulting equations are (check it)

$$\ddot{\bar{\varphi}}(\vec{x},t) + \frac{1}{4m^2}\triangle\triangle\bar{\varphi}(\vec{x},t) = 0 \qquad\qquad \ddot{\bar{\varphi}}(\vec{x},t) + \triangle\bar{\varphi}(\vec{x},t) + m^2\bar{\varphi}(\vec{x},t) = 0$$

in non-relativistic and relativistic cases respectively. So we have demonstrated that the new classical quantity behaves like a classical field satisfying some specific partial differential equation.[37]

Do we know such fields from classical physics? Well, in fact not, because of simple reason: there are no stable spinless particles in our nature and therefore there are no corresponding classical fields. Nevertheless these

But what if there were stable spinless bosons in our world. Then we would observe the corresponding classical fields. The relativistic case for massless particles, however, turns out to be familiar the notoriously known wave equation.

[37] The non-relativistic field equation does not look very familiar, the relativistic one is much more common, especially in the case of massless particles (where it is just the well known wave equation). But this is not the point. These very classical fields may not exist in our nature for a simple reason. There are no stable spinless particles in nature and so there is no corresponding classical field. The situation is different for particles with spin 1, where massless photons do exist.

1.3 Relativity and Quantum Theory

Actually, as to the quantum fields, the keyword is relativity. Even if QFT is useful also in the nonrelativistic context (see the previous section and the Appendix), the very notion of the quantum field originated from an endeavor to fit together relativity and quantum theory. This is a nontrivial task: to formulate a relativistic quantum theory is significantly more complicated than it is in a nonrelativistic case. The reason is that specification of a Hamiltonian, the crucial operator of any quantum theory, is much more restricted in relativistic theories.

To understand the source of difficulties, it is sufficient to realize that to have a relativistic quantum theory means to have a quantum theory with measurable predictions which remain unchanged by the relativistic transformations. The relativistic transformations at the classical level are the space-time transformations conserving the interval $ds = \eta_{\mu\nu}dx^\mu dx^\nu$, i.e. boosts, rotations, translations and space-time inversions (the list is exhaustive). They constitute a group.

Now to the quantum level: whenever macroscopic measuring and/or preparing devices enjoy a relativistic transformation, the Hilbert (Fock) space should transform according to a corresponding linear[38] transformation. It is also almost self-evident that the quantum transformation corresponding to a composition of classical transformations (at the level of macroscopic devices) should be equal to the composition of quantum transformations corresponding to the individual classical transformations. So at the quantum level the relativistic transformations should constitute a representation of the group of classical transformations.

The point now is that the Hamiltonian, as the time-translations generator, is just one of ten generators of the Poincaré group (boosts, rotations, translations). Consequently, unlike in a nonrelativistic QM, in a relativistic quantum theory one cannot specify the Hamiltonian alone, one has rather to specify it within a complete representation of the Poincaré algebra. This is the starting point of any effort to get a relativistic quantum theory, even if it is not always stated explicitly. The outcome of such efforts are quantum fields. Depending on the philosophy adopted, they use to emerge in at least two different ways. We will call them particle-focused and field-focused.

The particle-focused approach, represented mostly by the Weinberg's book, is very much in the spirit of the previous section. One starts from the Fock space, which is systematically built up from the 1-particle Hilbert space. Creation and annihilation operators are then defined as very natural objects, namely as maps from the n-particle subspace into $(n \pm 1)$-particle ones, and only afterwards quantum fields are built from these operators in a bit sophisticated way (keywords being cluster decomposition principle and relativistic invariance).

The field-focused approach (represented by Peskin–Schroeder, and shared by the majority of textbooks on the subject) starts from the quantization of a classical field, introducing the creation and annihilation operators in this way as quantum incarnations of the normal mode expansion coefficients, and finally providing Fock space as the world for these operators to live in. The logic behind this formal development is nothing else but construction of the Poincaré group generators. So, again, the corner-stone is the relativistic invariance.

[38] Linearity of transformations at the quantum level is necessary to preserve the superposition principle. The symmetry transformations should not change measurable things, which would not be the case if the superposition $|\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle$ would transform to $T|\psi\rangle \neq c_1 T|\psi_1\rangle + c_2 T|\psi_2\rangle$.

1.3.1 Lorentz and Poincaré groups

This is by no means a systematic exposition to the Lorentz and Poincaré groups and their representations. It is rather a summary of important relations, some of which should be familiar (at some level of rigor) from the previous courses.

the groups

The classical relativistic transformations constitute a group, the corresponding transformations at the quantum level constitute a representation of this group. The group transformations are

$$x^\mu \to \Lambda^\mu_{\ \nu} x^\nu + a^\mu$$

where $\Lambda^\mu_{\ \nu}$ are combined space rotations, boosts and space-time inversions, while a^μ describe space-time translations. The rotations around (and the boosts along) the space axes are [39]

$$R_1(\vartheta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\vartheta & -\sin\vartheta \\ 0 & 0 & \sin\vartheta & \cos\vartheta \end{pmatrix} \qquad B_1(\beta) = \begin{pmatrix} \mathrm{ch}\,\beta & \mathrm{sh}\,\beta & 0 & 0 \\ \mathrm{sh}\,\beta & \mathrm{ch}\,\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_2(\vartheta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\vartheta & 0 & \sin\vartheta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\vartheta & 0 & \cos\vartheta \end{pmatrix} \qquad B_2(\beta) = \begin{pmatrix} \mathrm{ch}\,\beta & 0 & \mathrm{sh}\,\beta & 0 \\ 0 & 1 & 0 & 0 \\ \mathrm{sh}\,\beta & 0 & \mathrm{ch}\,\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$R_3(\vartheta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\vartheta & -\sin\vartheta & 0 \\ 0 & \sin\vartheta & \cos\vartheta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad B_3(\beta) = \begin{pmatrix} \mathrm{ch}\,\beta & 0 & 0 & \mathrm{sh}\,\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \mathrm{sh}\,\beta & 0 & 0 & \mathrm{ch}\,\beta \end{pmatrix}$$

where ϑ is the rotation angle (in counterclockwise direction) and $\tanh\beta = v/c$ where v is the velocity of the boost. They constitute the Lorentz group. It is a non-compact (since $\beta \in (-\infty, \infty)$) Lie group. The translations along the space-time axes are

$$T_0(\alpha) = \begin{pmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad T_1(\alpha) = \begin{pmatrix} 0 \\ \alpha \\ 0 \\ 0 \end{pmatrix} \quad T_2(\alpha) = \begin{pmatrix} 0 \\ 0 \\ \alpha \\ 0 \end{pmatrix} \quad T_3(\alpha) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \alpha \end{pmatrix}$$

Together with the boosts and rotations they constitute the Poincaré group. It is a non-compact Lie group (on top of the non-compactness of the Lorentz subgroup one has $\alpha \in (-\infty, \infty)$). The space-time inversions are three different diagonal matrices. The time inversion is given by $T = \mathrm{diag}(-1, 1, 1, 1)$, the space inversion is given by $P = \mathrm{diag}(1, -1, -1, -1)$ and their product is $PT = \mathrm{diag}(-1, -1, -1, -1)$.

[39] A comment on sings seems to be appropriate here. Transformations can be understood as active ones (transformations of objects) or passive ones (transformations of viewpoints). In the quantum mechanics we usually understand rotations and translations as active transformations within one reference frame. In the special relativity, however, we usually understand Lorentz transformations as passive transformations between two different reference frames. If the primed frame is rotated or boosted with respect to the unprimed frame then we can express the unprimed coordinates as functions of the primed ones: $x^\mu = \Lambda^\mu_{\ \nu'} x^{\nu'}$. Our choice of signs in front of $\sin(\vartheta)$ and $\mathrm{sh}(\beta)$ corresponds to this choice (which, by the way, is equivalent to active transformations). If we decide to express primed coordinates as functions of unprimed, we would get opposite signs.

the Lie algebra

The standard technique of finding representations of a Lie group is to find representations of the corresponding Lie algebra (the commutator algebra of the generators). The standard choice of the generators corresponds to the 10 types of (infinitesimal) transformations listed above: rotations $R_i(\varepsilon) = \mathbb{1} - i\varepsilon J_i + \mathcal{O}(\varepsilon^2)$, boosts $B_i(\varepsilon) = \mathbb{1} - i\varepsilon K_i + \mathcal{O}(\varepsilon^2)$, space translations[40] $T_i(\varepsilon) = -i\varepsilon P_i + \mathcal{O}(\varepsilon^2)$ and time translation[41] $T_0(\varepsilon) = i\varepsilon P_0 + \mathcal{O}(\varepsilon^2)$.

$$J_1 = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad K_1 = i\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$J_2 = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \qquad K_2 = i\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$J_3 = i\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad K_3 = i\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$P_0 = i\begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad P_1 = i\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad P_2 = i\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad P_3 = i\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Calculation of the commutators is straightforward[42] (even if not very exciting)

$$[J_i, J_j] = i\varepsilon_{ijk}J_k \qquad [J_i, P_0] = 0$$

$$[J_i, K_j] = i\varepsilon_{ijk}K_k \qquad [J_i, P_j] = i\varepsilon_{ijk}P_k$$

$$[K_i, K_j] = -i\varepsilon_{ijk}J_k \qquad [K_i, P_0] = -iP_i$$

$$[P_\mu, P_\nu] = 0 \qquad [K_i, P_j] = -iP_0\delta_{ij}$$

[40] For additive group the unit element is zero (no translation at all) and therefore we have 0 instead of 1 in the definition of the translation generator.

[41] The sign in the definition of this generator is plus instead of the usual minus. Such a choice of sign is a common habit in quantum mechanics and it leads to the standard relations for the space translation generators $P_i = -i\partial_i$ (the momentum operator in the x-representation) and the time translation generator $P_0 = i\partial_0$ i.e. $H = i\partial_t$ (the Schrödinger equation). In the non-relativistic QM this choice of signs may look a bit weird (even if we are all used to it), but in relativity it is quite natural, since it corresponds to the relativistic relation $P_\mu = i\partial_\mu$. Recall that the sign of space components of a 4-vector and the 4-gradient are opposite: $P_\mu = (P_0, -\vec{P})$, $\partial_\mu = (\partial_0, \nabla)$.

[42] Commutators of J_i (or K_i) and P_μ follow from $\Lambda^\mu_{\ \nu}(x^\nu + a^\nu) - \left(\Lambda^\mu_{\ \nu}x^\nu + a^\mu\right) = \Lambda^\mu_{\ \nu}a^\nu - a^\mu = (\Lambda - 1)^\mu_{\ \nu}\, a^\nu$, i.e. one obtains the commutator by acting of the corresponding generator J_i or K_i on the (formal) vector P_μ.

Remark: *As the matter of fact (a sad one), there are multiple possibilities for choice of sign when speaking about symmetry groups and their representations. One can understand transformations in either active or passive sense (the difference being just the sign). One can use either plus or minus sign in the definition of generators* $U(\varepsilon) = \mathbb{1} \pm i\varepsilon G + \mathcal{O}(\varepsilon^2)$ *(and we have already used both – one for time translations and the other one for all the others). One can, but is not obliged to, use the explicit* i *in this definition, which change the sign of commutators (let us note that the standard physicists's argument that with the* i *the generators become hermitian does not work for the Lorentz boosts, as we could notice). One can define transformations of operators either as* $A \to UAU^{-1}$ *or as* $A \to U^{-1}AU$*. The first choice corresponds to simultaneous transformations of state vectors and operators, the second choice corresponds to the Heisenberg picture of transforming operators and untouched states. For infinitesimal transformations the difference is again just the sing. And one can use different signatures for Minkowski metric tensor, i.e. either* $\eta_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$ *or* $\eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$*. In this text we will (try to) use: active transformations,* $U(\varepsilon) = \mathbb{1} - i\varepsilon G + \mathcal{O}(\varepsilon^2)$ *(with exception of the time translations),* $A \to UAU^{-1}$ *and* $\eta_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$*.*

Remark: *The commutation relations hold for the Lie algebra as well as for its representations. For representations in Hilbert space of quantum states, the generators of time-translations, space-translations and space-rotations are energy, momentum and angular momentum operators respectively. From their commutation relations one can conclude that*
1. Representations of J_j*,* K_j*,* P_j *transform as components of 3-vectors.*[43]
2. Representations of P_0*,* P_j*,* J_j *are conserved quantities (they commute with the hamiltonian).*
3. Representations of K_j *are not conserved (they do not commute with the hamiltonian).*
4, Representations of P_μ *transform as components of 4-vector.*[44]
5. Representations of J_i *and* K_i *transform as components of antisymmetric 4-tensor.*[45]

Remark: *The generators of the Lorentz group are* 4×4 *matrices with only non-zero elements being* $\pm i$*. Such matrices can be written using just the metric tensor* η*. As an example let us write the matrix* $(K_1)^\mu_{\ \nu}$ *as* $i(\eta^\mu_1\eta_{0\nu} - \eta^\mu_0\eta_{1\nu})$*. In general one can write (as the reader should check)*

$$(M_{\alpha\beta})^\mu_{\ \nu} = i\left(\eta^\mu_\alpha\eta_{\beta\nu} - \eta^\mu_\beta\eta_{\alpha\nu}\right) \qquad J_k = \frac{1}{2}\varepsilon_{ijk}M_{ij} \qquad K_i = M_{i0}$$

The matrices $M_{\alpha\beta}$ *are numbered by the indices* α, β *and are antisymmetric in them. One can therefore write the general infinitesimal Lorentz transformation as* $\Lambda = \mathbb{1} - \frac{i}{2}\epsilon^{\alpha\beta}M_{\alpha\beta}$ *with an antisymmetric* $\epsilon^{\alpha\beta}$*. The advantage of the matrices* $M^{\alpha\beta}$ *is twofold. First, it provides a compact notation, and second, it reveals clearly the 4-tensor character of the Lorentz generators. Indeed, since* η *is the (metric) 4-tensor, the* M *is by definition 4-tensor with four indices (rank 4), i.e. it is a second rank tensor with respect to the indices* α *and* β*.*

---

[43] Infinitesimal rotations of three operators A_j are given by $(1 - i\varepsilon J_i)A_j(1 + i\varepsilon J_i) = A_j - i\varepsilon[J_i, A_j] + \mathcal{O}(\varepsilon^2)$, which for $A_i = J_i, K_i, P_i$ is equal to $A_j + \varepsilon\,\varepsilon_{ijk}A_k + \mathcal{O}(\varepsilon^2)$ (due to the commutation relations). Components of a 3-vector $\vec{a}$, on the other hand, are transformed according to the defining representation of the rotation group $\vec{a} \to (1 - i\varepsilon J_i)\,\vec{a}$, which for $(J_i)_{jk} = -i\varepsilon_{ijk}$ (in this representation), equals to $a_j \to a_j - \varepsilon\,\varepsilon_{ijk}a_k + \mathcal{O}(\varepsilon^2)$.
So up to the sign the operators $A_i = J_i, K_i, P_i$ transform as the components of a 3-vector. And since in this case the difference in the sign is just the difference between e.g. active and passive transformations, we may conclude that operators J_i, K_i, P_i do indeed transform as components of 3-vectors.

[44] Infinitesimal Lorentz transformations of four operators A_j are $(1 - i\varepsilon J_i)P_\mu(1 + i\varepsilon J_i) = P_\mu - i\varepsilon[J_i, P_\mu] + \mathcal{O}(\varepsilon^2)$ and $(1 - i\varepsilon K_i)P_\mu(1 + i\varepsilon K_i) = P_\mu - i\varepsilon[K_i, P_\mu] + \mathcal{O}(\varepsilon^2)$. Components of a 4-vector a_μ are transformed according to the defining representation of the Lorentz group $a \to (1 - i\varepsilon J_i)\,a$ and $a \to (1 - i\varepsilon K_i)\,a$. The commutators (in general) and the generators in the defining representation are given explicitly at the previous page, from where one can conclude (just like in the previous footnote) that indeed up to the sign the operators P_μ transform as the components of a 4-vector.

[45] This should be clear from the next remark.

the scalar representation of the Poincaré group

Representations of the Poincaré group are of vital importance to any serious attempt to discuss relativistic quantum theory. It is, however, not our task right now. For quite some time we will need only the simplest representation – the so-called scalar one, which is quite easy to guess (and that's what we are going to do now). Systematic analysis of representations of the Lorentz and Poincaré groups and all the complications introduced by non-scalar representations are postponed to next chapters.

Let us consider a Hilbert space in which some representation of the Poincaré group is defined. Perhaps the most convenient basis of such a space is the one defined by the eigenvectors of the representations of the translation generators $\mathcal{P}_\mu$, commuting with each other. The eigenvectors are usually denoted as $|p,\sigma\rangle$, where $p=(p_0,p_1,p_2,p_3)$ stands for eigenvalues of $\mathcal{P}$ and σ stands for any other quantum numbers. In this notation $\mathcal{P}_\mu|p,\sigma\rangle = p_\mu|p,\sigma\rangle$.

At this point we are going to consider only the simplest case, the so-called scalar representation, in which no σ is involved. The states in this representation are characterized completely by the 4-momentum p, which is exactly what one expects to be the case for one spinless particle. For the translation generators in this representation one has

$$\mathcal{P}_\mu|p\rangle = p_\mu|p\rangle$$

The representation of non-infinitesimal space-time translations is obtained by exponentiation of the generators. If an operator $U(\Lambda,a)$ is the element of the representation, corresponding to the Poincaré transformation $x\to\Lambda x+a$, then $U(1,a)=e^{-ia\mathcal{P}}$. And since $|p\rangle$ is an eigenstate of $\mathcal{P}_\mu$, one has

$$U(1,a)|p\rangle = e^{-ipa}|p\rangle$$

where $pa=p_\mu a^\mu$ is a scalar product of 4-vectors.

How does the representation of the Lorentz transformations look like? Since p is a 4-vector, the obvious first guess for generators is

$$\mathcal{J}_i|p\rangle = |J_i p\rangle$$
$$\mathcal{K}_i|p\rangle = |K_i p\rangle$$

or, at the level of non-infinitesimal Lorentz transformations

$$U(\Lambda,0)|p\rangle = |\Lambda p\rangle$$

It is straightforward to check (do it) that this, indeed, is a representation of the Lorentz group.

Finally one puts the Lorentz transformations and translations together to get

$$U(\Lambda,a)|p\rangle = e^{-i(\Lambda p)a}|\Lambda p\rangle$$

where $(\Lambda p)a=\Lambda_\mu{}^\nu p_\nu a^\mu$. It is again easy to check directly (do it) that this defines a representation of the Poincaré group.[46]

[46] Let us remark that once the spin is involved, the parameter σ enters the game and the transformation is a bit more complicated. It goes like this: $U(\Lambda,a)|p,\sigma\rangle = \sum_{\sigma'} e^{-i(\Lambda p)a}C_{\sigma\sigma'}|\Lambda p,\sigma'\rangle$ and the coefficients $C_{\sigma\sigma'}$ define the particular representation of the Lorentz group. These complications, however, do not concern us now. For our present purposes the simplest scalar representation will be sufficient.

The next step is to realize that the operator $P^2 = P_\mu P^\mu$ commutes with all the generators of the Poincaré group (check it), i.e. it is a Casimir operator of this group. If we use the symbol m^2 to denote the eigenvalue of this operator then the irreducible representations of the Poincaré group can be classified by the value of the m^2. The relation between the energy and the 3-momentum of the state $|p\rangle$ is

$$E^2 - \vec{p}^2 = m^2$$

i.e. for each value of m^2 one has the Hilbert space of states of a free spinless relativistic particle with the mass m. The question now is how to add some interaction and this turns out to be a surprisingly difficult task.

Nevertheless, this rather trivial representation is a very important one — it becomes the starting point of what we call the particle-focused approach to the quantum field theory. We shall comment on this briefly in the next paragraph.

the scalar representation of the Poincaré group. OLD

Investigations of the Poincaré group representations are of vital importance to any serious attempt to discuss relativistic quantum theory. It is, however, not our task right now. For quite some time we will need only the simplest representation, the so-called scalar one. All complications introduced by higher representations are postponed to the next parts of the text.

Let us consider a Hilbert space in which some representation of the Poincaré group is defined. Perhaps the most convenient basis of such a space is the one defined by the eigenvectors of the translation generators P_μ, commuting with each other. The eigenvectors are usually denoted as $|p,\sigma\rangle$, where $p=(p_0,p_1,p_2,p_3)$ stands for eigenvalues of P and σ stands for any other quantum numbers. In this notation one has

$$P_\mu |p,\sigma\rangle = p_\mu |p,\sigma\rangle$$

The representation of space-time translations are obtained by exponentiation of generators. If an operator $U(\Lambda,a)$ is the element of the representation, corresponding to the Poincaré transformation $x\to\Lambda x+a$, then $U(1,a)=e^{-iaP}$. And since $|p,\sigma\rangle$ is an eigenstate of P_μ, one has

$$U(1,a)|p,\sigma\rangle = e^{-ipa}|p,\sigma\rangle$$

where $pa=p_\mu a^\mu$ is a scalar product of 4-vectors.

And how does the representation of the Lorentz subgroup look like? Since the p_μ is a 4-vector, one may be tempted to try $U(\Lambda,0)|p,\sigma\rangle \stackrel{?}{=} |\Lambda p,\sigma\rangle$. This really works, but only in the simplest, the so-called scalar representation, in which no σ is involved. It is straightforward to check that in such a case the relation

$$U(\Lambda,a)|p\rangle = e^{-i(\Lambda p)a}|\Lambda p\rangle$$

defines indeed a representation of the Poincaré group (check it).[47]

Now it looks like if we had reached our goal — we have a Hilbert space with a representation of the Poincaré group acting on it, i.e. we have a relativistic quantum theory. A short inspection reveals, however, that this is just the rather trivial case of the free particle. To see this, it is sufficient to realize that the operator $P^2=P_\mu P^\mu$ commutes with all the generators of the Poincaré group (check it), i.e. it is a Casimir operator of this group. If we denote the eigenvalue of this operator by the symbol m^2 then the irreducible representations of the Poincaré group can be classified by the value of the m^2. The relation between the energy and the 3-momentum of the state $|p\rangle$ is $E^2-\vec{p}^2=m^2$, i.e. for each value of m^2 we really do have the Hilbert space of states of a free relativistic particle with the mass m. (The reader is encouraged to clarify him/herself how should the Hilbert space of the states of free relativistic particle look like. He/she should come to conclusion, that it has to be equal to what we have encountered just now.)

Nevertheless, this rather trivial representation is a very important one — it becomes the starting point of what we call the particle-focused approach to the quantum field theory. We shall comment on this briefly in the next paragraph.

[47] Let us remark that once the spin is involved, the parameter σ enters the game and the transformation is a bit more complicated. It goes like this: $U(\Lambda,a)|p,\sigma\rangle=\sum_{\sigma'}e^{-i(\Lambda p)a}C_{\sigma\sigma'}|\Lambda p,\sigma'\rangle$ and the coefficients $C_{\sigma\sigma'}$ define the particular representation of the Lorentz group. These complications, however, do not concern us now. For our present purposes the simplest scalar representation will be sufficient.

The Hilbert space spanned over the eigenvectors $|p,\sigma\rangle$ of the translation generators P_μ is not the only possible choice of a playground for the relativistic quantum theory. Another quite natural Hilbert space is provided by the functions $\varphi(x)$. A very simple representation of the Poincaré group is obtained by the mapping

$$\varphi(x) \to \varphi(\Lambda x + a)$$

This means that with any Poincaré transformation $x \to \Lambda x + a$ one simply pulls back the functions in accord with the transformation. It is almost obvious that this, indeed, is a representation (if not, check it in a formal way). Actually, this representation is equivalent to the scalar representation discussed above, as we shall see shortly. (Let us remark that more complicated representations can be defined on n-component functions, where the components are mixed by the transformation in a specific way.)

It is straightforward to work out the generators in this representation (and we shall need them later on). The changes of the space-time position x^μ and the function $\varphi(x)$ by the infinitesimal Poincaré transformation are δx^μ and $\delta x^\mu \partial_\mu \varphi(x)$ respectively, from where one can directly read out the Poincaré generators in this representation

$$\begin{aligned}
\mathcal{J}_i \varphi(x) &= (J_i x)^\mu \, \partial_\mu \varphi(x) \\
\mathcal{K}_i \varphi(x) &= (K_i x)^\mu \, \partial_\mu \varphi(x) \\
\mathcal{P}_\mu \varphi(x) &= i\partial_\mu \varphi(x)
\end{aligned}$$

Using the explicit knowledge of the generators J_i and K_i one obtains

$$\begin{aligned}
\mathcal{J}_i \varphi(x) &= \frac{i}{2}\varepsilon_{ijk}\left(\delta^\mu_j \eta_{k\nu} - \delta^\mu_k \eta_{j\nu}\right) x^\nu \partial_\mu \varphi(x) = -i\varepsilon_{ijk} x_j \partial_k \varphi(x) \\
\mathcal{K}_i \varphi(x) &= i\left(\delta^\mu_i \eta_{0\nu} - \delta^\mu_0 \eta_{i\nu}\right) x^\nu \partial_\mu \varphi(x) = i x_0 \partial_i \varphi(x) - i x_i \partial_0 \varphi(x)
\end{aligned}$$

or even more briefly $\vec{\mathcal{J}}\,\varphi(x) = -i\vec{x} \times \nabla\varphi(x)$ and $\vec{\mathcal{K}}\,\varphi(x) = it\nabla\varphi(x) - i\vec{x}\dot{\varphi}(x)$.

At this point the reader may be tempted to interpret $\varphi(x)$ as a wave-function of the ordinary quantum mechanics. There is, however, an important difference between what have now and the standard quantum mechanics. In the usual formulation of the quantum mechanics in terms of wave-functions, the Hamiltonian is specified as a differential operator (with space rather than space-time derivatives) acting on the wave-function $\varphi(x)$. Our representation of the Poincaré algebra did not provide any such Hamiltonian, it just states that the Hamiltonian is the generator of the time translations.

However, if one is really keen to interpret $\varphi(x)$ as the wave-function, one is allowed to do so. Then one may try to specify the Hamiltonian for this irreducible representation by demanding $p^2 = m^2$ for any eigenfunction e^{-ipx}. In this way, one is lead to some specific differential equation for $\varphi(x)$, e.g. to the equation

$$i\partial_t \varphi(x) = \sqrt{m^2 - \partial_i \partial_i}\;\varphi(x)$$

Because of the square root, however, this is not very convenient equation to work with. First of all, it is not straightforward to check, if this operator obeys all the commutation relations of the Poincaré algebra. Second, after the Taylor expansion of the square root one gets infinite number of derivatives, which corresponds to a non-local theory (which is usually quite non-trivial to be put in accord with special relativity). Another ugly feature of the proposed equation is that it treated the time and space derivatives in very different manner, which is at least strange in a

would-be relativistic theory. The awkwardness of the square root becomes even more apparent once the interaction with electromagnetic field is considered, but we are not going to penetrate in such details here.

For all these reasons it is a common habit to abandon the above equation and rather to consider the closely related so-called Klein–Gordon equation

$$\left(\partial_\mu \partial^\mu + m^2\right) \varphi(x) = 0$$

as a kind of a relativistic version of the Schrödinger equation (even if the order of the Schrödinger and Klein–Gordon equations are different).

Note, however, that the Klein–Gordon equation is only related, but not equivalent to the equation with the square root. One of the consequences of this non-equivalence is that the solutions of the Klein-Gordon equation may have both positive and negative energies. This does not pose an immediate problem, since the negative energy solutions can be simply ignored, but it becomes really puzzling, once the electromagnetic interactions are switched on.

Another unpleasant feature is that one cannot interpret $|\varphi(x)|^2$ as a probability density, because this quantity is not conserved. For the Schrödinger equation one was able to derive the continuity equation for the density $|\varphi(x)|^2$ and the corresponding current, but for the Klein–Gordon equation the quantity $|\varphi(x)|^2$ does not obey the continuity equation any more. One can, however, perform with the Klein–Gordon equation a simple massage analogous to the one known from the treatment of the Schrödinger equation, to get another continuity equation with the density $\varphi^* \partial_0 \varphi - \varphi \partial_0 \varphi^*$. But this density has its own drawback — it can be negative. It cannot play, therefore the role of the probability density.

All this was well known to the pioneers of the quantum theory and eventually led to rejection of wave-function interpretation of $\varphi(x)$ in the Klein–Gordon equation. Strangely enough, the field $\varphi(x)$ equation remained one of the cornerstones of the quantum field theory. The reason is that it was not the function and the equation which were rejected, but rather only their wave-function interpretation.

The function $\varphi(x)$ satisfying the Klein–Gordon is very important — it becomes the starting point of what we call the field-focused approach to the quantum field theory. In this approach the function $\varphi(x)$ is treated as a classical field (transforming according to the considered representation of the Poincaré group) and starting from it one develops step by step the corresponding quantum theory. The whole procedure is discussed in quite some detail in the following chapters.

1.3.2 The logic of the particle-focused approach to QFT

The relativistic quantum theory describes, above all, the physics of elementary particles. Therefore the particle-focused approach looks like the most natural. Nevertheless, it is by far not the most common, for reasons which are mainly historical. Now we have to confess (embarrassed) that in these lectures we are going to follow the less natural, but more wide-spread field-focused approach.[48] The particle-focused approach is only very briefly sketched in this paragraph. *Not everything should and could be understood here, it is sufficient just to catch the flavor.* If too dense and difficult (as it is) the paragraph should be skipped.

One starts with an irreducible representation of the Poincaré group on some 1-particle Hilbert space. The usual basis vectors in the Hilbert space are of the form $|p,\sigma\rangle$, where p is the (overall) momentum of the state and all the other characteristics are included in σ. For a multiparticle state, the σ should contain a continuous spectrum of momenta of particular particles. This provides us with a natural definition of 1-particle states as the ones with discrete σ. In this case it turns out that values of σ correspond to spin (helicity) projections.

Irreducible representations are characterized by eigenvalues of two Casimir operators (operators commuting with all generators), one of them being m^2, the eigenvalue of the Casimir operator P^2, and the second one having to do with the spin. The states in the Hilbert space are therefore characterized by eigenvalues of 3-momentum, i.e. the notation $|\vec{p},\sigma\rangle$ is more appropriate than $|p,\sigma\rangle$ (nevertheless, when dealing with Lorentz transformations, the $|p,\sigma\rangle$ notation is very convenient). The $|\vec{p},\sigma\rangle$ states are still eigenstates of the Hamiltonian, with the eigenvalues $E_{\vec{p}}=\sqrt{\vec{p}^2+m^2}$.

Once a representation of the Poincaré group on a 1-particle Hilbert space is known, one can systematically build up the corresponding Fock space from direct products of the Hilbert ones. The motivation for such a construction is that this would be a natural framework for processes with nonconserved numbers of particles, and such processes are witnessed in the nature. This Fock space benefits from having a natural representation of Poincaré group, namely the one defined by the direct products of the representations of the original 1-particle Hilbert space. The Hamiltonian constructed in this way, as well as all the other generators, correspond to a system of noninteracting particles. In terms of creation and annihilation operators, which are defined as very natural operators in the Fock space the free Hamiltonian has a simple form $H_0=\int \frac{d^3p}{(2\pi)^3}E_{\vec{p}}a^+_{\vec{p}}a_{\vec{p}}$.

The next step, and this is the hard one, is to find another Hamiltonian which would describe, in a relativistic way, a system of interacting particles. One does not start with a specific choice, but rather with a perturbation theory for a generic Hamiltonian $H=H_0+H_{\rm int}$. The perturbation theory is neatly formulated in the interaction picture, where $|\psi_I(t)\rangle = U(t,0)|\psi_I(0)\rangle$, with $U(t,0)$ satisfying $i\partial_t U(t,0)=H_{{\rm int},I}(t)\,U(t,0)$ with the initial condition

[48] The explanation for this is a bit funny.

As to what we call here the particle-centered approach, the textbook is the Weinberg's one. We strongly recommend it to the reader, even if it would mean that he/she will quit these notes. The present author feels that he has nothing to add to the Weinberg's presentation.

But even if the approach of the Weinberg's book is perhaps more natural than any other, it is certainly not a good idea to ignore the traditional development, which we call here the field-centered approach. If for nothing else, then simply because it is traditional and therefore it became a part of the standard background of the majority of particle physicists.

Now as to the textbooks following the traditional approach, quite a few are available. But perhaps in all of them there are points (and unfortunately not just one or two) which are not explained clearly enough, and are therefore not easy to grasp. The aim of the present notes is to provide the standard material with perhaps a bit more emphasis put on some points which are often only glossed over. The hope is, that this would enable reader to penetrate into the subject in combination of a reasonable depth with a relative painlessness.

Nevertheless, beyond any doubt, this hope is not to be fulfilled. The reader will surely find a plenty of disappointing parts in the text.

$U(0,0) = 1$. The perturbative solution of this equation leads to the sum of integrals of the form $\int_{t_0}^{t} dt_1 \dots dt_n \; T\, H_{\mathrm{int},I}(t_1) \dots H_{\mathrm{int},I}(t_n)$, where T orders the factors with respect to decreasing time. For a relativistic theory, these integrals should better be Lorentz invariant, otherwise the scalar products of the time-evolved states would be frame dependent. This presents nontrivial restrictions on the interaction Hamiltonian H_{int}. First of all, the space-time variables should be treated on the same footing, which would suggest an interaction Hamiltonian of the form $H_{\mathrm{int}} = \int d^3x \; \mathcal{H}_{\mathrm{int}}$ and $\mathcal{H}_{\mathrm{int}}$ should be a Lorentz scalar. Furthermore, the T-ordering should not change the value of the product when going from frame to frame, which would suggest $[\mathcal{H}_{\mathrm{int}}(x), \mathcal{H}_{\mathrm{int}}(y)] = 0$ for $(x-y)^2 \leq 0$ (for time-like intervals, the ordering of the Hamiltonians in the T-product is the same in all the reference frames, for space-like intervals the ordering is frame-dependent, but becomes irrelevant for Hamiltonians commuting with each other).

All these requirements do not have a flavor of rigorous statements, they are rather simple observations about how could (would) a relativistic quantum theory look like. It comes as a kind of surprise, that the notion of quantum fields is a straightforward outcome of these considerations. Without going into details, let us sketch the logic of the derivation:

1. As any linear operator, the Hamiltonian can be written as a sum of products of the creation and annihilation operators. The language of the $a^+_{\vec{p}}$ and $a_{\vec{p}}$ operators is technicaly advantageous, e.g. in the formulation of the so-called cluster decomposition principle, stating that experiments which are sufficiently separated in space, have unrelated results.

2. Poincaré transformations of $a^+_{\vec{p}}$ and $a_{\vec{p}}$ (inherited from the transformations of states) are given by $\vec{p}$-dependent matrices, and so the products of such operators (with different momenta) have in general complicated transformation properties. One can, however, combine the $a^+_{\vec{p}}$ and $a_{\vec{p}}$ operators into simply transforming quantities called the creation and annihilation fields $\varphi^+_l(x) = \sum_\sigma \int d^3p \, u_l(x,\vec{p},\sigma) a^+_{\vec{p}}$ and $\varphi^-_l(x) = \sum_\sigma \int d^3p \, v_l(x,\vec{p},\sigma) a_{\vec{p}}$, which are much more suitable for a construction of relativistic quantum theories.

3. The required simple transformation properties of $\varphi^\pm_l$ are the ones independent of any x or $\vec{p}$, namely $\varphi^\pm_l(x) \to \sum_{l'} D_{ll'}(\Lambda^{-1}) \varphi^\pm_{l'}(x)(\Lambda x + a)$, where the D matrices furnish a representation of the Lorentz group. The coefficients u_l and v_l are calculable for any such representation, e.g. for the trivial one $D_{ll'}(\Lambda^{-1}) = 1$ one gets $u(x,\vec{p},\sigma) = \frac{1}{(2\pi)^3\sqrt{2E_{\vec{p}}}} e^{ipx}$ and $v(x,\vec{p},\sigma) = \frac{1}{(2\pi)^3\sqrt{2E_{\vec{p}}}} e^{-ipx}$.

4. One can easily construct a scalar $\mathcal{H}_{\mathrm{int}}(x)$ from the creation and annihilation fields. The vanishing commutator of two such $\mathcal{H}_{\mathrm{int}}(x)$ for time-like intervals, however, is not automatically guaranteed. But if $\mathcal{H}_{\mathrm{int}}(x)$ is constructed from a specific linear combination of the creation and annihilation fields, namely form the fields $\varphi_l(x) = \varphi^+_l(x) + \varphi^-_l(x)$, then the commutator is really zero for time-like intervals. This is the way how the quantum fields are introduced in the Weinberg's approach — as (perhaps the only) the natural objects for construction of interaction Hamiltonians leading to relativistic quantum theories.

1.3.3 The logic of the field-focused approach to QFT

The basic idea of the field-focused approach to quantum fields is to take a classical relativistic field theory and to quantize it canonically (the exact meaning of this statement is to be explained in the next chapter). This makes a perfect sense in case of the electromagnetic field, since the primary task of the canonical quantization is to provide a quantum theory with a given classical limit. If the field is classically well known, but one suspects that there is some underlying quantum theory, then the canonical quantization is a handy tool.

This tool, however, is used also for quantum fields for which there is no such thing as the corresponding classical fields, at least not one observed normally in the nature (the electron-positron field is perhaps the prominent example). This may sound even more surprising after one realizes that there is a well known classical counterpart to the quantum electron, namely the classical electron. So if one is really keen on the canonical quantization, it seems very natural to quantize the (relativistic) classical mechanics of the electron particle, rather than a classical field theory of non-existing classical electron field. But still, what is quantized is indeed the classical field. What is the rationale for this?

First, let us indicate why one avoids the quantization of relativistic particles. Actually even for free particles this would be technically more demanding than it is for free fields. But this is not the main reason in favor of field quantization. The point is that we are not interested in free (particles or field) theory, but rather in a theory with interaction. And while it is straightforward to generalize a relativistic classical free field theory to a relativistic classical field theory with interaction (and then to quantize it), it is quite non-trivial to do so for particles.

Second, it should be perhaps emphasized that the nickname "second quantization", which is sometimes used for the canonical quantization of fields, provides absolutely no clue as to any real reasons for the procedure. On the contrary, the nickname could be very misleading. It suggests that what is quantized is not a classical field, but rather a wave-function, which may be regarded to be the result of (the first) quantization. This point of view just obscures the whole problem and is of no relevance at all (except of, perhaps, the historical one).

So why are the fields quantized? The reason is this: In the non-relativistic quantum theories the dynamics is defined by the Hamiltonian. Important point is that any decent Hamiltonian will do the job. In the relativistic quantum theories, on the other hand, the Hamiltonian, as the time-translations generator, comes in the unity of ten generators of the Poincaré group. Not every decent Hamiltonian defines a relativistic dynamics. The reason is that for a given Hamiltonian, one cannot always supply the nine friends to furnish the Poincaré algebra. As a matter of fact, it is in general quite difficult, if not impossible, to find such nine friends. Usually the most natural way is not to start with the Hamiltonian and then to try to find the corresponding nine generators, but to define the theory from the beginning by presenting the whole set of ten generators[49]. This is definitely much easier to say than to really provide. And here comes the field quantization, a clever trick facilitating simultaneous construction of all ten Poincaré generators.

The starting point is a relativistic classical field theory. This means, first of all, that the Poincaré transformations of the field are well defined (as an example we may take the function $\varphi(x)$ discussed on p. 51, which is now treated as a classical field transforming according to the scalar representation[50] of the Poincaré group). Then only theories which are symmetric under

[49] From this it should be clear that even the formulation, not to speak about solution, of the relativistic quantum theory is about 10 times more difficult than that of non-relativistic quantum theory. The situation is similar to the one in general relativity with 10 components of the metric tensor as opposed to one potential describing the Newtonian gravity.

[50] Why representation, why not any (possibly non-linear) realization? We have answered this question (the

these transformations are considered. Now one could expect that, after the canonical quantization, the Poincaré transformations of classical fields become somehow the desired Poincaré transformations of the Hilbert space of states. The reality, however, is a bit more sophisticated. Here we are going to sketch it only very briefly, details are to be found in the next chapter

At the classical level, to each symmetry there is a conserved charge (Noether's theorem). When formulated in the Hamiltonian formalism, the Poisson brackets of these charges obey the same algebra, as do the Poincaré generators. After canonical quantization, the Noether charges become operators (in the Hilbert space of states), the Poisson brackets become commutators, and the Poisson bracket algebra becomes the Poincaré algebra itself (or, strictly speaking, some representation of the Poincaré algebra). Consequently, the Noether charges become, in the process of the canonical quantization, the generators of the symmetry at the quantum level.

Precisely this is going to be the logic behind the field quantization adopted in these lecture notes: field quantization is a procedure leading in a systematic way to a quantum theory with a consistent package of the ten Poincaré generators.

Let us emphasize once more that another important aspect of canonical quantization, namely that it leads to a quantum theory with a given classical limit, is not utilized here. We ignore this aspect on purpose. In spite of the immense role it has played historically and in spite of the undisputed importance of this aspect in the case of the electromagnetic field, for other fields it is illusory and may lead to undue misconceptions.

To summarize: Enlightened by the third introduction (Relativity and Quantum Theory) we are now going to penetrate a bit into the technique of the canonical quantization of relativistic classical fields. The result will be a relativistic quantum theory in terms of creation and annihilation operators familiar from the second introduction (Many-Body Quantum Mechanics). Clever version of the perturbation theory formulated within the obtained theories will then lead us to the Feynman rules discussed in the first introduction (Conclusions).

Remark: *For the sake of completeness let us mention yet another approach to the quantum field theory — the one which can be naturally called the path integral-focused approach. We will have much to say about it in the chapter* **??**

keyword was the superposition principle) supposing realization in the Hilbert space of quantum states. Now $\varphi(x)$ does not correspond to the quantum state, so it is legitimate to raise the question again.

The answer (pretty unclear at the moment) is that non-linear transformations of classical fields would lead, after quantization, to transformations not conserving the number of particles, which is usually "unphysical" in a sense that one could discriminate between two inertial systems by counting particles.

Chapter 2

Free Scalar Quantum Field

In this chapter the simplest QFT, namely the theory of the free scalar field, is developed along the lines described at the end of the previous chapter. The keyword is the canonical quantization (of the corresponding classical field theory).

2.1 Elements of Classical Field Theory

2.1.1 Lagrangian Field Theory

mass points	nonrelativistic fields	relativistic fields
$q_a(t) \quad a = 1, 2, \ldots$	$\varphi(\vec{x}, t) \qquad \vec{x} \in R^3$	$\varphi(x) \qquad x \in R(3,1)$
$S = \int dt\, L(q, \dot{q})$	$S = \int d^3x\, dt\, \mathcal{L}(\varphi, \nabla\varphi, \dot{\varphi})$	$S = \int d^4x\, \mathcal{L}(\varphi, \partial_\mu\varphi)$
$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_a} - \frac{\partial L}{\partial q_a} = 0$	$\longrightarrow$	$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu\varphi)} - \frac{\partial \mathcal{L}}{\partial \varphi} = 0$

The third column is just a straightforward generalization of the first one, with the time variable replaced by the corresponding space-time analog. The purpose of the second column is to make such a formal generalization easier to digest[1]. For more than one field $\varphi_a(x)$ $a = 1, \ldots, n$ one has$\mathcal{L}(\varphi_1, \partial_\mu\varphi_1, \ldots, \varphi_n, \partial_\mu\varphi_n)$ and there are n Lagrange-Euler equations

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu\varphi_a)} - \frac{\partial \mathcal{L}}{\partial \varphi_a} = 0$$

[1]The dynamical variable is renamed ($q \to \varphi$) and the discrete index is replaced by the continuous one ($q_a \to \varphi_x = \varphi(x)$). The kinetic energy T, in the Lagrangian $L = T - U$, is written as the integral ($\sum_a T(\dot{q}_a) \to \int d^3x\ T(\dot{\varphi}(x))$). The potential energy is in general the double integral ($\sum_{a,b} U(q_a, q_b) \to \int d^3x\, d^3y\, U(\varphi(x), \varphi(y))$), but for the continuous limit of the nearest neighbor interactions (e.g. for an elastic continuum) the potential energy is the function of $\varphi(x)$ and its gradient, with the double integral reduced to the single one ($\int d^3x\, d^3y\, U(\varphi(x), \varphi(y)) \to \int d^3x\, u(\varphi(x), \nabla\varphi(x))$)

The fundamental quantity in the Lagrangian theory is the variation of the Lagrangian density with a variation of fields

$$\begin{aligned}\delta\mathcal{L} &= \mathcal{L}\left(\varphi+\varepsilon\delta\varphi, \partial_\mu\varphi+\varepsilon\partial_\mu\delta\varphi\right) - \mathcal{L}\left(\varphi,\partial_\mu\varphi\right)\\ &= \varepsilon\left[\frac{\partial\mathcal{L}\left(\varphi,\partial_\mu\varphi\right)}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}\left(\varphi,\partial_\mu\varphi\right)}{\partial(\partial_\mu\varphi)}\partial_\mu\delta\varphi\right] + \mathcal{O}(\varepsilon^2)\\ &= \varepsilon\left[\delta\varphi\left(\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right) + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi\right)\right] + \mathcal{O}(\varepsilon^2)\end{aligned}$$

It enters the variation of the action

$$\begin{aligned}\delta S &= \int \delta\mathcal{L}\, d^4x\\ &= \varepsilon\int\left[\delta\varphi\left(\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right) + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi\right)\right] d^4x + \mathcal{O}(\varepsilon^2)\end{aligned}$$

which in turn defines the equations of motion, i.e. the Lagrange–Euler equations for extremal action S (for $\delta\varphi$ vanishing at space infinity always and for initial and final time everywhere)

$$\delta S = 0 \qquad \Rightarrow \qquad \int\delta\varphi\left(\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right) d^4x + \int\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi\, d^3\Sigma = 0$$

The second term vanishes for $\delta\varphi$ under consideration, the first one vanishes for any allowed $\delta\varphi$ iff $\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} = 0$.

Example: *Free real Klein-Gordon field* $\mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2$

$$\partial_\mu\partial^\mu\varphi + m^2\varphi = 0$$

This Lagrange-Euler equation is the so-called Klein-Gordon equation. It entered physics as a relativistic generalization of the Schrödinger equation ($\vec{p} = -i\nabla$, $E = i\partial_t$, $\left(p^2 - m^2\right)\varphi = 0$, recall $\hbar = c = 1$). Here, however, it is the equation of motion for some classical field φ.

Example: *Interacting Klein-Gordon field* $\mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{1}{4!}g\varphi^4$

$$\partial_\mu\partial^\mu\varphi + m^2\varphi + \frac{1}{3!}g\varphi^3 = 0$$

Nontrivial interactions lead, as a rule, to nonlinear equations of motion.

Example: *Free complex Klein-Gordon field* $\mathcal{L}[\varphi] = \partial_\mu\varphi^*\partial^\mu\varphi - m^2\varphi^*\varphi$
The fields φ^* and φ are treated as independent.[2]

$$\left(\partial_\mu\partial^\mu + m^2\right)\varphi = 0$$
$$\left(\partial_\mu\partial^\mu + m^2\right)\varphi^* = 0$$

The first (second) equation is the Lagrange-Euler equation for $\varphi^*(\varphi)$ respectively.

[2]It seems more natural to write $\varphi = \varphi_1 + i\varphi_2$, where φ_i are real fields and treat these two real fields as independent variables. However, one can equally well take their linear combinations $\varphi_i' = c_{ij}\varphi_j$ as independent variables, and if complex c_{ij} are allowed, then φ^* and φ can be viewed as a specific choice of such linear combinations.

Noether's Theorem

Symmetries imply conservation laws. A symmetry is an (infinitesimal, local) field transformation

$$\varphi(x) \to \varphi(x) + \varepsilon\delta\varphi(x)$$

leaving unchanged (maybe up to a 4-divergence) either the Lagrangian density $\mathcal{L}$, or (maybe up to a total time derivative) the Lagrangian $L = \int \mathcal{L}\, d^3x$. Conservation laws are either local $\partial_\mu j^\mu = 0$ or global $\partial_t Q = 0$, and they hold only for fields satisfying the Lagrange-Euler equations.

symmetry	conservation law	current or charge
$\delta\mathcal{L} = 0$	$\partial_\mu j^\mu = 0$	$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi$
$\delta\mathcal{L} = \varepsilon\partial_\mu\mathcal{J}^\mu(x)$	$\partial_\mu j^\mu = 0$	$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi - \mathcal{J}^\mu$
$\delta L = 0$	$\partial_t Q = 0$	$Q = \int d^3x\, \frac{\partial\mathcal{L}}{\partial(\partial_0\varphi)}\delta\varphi$
$\delta L = \varepsilon\partial_t\mathcal{Q}(t)$	$\partial_t Q = 0$	$Q = \int d^3x\, \frac{\partial\mathcal{L}}{\partial(\partial_0\varphi)}\delta\varphi - \mathcal{Q}$

The first two lines (the stronger version of the Noether's theorem) follow directly from $\delta\mathcal{L}$ given above (supposing the untransformed field φ obeys the Lagrange-Euler equations). The next two lines (the weaker version of the Noether's theorem) follow from the $\delta\mathcal{L}$ integrated through the space $\int \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi\right)\, d^3x = 0$, which can be written as

$$\partial_0 \int d^3x\, \frac{\partial\mathcal{L}}{\partial(\partial_0\varphi)}\delta\varphi = -\int d^3x\, \partial_i\left(\frac{\partial\mathcal{L}}{\partial(\partial_i\varphi)}\delta\varphi\right) = -\int dS_i\, \frac{\partial\mathcal{L}}{\partial(\partial_i\varphi)}\delta\varphi = 0$$

for the fields obeying the Lagrange-Euler equations and vanishing in the spatial infinity. The conserved quantity has a form of the spatial integral of some "density", but this is not necessary a time-component of a conserved current.

Remark: *For more than one field in the Lagrangian and for the symmetry transformation $\varphi_a(x) \to \varphi_a(x) + \varepsilon\delta_a\varphi(x)$, the conserved current is given by*

$$j^\mu = \sum_a \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta\varphi_a$$

Proof: $\delta\mathcal{L} = \varepsilon\sum_a \left[\delta\varphi_a\left(\frac{\partial\mathcal{L}}{\partial\varphi_a} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\right) + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta\varphi_a\right)\right] + \mathcal{O}(\varepsilon^2)$.

On the other hand, if more symmetries are involved, i.e. if the Lagrangian density is symmetric under different transformations $\varphi(x) \to \varphi(x) + \varepsilon\delta_k\varphi(x)$, then there is one conserved current for every such transformation

$$j_k^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta_k\varphi$$

Proof: $\delta\mathcal{L} = \varepsilon\left[\delta_k\varphi\left(\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right) + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta_k\varphi\right)\right] + \mathcal{O}(\varepsilon^2)$.

Example: *Phase change — field transformations* $\varphi \to \varphi e^{-i\alpha}$, $\varphi^* \to \varphi^* e^{i\alpha}$. *Infinitesimal form* $\varphi \to \varphi - i\varepsilon\varphi$, $\varphi^* \to \varphi^* + i\varepsilon\varphi^*$, *i.e.* $\delta\varphi = -i\varphi$ *and* $\delta\varphi^* = i\varphi^*$. *Lagrangian density* $\mathcal{L}[\varphi] = \partial_\mu\varphi^*\partial^\mu\varphi - m^2\varphi^*\varphi - \frac{1}{4}g\left(\varphi^*\varphi\right)^2$. *Symmetry* $\delta\mathcal{L} = 0$

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^*)}\delta\varphi^* = -i\varphi\partial^\mu\varphi^* + i\varphi^*\partial^\mu\varphi$$
$$Q = \int d^3x\, j^0(x) = i\int d^3x\, \left(\varphi^*\partial^0\varphi - \varphi\partial^0\varphi^*\right)$$

Once the interaction with the electromagnetic field is turned on, this happens to be the electromagnetic current of the Klein-Gordon field.

Example: *Internal symmetries — field transformations* $\varphi_i \to T_{ij}\varphi_j$ $(i,j = 1,\ldots,N)$, *where* $T \in G$ *and* G *is some Lie group of linear transformations. Infinitesimal form* $\varphi_i \to \varphi_i - i\varepsilon_k\left(t_k\right)_{ij}\varphi_j$, *i.e.* $\delta_k\varphi_i = -i\left(t_k\right)_{ij}\varphi_j$. *Lagrangian density* $\mathcal{L}[\varphi_1,\ldots,\varphi_N] = \frac{1}{2}\partial_\mu\varphi_i\partial^\mu\varphi_i - \frac{1}{2}m^2\varphi_i^2 - \frac{1}{4}g\left(\varphi_i\varphi_i\right)^2$. *Symmetry* $\delta\mathcal{L} = 0$

$$j_k^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_i)}\delta_k\varphi_i = -i\left(\partial^\mu\varphi_i\right)\left(t_k\right)_{ij}\varphi_j$$
$$Q_k = \int d^3x\, j_k^0(x) = -i\int d^3x\, \dot{\varphi}_i\left(t_k\right)_{ij}\varphi_j$$

Example: *Space-time translations — field transformations* $\varphi(x) \to \varphi(x+a)$ *(four independent parameters* a_ν *will give four independent conservation laws). Infinitesimal transformations* $\varphi(x) \to \varphi(x) + \varepsilon_\nu\partial^\nu\varphi(x)$, *i.e.* $\delta^\nu\varphi(x) = \partial^\nu\varphi(x)$. *The Lagrangian density* $\mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{1}{4!}g\varphi^4$ *as a specific example, but everything holds for any scalar Lagrangian density. Symmetry* $\delta\mathcal{L} = \varepsilon_\nu\partial^\nu\mathcal{L}$ *(note that a scalar Lagrangian density is transformed as* $\mathcal{L}(x) \to \mathcal{L}(x+a)$, *since this holds for any scalar function of* x*). Technically more suitable form of the symmetry* $\delta\mathcal{L} = \varepsilon_\nu\partial_\mu\mathcal{J}^{\mu\nu}(x) = \varepsilon_\nu\partial_\mu\eta^{\mu\nu}\mathcal{L}(x)$.

$$j^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta^\nu\varphi - \mathcal{J}^{\mu\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\partial^\nu\varphi - \eta^{\mu\nu}\mathcal{L}$$
$$Q^\nu = \int d^3x\, j^{0\nu}(x) = \int d^3x\, \left(\frac{\partial\mathcal{L}}{\partial(\partial_0\varphi)}\partial^\nu\varphi - \eta^{0\nu}\mathcal{L}\right)$$

The conserved quantity is the energy-momentum

$$Q^0 = E = \int d^3x\, \left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\dot{\varphi} - \mathcal{L}\right) \qquad Q^i = P^i = \int d^3x\, \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\,\partial^i\varphi$$

which in our specific example gives

$$Q^0 = E = \frac{1}{2}\int d^3x\, \left(\dot{\varphi}^2 + |\nabla\varphi|^2 + m^2\varphi^2 + \frac{1}{12}g\varphi^4\right)$$
$$\vec{Q} = \vec{P} = \int d^3x\, \dot{\varphi}\,\nabla\varphi$$

Example: *Lorentz transformations — field transformations* $\varphi(x) \to \varphi(\Lambda x)$. *Infinitesimal transformations* $\varphi(x) \to \varphi(x) - \frac{i}{2}\omega^{\rho\sigma}\left(M_{\rho\sigma}\right)^\mu{}_\nu\, x^\nu\partial_\mu\varphi(x)$ *(sorry)*[3], *i.e.* $\delta_{\rho\sigma}\varphi = -i\left(M_{\rho\sigma}\right)^\mu{}_\nu\, x^\nu\partial_\mu\varphi =$

[3]For an explanation of this spooky expression see 1.3.1. Six independent parameters $\omega^{\rho\sigma}$ correspond to 3 rotations (ω^{ij}) and 3 boosts (ω^{0i}). The changes in φ due to these six transformations are denoted as $\delta_{\rho\sigma}\varphi$ with $\rho < \sigma$.

$\left(\delta^\mu_\rho \eta_{\sigma\nu} - \delta^\mu_\sigma \eta_{\rho\nu}\right) x^\nu \partial_\mu \varphi = \left(x_\rho \partial_\sigma - x_\sigma \partial_\rho\right) \varphi$. *The Lagrangian density* $\mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{1}{4!}g\varphi^4$ *again only as a specific example, important is that everything holds for any scalar Lagrangian density. Symmetry* $\delta\mathcal{L} = -\frac{i}{2}\omega^{\rho\sigma}\left(M_{\rho\sigma}\right)^\lambda_{\ \nu} x^\nu \partial_\lambda \mathcal{L} = \omega^{\lambda\nu} x_\nu \partial_\lambda \mathcal{L}$, *is processed further to* $\delta\mathcal{L} = \omega^{\lambda\nu}\partial_\lambda\left(x_\nu\mathcal{L}\right) - \omega^{\lambda\nu}\eta_{\lambda\nu}\mathcal{L} = \omega^{\lambda\nu}\partial_\lambda\left(x_\nu\mathcal{L}\right) = \omega^{\lambda\nu}\partial^\mu\left(g_{\lambda\mu}x_\nu\mathcal{L}\right)$. *So one can write* $\delta\mathcal{L} = \omega_{\lambda\nu}\partial_\mu\mathcal{J}^{\mu\lambda\nu}$ *where*[4] $\mathcal{J}^{\mu\lambda\nu} = g^{\lambda\mu}x^\nu\mathcal{L}$

$$j^{\mu\lambda\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\left(x^\nu\partial^\lambda - x^\lambda\partial^\nu\right)\varphi - \eta^{\lambda\mu}x^\nu\mathcal{L}$$

rotations

$$j^{\mu ij} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\left(x^j\partial^i - x^i\partial^j\right)\varphi - \eta^{i\mu}x^j\mathcal{L}$$
$$Q^{ij} = -\int d^3x\,\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\left(x^i\partial^j - x^j\partial^i\right)\varphi$$

boosts

$$j^{\mu 0i} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\left(x^i\partial^0 - x^0\partial^i\right)\varphi - \eta^{0\mu}x^i\mathcal{L}$$
$$Q^{0i} = -\int d^3x\,\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\left(x^0\partial^i - x^i\partial^0\right)\varphi + x^i\mathcal{L}$$

In a slightly different notation

$$\textit{rotations}\qquad \vec{Q}_R = \int d^3x\,\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\vec{x}\times\nabla\varphi$$
$$\textit{boosts}\qquad \vec{Q}_B = t\int d^3x\,\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\nabla\varphi - \int d^3x\;\vec{x}\;\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\dot{\varphi} - \mathcal{L}\right)$$

and finally in our specific example

$$\textit{rotations}\qquad \vec{Q}_R = -\int d^3x\;\dot{\varphi}\;\vec{x}\times\nabla\varphi$$
$$\textit{boosts}\qquad \vec{Q}_B = -t\int d^3x\;\dot{\varphi}\;\nabla\varphi$$
$$+\frac{1}{2}\int d^3x\;\vec{x}\;(\dot{\varphi}^2 + |\nabla\varphi|^2 + m^2\varphi^2 + \frac{1}{12}g\varphi^4)$$

These bunches of letters are not very exciting. The only purpose of showing them is to demonstrate how one can obtain conserved charges for all 10 generators of the Poincarè group. After quantization, these charges will play the role of the generators of the group representation in the Fock space.

[4]The index μ is the standard Lorentz index from the continuity equation, the pair $\lambda\nu$ specifies the transformation, and by coincidence in this case it has the form of Lorentz indices.

2.1.2 Hamiltonian Field Theory

mass points	relativistic fields
$q_a(t)\ ,\ p_a(t) = \frac{\partial L}{\partial \dot{q}_a}$	$\varphi(x)\ ,\ \pi(x) = \frac{\delta L}{\delta \dot{\varphi}(x)} = \frac{\partial \mathcal{L}(x)}{\partial \dot{\varphi}(x)}$
$H = \sum_a \dot{q}_a p_a - L$	$H = \int d^3x\ \left(\dot{\varphi}(x)\,\pi(x) - \mathcal{L}(x)\right)$
$\frac{d}{dt} f = \{H, f\} + \frac{\partial}{\partial t} f$	$\frac{d}{dt} F = \{H, F\} + \frac{\partial}{\partial t} F$
$\{f, g\} = \sum_a \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q_a} - \frac{\partial g}{\partial p_a}\frac{\partial f}{\partial q_a}$	$\{F, G\} = \int d^3z\ \left(\frac{\delta F}{\delta \pi(z)}\frac{\delta G}{\delta \varphi(z)} - \frac{\delta G}{\delta \pi(z)}\frac{\delta F}{\delta \varphi(z)}\right)$

The only non-trivial issue is the functional derivative $\delta F/\delta\varphi(x)$ which is the generalization of the partial derivative $\partial f/\partial x_n$ (note that in the continuous case φ plays the role of the variable and x plays the role of index, while in the discrete case x is the variable and n is the index). For functions of n variables one has $\delta f[\vec{x}] = f[\vec{x} + \varepsilon\delta\vec{x}] - f[\vec{x}] = \varepsilon\delta\vec{x}.\,\mathrm{grad}\, f + \mathcal{O}(\varepsilon^2) = \sum_n \varepsilon\delta x_n . \partial f/\partial x_n + \mathcal{O}(\varepsilon^2)$. For functionals[5], i.e. for functions with continuous infinite number of variables

$$\delta F[\varphi] = F[\varphi + \varepsilon\delta\varphi] - F[\varphi] = \int dx\ \varepsilon\delta\varphi(x)\ \frac{\delta F[\varphi]}{\delta\varphi(x)} + \mathcal{O}(\varepsilon^2)$$

Clearly $\frac{\delta FG}{\delta\varphi(x)} = \frac{\delta F}{\delta\varphi(x)}G + F\frac{\delta G}{\delta\varphi(x)}$ and $\frac{\delta f(G[\varphi])}{\delta\varphi(x)} = \frac{df[G]}{dG}\frac{\delta G}{\delta\varphi(x)}$, which are the basic properties of anything deserving the name derivative.

For our purposes, the most important functionals are going to be of the form $\int dy\ f(\varphi, \partial_y\varphi)$, where $\partial_y \equiv \frac{\partial}{\partial y}$. In such a case one has[6]

$$\frac{\delta}{\delta\varphi(x)}\int dy\ f(\varphi(y), \partial_y\varphi(y)) = \frac{\partial f(\varphi(x), \partial_y\varphi(x))}{\partial\varphi(x)} - \partial_y \frac{\partial f(\varphi(x), \partial_y\varphi(x))}{\partial(\partial_y\varphi(x))}$$

For 3-dimensional integrals in field Lagrangians this reads

$$\frac{\delta}{\delta\varphi(x)}\int d^3y\ f(\varphi, \dot{\varphi}, \nabla\varphi) = \frac{\partial f(\varphi, \dot{\varphi}, \nabla\varphi)}{\partial\varphi} - \partial_i \frac{\partial f(\varphi, \dot{\varphi}, \nabla\varphi)}{\partial(\partial_i\varphi)}$$
$$\frac{\delta}{\delta\dot{\varphi}(x)}\int d^3y\ f(\varphi, \dot{\varphi}, \nabla\varphi) = \frac{\partial f(\varphi, \dot{\varphi}, \nabla\varphi)}{\partial\dot{\varphi}}$$

where RHS are evaluated at the point x.

As illustrations one can take $\frac{\delta L}{\delta\dot{\varphi}(x)} = \frac{\partial\mathcal{L}(x)}{\partial\dot{\varphi}(x)}$ used in the table above, and $\frac{\delta}{\delta\varphi}\int d^3x\,|\nabla\varphi|^2 = -2\nabla\nabla\varphi = -2\triangle\varphi$ used in the example below.

---

[5] Functional is a mapping from the set of functions to the set of numbers (real or complex).

[6] $\delta\int dy\ f(\varphi, \partial_y\varphi) = \int dy\ f(\varphi + \varepsilon\delta\varphi, \partial_y\varphi + \varepsilon\partial_y\delta\varphi) - f(\varphi, \partial_y\varphi)$
$= \varepsilon\int dy \frac{\partial f(\varphi, \partial_y\varphi)}{\partial\varphi}\delta\varphi + \frac{\partial f(\varphi, \partial_y\varphi)}{\partial(\partial_y\varphi)}\partial_y\delta\varphi + \mathcal{O}(\varepsilon^2)$
$= \varepsilon\int dy \left(\frac{\partial f(\varphi, \partial_y\varphi)}{\partial\varphi} - \partial_y\frac{\partial f(\varphi, \partial_y\varphi)}{\partial(\partial_y\varphi)}\right)\delta\varphi + \begin{matrix}\text{vanishing}\\ \text{surface term}\end{matrix} + \mathcal{O}(\varepsilon^2)$

Example: *Klein-Gordon field* $\mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 \quad \left(-\frac{1}{4!}g\varphi^4\right)$

$$\pi(x) = \frac{\partial\mathcal{L}(x)}{\partial\dot{\varphi}(x)} = \dot{\varphi}(x) \qquad H = \int d^3x\,\mathcal{H}(x)$$
$$\mathcal{H}(x) = \dot{\varphi}(x)\,\pi(x) - \mathcal{L}(x) = \frac{1}{2}\pi^2 + \frac{1}{2}|\nabla\varphi|^2 + \frac{1}{2}m^2\varphi^2$$
$$\dot{\varphi} = \{H,\varphi\} = \pi \qquad \dot{\pi} = \{H,\pi\} = \triangle\varphi - m^2\varphi$$

Inserting the last relation to the time derivative of the second last one obtains the Klein-Gordon equation $\ddot{\varphi} - \triangle\varphi + m^2\varphi = 0$

Poisson brackets of Noether charges

The conserved Noether charges have an important feature, which turns out to be crucial for our development of QFT: their Poisson brackets obey the Lie algebra of the symmetry group. That is why after the canonical quantization, which transfers functions to operators in a Hilbert space and Poisson brackets to commutators, the Noether charges become operators obeying the Lie algebra of the symmetry group. As such they are quite natural choice for the generators of the group representation in the Hilbert space.

The proof of the above statement is straightforward for internal symmetries. The infinitesimal internal transformations are $\delta_k\varphi_i = -i\,(t_k)_{ij}\,\varphi_j$, where t_k are the generators of the group, satisfying the Lie algebra $[t_i,t_j] = if_{ijk}t_k$. The Poisson brackets of the Noether charges are

$$\begin{aligned}
\{Q_k,Q_l\} &= \left\{\int d^3x\;\pi_i(x)\delta_k\varphi_i(x), \int d^3y\;\pi_m(y)\delta_l\varphi_m(y)\right\} \\
&= -\,(t_k)_{ij}\,(t_l)_{mn}\left\{\int d^3x\;\pi_i(x)\varphi_j(x), \int d^3y\;\pi_m(y)\varphi_n(y)\right\} \\
&= -\int d^3z\sum_a (t_k)_{ij}\,(t_l)_{mn}\,(\delta_{ia}\varphi_j(z)\pi_m(z)\delta_{na} - \delta_{ma}\varphi_n(z)\pi_i(z)\delta_{ja}) \\
&= -\int d^3z\;\pi(z)\,[t_l,t_k]\,\varphi(z) = \int d^3z\;\pi(z) if_{klm}t_m\varphi(z) \\
&= if_{klm}Q_m
\end{aligned}$$

For the Poincarè symmetry the proof is a bit more involved. First of all, the generators now contain derivatives, but this is not a serious problem. For the space translations and rotations, e.g., the proof is just a continuous index variation of the discrete index proof for the internal symmetries, one just writes $P^i = \int d^3x\;d^3y\;\pi(y)\,t^i(y,x)\varphi(x)$ and $Q^{ij} = \int d^3x\;d^3y\;\pi(y)\,t^{ij}(y,x)\varphi(x)$ where $t^i(x,y) = \delta^3(x-y)\partial^i$ and $t^{ij}(x,y) = \delta^3(x-y)(x^j\partial^i - x^i\partial^j)$.

For the time-translation and boosts the generators are not linear in π and φ, the universal proof for such generators is a bit tricky[7]. But for any particular symmetry one may prove the statement just by calculating the Poisson brackets for any pair of generators. In the case of the Poincarè group this "exciting" exercise is left to the reader[8].

[7] P. Ševera, private communication.

[8] Just kidding, the reader has probably better things to do and he or she is invited to take the statement for granted even for the Poincarè group.

2.2 Canonical Quantization

2.2.1 The procedure

The standard way of obtaining a quantum theory with a given classical limit[9]:

$$\begin{array}{c}
\text{classical mechanics in any formalism} \\
\downarrow \\
\text{classical mechanics in the Hamiltonian formalism} \\
\text{(with Poisson brackets)} \\
\Downarrow \\
\text{replacement of canonical variables by linear operators} \\
\text{replacement of Poisson brackets by commutators } \{f,g\} \to \frac{i}{\hbar}\left[\hat{f},\hat{g}\right] \\
\downarrow \\
\text{explicit construction of a Hilbert space } \mathbb{H} \\
\text{explicit construction of the operators} \\
\downarrow \\
\text{quantum mechanics in the Heisenberg picture}
\end{array}$$

Example: *A particle in a potential* $U(x)$

$$\begin{array}{ccc}
 & L = \frac{m\dot{x}^2}{2} - U(x) & \\
 & \downarrow & \\
 & H = \frac{p^2}{2m} + U(x) & \\
\{p,x\} = 1 & & \frac{dF(x,p)}{dt} = \{H,F\} \\
 & \Downarrow & \\
 & \hat{H} = \frac{\hat{p}^2}{2m} + U(\hat{x}) & \\
[\hat{p},\hat{x}] = -i\hbar & & \frac{dF(\hat{x},\hat{p})}{dt} = \frac{i}{\hbar}\left[\hat{H},F\right] \\
 & \downarrow & \\
 & \mathbb{H} = L^2 & \\
\hat{x}\psi(x) = x\psi(x) & & \hat{p}\psi(x) = -i\hbar\partial_x\psi(x) \\
 & \downarrow & \\
\frac{d\hat{p}}{dt} = \frac{i}{\hbar}\left[\hat{H},\hat{p}\right] = -\partial_x U(x) & & \frac{d\hat{x}}{dt} = \frac{i}{\hbar}\left[\hat{H},\hat{x}\right] = \frac{1}{m}\partial_x
\end{array}$$

When written in the Schrödinger picture, the Schrödinger equation is obtained.

The above example is not very impressive, since the final result (in the Schrödinger picture) is the usual starting point of any textbook on quantum mechanics. More instructive examples are provided by a particle in a general electromagnetic field or by the electromagnetic field itself. The latter has played a key role in the development of the quantum field theory, and is going to be discussed thoroughly later on (QFT II, summer term). Here we are going to concentrate on the scalar field quantization. Even this simpler case is sufficient to illustrate some new conceptual problems, not present in the quantization of ordinary mechanical systems with a finite number of degrees of freedom.

[9] As already mentioned in the first chapter, for our purposes, the classical limit is not the issue. Nevertheless, the technique (of the canonical quantization) is going to be very useful.

Example: *The scalar field*

$$L = \int d^3x \, \tfrac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \tfrac{1}{2}m^2\varphi^2$$

$$\downarrow$$

$$H = \int d^3x \, \tfrac{1}{2}\pi^2 + \tfrac{1}{2}\left|\nabla\varphi\right|^2 + \tfrac{1}{2}m^2\varphi^2$$

$$\{\pi(\vec{x},t), \varphi(\vec{y},t)\} = \delta^3(\vec{x}-\vec{y}) \qquad \frac{dF[\varphi,\pi]}{dt} = \{H,F\}$$

$$\Downarrow\Downarrow$$

$$H = \int d^3x \, \tfrac{1}{2}\hat{\pi}^2 + \tfrac{1}{2}\left|\nabla\hat{\varphi}\right|^2 + \tfrac{1}{2}m^2\hat{\varphi}^2$$

$$[\hat{\pi}(\vec{x},t), \hat{\varphi}(\vec{y},t)] = -i\hbar\delta^3(\vec{x}-\vec{y}) \qquad \frac{dF[\hat{\varphi},\hat{\pi}]}{dt} = \frac{i}{\hbar}\left[\hat{H},F\right]$$

$$\downarrow$$

$$\mathbb{H} = ???$$

to be continued

The problem with the example (the reason why it is not finished): $\mathbb{H}$ is in general a non-separable Hilbert space. Indeed: for one degree of freedom (DOF) one gets a separable Hilbert space, for finite number of DOF one would expect still a separable Hilbert space (e.g. the direct product of Hilbert spaces for one DOF), but for infinite number of DOF there is no reason for the Hilbert space to be separable. Even for the simplest case of countable infinite many spins 1/2 the cardinality of the set of orthogonal states is $2^{\aleph_0} = c$. For a field, being a system with continuously many DOF (with infinitely many possible values each) the situation is to be expected at least this bad.

The fact that the resulting Hilbert space comes out non-separable is, on the other hand, a serious problem. The point is that the QM works the way it works due to the beneficial fact that many useful concepts from linear algebra survive a trip to the countable infinite number of dimensions (i.e. the functional analysis resembles in a sense the linear algebra, even if it is much more sophisticated). For continuously many dimensions this is simply not true any more.

Fortunately, there is a way out, at least for the free fields. The point is that the free quantum field can be, as we will see shortly, naturally placed into a separable playground — the Fock space[10]. This is by no means the only possibility, there are other non-equivalent alternatives in separable spaces and yet other alternatives in non-separable ones. However, it would be everything but wise to ignore this nice option. So we will, together with the rest of the world, try to stick to this fortunate encounter and milk it as much as possible.

The Fock space enters the game by changing the perspective a bit and viewing the scalar field as a system of coupled harmonic oscillators. This is done in the next section. The other possibilities and their relation to the Fock space are initially ignored, to be discussed afterwards.

Scalar Field as Harmonic Oscillators

The linear harmonic oscillator is just a special case of the already discussed example, namely a particle in the potential $U(x)$. One can therefore quantize the LHO just as in the above general

[10]For interacting fields the Fock space is not so natural and comfortable choice any more, but one usually tries hard to stay within the Fock space, even if it involves quite some benevolence as to the rigor of mathematics in use. These issues are the subject-matter of the following chapter.

example (with the particular choice $U(x) = m\omega^2 x^2/2$), but this is not the only possibility.

Let us recall that search of the solution of the LHO in the QM (i.e. the eigenvalues and eigenvectors of the Hamiltonian) is simplified considerably by introduction of the operators a and a^+. Analogous quantities can be introduced already at the classical level[11] simply as

$$a = x\sqrt{\frac{m\omega}{2}} + p\frac{i}{\sqrt{2m\omega}}$$
$$a^+ = x\sqrt{\frac{m\omega}{2}} - p\frac{i}{\sqrt{2m\omega}}$$

The point now is that the canonical quantization can be performed in terms of the variables a and a^+

$$L = \tfrac{m\dot{x}^2}{2} - \tfrac{m\omega^2 x^2}{2}$$

$$\downarrow$$

$$H = \tfrac{p^2}{2m} + \tfrac{m\omega^2 x^2}{2} = \omega a^+ a \quad \text{or} \quad H = \tfrac{\omega}{2}\left(a^+ a + a a^+\right)$$

$$\{a, a^+\} = i \qquad \dot{a} = -i\omega a \qquad \dot{a}^+ = i\omega a^+$$

$$\Downarrow\Downarrow$$

$$H = \omega a^+ a \quad \text{or} \quad H = \tfrac{\omega}{2}\left(a^+ a + a a^+\right)$$

$$[a, a^+] = 1 \qquad \dot{a} = -i\omega a \qquad \dot{a}^+ = i\omega a^+$$

$$\downarrow$$

$$\mathbb{H} = \text{space spanned by } |0\rangle, |1\rangle, \ldots$$

$$a|n\rangle = |n-1\rangle \qquad\qquad a^+|n\rangle = |n+1\rangle$$

Note that we have returned back to the convention $\hbar = c = 1$ and refrained from writing the hat above operators. We have considered two (out of many possible) Hamiltonians equivalent at the classical level, but non-equivalent at the quantum level (the standard QM choice being $H = \omega\left(a^+ a + 1/2\right)$). The basis $|n\rangle$ is orthogonal, but not orthonormal.

The relevance of the LHO in the context of the QFT is given by a "miracle" furnished by the 3D Fourier expansion

$$\varphi(\vec{x}, t) = \int \frac{d^3 p}{(2\pi)^3} e^{i\vec{p}.\vec{x}} \varphi(\vec{p}, t)$$

which when applied to the Klein-Gordon equation $\partial_\mu \partial^\mu \varphi(\vec{x}, t) + m^2 \varphi(\vec{x}, t) = 0$ leads to

$$\ddot{\varphi}(\vec{p}, t) + (\vec{p}^2 + m^2)\varphi(\vec{p}, t) = 0$$

[11] The complex linear combinations of $x(t)$ and $p(t)$ are not as artificial as they may appear at the first sight. It is quite common to write the solution of the classical equation of motion for LHO in the complex form as $x(t) = \frac{1}{2}(Ae^{-i\omega t} + Be^{i\omega t})$ and $p(t) = -\frac{im\omega}{2}(Ae^{-i\omega t} - Be^{i\omega t})$. Both $x(t)$ and $p(t)$ are in general complex, but if one starts with real quantities, then $B = A^*$, and they remain real forever. The $a(t)$ is just a rescaled $Ae^{-i\omega t}$: $a(t) = \sqrt{m\omega/2}Ae^{-i\omega t}$.

for any $\vec{p}$. Conclusion: the free classical scalar field is equivalent to the (infinite) system of decoupled LHOs[12], where $\varphi(\vec{p},t)$ plays the role of the coordinate (not necessarily real, even if $\varphi(\vec{x},t)$ is real), $\pi=\dot{\varphi}$ the role of the momentum and $\vec{p}$ the role of the index. Note that m has nothing to do with the mass of the oscillators which all have unit mass and

$$\omega_{\vec{p}}^2=\vec{p}^2+m^2$$

Quantization of each mode proceeds in the standard way described above. At the classical level we define

$$a_{\vec{p}}(t)=\varphi(\vec{p},t)\sqrt{\tfrac{\omega_{\vec{p}}}{2}}+\pi(\vec{p},t)\tfrac{i}{\sqrt{2\omega_{\vec{p}}}}$$
$$A_{\vec{p}}^{+}(t)=\varphi(\vec{p},t)\sqrt{\tfrac{\omega_{\vec{p}}}{2}}-\pi(\vec{p},t)\tfrac{i}{\sqrt{2\omega_{\vec{p}}}}$$

We have used the symbol $A_{\vec{p}}^{+}$ instead of the usual $a_{\vec{p}}^{+}$, since we want to reserve the symbol $a_{\vec{p}}^{+}$ for the complex conjugate to $a_{\vec{p}}$. It is essential to realize that for the "complex oscillators" $\varphi(\vec{p},t)$ there is no reason for $A_{\vec{p}}^{+}(t)$ to be equal to the complex conjugate $a_{\vec{p}}^{+}(t)=\varphi^*(\vec{p},t)\sqrt{\omega_{\vec{p}}/2}-i\pi^*(\vec{p},t)/\sqrt{2\omega_{\vec{p}}}$.

For the real classical field, however, the condition $\varphi(\vec{x},t)=\varphi^*(\vec{x},t)$ implies $\varphi(-\vec{p},t)=\varphi^*(\vec{p},t)$ (check it) and the same holds also for the conjugate momentum $\pi(\vec{x},t)$. As a consequence $a_{\vec{p}}^{+}(t)=A_{-\vec{p}}^{+}(t)$ and therefore one obtains

$$\varphi(\vec{x},t)=\int\frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(a_{\vec{p}}(t)\,e^{i\vec{p}.\vec{x}}+a_{\vec{p}}^{+}(t)\,e^{-i\vec{p}.\vec{x}}\right)$$
$$\pi(\vec{x},t)=\int\frac{d^3p}{(2\pi)^3}(-i)\sqrt{\frac{\omega_{\vec{p}}}{2}}\left(a_{\vec{p}}(t)\,e^{i\vec{p}.\vec{x}}-a_{\vec{p}}^{+}(t)\,e^{-i\vec{p}.\vec{x}}\right)$$

Now comes the quantization, leading to commutation relations

$$\left[a_{\vec{p}}(t),a_{\vec{p}'}^{+}(t)\right]=(2\pi)^3\delta(\vec{p}-\vec{p}')\qquad[a_{\vec{p}}(t),a_{\vec{p}'}(t)]=\left[a_{\vec{p}}^{+}(t),a_{\vec{p}'}^{+}(t)\right]=0$$

The reader may want to check that these relations are consistent with another set of commutation relations, namely with $[\varphi(\vec{x},t),\pi(\vec{y},t)]=i\delta^3(\vec{x}-\vec{y})$ and $[\varphi(\vec{x},t),\varphi(\vec{y},t)]=[\pi(\vec{x},t),\pi(\vec{y},t)]=0$ (hint: $\int\frac{d^3x}{(2\pi)^3}e^{-i\vec{k}.\vec{x}}=\delta^3(\vec{k})$).

The "miracle" is not over yet. The free field have turned out to be equivalent to the system of independent oscillators, and this system will now turn out to be equivalent to still another system, namely to the system of free non-teracting relativistic particles. Indeed, the free Hamiltonian

[12]What is behind the miracle: The free field is equivalent to the continuous limit of a system of linearly coupled oscillators. Any such system can be "diagonalized", i.e. rewritten as an equivalent system of decoupled oscillators. For a system with translational invariance, the diagonalization is provided by the Fourier transformation. The keyword is diagonalization, rather than Fourier.

written in terms of $a_{\vec{p}}(t)$ and $a^+_{\vec{p}}(t)$ becomes[13]

$$H = \int d^3x \left(\frac{1}{2}\pi^2 + \frac{1}{2}|\nabla\varphi|^2 + \frac{1}{2}m^2\varphi^2\right)$$
$$= \int \frac{d^3p}{(2\pi)^3}\,\omega_{\vec{p}}\left(a^+_{\vec{p}}(t)\,a_{\vec{p}}(t) + \frac{1}{2}\left[a_{\vec{p}}(t), a^+_{\vec{p}}(t)\right]\right)$$

where the last term is an infinite constant (since $[a_{\vec{p}}(t), a^+_{\vec{p}}(t)] = (2\pi)^3\delta^3(\vec{p}-\vec{p})$). This is our first example of the famous (infinite) QFT skeletons in the cupboard. This one is relatively easy to get rid of (to hide it away) simply by subtracting the appropriate constant from the overall energy, which sounds as a legal step.

Another way leading to the same result is to realize that the canonical quantization does not fix the ordering in products of operators. One can obtain different orderings at the quantum level (where the ordering does matter) starting from the different orderings at clasical level (where it does not). One may therefore choose any of the equivalent orderings at the classical level to get the desired ordering at the quantum level. Then one can postulate that the correct ordering is the one leading to the decent Hamiltonian. Anyway, the standard form of the free scalar field Hamiltonian in terms of creation and annihilation operators is

$$H = \int \frac{d^3p}{(2\pi)^3}\,\omega_{\vec{p}}\,a^+_{\vec{p}}(t)\,a_{\vec{p}}(t)$$

This looks pretty familiar. Was it not for the explicit time dependence of the creation and annihilation operators, this would be the Hamiltonian of the ideal gas of free relativistic particles (relativistic because of the relativistic energy $\omega_{\vec{p}} = \sqrt{\vec{p}^2 + m^2}$). The explicit time dependence of the operators, however, is not an issue — the hamiltonian is in fact time-independent, as we shall see shortly (the point is that the time dependence of the creation and annihilation operators turns out to be $a^+_{\vec{p}}(t) = a^+_{\vec{p}}e^{i\omega_{\vec{p}}t}$ and $a_{\vec{p}}(t) = a_{\vec{p}}e^{-i\omega_{\vec{p}}t}$ respectively).

Still, it is not the proper Hamiltonian yet, since it has nothing to act on. But once we hand over an appropriate Hilbert space, it will indeed become the old friend.

For the relativistic quantum theory (and this is what we are after) the Hamiltonian is not the whole story, one rather needs all 10 generators of the Poincarè group. For the space-translations $\vec{P} = \int d^3x\ \pi\ \nabla\varphi$ one obtains[14]

$$\vec{P} = \int \frac{d^3p}{(2\pi)^3}\,\vec{p}\,a^+_{\vec{p}}(t)\,a_{\vec{p}}(t)$$

[13] The result is based on the following algebraic manipulations

$$\int d^3x\ \pi^2 = -\int \frac{d^3x d^3p d^3p'}{(2\pi)^6}\frac{\sqrt{\omega_{\vec{p}}\omega_{\vec{p}'}}}{2}\left(a_{\vec{p}}(t) - a^+_{-\vec{p}}(t)\right)\left(a_{\vec{p}'}(t) - a^+_{-\vec{p}'}(t)\right)e^{i(\vec{p}+\vec{p}').\vec{x}}$$
$$= -\int \frac{d^3p d^3p'}{(2\pi)^3}\frac{\sqrt{\omega_{\vec{p}}\omega_{\vec{p}'}}}{2}\left(a_{\vec{p}}(t) - a^+_{-\vec{p}}(t)\right)\left(a_{\vec{p}'}(t) - a^+_{-\vec{p}'}(t)\right)\delta^3(\vec{p}+\vec{p}')$$
$$= -\int \frac{d^3p}{(2\pi)^3}\frac{\omega_{\vec{p}}}{2}\left(a_{\vec{p}}(t) - a^+_{-\vec{p}}(t)\right)\left(a_{-\vec{p}}(t) - a^+_{\vec{p}}(t)\right)$$

$$\int d^3x\ |\nabla\varphi|^2 + m^2\varphi^2 = \int \frac{d^3x d^3p d^3p'}{(2\pi)^6}\frac{-\vec{p}.\vec{p}' + m^2}{2\sqrt{\omega_{\vec{p}}\omega_{\vec{p}'}}}\left(a_{\vec{p}}(t) + a^+_{-\vec{p}}(t)\right)\left(a_{\vec{p}'}(t) + a^+_{-\vec{p}'}(t)\right)e^{i(\vec{p}+\vec{p}').\vec{x}}$$
$$= \int \frac{d^3p}{(2\pi)^3}\frac{\omega_{\vec{p}}}{2}\left(a_{\vec{p}}(t) + a^+_{-\vec{p}}(t)\right)\left(a_{-\vec{p}}(t) + a^+_{\vec{p}}(t)\right)$$

where in the last line we have used $(\vec{p}^2 + m^2)/\omega_{\vec{p}} = \omega_{\vec{p}}$.

Putting everything together, one obtains the result.

[14] $\vec{P} = \int \frac{d^3p}{(2\pi)^3}\frac{\vec{p}}{2}(a_{\vec{p}}(t) - a^+_{-\vec{p}}(t))(a_{-\vec{p}}(t) + a^+_{\vec{p}}(t))$
$= \int \frac{d^3p}{(2\pi)^3}\frac{\vec{p}}{2}(a_{\vec{p}}(t)\,a_{-\vec{p}}(t) - a^+_{-\vec{p}}(t)\,a_{-\vec{p}}(t) + a_{\vec{p}}(t)\,a^+_{\vec{p}}(t) - a^+_{-\vec{p}}(t)\,a^+_{\vec{p}}(t))$

while for the rotations and the boosts, i.e. for $Q^{ij} = \int d^3x \ \pi \left(x^j \partial^i - x^i \partial^j\right) \varphi$ and $Q^{0i} = \int d^3x \ \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \left(x^i \partial^0 - x^0 \partial^i\right) \varphi - x^i \mathcal{L}$ the result is[15]

$$Q^{ij} = i \int \frac{d^3p}{(2\pi)^3} \ a^+_{\vec{p}}(t) \left(p^i \partial^j - p^j \partial^i\right) a_{\vec{p}}(t)$$
$$Q^{0i} = i \int \frac{d^3p}{(2\pi)^3} \ \omega_{\vec{p}} a^+_{\vec{p}}(t) \, \partial^i a_{\vec{p}}(t)$$

where $\partial^i = \partial/\partial p_i$ in these formulae.

The reason why these operators are regarded as good candidates for the Poincarè group generators is that at the classical level their Poisson brackets obey the corresponding Lie algebra. After the canonical quantization they are supposed to obey the algebra as well. This, however, needs a check.

The point is that the canonical quantization does not fix the ordering of terms in products. One is free to choose any ordering, each leading to some version of the quantized theory. But in general there is no guarantee that a particular choice of ordering in the 10 generators will preserve the Lie algebra. Therefore one has to check if his or her choice of ordering did not spoil the algebra. The alert reader may like to verify the Lie algebra of the Poincarè group for the generators as given above. The rest of us may just trust the printed text.

Now both the first and the last term vanish since they are integrals of odd functions, so
$\vec{P} = \int \frac{d^3p}{(2\pi)^3} \frac{\vec{p}}{2}(a^+_{\vec{p}}(t) \, a_{\vec{p}}(t) + a_{\vec{p}}(t) \, a^+_{\vec{p}}(t)) = \int \frac{d^3p}{(2\pi)^3} \vec{p}(a^+_{\vec{p}}(t) \, a_{\vec{p}}(t) + \frac{1}{2}[a_{\vec{p}}(t) \, , a^+_{\vec{p}}(t)])$
and the last term again vanishes as a symmetric integral of an odd function.

[15] $Q_{ij} = \int \frac{d^3x d^3p d^3p'}{2(2\pi)^6} \sqrt{\omega_{\vec{p}}/\omega_{\vec{p}'}} (a_{\vec{p}}(t) - a^+_{-\vec{p}}(t))(a_{\vec{p}'}(t) + a^+_{-\vec{p}'}(t)) x^j p'^i e^{i(\vec{p}+\vec{p}').\vec{x}} - i \leftrightarrow j$
We start with $a_{\vec{p}} a^+_{-\vec{p}'} = [a_{\vec{p}}, a^+_{-\vec{p}'}] + a^+_{-\vec{p}'} a_{\vec{p}} = (2\pi)^3 \delta(\vec{p} + \vec{p}') + a^+_{-\vec{p}'} a_{\vec{p}}$, where $a_{\vec{p}}$ stands for $a_{\vec{p}}(t)$ etc. The term with the δ-function vanishes after trivial integration. Then we write $x^j p'^i e^{i(\vec{p}+\vec{p}').\vec{x}}$ as $-ip'^i \partial'^j e^{i(\vec{p}+\vec{p}').\vec{x}}$, after which the d^3x integration leads to $\partial'^j \delta^3(\vec{p} + \vec{p}')$ and then using $\int dk \, f(k) \, \partial_k \delta(k) = -\partial_k f(k)|_{k=0}$ one gets
$Q_{ij} = -i \int \frac{d^3p}{2(2\pi)^3} \partial'^j \sqrt{\omega_{\vec{p}}/\omega_{\vec{p}'}} (a^+_{-\vec{p}'} a_{\vec{p}} - a^+_{-\vec{p}} a_{\vec{p}'} + a_{\vec{p}} a_{\vec{p}'} - a^+_{-\vec{p}} a^+_{-\vec{p}'}) p'^i|_{\vec{p}'=-\vec{p}} + i \leftrightarrow j$
Now using the Leibniz rule and symetric×antisymmetric cancelations one obtains
$Q_{ij} = -i \int \frac{d^3p}{2(2\pi)^3} p^i ((\partial^j a^+_{\vec{p}}) a_{\vec{p}} - a^+_{-\vec{p}} \partial^j a_{-\vec{p}} + a_{\vec{p}} \partial^j a_{-\vec{p}} - a^+_{-\vec{p}} \partial^j a^+_{\vec{p}}) + i \leftrightarrow j$
At this point one uses per partes integration for the first term, the substitution $\vec{p} \to -\vec{p}$ for the second term, and the commutation relations (and $\vec{p} \to -\vec{p}$) for the last two terms, to get
$Q_{ij} = i \int \frac{d^3p}{(2\pi)^3} a^+_{\vec{p}} (p^i \partial^j - p^j \partial^i) a_{\vec{p}} - i \int \frac{d^3p}{4(2\pi)^3} (p^i \partial^j - p^j \partial^i)(2 a^+_{\vec{p}} a_{\vec{p}} + a_{\vec{p}} a_{-\vec{p}} - a^+_{\vec{p}} a^+_{-\vec{p}})$
In the second term the dp^i or dp^j integral is trivial, leading to momenta with some components infinite. This term would be therefore important only if states containing particles with infinite momenta are allowed in the game. Such states, however, are considered unphysical (having e.g. infinite energy in the free field case). Therefore the last term can be (has to be) ignored. (One can even show, that this term is equal to the surface term in the x-space, which was set to zero in the proof of the Noether's theorem, so one should set this term to zero as well.)

Boost generators are left as an exercise for the reader.

Side remark on complex fields

For the introductory exposition of the basic ideas and techniques of QFT, the real scalar field is an appropriate and sufficient tool. At this point, however, it seems natural to say a few words also about complex scalar fields. If nothing else, the similarities and differences between the real and the complex scalar fields are quite illustrative. The content of this paragraph is not needed for the understanding of what follows, it is presented here rather for sake of future references.

The Lagrangian density for the free complex scalar fields reads

$$\mathcal{L}\left[\varphi^*,\varphi\right]=\partial_\mu\varphi^*\partial^\mu\varphi-m^2\varphi^*\varphi$$

where $\varphi=\varphi_1+i\varphi_2$. The complex field φ is a (complex) linear combination of two real fields φ_1 and φ_2. One can treat either φ_1 and φ_2, or φ and φ^* as independent variables, the particular choice is just the mater of taste. Usually the pair φ and φ^* is much more convenient.

It is straightforward to check that the Lagrange-Euler equation for φ and φ^* (as well as for φ_1 and φ_2) is the Klein-Gordon equation. Performing now the 3D Fourier transformation of both $\varphi\left(\vec{x},t\right)$ and $\varphi^*\left(\vec{x},t\right)$, one immediately realizes (just like in the case of the real scalar field) that $\varphi\left(\vec{p},t\right)$ and $\varphi^*\left(\vec{p},t\right)$ play the role of the coordinate of a harmonic oscillator

$$\ddot{\varphi}\left(\vec{p},t\right)+(\vec{p}^2+m^2)\varphi\left(\vec{p},t\right)=0$$
$$\ddot{\varphi}^*\left(\vec{p},t\right)+(\vec{p}^2+m^2)\varphi^*\left(\vec{p},t\right)=0$$

while $\pi\left(\vec{p},t\right)=\dot{\varphi}^*\left(\vec{p},t\right)$ and $\pi^*\left(\vec{p},t\right)=\dot{\varphi}\left(\vec{p},t\right)$ play the role of the corresponding momenta. The variable $\vec{p}$ plays the role of the index and the frequency of the oscillator with the index $\vec{p}$ is $\omega_{\vec{p}}^2=\vec{p}^2+m^2$

Quantization of each mode proceeds again just like in the case of the real field. For the φ field one obtains[16]

$$a_{\vec{p}}\left(t\right)=\varphi\left(\vec{p},t\right)\sqrt{\tfrac{\omega_{\vec{p}}}{2}}+\pi^*\left(\vec{p},t\right)\tfrac{i}{\sqrt{2\omega_{\vec{p}}}}$$
$$A_{\vec{p}}^+\left(t\right)=\varphi\left(\vec{p},t\right)\sqrt{\tfrac{\omega_{\vec{p}}}{2}}-\pi^*\left(\vec{p},t\right)\tfrac{i}{\sqrt{2\omega_{\vec{p}}}}$$

but now, on the contrary to the real field case, there is no relation between $A_{\vec{p}}^+$ and $a_{\vec{p}}^+$. It is a common habit to replace the symbol $A_{\vec{p}}^+$ by the symbol $b_{-\vec{p}}^+=A_{\vec{p}}^+$ and to write the fields as[17]

$$\varphi\left(\vec{x},t\right)=\int\frac{d^3p}{\left(2\pi\right)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(a_{\vec{p}}\left(t\right)e^{i\vec{p}.\vec{x}}+b_{\vec{p}}^+\left(t\right)e^{-i\vec{p}.\vec{x}}\right)$$
$$\varphi^*\left(\vec{x},t\right)=\int\frac{d^3p}{\left(2\pi\right)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(a_{\vec{p}}^+\left(t\right)e^{-i\vec{p}.\vec{x}}+b_{\vec{p}}\left(t\right)e^{i\vec{p}.\vec{x}}\right)$$

Now comes the quantization, leading to the commutation relations

$$\left[a_{\vec{p}}\left(t\right),a_{\vec{p}'}^+\left(t\right)\right]=\left(2\pi\right)^3\delta\left(\vec{p}-\vec{p}'\right)$$
$$\left[b_{\vec{p}}\left(t\right),b_{\vec{p}'}^+\left(t\right)\right]=\left(2\pi\right)^3\delta\left(\vec{p}-\vec{p}'\right)$$

[16] For the φ^* field one has the complex conjugated relations
$a_{\vec{p}}^+\left(t\right)=\varphi^*\left(\vec{p},t\right)\sqrt{\omega_{\vec{p}}/2}+i\pi\left(\vec{p},t\right)/\sqrt{2\omega_{\vec{p}}}$ and $A_{\vec{p}}\left(t\right)=\varphi^*\left(\vec{p},t\right)\sqrt{\omega_{\vec{p}}/2}-i\pi\left(\vec{p},t\right)/\sqrt{2\omega_{\vec{p}}}$

[17] For the momenta one has
$\pi\left(\vec{x},t\right)=\int\frac{d^3p}{\left(2\pi\right)^3}\left(-i\right)\sqrt{\frac{\omega_{\vec{p}}}{2}}\left(a_{\vec{p}}\left(t\right)e^{i\vec{p}.\vec{x}}-b_{\vec{p}}^+\left(t\right)e^{-i\vec{p}.\vec{x}}\right)$
$\pi^*\left(\vec{x},t\right)=\int\frac{d^3p}{\left(2\pi\right)^3}i\sqrt{\frac{\omega_{\vec{p}}}{2}}\left(a_{\vec{p}}^+\left(t\right)e^{-i\vec{p}.\vec{x}}-b_{\vec{p}}\left(t\right)e^{i\vec{p}.\vec{x}}\right)$

while all the other commutators vanish.

The standard form of the Hamiltonian becomes

$$H = \int \frac{d^3p}{(2\pi)^3}\omega_{\vec{p}} \left(a^+_{\vec{p}}(t)\, a_{\vec{p}}(t) + b^+_{\vec{p}}(t)\, b_{\vec{p}}(t)\right)$$

which looks very much like the Hamiltonian of the ideal gas of two types (a and b) of free relativistic particles. The other generators can be obtained just like in the case of the real field.

The main difference with respect to the real field is that now there are two types of particles in the game, with the creation operators $a^+_{\vec{p}}$ and $b^+_{\vec{p}}$ respectively. Both types have the same mass and, as a rule, they correspond to a particle and its antiparticle. This becomes even more natural when the interaction with the electromagnetic field is introduced in the standard way (which we are not going to discuss now). It turns out that the particles created by $a^+_{\vec{p}}$ and $b^+_{\vec{p}}$ have strictly opposite electric charge.

Remark: *At the end of the Introduction No. 2 (many-body quantum mechanics) we have seen that lines in Feynman diagrams originate form reshuffling of creation and annihilation operators in computation of vacuum expectation values of specific operators. Once there are two sets of creation and annihilation operators in the game (like the $a^+_{\vec{p}}$, $a_{\vec{p}}$ and $b^+_{\vec{p}}$, $b_{\vec{p}}$ for the complex scalar field) one should perhaps use two different types of lines (plain, dashed, wiggly, etc.) in diagrams. It is a common habit to use the same type of line for a- and b-operators and to distinguish between them using an arrow. A line with one orientation of the arrow coresponds to the a-operators, while the opposite orientation corresponds to the b-operators.*
So far we have mentioned only one type of lines with arrows in Feynman diagrams, namely the lines corresponding to electrons and positrons. These arrows are present due to the fact that the corresponding field is complex (the presence of two types of creation and annihilation operators $a^+_{\vec{p}}$, $a_{\vec{p}}$ and $b^+_{\vec{p}}$, $b_{\vec{p}}$ is a typical feature of any complex field).
Strangely enough (at least at the first sight), for the scalar complex field the arrows are not used. The reason is that the factor corresponding to the line coming from reshuffling of $a^+_{\vec{p}}$ and $a_{\vec{p}}$ is exactly the same as the factor for the line corresponding to reshuffling of $b^+_{\vec{p}}$ and $b_{\vec{p}}$. So for the scalar field there is no need to distinguish these two cases by arrows or by any other means.

Time dependence of free fields

Even if motivated by the free field case, the operators $a_{\vec{p}}(t)$, $a^+_{\vec{p}}(t)$ can be introduced equally well in the case of interacting fields. The above (so-called equal-time) commutation relations would remain unchanged. Nevertheless, in the case of interacting fields, one is faced with very serious problems which are, fortunately, not present in the free field case.

The crucial difference between the free and interacting field lies in the fact that for the free fields the time dependence of these operators is explicitly known. At the classical level, the independent oscillators enjoy the simple harmonic motion, with the time dependence $e^{\pm i\omega_{\vec{p}}t}$. At the quantum level the same is true, as one can see immediately by solving the equation of motion

$$\dot{a}^+_{\vec{p}}(t) = i\left[H, a^+_{\vec{p}}(t)\right] = i\left[\int \frac{d^3p'}{(2\pi)^3}\omega_{\vec{p}'}\; a^+_{\vec{p}'}(t)\, a_{\vec{p}'}(t), a^+_{\vec{p}}(t)\right] = i\omega_{\vec{p}}\, a^+_{\vec{p}}(t)$$

and $\dot{a}_{\vec{p}}(t) = -i\omega_{\vec{p}}\, a_{\vec{p}}(t)$ along the same lines. From now on, we will therefore write for the free fields

$$a^+_{\vec{p}}(t) = a^+_{\vec{p}} e^{i\omega_{\vec{p}}t}$$
$$a_{\vec{p}}(t) = a_{\vec{p}} e^{-i\omega_{\vec{p}}t}$$

where $a^+_{\vec{p}}$ and $a_{\vec{p}}$ are time-independent creation and annihilation operators (they coincide with $a^+_{\vec{p}}(0)$ and $a_{\vec{p}}(0)$). This enables us to write the free quantum field in a bit nicer way as

$$\varphi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(a_{\vec{p}}e^{-ipx} + a^+_{\vec{p}}e^{ipx}\right)$$

where $p^0 = \omega_{\vec{p}}$. For interacting fields there is no such simple expression and this very fact makes the quantum theory of interacting fields such a complicated affair.

Remark: *The problem with the interacting fields is not only merely that we do not know their time dependence explicitly. The problem is much deeper and concerns the Hilbert space of the QFT. In the next section we are going to build the separable Hilbert space for the free fields, the construction is based on the commutation relations for the operators $a_{\vec{p}}(t)$ and $a^+_{\vec{p}}(t)$. Since these commutation relations hold also for the interacting fields, one may consider this construction as being valid for both cases. This, however, is not true.*
The problem is that once the Hilbert space is specified, one has to check whether the Hamiltonian is a well defined operator in this space, i.e. if it defines a decent time evolution. The explicitly known time-dependence of the free fields answers this question for free fields. For the interacting fields the situation is much worse. Not only we do not have a proof for a decent time evolution, on contrary, in some cases we have a proof that the time evolution takes any initial state away from this Hilbert space. We will come back to these issues in the next chapter. Until then, let us enjoy the friendly (even if not very exciting) world of the free fields.

Hilbert Space

The construction of the Hilbert space for any of the infinitely many LHOs representing the free scalar field is straightforward, as described above. Merging all these Hilbert spaces together is also straightforward, provided there is a *common ground state* $|0\rangle$. Once such a state is postulated[18], the overall Hilbert space is built as an infinite direct sum of Hilbert spaces of individual LHOs.

Such a space is simply the space spanned by the basis

$$\begin{aligned} |0\rangle & \\ |\vec{p}\rangle &= \sqrt{2\omega_{\vec{p}}}\, a^+_{\vec{p}}\ |0\rangle \\ |\vec{p},\vec{p}\,'\rangle &= \sqrt{2\omega_{\vec{p}}}\, a^+_{\vec{p}}\ |\vec{p}\,'\rangle \\ |\vec{p},\vec{p}\,',\vec{p}\,''\rangle &= \sqrt{2\omega_{\vec{p}}}\, a^+_{\vec{p}}\ |\vec{p}\,',\vec{p}\,''\rangle \\ &\vdots \end{aligned}$$

where all creation operators are taken at a fixed time, say $t = 0$, i.e. $a^+_{\vec{p}} \equiv a^+_{\vec{p}}(0)$. The normalization (with notation $E_{\vec{p}} = \omega_{\vec{p}} = \sqrt{\vec{p}^2 + m^2}$)

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

is Lorentz invariant (without $\sqrt{2E_{\vec{p}}}$ in the definition of $|\vec{p}\rangle$ it would not be). Reason: The integral $\int d^3p\ \delta^3(\vec{p}) = 1$ is Lorentz invariant, while d^3p and $\delta^3(\vec{p})$ individually are not. The ratio E/d^3p, on the other hand, is invariant[19] and so is the $d^3p\ \delta^3(\vec{p})\ E/d^3p = E\ \delta^3(\vec{p})$.

The Hilbert space constructed in this way is nothing else but the Fock space introduced in the first chapter. This leads to another shift in perspective: first we have viewed a classical field as a system of classical oscillators, now it turns out that the corresponding system of quantum oscillators can be viewed as a system of particles. Perhaps surprising, and very important.

Let us remark that the Fock space is a separable Hilbert space. Originally our system looked like having continuously infinite dimensional space of states, nevertheless now the number of dimensions seems to be countable. How come? This question is definitely worth discussing, but let us postpone it until the last section of this chapter.

At this point we can continue with the scalar field quantization. It was interrupted at the point $\mathbb{H} =$???, where one now takes $\mathbb{H} =$ the Fock space. Once the Hilbert space is given explicitly, the last step is the explicit construction of the relevant operators. As to the Hamiltonian, we know it in terms of creation and annihilation operators already, and so it happened that it is just the Hamiltonian of a system of free noninteracting relativistic particles.

An important consequence of the explicit form of the Poincarè generators are the transformation properties of the basis vectors $|\vec{p}\rangle$. For this purpose the suitable notation is the 4-vector one: instead of $|\vec{p}\rangle$ one writes $|p\rangle$ where $p^0 = \omega_{\vec{p}}$ (dependent variable). The Lorentz transformation takes a simple form

$$|p\rangle \overset{\Lambda}{\to} |\Lambda p\rangle$$

[18] For infinite number of oscillators (unlike for the finite one) the existence of such a state is not guaranteed. One is free to assume its existence, but this is an independent assumption, not following from the commutation relations. We will search more into this issue in a while.

[19] Indeed, let us consider the boost along the x_3 axis, with the velocity β. The Lorentz transformation of a 4-momentum $p = (E, \vec{p})$ is $E \to \gamma E + \gamma\beta p_3$, $p_1 \to p_1$, $p_2 \to p_2$ and $p_3 \to \gamma p_3 + \gamma\beta E$ (where $\gamma = \sqrt{1-\beta^2}$), and the same transformation holds for an infinitesimal 4-vector dp. Clearly, d^3p is not invariant $d^3p \to dp_1 dp_2\,(\gamma dp_3 + \gamma\beta dE) = d^3p\,(\gamma + \gamma\beta dE/dp_3)$. For $dE = 0$ this would be just a Lorentz contraction, but if both p and $p + dp$ correspond to the same mass m, then $E = \sqrt{m^2 + \vec{p}^2}$ and $dE/dp_3 = p_3/\sqrt{m^2 + \vec{p}^2} = p_3/E$. Therefore $d^3p \to d^3p\,(\gamma + \gamma\beta p_3/E)$ and finally $\frac{d^3p}{E} \to \frac{d^3p(\gamma + \gamma\beta p_3/E)}{\gamma E + \gamma\beta p_3} = \frac{d^3p}{E}$

This may seem almost self-evident, but it is not. As we will see in a moment, for another basis (x-representation) where this transformation rule looks equally self-evident, it simply does not hold. The proof of the transformation rule, i.e. the calculation of how the generators act on the states $|p\rangle$, is therefore mandatory. For rotations at, say, $t = 0$ one has

$$\begin{aligned} -iQ^{ij}\,|p\rangle &= \int \frac{d^3p'}{(2\pi)^3} a^+_{\vec{p}'} \left(p'^i\partial'^j - p'^j\partial'^i\right) a_{\vec{p}'} \sqrt{2\omega_{\vec{p}}}\, a^+_{\vec{p}}\,|0\rangle \\ &= \sqrt{2\omega_{\vec{p}}} \int d^3p'\, a^+_{\vec{p}'} \left(p'^i\partial'^j - p'^j\partial'^i\right) \delta\left(\vec{p}' - \vec{p}\right)|0\rangle \\ &= -\sqrt{2\omega_{\vec{p}}}\left(p^i\partial^j - p^j\partial^i\right) a^+_{\vec{p}}\,|0\rangle = -\left(p^i\partial^j - p^j\partial^i\right)|p\rangle \end{aligned}$$

where in the last step we have used $(p^i\partial^j - p^j\partial^i)\sqrt{2\omega_{\vec{p}}} = 0$. Now the derivative of $|p\rangle$ in a direction k is defined by $|p + \epsilon k\rangle = |p\rangle + \epsilon k_\mu \partial^\mu |p\rangle$. For rotations $k = -iJ^k p$ ($k_i = -i(J^k)_{ij}p_j = -\varepsilon_{ijk}p_j$) $\Rightarrow \left|p - i\epsilon J^k p\right\rangle = |p\rangle - \epsilon.\varepsilon_{ijk}p_j\partial_i\,|p\rangle$, i.e.

$$\left(1 - i\epsilon Q^{ij}\right)|p\rangle = \left|\left(1 - i\epsilon J^k\right)p\right\rangle$$

which is an infinitesimal form of the transformation rule for rotations.

For boosts one gets along the same lines

$$\begin{aligned} -iQ^{0i}\,|p\rangle &= \int \frac{d^3p'}{(2\pi)^3}\omega_{\vec{p}'}\, a^+_{\vec{p}'}\partial'^i a_{\vec{p}'} \sqrt{2\omega_{\vec{p}}}\, a^+_{\vec{p}}\,|0\rangle \\ &= \sqrt{2\omega_{\vec{p}}} \int d^3p'\,\omega_{\vec{p}'}\, a^+_{\vec{p}'}\partial'^i \delta\left(\vec{p}' - \vec{p}\right)|0\rangle \\ &= -\sqrt{2\omega_{\vec{p}}}\partial^i\omega_{\vec{p}} a^+_{\vec{p}}\,|0\rangle = -\frac{p^i}{2\omega_{\vec{p}}}\,|\vec{p}\rangle - \omega_{\vec{p}}\partial^i\,|p\rangle \end{aligned}$$

and since $\left|p - i\varepsilon K^i p\right\rangle = |p\rangle - i\epsilon\left(K^i\right)_{jk} p_k\partial_j\,|p\rangle = |p\rangle - \epsilon p_i\partial_0\,|p\rangle - \epsilon p_0\partial_i\,|p\rangle$, one finally obtains (realizing that $\partial_0\,|p\rangle = \partial_0\sqrt{2p_0}a^+_{\vec{p}}\,|0\rangle = \frac{1}{\sqrt{2p_0}}a^+_{\vec{p}}\,|0\rangle = \frac{1}{2p_0}\,|p\rangle$)

$$\left(1 - i\epsilon Q^{0i}\right)|p\rangle = \left|\left(1 - i\epsilon K^i\right)p\right\rangle$$

which is an infinitesimal form of the transformation rule for boosts.

As to the translations, the transformation rule is even simpler

$$|p\rangle \xrightarrow{a} e^{-ipa}\,|p\rangle$$

as follows directly from the explicit form of the translation generators, which implies $P\,|p\rangle = p\,|p\rangle$ (where $P^0 = H$).

So far everything applied only to $t = 0$. However, once the explicit time dependence of the creation and annihilation operators in the free field case is found in the next section, the proof is trivially generalized for any t.

Quasilocalized states

So far, the quantum field have played a role of merely an auxiliary quantity, appearing in the process of the canonical quantization. The creation and annihilation operators look more "physical", since they create or annihilate physically well defined states (of course, only to the extent to which we consider the states with the sharp momentum being well defined). Nevertheless the

fields will appear again and again, so after a while one becomes so accustomed to them, that one tends to consider them to be quite natural objects. Here we want to stress that besides this psychological reason there is also a good physical reason why the quantum fields really deserve to be considered "physical".

Let us consider the Fourier transform of the creation operator

$$a^{+}(x)=\int\frac{d^{3}p}{(2\pi)^{3}}e^{-i\vec{p}.\vec{x}}a_{\vec{p}}^{+}(t)=\int\frac{d^{3}p}{(2\pi)^{3}}e^{ipx}a_{\vec{p}}^{+}$$

Acting on the vacuum state one obtains $a^{+}(x)|0\rangle=\int\frac{d^{3}p}{(2\pi)^{3}}e^{ipx}\frac{1}{\sqrt{2\omega_{-\vec{p}}}}|\vec{p}\rangle$, which is the superposition (of normalized momentum eigenstates) corresponding to the state of the particle localized at the point x. This would be a nice object to deal with, was there not for the unpleasant fact that it is not covariant. The state localized at the point x is in general not Lorentz transformed[20] to the state localized at the point Λx. Indeed

$$a^{+}(x)|0\rangle\rightarrow\int\frac{d^{3}p}{(2\pi)^{3}}\frac{1}{\sqrt{2\omega_{\vec{p}}}}e^{ipx}|\Lambda p\rangle$$

and this is in general not equal to $a^{+}(\Lambda x)|0\rangle$. The problem is the non-invariance of $d^{3}p/\sqrt{2\omega_{\vec{p}}}$. Were it invariant, the substitution $p\rightarrow\Lambda^{-1}p$ would do the job.

Now let us consider $\varphi(x)|0\rangle=\int\frac{d^{3}p}{(2\pi)^{3}}\frac{1}{2E_{p}}e^{ipx}|p\rangle$ where $E_{p}=\omega_{\vec{p}}=p^{0}$

$$\begin{aligned}\varphi(x)|0\rangle&\rightarrow\int\frac{d^{3}p}{(2\pi)^{3}}\frac{1}{2E_{p}}e^{ipx}|\Lambda p\rangle\overset{p\rightarrow\Lambda^{-1}p}{=}\int\frac{d^{3}\Lambda^{-1}p}{(2\pi)^{3}}\frac{1}{2E_{\Lambda^{-1}p}}e^{i\left(\Lambda^{-1}p\right)x}\left|\Lambda\Lambda^{-1}p\right\rangle\\&=\int\frac{d^{3}p}{(2\pi)^{3}}\frac{1}{2E_{p}}e^{i\left(\Lambda^{-1}p\right)\left(\Lambda^{-1}\Lambda x\right)}|p\rangle=\int\frac{d^{3}p}{(2\pi)^{3}}\frac{1}{2E_{p}}e^{ip(\Lambda x)}|p\rangle=\varphi(\Lambda x)|0\rangle\end{aligned}$$

so this object is covariant in a well defined sense. On the other hand, the state $\varphi(x)|0\rangle$ is well localized, since

$$\langle0|a(x')\varphi(x)|0\rangle=\langle0|\int\frac{d^{3}p}{(2\pi)^{3}}\frac{1}{\sqrt{2\omega_{\vec{p}}}}e^{ip(x-x')}|0\rangle$$

and this integral decreases rapidly for $|\vec{x}-\vec{x}'|$ greater than the Compton wavelength of the particle $\hbar/mc$, i.e.$1/m$. (Exercise: convince yourself about this. Hint: use Mathematica or something similar.)

Conclusion: $\varphi(x)|0\rangle$ is a reasonable relativistic generalization of a state of a localized particle. Together with the rest of the world, we will treat $\varphi(x)|0\rangle$ as a handy compromise between covariance and localizability.

[20] We are trying to avoid the natural notation $a^{+}(x)|0\rangle=|x\rangle$ here, since the symbol $|x\rangle$ is reserved for a different quantity in Peskin-Schroeder. Anyway, we want to use it at least in this footnote, to stress that in this notation $|x\rangle\nrightarrow|\Lambda x\rangle$ in spite of what intuition may suggest. This fact emphasizes a need for proof of the transformation $|p\rangle\nrightarrow|\Lambda p\rangle$ which is intuitively equally "clear".

2.2.2 Contemplations and subtleties

Let us summarize our achievements: we have undergone a relatively exhaustive journey to come to almost obvious results. The (relativistic quantum) theory (of free particles) is formulated in the Fock space, which is something to be expected from the very beginning. The basis vectors of this space transform in the natural way. Hamiltonian of the system of free particles is nothing else but the well known beast, found easily long ago (see Introductions).

Was all this worth the effort, if the outcome is something we could guess with almost no labor at all? Does one get anything new? One new thing is that now we have not only the Hamiltonian, but all 10 Poincarè generators — this is the "leitmotiv" of our development of the QFT. All generators are expressible in terms of $a^+_{\vec{p}}(t)$ and $a_{\vec{p}}(t)$ but, frankly, for the free particles this is also relatively straightforward to guess.

The real yield of the whole procedure remains unclear until one proceeds to the interacting fields or particles. The point is that, even if being motivated by the free field case, the Fourier expansion of fields and quantization in terms of the Fourier coefficients turns out to be an efficient tool also for interacting fields. Even in this case the canonical quantization provides the 10 Poincarè generators in terms of the fields $\varphi(\vec{x},t)$, i.e. in terms of $a^+_{\vec{p}}(t)$ and $a_{\vec{p}}(t)$, which again have (in a sense) a physical meaning of creation and annihilation operators.

Unfortunately, all this does not go smoothly. In spite of our effort to pretend the opposite, the canonical quantization of systems with infinitely many DOF is much more complex than of those with a finite number of DOF. The only reason why we were not faced with this fact hitherto, is that for the free fields the difficulties are not inevitably manifest. More precisely, there is a representation (one among infinitely many) of canonical commutation relations which looks almost like if the system has a finite number of DOF. Not surprisingly, this is the Fock representation — the only one discussed so far. For interacting fields, however, the Fock space is not the trouble-free choice any more. In this case neither the Fock space, nor any other explicitly known representation, succeeds in avoiding serious difficulties brought in by the infinite number of DOF.

In order to understand, at least to some extent, the problems with the quantization of interacting fields, the said difficulties are perhaps worth discussion already for the free fields. So are the reasons why these difficulties are not so serious in the free field case.

Let us start with recollections of some important properties of systems defined by a finite number of canonical commutation relations $[p_i,q_j]=-i\delta_{ij}$ and $[p_i,p_j]=[q_i,q_j]=0$, where $i,j=1,\ldots,n$. One can always introduce operators $a_i=q_ic_i/2+ip_i/c_i$ and $a_i^+=q_ic_i/2-ip_i/c_i$ where c_i is a constant (for harmonic oscillators the most convenient choice is $c_i=\sqrt{2m_i\omega_i}$) satisfying $\left[a_i,a_j^+\right]=\delta_{ij}$ and $[a_i,a_j]=\left[a_i^+,a_j^+\right]=0$. The following holds:

- a state $|0\rangle$ annihilated by all a_i operators does exist $(\exists\,|0\rangle\,\forall i a_i\,|0\rangle=0)$
- the Fock representation of the canonical commutation relations does exist
- all irreducible representations are unitary equivalent to the Fock one
- Hamiltonians and other operators are usually well-defined

For infinite number of DOF, i.e. for the same set of commutation relations, but with $i,j=1,\ldots,\infty$ the situation is dramatically different:

- existence of $|0\rangle$ annihilated by all a_i operators is not guaranteed
- the Fock representation, nevertheless, does exist

- there are infinitely many representations non-equivalent to the Fock one
- Hamiltonians and other operators are usually ill-defined in the Fock space

Let us discuss these four point in some detail.

As to the existence of $|0\rangle$, for one oscillator the proof is notoriously known from QM courses. It is based on well-known properties of the operator $N = a^+a$: 1. $N|n\rangle = n|n\rangle \Rightarrow Na|n\rangle = (n-1)a|n\rangle$ $(a^+aa = [a^+, a]\,a + aa^+a = -a + aN)$
2. $0 \leq \|a|n\rangle\|^2 = \langle n|a^+a|n\rangle = n\langle n|n\rangle$ implying $n \geq 0$, which contradicts 1 unless the set of eigenvalues n contains 0. The corresponding eigenstate is $|0\rangle$.

For a finite number of independent oscillators the existence of the *common* $|0\rangle$ $(\forall i\ a_i|0\rangle = 0)$ is proven along the same lines. One considers the set of commuting operators $N_i = a_i^+ a_i$ and their sum $N = \sum_i a_i^+ a_i$. The proof is basically the same as for one oscillator.

For infinite number of oscillators, however, neither of these two approaches (nor anything else) really works. The step by step argument proves the statement for any finite subset of N_i, but fails to prove it for the whole infinite set. The proof based on the operator N refuses to work once the convergence of the infinite series is discussed with a proper care.

Instead of studying subtleties of the breakdown of the proofs when passing from finite to infinite number of oscillators, we will demonstrate the existence of the so-called strange representations of a_i^+, a_i (representations for which there is no vacuum state $|0\rangle$) by an explicit construction (Haag 1955). Let a_i^+, a_i be the creation and annihilation operators in the Fock space with the vacuum state $|0\rangle$. Introduce their linear combinations $b_i = a_i \cosh\alpha + a_i^+ \sinh\alpha$ and $b_i^+ = a_i \sinh\alpha + a_i^+ \cosh\alpha$. Commutation relations for the b-operators are the same as for the a-operators (check it). Now let us assume the existence of a state vector $|0_\alpha\rangle$ satisfying $\forall i$ $b_i|0_\alpha\rangle = 0$. For such a state one would have

$$0 = \langle i|b_j|0_\alpha\rangle = \langle i,j|0_\alpha\rangle \cosh\alpha + \langle 0|0_\alpha\rangle\,\delta_{ij}\sinh\alpha$$

which implies $\langle i,i|0_\alpha\rangle = \text{const}$ (no i dependence). Now for i being an element of an *infinite* index set this constant must vanish, because otherwise the norm of the $|0_\alpha\rangle$ state comes out infinite $(\langle 0_\alpha|0_\alpha\rangle \geq \sum_{i=1}^\infty |\langle i,i|0_\alpha\rangle|^2 = \sum_{i=1}^\infty \text{const}^2)$. And since $\text{const} = -\langle 0|0_\alpha\rangle \tanh\alpha$ the zero value of this constant implies $\langle 0|0_\alpha\rangle = 0$. Moreover, vanishing $\langle 0|0_\alpha\rangle$ implies $\langle i,j|0_\alpha\rangle = 0$.

It is straightforward to show that also $\langle i|0_\alpha\rangle = 0$ $(0 = \langle 0|b_i|0_\alpha\rangle = \langle i|0_\alpha\rangle\cosh\alpha)$ and finally $\langle i,j,\ldots|0_\alpha\rangle = 0$ by induction

$$\underbrace{\langle i,j,\ldots|}_{n} b_k|0_\alpha\rangle = \underbrace{\langle i,j,\ldots|}_{n+1} 0_\alpha\rangle\cosh\alpha + \underbrace{\langle i,j,\ldots|}_{n-1} 0_\alpha\rangle\sinh\alpha$$

But $\langle i,j,\ldots|$ form a basis of the Fock space, so we can conclude that within the Fock space there is no vacuum state, i.e. a non-zero normalized vector $|0_\alpha\rangle$ satisfying $\forall i$ $b_i|0_\alpha\rangle = 0$.

Representations of the canonical commutation relations without the vacuum vector are called *the strange representations*. The above example[21] shows not only that such representations exist,

[21] Another instructive example (Haag 1955) is provided directly by the free field. Here the standard annihilation operators are given by $a_{\vec p} = \varphi(\vec p, 0)\sqrt{\omega_{\vec p}/2} + i\pi(\vec p, 0)/\sqrt{2\omega_{\vec p}}$, where $\omega_{\vec p} = \vec p^{\,2} + m^2$. But one can define another set $a'_{\vec p}$ in the same way, just with m replaced by some $m' \neq m$. Relations between the two sets are (check it) $a'_{\vec p} = c_+ a_{\vec p} + c_- a^+_{-\vec p}$ and $a'^+_{\vec p} = c_- a^+_{\vec p} + c_+ a_{-\vec p}$, where $2c_\pm = \sqrt{\omega'_{\vec p}/\omega_{\vec p}} \pm \sqrt{\omega_{\vec p}/\omega'_{\vec p}}$. The commutation relations become $\left[a'_{\vec p}, a'^+_{\vec k}\right] = (2\pi)^3\,\delta(\vec p - \vec k)(\omega'_{\vec p} - \omega_{\vec p})/2\omega'_{\vec p}$ and $\left[a'_{\vec p}, a'_{\vec k}\right] = \left[a'^+_{\vec p}, a'^+_{\vec k}\right] = 0$. The rescaled operators $b_{\vec p} = r_{\vec p} a'_{\vec p}$ and $b^+_{\vec p} = r_{\vec p} a^{+\prime}_{\vec p}$ where $r^2_{\vec p} = 2\omega'_{\vec p}/(\omega'_{\vec p} - \omega_{\vec p})$ constitutes a representation of the canonical commutation relations.

If there is a vacuum vector for a-operators, i.e. if $\exists\,|0\rangle\ \forall \vec p\ a_{\vec p}|0\rangle = 0$, then there is no $|0'\rangle$ satisfying $\forall \vec p\ b_{\vec p}|0'\rangle = 0$

but that one can obtain (some of) them from the Fock representation by very simple algebraic manipulations.

As to the Fock representation, it is always available. One just has to postulate the existence of the vacuum $|0\rangle$ and then to build the basis of the Fock space by repeated action of a_i^+ on $|0\rangle$. Let us emphasize that even if we have proven that existence of such a state does not follow from the commutation relations in case of infinite many DOF, we are nevertheless free to postulate its existence and investigate the consequences. The very construction of the Fock space guarantees that the canonical commutation relations are fulfilled.

Now to the (non-)equivalence of representations. Let us consider two representations of canonical commutation relations, i.e. two sets of operators a_i, a_i^+ and $a'_i, a_i'^+$ in Hilbert spaces $\mathbb{H}$ and $\mathbb{H}'$ correspondingly. The representations are said to be equivalent if there is an unitary mapping $\mathbb{H} \stackrel{U}{\to} \mathbb{H}'$ satisfying $a'_i = Ua_iU^{-1}$ and $a_i'^+ = Ua_i^+U^{-1}$.

It is quite clear that the Fock representation cannot be equivalent to a strange one. Indeed, if the representations are equivalent and the non-primed one is the Fock representation, then defining $|0'\rangle = U|0\rangle$ one has $\forall i\ a'_i|0'\rangle = Ua_iU^{-1}U|0\rangle = Ua_i|0\rangle = 0$, i.e. there is a vacuum vector in the primed representation, which cannot be therefore a strange one.

Perhaps less obvious is the fact that as to the canonical commutation relations, any *irreducible representation* (no invariant subspaces) *with the vacuum vector* is equivalent to the Fock representation. The proof is constructive. The considered space $\mathbb{H}'$ contains a subspace $\mathbb{H}_1 \subset \mathbb{H}'$ spanned by the basis $|0'\rangle$, $a_i'^+|0'\rangle$, $a_i'^+a_j'^+|0'\rangle$, ... One defines a linear mapping U from the subspace $\mathbb{H}_1$ on the Fock space $\mathbb{H}$ as follows: $U|0'\rangle = |0\rangle$, $Ua_i^{+\prime}|0'\rangle = a_i^+|0\rangle$, $Ua_i'^+a_j'^+|0'\rangle = a_i^+a_j^+|0\rangle$, ... The mapping U is clearly invertible and preserves the scalar product, which implies unitarity (Wigner's theorem). It is also straightforward that operators are transformed as $Ua_i'^+U^{-1} = a_i^+$ and $Ua'_iU^{-1} = a_i$. The only missing piece is to show that $\mathbb{H}_1 = \mathbb{H}'$ and this follows, not surprisingly, form the irreducibility assumption[22].

An immediate corollary of the above considerations is that all irreducible representations of the canonical commutation relations for finite number of DOF are equivalent (Stone–von Neumann). Indeed, having a finite number of DOF they are obliged to have a vacuum state, and having a vacuum state they are necessarily equivalent. As to the (non-)equivalence of various strange representations, we are not going to discuss the subject here. Let us just remark that a complete classification of strange representations of the canonical commutation relations is not known yet.

Before going further, we should mention an important example of a reducible representation with a vacuum state. Let us consider perhaps the most natural (at least at the first sight) representation of a quantized system with infinite many DOF — the one in which a state is represented by a function $\psi(q_1, q_2, \ldots)$ of infinitely many variables[23]. The function $\psi_0(q_1, q_2, \ldots) = \prod_{i=1}^{\infty} \varphi_0(q_i)$, where φ_0 is a wavefunction of the ground state of LHO, is killed by all annihilation

(the proof is the same as in the example in the main text). In other words at least one of the representations under consideration is a strange one.

Yet another example is provided by an extremely simple prescription $b_{\vec{p}} = a_{\vec{p}} + \alpha(\vec{p})$, where $\alpha(\vec{p})$ is a complex-valued function. For $\int |\alpha(\vec{p})|^2 = \infty$ this representation is a strange one (the proof is left to the reader as an exercise)

[22] As always with this types of proofs, if one is not quite explicit about definition domains of operators, the "proof" is a hint at best. For the real, but still not complicated, proof along the described lines see Berezin, Metod vtornicnovo kvantovania, p.24.

[23] Of course, not any such function can represent a state. Recall that for one variable, only functions from L^2 qualify for states. To proceed systematically, one has to define a scalar product, which can be done for specific functions of the form $\psi(q_1, q_2, \ldots) = \prod_{i=1}^{\infty} \psi_i(q_i)$ in a simple way as $\psi.\psi' = \prod_{i=1}^{\infty} \int dq_i \psi_i^*(q_i)\psi'_i(q_i)$. This definition can be extended to the linear envelope of the "quadratically integrable specific functions" and the

operators, so it represents the vacuum state. Nevertheless, the Hilbert space $\mathbb{H}$ of such functions cannot be unitary mapped on the Fock space $\mathbb{H}_B$, because of different dimensionalities (as already discussed, $\mathbb{H}$ is non-separable, while $\mathbb{H}_B$ is separable). The Fock space can be constructed from this vacuum, of course, and it happens to be a subspace of $\mathbb{H}$ (invariant with respect to creation and annihilation operators). This Fock space, however, does not cover the whole $\mathbb{H}$. What is missing are states with actually infinite number of particles. The point is that only states with finite, although arbitrarily large, number of particles are accessible by repeated action of the creator operators on the vacuum vector[24].

This brings us back to the question of how does it come that for infinitely many oscillators we got a separable, rather then a non-separable, Hilbert space. It should be clear now that this is just a matter of choice — the Fock space is not the only option, we could have chosen a non-separable Hilbert space (or a separable strange representation) as well. The main advantage of the Fock space is that the relevant mathematics is known. On the other hand, the Fock space also seems to be physically acceptable, as far as all physically relevant states do not contain infinitely many particles. It would be therefore everything but wise to ignore the fact that thanks to the Fock space we can proceed further without a development of a new and difficult mathematics. So we will, like everybody does, try to stick to this fortunate encounter and milk it as much as possible.

Anyway, the choice of the Fock space as the playground for QFT does not close the discussion. It may turn out that the Hamiltonian and other Poincarè generators are ill-defined in the Fock space. For the free field, fortunately, the generators turn out to be well defined. But the reader should make no mistake, this is an exception rather than a rule.

The last point from the above lists of characteristic features of systems of finite and infinite DOF concern definitions of operators. This is a subtle point already for a finite number of DOF[25]. For systems with an infinite number of DOF the situation is usually even worse. The reason is that many "natural" operators are of the form O^n where $O = \sum_{i=1}^{\infty} c_i a_i^+ + c_i^* a_i$. The trouble now is that for *infinite* sum one can have $\sum_{i=1}^{\infty} |c_i|^2 = \infty$, the quantum field $\varphi(x)$ is a prominent example. Such an operator leads to a state with an infinite norm acting on any standard basis vector in the Fock space (convince yourself). But this simply means that such operators are not defined within the Fock space.

Nevertheless, the operator O, as well as the quantum field $\varphi(x)$, has a finite matrix elements between any two standard basis vectors. This enables us to treat them as objects having not a well defined meaning as they stand, but only within scalar products — a philosophy similar to

space of normalized functions is to be checked for completeness. But as to the mathematical rigor, this remark represents the utmost edge of our exposition.

[24]This may come as a kind of surprise, since due to the infinite direct sum $\oplus_{n=0}^{\infty}\mathbb{H}^n$ in the definition of the Fock space, one may expect (incorrectly) that it also contains something like $\mathbb{H}^{\infty}$. This symbol, however, is just an abuse of notation — it does not correspond to any many-particle subspace of the Fock space. An analogy may be of some help in clarifying this issue: the set of natural numbers $\mathbb{N}$ does not contain an infinite number ∞, even if it contains every n where $n = 1, \ldots, \infty$.

[25]The point is that (unbounded) operators in quantum theory usually enter the game in the so-called formal way, i.e. without a precise specification of domains. Precise domains, on the other hand, are of vital importance for such attributes as selfadjointness, which in turn is a necessary condition for a Hamiltonian to define a dynamics, i.e. a unitary time evolution (Stone theorem). Formal Hamiltonians are usually Hermitian (symmetric) on a dense domain in the Hilbert space, and for some (but not for all) such symmetric operators the selfadjoint extensions do exist. If so, the Hamiltonian is considered to be well-defined.

For our present purposes the important thing is that the Hamiltonian of the LHO is well-defined in this strict sense. One can even show, using sophisticated techniques of modern mathematical physics, that the Hamiltonian of an anharmonic oscillator $H = p^2/2m + q^2 + q^4$ is well defined (see e.g. Reed-Simon, volume 2, for five proofs of this statement) and this holds for any finite number of oscillators. On the other hand, some formal Hamiltonians are doomed to be ill-defined and lead to no dynamics whatsoever.

that of distributions like the δ-function. Operators which can be defined only in this sense are sometimes called improper operators.

But the main problem is yet to come. Were all operators appearing in QFT proper or improper ones, the QFT would be perhaps much easier and better understood then it actually is. Unfortunately, for many "natural" operators even the scalar products are infinite. Such objects are neither proper, nor improper operators, they are simply senseless expressions.

Nevertheless, the free field Hamiltonian $H = \int d^3p \; a^+_{\vec{p}}\,(t)\, a_{\vec{p}}\,(t)\, \omega_{\vec{p}} / \,(2\pi)^3$ is a proper (even if unbounded) operator in the Fock space, since it maps an n-particle basis state to itself, multiplied by a finite number[26]. The other generators map n-particle basis states to normalized n-particle states, so all these operators are well defined. That is why all difficulties discussed in this section remain hidden in the free field case. But they will reappear quickly, once the interacting fields are considered.

[26]The remaining question is if this unbounded Hamiltonian has a selfadjoint extension. The answer is affirmative, the proof, however, is to be looked for in books on modern mathematical physics rather than in introductory texts on QFT. One may raise an objection that we have demonstrated selfadjointness of the free field Hamiltonian indirectly by finding the explicit unitary time evolution of states (which follows from the time evolution of the creation and annihilation operators). This, however, was found by formal manipulations, without bothering about if the manipulated objects are well defined. Needless to say, such an approach can lead to contradictions. Anyway, for the free fields the formal manipulations are fully supported by more careful analysis.

All this applies, of course, only for one particular ordering of creation and annihilation operators — not surprisingly the one we have adopted. Other orderings are, strictly speaking, the above mentioned senseless expressions with infinite matrix elements between basis vectors.

Chapter 3

Interacting Quantum Fields

3.1 Naive approach

In the last section of the previous chapter we have discussed several unpleasant features which may appear in a theory of interacting fields (strange representations, ill-defined operators, no dynamics in a sense of unitary time evolution). In the first two sections of the present chapter we are going to ignore all this completely. On top of that, in the first section we will oversimplify the matters even more than is the common habit.

The reason for this oversimplification is purely didactic. As we will see, one can get pretty far using a bit simple-minded approach and almost everything developed in this framework will survive, with necessary modifications, later critical reexamination. The said modifications are, on the other hand, quite sophisticated and both technically and conceptually demanding. We prefer, therefore, to postpone their discussion until the basic machinery of dealing with interacting fields is developed in the simplified naive version[1].

As the matter of fact, the naive approach is the most natural one. It is based on the assumption that the free field lagrangian defines what particles are[2], and the interaction lagrangian defines how do these particles interact with each other. Life, however, turns out to be surprisingly more complex.

So it happens that by switching on the interaction, one in fact redefines what particles are. This rather non-trivial and surprising fact has to be taken into account — otherwise one is, sooner or later, faced with serious inconsistencies in the theory. In the standard approach one indeed develops the theory of interacting quantum fields having in mind from the very beginning that "particle content" of the free and interacting theories may differ significantly.

In our naive approach we will ignore all this and move on happily until we will understand almost completely where the Feynman rules come from. The few missing ingredients will be obtained afterwards within the standard approach.

[1] It should be stressed that even after all known modifications (see section 3.2) the resulting theory of interacting quantum fields is still not satisfactory in many respects (see section **??**). The difference between the oversimplified and the standard approach is not that they are incorrect and correct respectively, but rather that they are incorrect to different degrees.

[2] According to the naive approacg, one-particle states are those obtained by acting of the creation operator $a^+_{\vec{p}}$ on the vacuum state $|0\rangle$, two-particle states are those obtained by acting of two creation operators on the vacuum state, etc.

canonical quantization of interacting fields

For interacting fields the quantization proceeds basically along the same lines as for the free fields. A particular theory is defined by a Lagrangian density in the form of a sum of the free and the interaction Lagrangian densities. For relativistic theory this Lagrangian density is a Lorentz (as well as Poincaré) scalar. Invariance of the Lagrangian density with respect to Poincaré transformations provides us with 10 Noether's charges.
Example: *φ^4-theory (we obtained the following results already in 2.1.1)*

$$\begin{aligned}
\textit{t-translations} \qquad Q^0 &= \frac{1}{2}\int d^3x\ (\dot{\varphi}^2 + |\nabla\varphi|^2 + m^2\varphi^2) \\
&\quad + \frac{1}{2}\int d^3x\ \frac{1}{12}g\varphi^4 \\
\vec{x}\textit{-translations} \qquad \vec{Q} &= \int d^3x\ \dot{\varphi}\,\nabla\varphi \\
\textit{rotations} \qquad \vec{Q}_R &= -\int d^3x\ \dot{\varphi}\,\vec{x}\times\nabla\varphi \\
\textit{boosts} \qquad \vec{Q}_B &= -t\int d^3x\ \dot{\varphi}\,\nabla\varphi + \frac{1}{2}\int d^3x\ \vec{x}\,(\dot{\varphi}^2 + |\nabla\varphi|^2 + m^2\varphi^2) \\
&\quad + \frac{1}{2}\int d^3x\ \vec{x}\ \frac{1}{12}g\varphi^4
\end{aligned}$$

After Hamiltonian reformulation of the theory, the Poisson algebra of these charges is isomorphic to the Lie algebra of the corresponding Poincaré generators. And after subsequent quantization, this Poisson algebra leads to the Lie algebra of the quantized charges, which defines a relativistic quantum theory (a representation of the Poincaré group in a Hilbert space).
Example: *φ^4-theory*

$$\begin{gathered}
L = \textstyle\int d^3x\ \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{1}{4!}g\varphi^4 \\
\downarrow \\
H = \textstyle\int d^3x\ \frac{1}{2}\pi^2 + \frac{1}{2}|\nabla\varphi|^2 + \frac{1}{2}m^2\varphi^2 + \frac{1}{4!}g\varphi^4 \\
\{\pi(\vec{x},t),\varphi(\vec{y},t)\} = \delta^3(\vec{x}-\vec{y}) \\
\{\pi(\vec{x},t),\pi(\vec{y},t)\} = \{\varphi(\vec{x},t),\varphi(\vec{y},t)\} = 0 \\
\frac{dF[\varphi,\pi]}{dt} = \{H,F\} \\
\Downarrow\Downarrow \\
\hat{H} = \textstyle\int d^3x\ \frac{1}{2}\hat{\pi}^2 + \frac{1}{2}|\nabla\hat{\varphi}|^2 + \frac{1}{2}m^2\hat{\varphi}^2 + \frac{1}{4!}g\hat{\varphi}^4 \\
[\hat{\pi}(\vec{x},t),\hat{\varphi}(\vec{y},t)] = -i\hbar\delta^3(\vec{x}-\vec{y}) \\
[\hat{\pi}(\vec{x},t),\hat{\pi}(\vec{y},t)] = [\hat{\varphi}(\vec{x},t),\hat{\varphi}(\vec{y},t)] = 0 \\
\frac{dF[\hat{\varphi},\hat{\pi}]}{dt} = \frac{i}{\hbar}\left[\hat{H},F\right] \\
\downarrow \\
\mathbb{H} = ???
\end{gathered}$$

As in the case of free fields, the problem now is that natural choice of $\mathbb{H}$ is a non-separable

Hilbert space. For free fields we managed to avoid the problem by taking a strange detour route, i.e. by rewriting the scalar field in terms of harmonic oscilators. This lead us to a Fock subspace of the non-separable space, which turned out to be a sufficient playground (when starting form this subspace, one was never expelled from it by the free field dynamics). It is quite natural to try the same trick also for interacting fields.

So the next step is the Fourier expansion of the classical fields and the conjugated momenta

$$\varphi(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left(a_{\vec{p}}(t)\, e^{i\vec{p}.\vec{x}} + a^+_{\vec{p}}(t)\, e^{-i\vec{p}.\vec{x}} \right)$$
$$\pi(\vec{x},t) = -i \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{\omega_{\vec{p}}}{2}} \left(a_{\vec{p}}(t)\, e^{i\vec{p}.\vec{x}} - a^+_{\vec{p}}(t)\, e^{-i\vec{p}.\vec{x}} \right)$$

which leads (after quatization) to the commutation relations

$$[a_{\vec{p}}(t), a^+_{\vec{p}'}(t)] = (2\pi)^3 \delta(\vec{p}-\vec{p}') \qquad [a_{\vec{p}}(t), a_{\vec{p}'}(t)] = [a^+_{\vec{p}}(t), a^+_{\vec{p}'}(t)] = 0$$

Let us emphasize that the commutation relations hold for arbitrary time t which, however, must be the same for both operators in the commutator — that is why they are known as "equal-time commutation relations". At any fixed time, these commutation relations can be represented by creation and annihilation operators in the Fock space.

So far, it looks like there is no serious difference between quantization of free fields and the interacting ones. Poincaré generators are, as a rule, more complicated for interacting fields, but otherwise the whole procedure looks pretty similar in both cases. For the free fields, appearence of the Fock space was the last important step which enabled us to complete the canonical quantization program. For the interacting fields, however, one does not have the Fock space, but rather Fock spaces.

The point is that for different times t the $a^+_{\vec{p}}(t)$ and $a_{\vec{p}'}(t)$ operators are, in principle, represented in different Fock subspaces of the "large" non-separable space. For the free fields all these Fock spaces coincide, they are in fact just one Fock space — we were able to demonstrate this due to the explicit knowledge of the time evolution of $a^+_{\vec{p}}$ and $a_{\vec{p}'}$. For interacting fields, however, such a knowledge is not at our disposal anymore. One of the main differences between the free and interacting fields is that the time evolution becomes highly nontrivial for the latter. In the Heisenberg picture, the equations for $a_{\vec{p}}(t)$ and $a^+_{\vec{p}'}(t)$ do not lead to simple harmonic time-dependence, nor do the equations for the basis states in the Schrödinger picture (let us epmhasize that basis vectors are eigenstates of the free, rather than the full Hamiltonian).

One of the consequences of the non-trivial and unknown time dependence of the interacting fields is that, frankly speaking, we do not understand our playground. For interacting fields the Fock spaces defined by $a^+_{\vec{p}}(t)$ and $a_{\vec{p}'}(t)$ at different times cannot be proven to coincide. And even if they did, we do not know the representation of the Poincaré algebra explicitly. The generators are defined in terms of $a^+_{\vec{p}}(t)$ and $a_{\vec{p}'}(t)$ and these operators are explicitly known only at one specific moment (let's say at $t = 0$)

How to proceed further in such circumstances? It is a common habit in quantum field theory to ignore the difficulties related to non-separability as long as possible (which is usualy rather long indeed, for most of us it is simply life-long ignorance). The said difficulties are evaded (to a certain degree) by a clever approximative scheme, namely by the perturbation theory in the so-called interaction picture. In this section, we will develop the scheme and learn how to use it in the simplified version. The scheme is valid also in the standard approach, but its usage is a bit different (as will be discussed thoroughly in the next section).

3.1.1 Interaction picture

Our main aim will be the development of some (approximate) techniques of solving the time evolution of interacting fields in the interaction picture of the time evolution in QFT. Operators and states in the interaction picture are defined as[3]

$$A_I(t) = e^{iH_0 t} e^{-iHt} A_H(t) e^{iHt} e^{-iH_0 t}$$
$$|\psi_I(t)\rangle = e^{iH_0 t} e^{-iHt} |\psi_H\rangle$$

where the operators H and H_0 are understood in the Schrödinger picture.

The time evolution of operators in the interaction picture is quite simple, it is equal to the time evolution of the free fields. Indeed, both time evolutions (the one of the free fields and the one of the interacting fields in the interaction picture) are controlled by the same Hamiltonian H_0.

Let us emphasize the similarities and the differences between the interacting fields in the Heisenberg and the interaction pictures. In both pictures one has identically looking expansions

$$\varphi_H(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left(a_{\vec{p},H}(t) + a^+_{-\vec{p},H}(t) \right) e^{i\vec{p}.\vec{x}}$$
$$\varphi_I(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left(a_{\vec{p},I}(t) + a^+_{-\vec{p},I}(t) \right) e^{i\vec{p}.\vec{x}}$$

However, the explicit time dependence of the creation and annihilation operators in the Heisenberg picture is unknown, while in the interaction picture it is known explicitly as $a^+_{\vec{p},I}(t) = a^+_{\vec{p}} e^{i\omega_{\vec{p}} t}$ and $a_{\vec{p},I}(t) = a_{\vec{p}} e^{-i\omega_{\vec{p}} t}$ (see section**??**). Using the free field results, one can therefore write immediately

$$\varphi_I(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\vec{p}}}} \left(a_{\vec{p}} e^{-ipx} + a^+_{\vec{p}} e^{ipx} \right)$$

where $a^+_{\vec{p}}$ and $a_{\vec{p}}$ are the creation and annihilation operators at $t = 0$ (in any picture, they all coincide at this moment). The explicit knowledge and the space-time structure (scalar products of 4-vectors) of the φ_I-fields are going to play an extremely important role later on.

The time evolution of states in the interaction picture is given by

$$i\partial_t |\psi_I\rangle = H_I(t) |\psi_I\rangle \qquad H_I(t) = e^{iH_0 t} (H - H_0) e^{-iH_0 t}$$

where H and H_0 are understood in the Schrödinger picture. The operator $H_I(t)$ is *the interaction Hamiltonian in the interaction picture.*

Needless to say, solving the evolution equation for states in the interaction picture is the difficult point. Nevertheless, we will be able to give the solution as a perturbative series in terms of $\varphi_I(x)$. To achieve this, however, we will need to express all quantities, starting with $H_I(t)$, in terms of $\varphi_I(x)$.

[3]Relations between the Schrödinger, Heisenberg and interaction pictures:

$$A_H(t) = e^{iHt} A_S e^{-iHt} \qquad |\psi_H\rangle = e^{iHt} |\psi_S(t)\rangle$$
$$A_I(t) = e^{iH_0 t} A_S e^{-iH_0 t} \qquad |\psi_I(t)\rangle = e^{iH_0 t} |\psi_S(t)\rangle$$

The operators H and H_0 are understood in the Schrödinger picture. Their subscripts are omitted for mainly esthetic reasons (to avoid too much make-up in the formulae). Anyway, directly from the definitions one has $H_H = H_S$ and $H_{0,I} = H_{0,S}$, therefore the discussed subscripts would be usually redundant.

The canonical quantization provides the Hamiltonian as a function of fields in the Heisenberg picture. What we will need is $H_I(t)$ expressed in terms of $\varphi_I(x)$. Fortunately, this is straightforward: one just replaces φ_H and π_H operators in the Heisenberg picture by these very operators in the interaction picture, i.e. by φ_I and π_I. Proof: one takes $t=0$ in the Heisenberg picture, and in thus obtained Schrödinger picture one simply inserts $e^{-iH_0t}e^{iH_0t}$ between any fields or conjugate momenta[4].

Example: *φ^4-theory* $\quad \mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{1}{4!}g\varphi^4$

$$H = \int d^3x(\frac{1}{2}\pi_H^2 + \frac{1}{2}\left|\nabla\varphi_H\right|^2 + \frac{1}{2}m^2\varphi_H^2 + \frac{1}{4!}g\varphi_H^4)$$

and taking $t=0$ one gets $H_{\text{int}} = \int d^3x\ \frac{1}{4!}g\varphi_S^4$, leading to

$$H_I = \int d^3x\ \frac{1}{4!}g\varphi_I^4$$

In what follows, the key role is going to be played by the operator $U(t,t')$, which describes the time evolution of states *in the interaction picture*

$$\left|\psi_I(t)\right\rangle = U(t,t')\left|\psi_I(t')\right\rangle$$

Directly from the definition one has

$$U(t,t'') = U(t,t')U(t',t'') \qquad\qquad U^{-1}(t,t') = U(t',t)$$

where the second relation follows from the first one and the obvious identity $U(t,t)=1$. Differentiating with respect to t one obtains $i\partial_t U(t,t')\left|\psi_I(t')\right\rangle = H_I(t)U(t,t')\left|\psi_I(t')\right\rangle$ for every $\left|\psi_I(t')\right\rangle$ and therefore

$$i\partial_t U(t,t') = H_I(t)U(t,t')$$

with the initial condition $U(t,t)=1$.

For $t'=0$ the solution of this equation is readily available[5]

$$U(t,0) = e^{iH_0t}e^{-iHt}$$

(H_0 and H in the Schrödinger picture). This particular solution shows that (in addition to providing the time evolution in the interaction picture) the $U(t,0)$ operator enters the relation between the field operators in the Heisenberg and interaction pictures[6]

$$\varphi_H(x) = U^{-1}\left(x^0,0\right)\varphi_I(x)U\left(x^0,0\right)$$

[4]Remark: the simple replacement $\varphi_H \to \varphi_I$, $\pi_H \to \pi_I$ works even for gradients of fields, one simply has to realize that $e^{-iH_0t}\nabla\varphi_S e^{iH_0t} = \nabla\left(e^{-iH_0t}\varphi_S e^{iH_0t}\right) = \nabla\varphi_I$, which holds because H_0 does not depend on the space coordinates.

[5]Indeed $\partial_t e^{iH_0t}e^{-iHt} = e^{iH_0t}(iH_0 - iH)e^{-iHt} = -ie^{iH_0t}H_{\text{int}}e^{-iH_0t}e^{iH_0t}e^{-iHt} =$
$= -iH_I(t)e^{iH_0t}e^{-iHt}$. Note that the very last equality requires $t'=0$ and therefore one cannot generalize the relation to any t'. In general $U(t,t') \neq e^{iH_0(t-t')}e^{-iH(t-t')}$.

[6]$A_H = e^{iHt}e^{-iH_0t}A_I e^{iH_0t}e^{-iHt} = U^{-1}(t,0)A_I U(t,0)$

3.1.2 Transition amplitudes

The dynamical content of a quantum theory is encoded in transition amplitudes, i.e. the probability amplitudes for the system to evolve from an initial state $|\psi_i\rangle$ at t_i to a given state $|\psi_f\rangle$ at t_f. These probability amplitudes are coefficients of the expansion of the final state (evolved from the given initial state) in the basis defined by vectors $|\psi_f\rangle$.

In the Schrödinger picture the initial and the final states are $|\psi_i\rangle$ and $U_S(t_f,t_i)|\psi_i\rangle$ respectively, where $U_S(t_f,t_i)=\exp\{-iH(t_f-t_i)\}$ is the time evolution operator in the Schrödinger picture. So in this picture one has

$$\text{transition amplitude} = \langle\psi_f| U_S(t_f,t_i)|\psi_i\rangle$$

It should be perhaps stressed that, in spite of what the notation might suggest, $|\psi_f\rangle$ does not define the final state (which is rather defined by $|\psi_i\rangle$ and the time evolution). Actually $|\psi_f\rangle$ just defines what component of the final state we are interested in.

In the Heisenberg picture, the time evolution of states is absent. Nevertheless, the transition amplitude can be easily written in this picture as[7]

$$\text{transition amplitude} = \langle\psi_{f,H}|\psi_{i,H}\rangle$$

where $|\psi_{i,H}\rangle = e^{iHt_i}|\psi_S(t_i)\rangle = e^{iHt_i}|\psi_i\rangle$ and $|\psi_{f,H}\rangle = e^{iHt_f}|\psi_S(t_f)\rangle = e^{iHt_f}|\psi_f\rangle$.

And last, but not least, in the interaction picture, one has[8]

$$\text{transition amplitude} = \langle\psi_{f,I}| U(t_f,t_i)|\psi_{i,I}\rangle$$

where $|\psi_{i,I}\rangle = e^{iH_0t_i}|\psi_S(t_i)\rangle = e^{iH_0t_i}|\psi_i\rangle$ and $|\psi_{f,I}\rangle = e^{iH_0t_f}|\psi_S(t_f)\rangle = e^{iH_0t_f}|\psi_f\rangle$. Note that the index I for the time evolution operator $U(t,t')$ is omitted. Throughout this text $U(t,t')$ always means the time evolution operator of states in the interaction picture.

In what follows we are going to encounter slightly generalized objects, namely transition amplitudes with the time evolution interupted by action of particular operators at particular times. Let us consider two such operators A_1 and A_2 acting at times t_1 and t_2 respectively (generalization to arbitrary number of operators is straightforward). In the Schrödinger picture the corrasponding transition amplitude is

$$\text{transition amplitude} = \langle\psi_{f,S}| U_S(t_f,t_1)A_{1,S}U_S(t_1,t_2)A_{2,S}U_S(t_2,t_i)|\psi_{i,S}\rangle$$

In the Heisenberg and interaction pictures one obtains[9]

$$\text{transition amplitude} = \langle\psi_{f,H}| A_{1,H}(t_1)A_{2,H}(t_2)|\psi_{i,H}\rangle$$

while in the interaction picture[10]

$$\text{transition amplitude} = \langle\psi_{f,I}| U(t_f,t_1)A_{1,I}(t_1)U(t_1,t_2)A_{2,I}(t_1)U(t_2,t_i)|\psi_{i,I}\rangle$$

---

[7] $\langle\psi_f| U_S(t_f,t_i)|\psi_i\rangle = \langle\psi_f| e^{-iHt_f}e^{iHt_i}|\psi_i\rangle = \langle\psi_{f,H}|\psi_{i,H}\rangle$

[8] $\langle\psi_f| U_S(t_f,t_i)|\psi_i\rangle = \langle\psi_{f,I}| e^{iH_0t_f}e^{-iH(t_f-t_i)}e^{-iH_0t_i}|\psi_{i,I}\rangle$ what can be written as $\langle\psi_{f,I}| U(t_f,0)U^{-1}(t_i,0)|\psi_{i,I}\rangle$ followed by $U(t_f,0)U^{-1}(t_i,0) = U(t_f,0)U(0,t_i) = U(t_f,t_i)$

[9] $\langle\psi_f| e^{-iHt_f}e^{iHt_1}A_{1,S}e^{-iHt_1}e^{iHt_2}A_{2,S}e^{-iHt_2}\ldots = \langle\psi_{f,H}| A_{1,H}(t_1)A_{2,H}(t_2)\ldots$

[10] $\langle\psi_f| e^{-iHt_f}e^{iHt_1}A_{1,S}\ldots = \langle\psi_{f,I}| e^{iH_0t_f}e^{-iHt_f}e^{iHt_1}e^{-iH_0t_1}A_{1,I}(t_1)e^{iH_0t_1}\ldots = \langle\psi_{f,I}| U(t_f,0)U^{-1}(t_1,0)A_{1,I}(t_1)e^{iH_0t_1}\ldots = \langle\psi_{f,I}| U(t_f,t_1)A_{1,I}(t_1)\ldots$

in and out states

There is a caveat hidden in the transition amplitude written in the Heisenberg picture. The point is that for a given initial state in the Schrödinger picture $|\psi_S(t_i)\rangle = |\psi_i\rangle$, the corresponding $|\psi_{i,H}\rangle$ is usually not known explicitly (and the same holds also for $|\psi_{f,H}\rangle$). Let us consider, e.g. $|\psi_f\rangle = |p_1, p_2\rangle$ and $|\psi_i\rangle = |p_3, p_4\rangle$. Then one may be tempted to write $|\psi_{i,H}\rangle = |p_3, p_4\rangle$, but this would be very misleading. The vector $|p_3, p_4\rangle$ in the Heisenberg picture describes the state which in the Schrödinger picture fulfils $|\psi_S(0)\rangle = |p_3, p_4\rangle$ rather than $|\psi_S(t_i)\rangle = |p_3, p_4\rangle$. So to avoid a notational mismatch, let us emphasize that

$$\begin{aligned}
\text{transition amplitude} &= \langle p_1, p_2 | \, U_S\left(t_f, t_i\right) | p_3, p_4 \rangle \neq \langle p_1, p_2 |_H \;\; | p_3, p_4 \rangle_H \\
\text{transition amplitude} &= \langle p_1, p_2 | \, U_S\left(t_f, t_1\right) A_{1,S} U_S\left(t_1, t_2\right) A_{2,S} U_S\left(t_2, t_i\right) | p_3, p_4 \rangle \\
&\neq \langle p_1, p_2 |_H \;\; A_{1,H}\left(t_1\right) A_{2,H}\left(t_2\right) | p_3, p_4 \rangle_H
\end{aligned}$$

In order to have a simple way to rewrite a transition amplitude from the Schrödinger picture to the Heisenberg one, so-called *in* and *out* states are introduced. They are both defined in the Schrödinger picture as states at $t = 0$, which at the time t_i or t_f are equal to $|\psi_i\rangle$ and $|\psi_f\rangle$ respectively

$$\begin{aligned}
|\psi_{i,\text{in}}\rangle &= |\psi_S(0)\rangle \qquad \text{where} \qquad |\psi_S(t_i)\rangle = |\psi_i\rangle \\
|\psi_{f,\text{out}}\rangle &= |\psi_S(0)\rangle \qquad \text{where} \qquad |\psi_S(t_f)\rangle = |\psi_f\rangle
\end{aligned}$$

It is obvious (using the Schrödinger picture) that

$$\begin{aligned}
|\psi_{i,\text{in}}\rangle &= e^{-iH(0-t_i)} \, |\psi_i\rangle = e^{iHt_i} \, |\psi_i\rangle \\
|\psi_{f,\text{out}}\rangle &= e^{-iH(0-t_f)} \, |\psi_f\rangle = e^{iHt_f} \, |\psi_f\rangle
\end{aligned}$$

and since these states are defined as states in the Schrödinger picture at $t = 0$, they are equal to the corresponding states in both Heisenberg and interaction pictures[11] $|\psi_{\text{in,out}}\rangle_S = |\psi_{\text{in,out}}\rangle_H = |\psi_{\text{in,out}}\rangle_I$. As to the above example, the correct formulae in the Hesisenberg and interaction pictures are

$$\begin{aligned}
\text{transition amplitude} &= \langle p_1, p_2 |_{\text{out}} \;\; | p_3, p_4 \rangle_{\text{in}} \\
&= \langle p_1, p_2 |_{\text{out}} \;\; U\left(t_f, t_i\right) | p_3, p_4 \rangle_{\text{in}} \\
\text{transition amplitude} &= \langle p_1, p_2 |_{\text{out}} \;\; A_{1,H}\left(t_1\right) A_{2,H}\left(t_2\right) | p_3, p_4 \rangle_{\text{in}} \\
&= \langle p_1, p_2 |_{\text{out}} \;\; U\left(t_f, t_1\right) A_{1,I} U\left(t_1, t_2\right) A_{2,I} U\left(t_2, t_i\right) | p_3, p_4 \rangle_{\text{in}}
\end{aligned}$$

Why to bother with the sophisticated notation in the Heisenberg and interaction pictures, if it anyway refers to the Schrödinger picture? The reason is, of course, that in the relativistic QFT it is preferable to use a covariant formalism, in which field operators depend on time and space-position on the same footing. It is simply preferable to deal with operators $\varphi(x)$ rather than $\varphi(\vec{x})$, which makes the Heisenberg picture more appropriate for relativistic QFT. The Schrödinger picture is most convenient for intuitive grasp of transition amlitudes, the Heisenberg picture is most convenient for formulation of relativistivc field theory and the interaction picture is most convenient for calculations.

[11] Note that the times t_i and t_f refer only to the Schrödinger picture states. Indeed, in spite of what the notation may suggest, the Heisenberg picture *in*- and *out*-states do not change in time. The *in*- and *out*- prefixes have nothing to do with the evolution of states in this picture (there is no such thing in the Heisenberg picture), they are simply labelling conventions (which have everything to do with the time evolution of the corresponding states in the Schrödinger picture).

green functions

For multiparticle systems, a particularly useful set of initial and final states is given by the states of particles simultaneously created at various positions, i.e. by the localized states. But as we have seen already, in the relativistic QFT the more appropriate states are the quasilocalized ones created by the field operators. The corresponding amplitude is apparently something like $\langle 0|\, \varphi_H(\vec{x}_1, t_f)\varphi_H(\vec{x}_2, t_f) \dots \varphi_H(\vec{x}_n, t_i)\, |0\rangle$. The vacuum state in this amplitude, however, is not exactly what it should be.

The time-independent state $|0\rangle$ in the Heisenberg picture corresponds to a particular time-evolving state $|\psi_0(t)\rangle$ in the Schrödinger picture, namely to the one for which $|\psi_0(0)\rangle = |0\rangle$. This state contains no particles at the time $t = 0$. But the fields $\varphi_H(\vec{x}, t_i)$ should rather act on a different state, namely the one which contains no particles at the time t_i. Such a state is nothing else than the previouslz defined $|0\rangle_{\text{in}}$

$$|0\rangle_{\text{in}} = e^{iHt_i}\, |0\rangle$$

In a complete analogy one has to replace the bra-vector $\langle 0|$ by

$$\langle 0|_{\text{out}} = \langle 0|\, e^{-iHt_f}$$

The quantity of interest is therefore given by the product of fields sandwiched not between $\langle 0|$ and $|0\rangle$, but rather

$$\langle 0|_{\text{out}}\, \varphi_H(\vec{x}_1, t_f)\varphi_H(\vec{x}_2, t_f) \dots \varphi_H(\vec{x}_n, t_i)\, |0\rangle_{\text{in}}$$

Because of relativity of simultaneity, however, this quantity looks differently for other observers, namely the time coordinates x_i^0 are not obliged to coincide. These time coordinates, on the other hand, are not completely arbitrary. To any observer the times corresponding to the simultaneous final state in one particular frame, must be all greater than the times corresponding to the simultaneous initial state in this frame. The more appropriate quantity would be a slightly more general one, namely the time-ordered T-product of fields[12] sandwiched between $\langle 0|_{\text{out}}$ and $|0\rangle_{\text{in}}$

$$\langle 0|_{\text{out}}\, T\{\varphi_H(x_1)\varphi_H(x_2) \dots \varphi_H(x_n)\}\, |0\rangle_{\text{in}}$$

The dependence on t_i and t_f is still present in $|0\rangle_{\text{in}}$ and $\langle 0|_{\text{out}}$. It is a common habit to get rid of this dependence by taking $t_i = -T$ and $t_f = T$ with $T \to \infty$

$$g(x_1, \dots, x_n) = \lim_{T\to\infty} \langle 0|\, e^{-iHT} T\, \{\varphi_H(x_1) \dots \varphi_H(x_n)\}\, e^{-iHT}\, |0\rangle$$

The exact reason for this rather arbitrary step remains unclear until the more serious treatment of the whole machinery becomes available in the next section).

The above matrix element is almost, but not quite, the Green function — one of the most prominent quantities in QFT. We shall call these functions the green functions (this notion is not common in literature, but this applies for the whole naive approach presented here). The genuine Green functions G are to be discussed later within the standard approach to the interacting fields (we will distinguish between analogous quantities in the naive and the standard approaches by using lowercase letters in the former and uppercase letter in the latter case).

Actual calculations of the green functions are performed, not surprisingly, in the interaction picture. The transition from the Heisenberg picture to the interaction one is provided by the

[12]For fields commuting at space-like intervals ($[\varphi_H(x), \varphi_H(y)] = 0$ for $(x-y)^2 < 0$) the time ordering is immaterial for times which coincide in a particular reference frame. For time-like intervals, on the other hand, the T-product gives the same ordering in all reference frames.

relations from the page 85. It is useful to start with

$$|0\rangle_{\text{in}} = e^{-iHT}|0\rangle = e^{-iHT}e^{iH_0T}|0\rangle = U^{-1}(-T,0)|0\rangle = U(0,-T)|0\rangle$$
$$\langle 0|_{\text{out}} = \langle 0|e^{-iHT} = \langle 0|e^{iH_0T}e^{-iHT} = \langle 0|U(T,0)$$

which holds for $H_0|0\rangle = 0$. The next step is to use $\varphi_H(x) = U^{-1}(x^0,0)\varphi_I(x)U(x^0,0)$ for every field in the green function, then to write $U(T,0)\varphi_H(x_1)\varphi_H(x_2)\ldots$ as

$$\underbrace{U(T,0)U^{-1}(x_1^0,0)}_{U(T,0)U(0,x_1^0)}\varphi_I(x_1)\underbrace{U(x_1^0,0)U^{-1}(x_2^0,0)}_{U(x_1^0,0)U(0,x_2^0)}\varphi_I(x_2)U(x_2^0,0)\ldots$$

and finally one utilizes $U(T,0)U(0,x_1^0) = U(T,x_1^0)$, etc. This means that for $x_1^0 \geq \ldots \geq x_n^0$ the green function $g(x_1,\ldots,x_n)$ is equal to

$$\lim_{T\to\infty}\langle 0|U(T,x_1^0)\varphi_I(x_1)U(x_1^0,x_2^0)\varphi_I(x_2)\ldots\varphi_I(x_n)U(x_n^0,-T)|0\rangle$$

and analogously for other orderings of times.

Let us now define a slightly generalized time ordered product as

$$T\{U(t,t')A(t_1)B(t_2)\ldots C(t_n)\} = U(t,t_1)A(t_1)U(t_1,t_2)B(t_2)\ldots C(t_n)U(t_n,t')$$

for $t \geq t_1 \geq t_2 \geq \ldots \geq t_n \geq t'$ and for other time orderings the order of operators is changed appropriately. With this definition we can finally write

$$g(x_1,\ldots,x_n) = \lim_{T\to\infty}\langle 0|T\{U(T,-T)\varphi_I(x_1)\ldots\varphi_I(x_n)\}|0\rangle$$

This form of the green function is what we were after. It has the form of the vacuum expectation value of the products of the field operators in the interaction picture and it is relatively straightforward to develop the technique for calculation of these objects. This technique will lead us directly to the Feynman rules. The rules were introduced in the introductory chapter, but they were not derived there. Now we are going to really derive them.

Remark: *The Feynman rules discussed in the Introductions/Conclusions concerned the scattering amplitude M_{fi}, while here we are dealing with the green functions. This, however, represents no contradiction. The green functions, as well as the genuine Green functions, are auxiliary quantities which are, as we will see briefly, closely related to the scattering amplitudes. It is therefore quite reasonable first to formulate the Feynman diagrams for the green or Green functions and only afterwards for the scattering amplitudes.*

perturbation theory

The practically useful, even if only approximate, solution for $U(t,t')$ is obtained by rewriting the differential equation to the integral one

$$U(t,t') = 1 - i\int_{t'}^{t} dt'' H_I(t'')\, U(t'',t')$$

which is then solved iteratively

0^{th} iteration $\quad U(t,t') = 1$

1^{st} iteration $\quad U(t,t') = 1 - i\int_{t'}^{t} dt_1 H_I(t_1)$

2^{nd} iteration $\quad U(t,t') = 1 - i\int_{t'}^{t} dt_1 H_I(t_1) + i^2\int_{t'}^{t} dt_1 H_I(t_1)\int_{t'}^{t_1} dt_2 H_I(t_2)$

etc.

Using a little artificial trick, the whole scheme can be written in a more compact form. The trick is to simplify the integration region in the multiple integrals $I_n = \int_{t'}^{t} dt_1 \int_{t'}^{t_1} dt_2 \ldots \int_{t'}^{t_{n-1}} dt_n\, H_I(t_1) H_I(t_2)\ldots$ Let us consider the n-dimensional hypercube $t' \le t_i \le t$ and for every permutation of the variables t_i take a region for which the first variable varies from t' up to t, the second variable varies from t' up to the first one, etc. There are $n!$ such regions ($n!$ permutations) and any point of the hypercube lies either inside exactly one of these regions, or at a common border of several regions. Indeed, for any point the ordering of the coordinates $(t_1,\ldots,t_n)$, from the highest to the lowest one, reveals unambiguously the region (or a border) within which it lies. The integral of the product of Hamiltonians over the whole hypercube is equal to the sum of integrals over the considered regions. In every region one integrates the same product of Hamiltonians, but with different ordering of the times. The trick now is to force the time ordering to be the same in all regions. This is achieved in a rather artificial way, namely by introducing the so-called time-ordered product $T\{A(t_1)B(t_2)\ldots\}$, which is the product with the terms organized from the left to the right with respect to decreasing time (the latest on the very left etc.). Integrals of this T-product over different regions are equal to each other, so we can replace the original integrals by the integrals over hypercubes $I_n = \frac{1}{n!}\int_{t'}^{t}\int_{t'}^{t} dt_1 \ldots dt_n\; T\{H_I(t_1)\ldots H_I(t_n)\}$ and consequently

$$U(t,t') = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{t'}^{t} \ldots \int_{t'}^{t} dt_1 \ldots dt_n\, T\{H_I(t_1)\ldots H_I(t_n)\}$$

which is usually written in a compact form as

$$U(t,t') = Te^{-i\int_{t'}^{t} dt''\, H_I(t'')}$$

where the definition of the RHS is the RHS of the previous equation.

Note that if the interaction Hamiltonian is proportional to some constant (e.g. a coupling constant) then this iterative solution represents the power expansion (perturbation series[13]) in this constant.

[13]The usual time-dependent perturbation theory is obtained by inserting the expansion $|\psi_I(t)\rangle = a_n(t)|\varphi_n\rangle$, where $|\varphi_n\rangle$ are eigenvectors of the free Hamiltonian H_0, into the original equation $i\partial_t|\psi_I\rangle = H_I(t)|\psi_I\rangle$. From here one finds a differential equation for $a_n(t)$, rewrites it as an integral equation and solves it iteratively.

Wick's theorem

The perturbative expansion of $U(T,-T)$ is a series in $H_I(t)$, which in turn is a functional of $\varphi_I(x)$, so our final expression for the green function gives them as a series of VEVs (vacuum expectation values) of products of φ_I-fields.

As we already know from the Introductions, the most convenient way of calculating the VEVs of products of creation and annihilation operators is to rush the creation and annihilation operators to the left and to the right respectively. We are now going to accommodate this technique to the VEVs of time-ordered products of φ_I-fields.

The keyword is *the normal product* of fields. First one writes $\varphi_I = \varphi_I^+ + \varphi_I^-$, where φ_I^+ and φ_I^- are parts of the standard expansion of $\varphi_I(x)$ containing only the annihilation and the creation operators respectively[14]. The normal product of fields, denoted as $N\{\varphi_I(x)\varphi_I(y)\ldots\}$ or $:\varphi_I(x)\varphi_I(y)\ldots:$, is defined as the product in which all φ_I^--fields are reshuffled by hand to the left of all φ_I^+-fields, e.g. $N\{\varphi_I(x)\varphi_I(y)\} = \varphi_I^-(x)\varphi_I^-(y) + \varphi_I^-(x)\varphi_I^+(y) + \varphi_I^-(y)\varphi_I^+(x) + \varphi_I^+(x)\varphi_I^+(y)$. Everybody likes normal products, because their VEVs vanish.

The trick, i.e. the celebrated Wick's theorem, concerns the relation between the time-ordered and normal products. For two fields one has $\varphi_I(x)\varphi_I(y) = N\{\varphi_I(x)\varphi_I(y)\} + \left[\varphi_I^+(x), \varphi_I^-(y)\right]$. It is straightforward to show (do it) that $\left[\varphi_I^+(x), \varphi_I^-(y)\right] = D(x-y)$ where

$$D(x-y) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\vec{p}}}e^{-ip(x-y)}$$

The relation between $T\{\varphi_I\varphi_I\}$ and $N\{\varphi_I\varphi_I\}$ is now straightforward

$$T\{\varphi_I(x)\varphi_I(y)\} = N\{\varphi_I(x)\varphi_I(y)\} + d_F(x-y)$$

where

$$d_F(\xi) = \vartheta\left(\xi^0\right)D(\xi) + \vartheta\left(-\xi^0\right)D(-\xi)$$

The function d_F is almost equal to the so-called Feynman propagator D_F (see p.97). Everybody likes $d_F(x-y)$, $D_F(x-y)$ and similar functions, because they are not operators and can be withdrawn out of VEVs.

For three fields one obtains in a similar way[15]

$$\begin{aligned}\varphi_I(x)\varphi_I(y)\varphi_I(z) &= N\{\varphi_I(x)\varphi_I(y)\varphi_I(z)\}\\ &\quad + D(x-y)\varphi_I(z) + D(x-z)\varphi_I(y) + D(y-z)\varphi_I(x)\\ T\{\varphi_I(x)\varphi_I(y)\varphi_I(z)\} &= N\{\varphi_I(x)\varphi_I(y)\varphi_I(z)\}\\ &\quad + d_F(x-y)\varphi_I(z) + d_F(x-z)\varphi_I(y) + d_F(y-z)\varphi_I(x)\end{aligned}$$

Now we can formulate and prove the Wick's theorem for n fields

$$\begin{aligned}T\{\varphi_I(x_1)\ldots\varphi_I(x_n)\} &= N\{\varphi_I(x_1)\ldots\varphi_I(x_n)\}\\ &\quad + d_F(x_1-x_2)\,N\{\varphi_I(x_3)\ldots\varphi_I(x_n)\} + \ldots\\ &\quad + d_F(x_1-x_2)\,d_F(x_3-x_4)\,N\{\varphi_I(x_5)\ldots\varphi_I(x_n)\} + \ldots\\ &\quad + \ldots\end{aligned}$$

---

[14] Strange, but indeed $\varphi_I^+(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}a_{\vec{p}}e^{-ipx}$ and $\varphi_I^-(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}a_{\vec{p}}^+e^{ipx}$. The superscript $\pm$ is not in honour of the creation and annihilation operators, but rather in honour of the sign of energy $E = \pm\omega_{\vec{p}}$.

[15] One starts with $\varphi_I(x)\varphi_I(y)\varphi_I(z) = \varphi_I(x)\varphi_I(y)\varphi_I^+(z) + \varphi_I(x)\varphi_I(y)\varphi_I^-(z)$, followed by $\varphi_I(x)\varphi_I(y)\varphi_I^-(z) = \varphi_I(x)[\varphi_I(y),\varphi_I^-(z)] + [\varphi_I(x),\varphi_I^-(z)]\varphi_I(y) + \varphi_I^-(z)\varphi_I(x)\varphi_I(y)$, so that $\varphi_I(x)\varphi_I(y)\varphi_I(z) = \varphi_I(x)\varphi_I(y)\varphi_I^+(z) + \varphi_I(x)D(y-z) + D(x-z)\varphi_I(y) + \varphi_I^-(z)\varphi_I(x)\varphi_I$. At this point one utilizes the previous result for two fields, and finally one has to realize that $N\{\varphi_I(x)\varphi_I(y)\}\varphi_I^+(z) + \varphi_I^-(z)N\{\varphi_I(x)\varphi_I(y)\} = N\{\varphi_I(x)\varphi_I(y)\varphi_I(z)\}$.

In each line the ellipsis stands for terms equivalent to the first term, but with the variables x_i permutated in all possible ways. The number of the Feynman propagators is increased by one when passing to the next line. The proof is done by induction, the method is the same as a we have used for three fields.

The most important thing is that except for the very last line, the RHS of the Wick's theorem has vanishing VEV (because of normal products). For n odd even the last line has vanishing VEV, for n even the VEV of the last line is an explicitly known number. This gives us the quintessence of the Wick's theorem: for n odd $\langle 0|\, T\left\{\varphi_I\left(x_1\right)\ldots\varphi_I\left(x_n\right)\right\}|0\rangle = 0$, while for n even

$$\langle 0|\, T\left\{\varphi_I\left(x_1\right)\ldots\varphi_I\left(x_n\right)\right\}|0\rangle = d_F\left(x_1-x_2\right)\ldots d_F\left(x_{n-1}-x_n\right) + \text{permutations}$$

Remark: *It is a common habit to economize a notation in the following way. Instead of writing down the products $d_F\left(x_1-x_2\right)\ldots d_F\left(x_{n-1}-x_n\right)$ one writes the product of fields and connects by a clip the fields giving the particular d_F. In this notation*

$$\begin{aligned}
d_F\left(x_1-x_2\right)d_F\left(x_3-x_4\right) &= \overbrace{\varphi_I\left(x_1\right)\varphi_I\left(x_2\right)}\,\overbrace{\varphi_I\left(x_3\right)\varphi_I\left(x_4\right)}\\
d_F\left(x_1-x_3\right)d_F\left(x_2-x_4\right) &= \overbrace{\varphi_I\left(x_1\right)\underbrace{\varphi_I\left(x_2\right)\varphi_I\left(x_3\right)}\varphi_I\left(x_4\right)}\\
d_F\left(x_1-x_4\right)d_F\left(x_2-x_3\right) &= \overbrace{\varphi_I\left(x_1\right)\underbrace{\varphi_I\left(x_2\right)\varphi_I\left(x_3\right)}\varphi_I\left(x_4\right)}
\end{aligned}$$

At this point we are practically done. We have expressed the green functions as a particular series of VEVs of time-ordered products of φ_I-operators and we have learned how to calculate any such VEV by means of the Wick's theorem. All one has to do now is to expand $U(T,-T)$ in the green function up-to a given order and then to calculate the corresponding VEVs.

Example: *$g\left(x_1,x_2,x_3,x_4\right)$ in the φ^4-theory*
notation: $\varphi_i := \varphi_I\left(x_i\right)$, $\varphi_x := \varphi_I\left(x\right)$, $d_{ij} := d_F(x_i-x_j)$, $d_{ix} := d_F(x_i-x)$

$$\begin{aligned}
g &= \lim_{T\to\infty}\langle 0|\, T\left\{U(T,-T)\varphi_1\varphi_2\varphi_3\varphi_4\right\}|0\rangle = g^{(0)}+g^{(1)}+\ldots\\
g^{(0)} &= \langle 0|\, T\left\{\varphi_1\varphi_2\varphi_3\varphi_4\right\}|0\rangle = d_{12}d_{34}+d_{13}d_{24}+d_{14}d_{23}\\
g^{(1)} &= -\frac{ig}{4!}\langle 0|\, T\left\{\int d^4x\,\varphi_x^4\varphi_1\varphi_2\varphi_3\varphi_4\right\}|0\rangle\\
&= -\frac{ig}{4!}\int d^4x\left\{24\times d_{1x}d_{2x}d_{3x}d_{4x}+12\times d_{12}d_{xx}d_{3x}d_{4x}+\ldots\right\}
\end{aligned}$$

where we have used $U(\infty,-\infty) = 1-i\int_{-\infty}^{\infty}dt\,H_I\left(t\right)+\ldots = 1-\frac{ig}{4!}\int d^4x\,\varphi_x^4+\ldots$

Feynman rules

The previous example was perhaps convincing enough in two respects: first that in principle the calculations are quite easy (apart from integrations, which may turn out to be difficult), and second that practically they become almost unmanageable rather soon (in spite of our effort to simplify the notation). Conclusion: further notational simplifications and tricks are called for urgently.

The most wide-spread trick uses a graphical representation of various terms in green function expansion. Each variable x is represented by a point labeled by x. So in the previous example we would have 4 points labeled by x_i $i = 1, 2, 3, 4$ and furthermore a new point $x, x', \ldots$ for each power of H_I. Note that the Hamiltonian density $\mathcal{H}_I$ contains several fields, but all at the same point — this is the characteristic feature of local theories. For each $d_F(y-z)$ the points labeled by y and z are connected by a line. If there are several different fields, there are several different Feynman propagators, and one has to use several different types of lines.

In this way one assigns a diagram to every term supplied by the team-work of the perturbative expansion of $U(T,-T)$ and the Wick's theorem. Such diagrams are nothing else but the famous Feynman diagrams. Their structure is evident from the construction. Every diagram has external points, given by the considered g-function, and internal points (vertices), given by $\mathcal{H}_I$. The number of internal points is given by the order in the perturbative expansion. The structure of the vertices (the number and types of lines entering the vertex) is given by the structure of H_I, each product of fields represent a vertex, each field in the product represent a line entering this vertex.

A diagram, by construction, represents a number. This number is a product of factors corresponding to lines and vertices. The factor corresponding to a line (internal or external) connecting x, y is $d_F(x-y)$. The factor corresponding to a vertex in the above example is $-\frac{ig}{4!}\int d^4x$, while in the full generality it is

$$-i \times \text{what remains of } \mathcal{H}_I \text{ after the fields are "stripped off"} \times \int d^4x$$

Further simplification concerns combinatorics. Our procedure, as described so-far, gives a separate diagram for each of the 24 terms $-\frac{ig}{4!}\int d^4x\, d_{1x}d_{2x}d_{3x}d_{4x}$ in $g^{(1)}$ in the above example. As should be clear from the example, this factor is purely combinatorial and since it is typical rather than exceptional, it is reasonable to include this 24 into the vertex factor (and to draw one diagram instead of 24 identical diagrams). This is achieved by doing the appropriate combinatorics already in the process of "stripping the fields off", and it amounts to nothing more than to the multiplication by $n!$ for any field appearing in $\mathcal{H}_I$ in the n-th power. An economic way of formalizing this "stripping off" procedure, with the appropriate combinatorics factors, is to use the derivatives of $\mathcal{H}_I$ with respect to the fields.

Having included the typical combinatorial factor into the vertex, we have to pay a special attention to those (exceptional) diagrams which do not get this factor. The 12 terms $-\frac{ig}{4!}\int d^4x\, d_{12}d_{xx}d_{3x}d_{4x}$ in $g^{(1)}$ in the example can serve as an illustration. Twelve identical diagrams are represented by one diagram according to our new viewpoint, but this diagram is multiplied by 24, hidden in the vertex factor, rather then by 12. To correct this, we have to divide by 2 — one example of the infamous explicit combinatorial factors of Feynman rules.
The rules can be summarized briefly as

the Feynman rules for the green functions in the x-representation

line (internal or external)	$d_F(x-y)$
vertex (n legs)	$-i\,\frac{\partial^n \mathcal{H}_I}{\partial \varphi_I^n}\Big\vert_{\varphi_I=0}\int d^4x$

These are not the Feynman rules from the Introductions yet, but we are on the right track.

The first step towards the rules from the Introductions concerns the relation between $\mathcal{H}_I$ and $\mathcal{L}_{\text{int}}$. For interaction Lagrangians with no derivative terms (like the φ^4-theory), the definition $H = \int d^3x\ (\dot\varphi\pi - \mathcal{L})$ implies immediately $\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}}$. And since $\mathcal{H}_{\text{int}}$ in the Heisenberg

picture is the same function of φ_H-fields, as $\mathcal{H}_I$ is of φ_I-fields (as we have convinced ourselves), one can replace $-\partial^n \mathcal{H}_I/\partial\varphi_I^n$ by $\partial^n \mathcal{L}_{\text{int}}/\partial\varphi^n$. Finally, for vertices one can replace $\mathcal{L}_{\text{int}}$ by $\mathcal{L}$ in the last expression, because the difference is the quadratic part of the Lagrangian, and vertices under consideration contain at least three legs. For interactions with derivative terms (say $\mathcal{L}_{\text{int}} \sim \varphi\partial_\mu\varphi\partial^\mu\varphi$) the reasoning is more complicated, but the result is the same. We will come back to this issue shortly (see p.94). For now let us proceed, as directly as possible, with the easier case.

Another step is the use of the Fourier expansion[16]

$$d_F(x-y) = \int \frac{d^4p}{(2\pi)^4} d_F(p)\, e^{-ip(x-y)}$$

This enables us to perform the vertex x-integrations explicitly, using the identity $\int d^4x\, e^{-ix(p+p'+\ldots)} = (2\pi)^4 \delta^4(p+p'+\ldots)$, what results in

the Feynman rules for the green functions in the p-representation

internal line	$\int \frac{d^4p}{(2\pi)^4} d_F(p)$
external line	$\int \frac{d^4p}{(2\pi)^4} d_F(p)\, e^{\pm ipx_i}$
vertex	$i \left.\frac{\partial^n \mathcal{L}}{\partial\varphi^n}\right\vert_{\varphi=0} (2\pi)^4 \delta^4(p+p'+\ldots)$

Let us remark that some authors prefer to make this table simpler-looking, by omitting the factors of $(2\pi)^4$ as well as the momentum integrations, and shifting them to the additional rule requiring an extra $(2\pi)^4$ for each vertex and $(2\pi)^{-4}\int d^4p$ for each line (internal or external). We have adopted such a convention in the Introductions.

derivative couplings

Now to the interaction Lagrangians with derivative terms. The prescription from the Introductions was quite simple: any ∂_μ in the interaction Lagrangian furnishes the $-ip^\mu$ factor for the corresponding vertex in the p-representation Feynman rules (p^μ being the momentum assigned to the corresponding leg, oriented toward the vertex). To understand the origin of this factor, it is (seemingly) sufficient to differentiate the Wick's theorem, e.g. for two fields

$$T\left\{\varphi_I(x)\,\partial'_\mu\varphi_I(x')\right\} = \partial'_\mu\left(N\left\{\varphi_I(x)\,\varphi_I(x')\right\} + d_F(x-x')\right)$$

When calculating the green function, the derivative can be withdrawn from VEV, and once $d_F(x-x')$ is Fourier expanded, it produces the desired factor (the reader is encouraged to make him/her-self clear about momentum orientations).

[16] One may be tempted to use $i(p^2-m^2)^{-1}$ or $i(p^2-m^2+i\varepsilon)^{-1}$ as $d_F(p)$ (see p. 97), but neither would be correct. Both choices lead to results differing from $d_F(x-y)$ by some functions of ε, which tend to disappear when $\varepsilon \to 0$. Nevertheless the ε-differences are the important ones, they determine even the seemingly ε-independent part of the result.

Anyway, apart from the $i\varepsilon$ subtleties, $d_F(p)$ comes out equal to what we have calculated in the Introductions (from quite different definition of propagator). This may seem like a coincidence, and one may suspect if one gets equal results even beyond the real scalar field example. The answer is affirmative, but we are not going to prove it here in the full generality. The reason is that the general statement is more transparent in another formulation of QFT, namely in the path integral formalism. So we prefer to discuss this issue within this formalism.

There is, however, a subtlety involved. The above identity is not straightforward, even if it follows from the relation $\varphi_I(x)\partial'_\mu\varphi_I(x') = \partial'_\mu(\varphi_I(x)\varphi_I(x')) = \partial'_\mu(N\{\varphi_I(x)\varphi_I(x')\} + D(x-x'))$. The point is that when combining two such identities to get the T-product at the LHS, one obtains $\vartheta(\xi^0)\partial'_\mu D(\xi) + \vartheta(-\xi^0)\partial'_\mu D(-\xi)$ instead of $\partial_\mu d_F(\xi)$ on the RHS (with $\xi = x - x'$). The extra term, i.e. the difference between what is desired and what is obtained, is $D(\xi)\partial_0\vartheta(\xi^0) + D(-\xi)\partial_0\vartheta(-\xi^0) = (D(\xi) - D(-\xi))\delta(\xi^0)$ and this indeed vanishes, as can be shown easily from the explicit form of $D(\xi)$ (see page 91).

Unfortunately, this is not the whole story. Some extra terms (in the above sense) are simply die-hard. They do not vanish as such, and one gets rid of them only via sophisticated cancellations with yet another extras entering the game in case of derivative couplings[17]. Attempting not to oppress the reader, we aim to outline the problem, without penetrating deeply into it.

The troublemaker is the T-product of several differentiated fields. An illustrative example is provided already by two fields, where one obtains

$$T\left\{\partial_\mu\varphi_I(x)\,\partial'_\nu\varphi_I(x')\right\} = \partial_\mu\partial'_\nu\left(N\left\{\varphi_I(x)\,\varphi_I(x')\right\} + d_F(x-x')\right) + \delta^0_\mu\delta^0_\nu\Delta(x-x')$$

with[18] $\Delta(\xi) = -i\delta^4(\xi)$. The same happens in products of more fields and the Wick's theorem is to be modified by the non-vanishing extra term $-i\delta^0_\mu\delta^0_\nu\delta^4(\xi)$, on top of the doubly differentiated standard propagator. In the "clip notation"

$$\partial_\mu\overbrace{\varphi_I(x)\,\partial'_\nu\varphi_I}(x') = \partial_\mu\partial'_\nu\overbrace{\varphi_I(x)\,\varphi_I}^{\text{standard}}(x') + \overbrace{\varphi_I(x)\,\varphi_I}^{\text{extra}}(x')$$

where $\overbrace{\varphi_I(x)\,\varphi_I}^{\text{standard}}(x') = d_F(x-x')$ and $\overbrace{\varphi_I(x)\,\varphi_I}^{\text{extra}}(x') = -i\delta^0_\mu\delta^0_\nu\delta^4(x-x')$

The rather unpleasant feature of this extra term is its non-covariance, which seems to ruin the highly appreciated relativistic covariance of the perturbation theory as developed so-far.

Because of the δ-function, the extra term in the propagator can be traded for an extra vertex. To illustrate this, let us consider as an example $\mathcal{L}[\varphi] = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 + \frac{g}{2}\varphi\partial_\mu\varphi\partial^\mu\varphi$. A typical term in the perturbative expansion of a green function contains $\mathcal{H}_{\text{int}}[\varphi(x)]\,\mathcal{H}_{\text{int}}[\varphi(x')]$ and clipping the fields together via the extra term gives[19]

$$\frac{-ig}{2}\varphi(x)\,\partial^\mu\varphi(x)\,\overbrace{\varphi_I(x)\,\varphi_I}^{\text{extra}}(x')\frac{-ig}{2}\varphi(x')\,\partial^\nu\varphi(x') = i\frac{g^2}{4}\varphi^2(x)\,\dot\varphi^2(x)$$

effectively contracting two original vertices into the new extra one. In this way one can get rid of the extra non-covariant term in the propagator, at the price of introduction of the non-covariant effective vertex. In our example this effective vertex corresponds to an extra term in the Lagrangian: $\mathcal{L}_{\text{extra}} = \frac{1}{2}g^2\varphi^2\dot\varphi^2$.

The factor $\frac{1}{2}$ follows from a bit of combinatorics. There are four possibilities for the extra clipping between the two $\mathcal{H}_{\text{int}}$, endowing the new effective vertex with the factor of 4. Less obvious is another factor of $\frac{1}{2}$, coming from the fact that interchange of the two contracted

[17] Similar problems (and similar solutions) haunt also theories of quantum fields with higher spins, i.e. they are not entirely related to derivative couplings.

[18] First one gets, along the same lines as above, $\Delta(\xi) = (D(\xi)-D(-\xi))\partial_0\delta(\xi^0)+2\delta(\xi^0)\partial_0(D(\xi)-D(-\xi))$. Due to the identity $f(x)\delta'(x) = -f'(x)\delta(x)$ this can be brought to the form $\Delta(\xi) = \delta(\xi^0)\partial_0(D(\xi)-D(-\xi))$ and plugging in the explicit form of $D(\xi)$ one obtains $\Delta(\xi) = \delta(\xi^0)\int\frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\vec p}}\partial_0(e^{-ip\xi}-e^{ip\xi}) = \delta(\xi^0)\int\frac{d^3p}{(2\pi)^3}\frac{-ip_0}{2\omega_{\vec p}}(e^{-ip\xi}+e^{ip\xi}) = -i\delta(\xi^0)\int\frac{d^3p}{(2\pi)^3}e^{i\vec p\vec\xi} = -i\delta^4(\xi)$.

[19] Here we pretend that $\mathcal{H}_{\text{int}} = -\mathcal{L}_{\text{int}}$, which is not the whole truth in the case at hand. We will correct this in the moment.

original vertices does not change the diagram. According to the rules for combinatoric factors (see section**??**) this requires the factor of $\frac{1}{2}$. Once the vertices are contracted, there is no (combinatoric) way to reconstruct this factor, so it has to be included explicitly.

The story is not over yet. There is another source of non-covariant vertices. The point is that once derivative couplings are present, the canonical momentum is not equal to the time derivative of the field any more. As an illustration let us consider our example again. Here one gets $\pi = \dot{\varphi} + g\varphi\dot{\varphi}$, i.e. $\dot{\varphi} = (1 + g\varphi)^{-1} \pi$. The corresponding Hamiltonian density can be written as $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\text{int}}$ where[20]

$$\mathcal{H}_0 = \frac{1}{2}\pi^2 + \frac{1}{2}\left|\nabla\varphi\right|^2 + \frac{1}{2}m^2\varphi^2$$
$$\mathcal{H}_{\text{int}} = \frac{g}{2}\varphi\left|\nabla\varphi\right|^2 - \frac{g}{2}\varphi\left(1 + g\varphi\right)^{-1}\pi^2$$

$\mathcal{H}_0$ corresponds to the Hamiltonian density of the free field, expressed in terms of conjugate quantities, obeying (after quantization) the standard commutation relation $\left[\varphi\left(x\right), \pi\left(y\right)\right] = i\delta^3\left(\vec{x} - \vec{y}\right)$. Using this $\mathcal{H}_0$ one can develop the perturbation theory in the standard way. Doing so it is convenient, as we have seen, to re-express the canonical momentum in terms of the field variables, leading to

$$\mathcal{H}_{\text{int}} = -\frac{1}{2}g\varphi\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}g^2\varphi^2\dot{\varphi}^2$$

As announced, this interaction Hamiltonian density contains, on top of the expected covariant term $-\mathcal{L}_{\text{int}}$, a non-covariant one. But now, the fanfares breaks out, and the non-covariant vertices originating from two different sources, cancel each other.[21] This miracle is not an exceptional feature of the example at hand, it is rather a general virtue of the canonical quantization: at the end of the day all non-covariant terms in vertices and propagators tend to disappear.

[20] $\mathcal{H} = \dot{\varphi}\pi - \mathcal{L} = \dot{\varphi}\pi - \frac{1}{2}\dot{\varphi}^2 + \frac{1}{2}\left|\nabla\varphi\right|^2 + \frac{1}{2}m^2\varphi^2 - \frac{g}{2}\varphi\partial_\mu\varphi\partial^\mu\varphi$
$= \frac{1}{2}\left(1 + g\varphi\right)^{-1}\pi^2 + \frac{1}{2}\left|\nabla\varphi\right|^2 + \frac{1}{2}m^2\varphi^2 + \frac{g}{2}\varphi\left|\nabla\varphi\right|^2$
$= \frac{1}{2}\pi^2 + \frac{1}{2}\left|\nabla\varphi\right|^2 + \frac{1}{2}m^2\varphi^2 + \frac{g}{2}\varphi\left|\nabla\varphi\right|^2 - \frac{g}{2}\varphi\left(1 + g\varphi\right)^{-1}\pi^2$

[21] One may worry about what happens to the non-covariant part of $\mathcal{H}_{\text{int}}$ contracted (with whatever) via the non-covariant part of the propagator. Indeed, we have not consider such contractions, but as should be clear from what was said so-far, for any such contraction there is a twin contraction with opposite sign, so all such terms cancels out.

propagator

In the Introductions/Conclusions we have learned that the propagator of the scalar field is equal to $i/(p^2 - m^2)$. Let us check now, whether this ansatz for $d_F(p)$ really leads to the correct expression for $d_F(\xi)$, i.e. if

$$d_F(\xi) \stackrel{?}{=} \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{-ip\xi}$$

where $d_F(\xi) = \vartheta\left(\xi^0\right) D(\xi) + \vartheta\left(-\xi^0\right) D(-\xi)$ and $D(\xi) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\vec{p}}} e^{-ip\xi}$.

It is very useful to treat the p_0-variable in this integral as a complex variable. Writing $p^2 - m^2 = (p_0 - \omega_{\vec{p}})(p_0 + \omega_{\vec{p}})$ (recall that $\omega_{\vec{p}} = \sqrt{\vec{p}^2 + m^2}$) one finds that the integrand has two simple poles in the p^0-variable, namely at $p^0 = \pm\omega_{\vec{p}}$ with the residua $\pm(2\omega_{\vec{p}})^{-1} e^{\mp i\omega_{\vec{p}}\xi_0} e^{i\vec{p}.\vec{\xi}}$. The integrand is, on the other hand, sufficiently small at the lower (upper) semicircle in the p^0-plane for $\xi^0 > 0$ ($\xi^0 < 0$), so that it does not contribute to the integral for the radius of the semicircle going to infinity. So it almost looks like if

$$\int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{-ip\xi} \stackrel{?}{=} \pm \int \frac{d^3p}{(2\pi)^3} \left(\frac{e^{-ip\xi}}{p^0 + \omega_{\vec{p}}}|_{p^0 = \omega_{\vec{p}}} + \frac{e^{-ip\xi}}{p^0 - \omega_{\vec{p}}}|_{p^0 = -\omega_{\vec{p}}} \right)$$

(the sign reflects the orientation of the contour) which would almost give the desired result after one inserts appropriate ϑ-functions and uses $\vec{p} \to -\vec{p}$ substitution in the last term.

Now, was that not for the fact that the poles lay on the real axis, one could perhaps erase the questionmarks safely. But since they do lay there, one can rather erase the equality sign.

It is quite interesting, however, that one can do much better if one shifts the poles off the real axis. Let us consider slightly modified ansatz for the propagator, namely $i/(p^2 - m^2 + i\varepsilon)$ with positive ε (see the footnote on the page **??**). The pole in the variable p_0^2 lies at $\omega_{\vec{p}}^2 - i\varepsilon$, i.e. the poles in the variable p_0 lie at $\omega_{\vec{p}} - i\varepsilon$ and $-\omega_{\vec{p}} + i\varepsilon$ and so the trick with the complex plane now works perfectly well, leading to (convince yourself that it really does)

$$\int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\varepsilon} e^{-ip\xi} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\vec{p}}} \left\{ \vartheta(\xi^0) e^{-ip\xi - \varepsilon\xi_0} + \vartheta(-\xi^0) e^{ip\xi + \varepsilon\xi_0} \right\}$$

At this point one may be tempted to send ε to zero and then to claim the proof of the identity $d_F(p) = i/(p^2 - m^2)$ being finished. This, however, would be very misleading. The limit $\varepsilon \to 0^+$ is quite non-trivial and one cannot simply replace ε by zero (that is why we were not able to take the integral in the case of $\varepsilon = 0$).

Within the naive approach one cannot move any further. Nevertheless, the result is perhaps sufficient to suspect the close relation between the result for the propagator as found in the Introductions/Conclusions and in the present chapter. Later on we will see that the $i\varepsilon$ prescription is precisely what is needed when passing from the naive approach to the standard one.

s-matrix

The green functions, discussed so-far, describe the time evolution between the initial and final states of particles created at certain positions. Most experimental setups in the relativistic particle physics correspond to different settings, namely to the initial and final states of particles created with certain momenta in the remote past and the remote future respectively. We will therefore investigate now the so-called s-matrix (almost, but not quite the famous S-matrix)

$$s_{fi} = \lim_{T\to\infty} \langle \vec{p}_1, \ldots, \vec{p}_m | \, U(T, -T) \, | \vec{p}_{m+1}, \ldots, \vec{p}_n \rangle$$

where f and i are abbreviations for $\vec{p}_1, \ldots, \vec{p}_m$ and $\vec{p}_{m+1}, \ldots, \vec{p}_n$ respectively. We have presented the definition directly in the interaction picture, which is most suitable for calculations. Of course, it can be rewritten in any other picture, as discussed on p.**??**.

The difference between the s-matrix and the genuine S-matrix (which is to be discussed within the standard approach) is in the states between which $U(T, -T)$ is sandwiched. Within our naive approach we adopt a natural and straightforward choice, based on the relation[22] $|\vec{p}\rangle = \sqrt{2\omega_{\vec{p}}} a^+_{\vec{p}} \, |0\rangle$, leading to

$$s_{fi} = \lim_{T\to\infty} \langle 0 | \sqrt{2\omega_{\vec{p}_1}} a_{\vec{p}_1,I}(T) \ldots U(T, -T) \ldots \sqrt{2\omega_{\vec{p}_n}} a^+_{\vec{p}_n,I}(-T) \, |0\rangle$$

Intuitively this looks quite acceptable, almost inevitable: the multi-particle state is created, by the corresponding creation operators, from the state with no particles at all. There is, however a loophole in this reasoning.

The main motivation for the s-matrix was how do the real experiments look like. The states entering the definition should therefore correspond to some typical states prepared by accelerators and detected by detectors. The first objection which may come to one's mind is that perhaps we should not use the states with sharp momenta (plane waves) but rather states with "well-defined, even if not sharp" momenta and positions (wave-packets). This, however, is not the problem. One can readily switch from plane-waves to wave-packets and vice versa, so the difference between them is mainly the difference in the language used, rather than a matter of principle.

The much more serious objection is this one: Let us suppose that we have at our disposal apparatuses for measurement of momenta. Then we can prepare states more or less close to the states with sharp energy and 3-momentum. The above considered states $|\vec{p}\rangle = \sqrt{2\omega_{\vec{p}}} a^+_{\vec{p}} \, |0\rangle$ are such states, but only for the theory with the free Hamiltonian. Once the interaction is switched on, the said $|\vec{p}\rangle$ states may differ significantly from what is prepared by the available experimental devices. Once more and aloud: typical experimentally accessible states in the worlds with and without interaction may differ considerably. And, as a rule, they really do.

One may, of course, ignore this difference completely. And it is precisely this ignorance, what constitutes the essence of our naive approach. Indeed, the core of this approach is the work with the s-matrix. defined in terms of explicitly known simple states, instead of dealing with the S-matrix, defined in terms of the states experimentally accessible in the real world (with interactions). The latter are usually not explicitly known, so the naivity simplifies life a lot.

What excuse do we have for such a simplification? Well, if the interaction may be viewed as only a small perturbation of the free theory — and we have adopted this assumption already,

[22] The relation is given in a bit sloppy way. In the Schrödinger picture, it is to be understood as $|\vec{p}, -T\rangle_S = \sqrt{2\omega_{\vec{p}}} a^+_{\vec{p},S} \, |0\rangle$, where $|0\rangle$ is just a particular state in the Fock space (no time dependence of $|0\rangle$ is involved in this relation). In the interaction picture the relation reads $|\vec{p}, -T\rangle_I = \sqrt{2\omega_{\vec{p}}} a^+_{\vec{p},I}(-T) \, |0\rangle$ (this is equivalent to the Schrödinger picture due to the fact that $H_0 \, |0\rangle = 0$). In the Heisenberg picture, however, one has $|\vec{p}\rangle_H = \sqrt{2\omega_{\vec{p}}} a^+_{\vec{p},H}(-T) \, e^{-iHT} \, |0\rangle \neq \sqrt{2\omega_{\vec{p}}} a^+_{\vec{p},H}(-T) \, |0\rangle$ (due to the fact that $|0\rangle$ is usually not an eigenstate of H).

namely in the perturbative treatment of $U(T,-T)$ operator — then one may hope that the difference between the two sets of states is negligible. To take this hope too seriously would be indeed naive. To ignore it completely would be a bit unwise. If nothing else, the s-matrix is the zeroth order approximation to the S-matrix, since the unperturbed states are the zeroth order approximation of the corresponding states in the full theory. Moreover, the developments based upon the naive assumption tend to be very useful, one can get pretty far using this assumption and almost everything will survive the more rigorous treatment of the standard approach.

As to the calculation of the s-matrix elements, it follows the calculation of the green functions very closely. One just uses the perturbative expansion $U(T,-T) = T\exp\{-i\int_{-T'}^{T} dt\, H_I(t)\}$ (see p.90) and the Wick's theorem, which is to be supplemented by[23]

$$\varphi_I(x)\, a^+_{\vec{p},I}(-T) = N\{\varphi_I(x)\, a^+_{\vec{p},I}(-T)\} + \frac{1}{\sqrt{2\omega_{\vec{p}}}} e^{-ipx} e^{-i\omega_{\vec{p}}T}$$
$$a_{\vec{p},I}(T)\, \varphi_I(x) = N\{a_{\vec{p},I}(T)\, \varphi_I(x)\} + \frac{1}{\sqrt{2\omega_{\vec{p}}}} e^{ipx} e^{-i\omega_{\vec{p}}T}$$

or in the "clipping notation"

$$\overbrace{\varphi_I(x)\, a_{\vec{p},I}}^{+}(-T) = \frac{1}{\sqrt{2\omega_{\vec{p}}}} e^{-ipx} e^{-i\omega_{\vec{p}}T}$$
$$\overbrace{a_{\vec{p},I}(T)\, \varphi_I}(x) = \frac{1}{\sqrt{2\omega_{\vec{p}}}} e^{ipx} e^{-i\omega_{\vec{p}}T}$$

Consequently, the s-matrix elements are obtained in almost the same way as are the green functions, i.e. by means of the Feynman rules. The only difference is the treatment of the external lines: instead of the factor $d_F(x-y)$, which was present in the case of the green functions, the external legs provide the factors $e^{\mp ipx} e^{-i\omega_{\vec{p}}T}$ for the s-matrix elements, where the upper and lower sign in the exponent corresponds to the ingoing and outgoing particle respectively. (Note that the $\sqrt{2\omega_{\vec{p}}}$ in the denominator is canceled by the $\sqrt{2\omega_{\vec{p}}}$ in the definition of s_{fi}.) In the Feynman rules the factor $e^{-i\omega_{\vec{p}}T}$ is usually omitted, since it leads to the pure phase factor $\exp\{-iT\sum_{i=1}^{n}\omega_{\vec{p}_n}\}$, which is redundant for probability densities, which we are interested in[24].

[23] Indeed, first of all one has $\varphi_I(x)\, a^+_{\vec{p},I}(t') = N\{\varphi_I(x) a^+_{\vec{p},I}(t')\} + [\varphi_I(x), a^+_{\vec{p},I}(t')]$ and then $[\varphi_I(x), a^+_{\vec{p},I}(t')] = \int \frac{d^3p'}{(2\pi)^3} \frac{e^{i\vec{p}'\cdot\vec{x}}}{\sqrt{2\omega_{\vec{p}'}}} [a_{\vec{p}',I}(t), a^+_{\vec{p},I}(t')] = \int \frac{d^3p'}{\sqrt{2\omega_{\vec{p}'}}} e^{i(\vec{p}'\cdot\vec{x} - \omega_{\vec{p}'}t + \omega_{\vec{p}}t')} \delta(\vec{p}-\vec{p}')$ and the same gymnastics (with an additional substitution $p' \to -p'$ in the integral) is performed for $[a_{\vec{p},I}(t'), \varphi_I(x)]$.

[24] Omission of the phase factor is truly welcome, otherwise we should bother about the ill-defined limit $T \to \infty$. Of course, avoiding problems by omitting the trouble-making pieces is at least nasty, but what we are doing here is not that bad. Our sin is just a sloppiness. We should consider, from the very beginning, the limit $T \to \infty$ for the probability and not for the amplitude.

In this way one obtains

the Feynman rules for the s-matrix elements in the x-representation

internal line	$d_F(x-y)$
ingoing external line	$e^{\mp ipx}$
vertex (n legs)	$-i\left.\frac{\delta^n \mathcal{L}}{\delta \varphi^n}\right\vert_{\varphi=0} \int d^4x$

The next step is the use of the Fourier expansion of $d_F(x-y)$ allowing for explicit x-integrations (see p.94), resulting in

the Feynman rules for the s-matrix elements in the p-representation

internal line	$\int \frac{d^4p}{(2\pi)^4}\, d_F(p)$
external line	1
vertex	$i\left.\frac{\delta^n \mathcal{L}}{\delta \varphi^n}\right\vert_{\varphi=0} (2\pi)^4 \delta^4(p+p'+\ldots)$

Omitting the factors of $(2\pi)^4$ as well as the momentum integrations, and shifting them to the additional rule requiring an extra $(2\pi)^4$ for each vertex and $(2\pi)^{-4}\int d^4p$ for each internal line, one obtains

another form of the Feynman rules
for the s-matrix elements in the p-representation

internal line	$d_F(p)$
external line	1
vertex	$i\left.\frac{\delta^n \mathcal{L}}{\delta \varphi^n}\right\vert_{\varphi=0} \delta^4(p+p'+\ldots)$

which is now really very close to our presentation of the Feynman rules in the Introductions/Conclusions[25].

[25] In the Introductions/Conclusions the Feynman rules were used to calculate the scattering amplitude M_{fi} rather than the S-matrix. These two quantities are, however, closely related: $S_{fi} = \mathbf{1} + iM_{fi}(2\pi)^4 \delta^{(4)}(P_f - P_i)$ or $s_{fi} = \mathbf{1} + im_{fi}(2\pi)^4 \delta^{(4)}(P_f - P_i)$.

connected diagrams

IIThe M_{fi}, or rather m_{fi} within our naive approach, is more appropriate for the discussion of the cross-sections and decay rates, which is our next task.At this point, there are only three differences left:

- presence of $d_F(p)$ instead of the genuine Feynman propagator $D_F(p)$
- no $\sqrt{Z}$ factors corresponding to external legs
- presence of disconnected diagrams like

in the perturbative expansions of the green function and the s-matrix[26], while in the Introductions/Conclusions only connected Feynman diagrams were accounted for.

The differences are due to the fact that we are dealing with the s-matrix rather than the S-matrix. In the next section we will learn how the so-far missed ingredients (replacement of d_F by D_F, appearance of $\sqrt{Z}$ and fadeaway of disconnected diagrams) will enter the game in the standard approach.

As to the comparison of the rules presented in the Introductions/Conclusions to the ones derived here, let us remark that in the Introductions/Conclusions we did not introduce the notion of the S-matrix explicitly. Neverthweless, it was present implicitly via the quantity M_{fi}, since S and M are very closely related

$$S_{fi} = \mathbf{1} + iM_{fi}(2\pi)^4 \delta^{(4)}(P_f - P_i)$$

The M_{fi}, or rather m_{fi} within our naive approach, is more appropriate for the discussion of the cross-sections and decay rates, which is our next task.

Remark: *As we have seen, the green functions g and the s-matrix elements s_{fi} are very closely related. The only differences are the external legs factors: $d_F(p)$ for the green function g and simply 1 (or something slightly more complicated in case of higher spins) for the s-matrix elements. This may be formulated in the following way: s_{fi} is obtained from the corresponding green function g by multiplication of each external leg by the inverse propagator. Another, even more popular, formulation: s_{fi} is obtained from the corresponding g by amputation of the external legs. Actually, the relation between s and g is of virtually no interest whatsoever. We are, after all, interested only in the s-matrix, so there is no reason to bother about the green functions. Indeed, we could ignore the whole notion of the green function and derive the rules directly for the s-matrix. Doing so, however, we would miss the nice analogy between the naive and the standard approaches.*
The point is that similar relation holds also for the genuine Green functions G and the S-matrix elements. Indeed, as we will see, S_{fi} is obtained from the corresponding Green function G by amputation of the external leg and multiplication by $\sqrt{Z}$. And in the standard approach, unlike in the naive one, one cannot easily avoid the Green functions when aiming at the S-matrix.

[26] Which term in the perturbative expansion of $U(T,-T)$ corresponds this diagram to?

remark on complex fields and arrows

The expansion of the complex scalar field in creation and annihilation operators reads

$$\varphi_I(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(a_{\vec{p}}(t)\,e^{i\vec{p}.\vec{x}} + b^+_{\vec{p}}(t)\,e^{-i\vec{p}.\vec{x}}\right)$$
$$\varphi^*_I(\vec{x},t) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_{\vec{p}}}}\left(a^+_{\vec{p}}(t)\,e^{-i\vec{p}.\vec{x}} + b_{\vec{p}}(t)\,e^{i\vec{p}.\vec{x}}\right)$$

with

$$\left[a_{\vec{p}}(t), a^+_{\vec{p}'}(t)\right] = (2\pi)^3\,\delta(\vec{p}-\vec{p}')$$
$$\left[b_{\vec{p}}(t), b^+_{\vec{p}'}(t)\right] = (2\pi)^3\,\delta(\vec{p}-\vec{p}')$$

and all other commutators of creation and annihilation operators equal to zero. It is now straightforward to show that

$$T\{\varphi_I(x)\,\varphi_I(y)\} = N\{\varphi_I(x)\,\varphi_I(y)\}$$
$$T\{\varphi^*_I(x)\,\varphi^*_I(y)\} = N\{\varphi^*_I(x)\,\varphi^*_I(y)\}$$
$$T\{\varphi_I(x)\,\varphi^*_I(y)\} = N\{\varphi_I(x)\,\varphi^*_I(y)\} + d_F(x-y)$$
$$T\{\varphi^*_I(x)\,\varphi_I(y)\} = N\{\varphi^*_I(x)\,\varphi_I(y)\} + d_F(x-y)$$

This means that the time ordered product of two φ_I-fields as well as of the two φ^*_I-fields is already in the normal form, i.e. the only contributions to the Feynman diagrams come from $T\{\varphi_I\varphi^*_I\}$ and $T\{\varphi^*_I\varphi_I\}$.

This result is typical for complex fields: they provide two types of propagators, corresponding to products of the field and the conjugate field in two possible orderings. In case of the complex scalar field the factors corresponding to the two different orderings are equal to each other, so there is no reason to use two different graphical representations. This, however, is not a general feature. In other cases (e.g. in case of the electron-positron field) the different orderings lead to different factors. It is therefore necessary to distinguish these two possibilities also in their graphical representation, and this is usually done by means of an arrow.

3.1.3 Cross-sections and decay rates

Because of the relativistic normalization of states $\langle\vec{p}|\vec{p}'\rangle = 2E_{\vec{p}}(2\pi)^3\delta^3(\vec{p}-\vec{p}')$ the S-matrix (s-matrix)[27] elements do not give directly the probability amplitudes. To take care of this, one has to use the properly normalized vectors $(2\pi)^{-3/2}(2E)^{-1/2}|\vec{p}\rangle$, which leads to the probability amplitude equal to S_{fi} multiplied by $\prod_{j=1}^{n}(2\pi)^{-3/2}(2E_j)^{-1/2}$. Taking the module squared one obtains

$$\text{probability density} = |S_{fi}|^2\prod_{j=1}^{n}\frac{1}{(2\pi)^3 2E_j}$$

This expression presents an unexpected problem. The point is that S_{fi} turns out to contain δ-functions, and so $|S_{fi}|^2$ involves the ill-defined square of the δ-function. The δ-functions originate from the normalization of states and they are potentially present in any calculation which involves states normalized to the δ-function. In many cases one is lucky enough not to encounter any such δ-function in the result, but sometimes one is faced with the problem of dealing with δ-functions in probability amplitudes. The problem, when present, is usually treated either by switching to the finite volume normalization, or by exploitation of the specific set of basis vectors ("wave-packets" rather than "plane waves").

There are two typical δ-functions occurring in S_{fi}. The first one is just the normalization δ-function, symbolically written as $\delta(f-i)$, which represents the complete result in the case of the free Hamiltonian (for initial and final states being eigenstates of the free Hamiltonian). In the perturbative calculations, this $\delta(f-i)$ remains always there as the lowest order result. It is therefore a common habit to split the S-matrix as (the factor i is purely formal)

$$S_{fi} = \delta(f-i) + iT_{fi}$$

and to treat the corresponding process in terms of T_{fi}, which contains the complete information on transition probability for any $|f\rangle \neq |i\rangle$. The probability for $|f\rangle = |i\rangle$ can be obtained from the normalization condition (for the probability). Doing so, one effectively avoids the square of $\delta(f-i)$ in calculations.

But even the T-matrix is not "δ-free", it contains the momentum conservation δ-function. Indeed, both the (full) Hamiltonian and the 3-momentum operator commute with the time-evolution operator $0 = [P^\mu, U(T,-T)]$. When sandwiched between some 4-momentum eigenstates[28] $\langle f|$ and $|i\rangle$, this implies $(p_f^\mu - p_i^\mu)\langle f|U(T,-T)|i\rangle = 0$, and consequently $(p_f^\mu - p_i^\mu)S_{fi} = 0$. The same, of course, must hold for T_{fi}. Now any (generalized) function of f and i, vanishing for $p_f^\mu \neq p_i^\mu$, is either a non-singular function (finite value for $p_f = p_i$), or a distribution proportional to $\delta^4(p_f - p_i)$, or even a more singular function (proportional to some derivative of $\delta^4(p_f-p_i)$). We will assume proportionality to the δ-function

$$T_{fi} = (2\pi)^4\delta^4(p_f - p_i)M_{fi}$$

where $(2\pi)^4$ is a commonly used factor. Such an assumption turns out to lead to the finite result. We will ignore the other two possibilities, since if the δ-function provides a final result, they would lead to either zero or infinite results.

Now back to the notational mismatch. When discussing the first δ-function $\delta(f-i)$, the states $\langle f|$ and $|i\rangle$ were eigenstates of the free Hamiltonian H_0. On the other hand, when discussing the second δ-function $\delta^4(p_f-p_i)$, the states $\langle f|$ and $|i\rangle$ were eigenstates of the full Hamiltonian

[27] All definitions of this paragraph are formulated for the S-matrix, but they apply equally well for the s-matrix, e.g. one have $s_{fi} = \delta(f-i) + it_{fi}$ and $t_{fi} = (2\pi)^4\delta^4(p_f-p_i)m_{fi}$.

[28] The notational mismatch, introduced at this very moment, is to be discussed in a while.

H. We should, of course, make just one unambiguous choice of notation, and in this naive approach the choice is: $\langle f|$ and $|i\rangle$ are eigenstates of the free Hamiltonian H_0. This choice invalidates a part of the above reasoning, namely the part leading to $\delta\left(E_f - E_i\right)$ (the part leading to $\delta^3\left(\vec{p}_f - \vec{p}_i\right)$ remains untouched, since H_0 commutes with $\vec{P}$ and so one can choose $\langle f|$ and $|i\rangle$ to be eigenstates of both H_0 and $\vec{P}$).

Nevertheless, we are going to use $\delta^4\left(p_f - p_i\right)$ (rather then $\delta^3\left(\vec{p}_f - \vec{p}_i\right)$)in the definition of M_{fi}. The point is that $\delta\left(E_f - E_i\right)$ can be present in T_{fi}, even if the above justification fails. That this is indeed the case (to any order of the perturbation theory) can be understood directly from the Feynman rules. Recall that in the p-representation every vertex contains the momentum δ-function. Every propagator, on the other hand, contains the momentum integration and after all these integrations are performed, one is left with just one remaining δ-function, namely $\delta^4\left(p_f - p_i\right)$.

Proof: take a diagram, ignore everything except for momentum δ-functions and integrations. Take any two vertices connected directly by an internal line and perform the corresponding integration, using one of the vertices δ-functions. After integration the remaining δ-function contains momenta of all legs of the both vertices. The result can be depicted as a diagram with two vertices shrunk into a new one, and the new vertex contains the appropriate momentum δ-function. If the considered vertices were directly connected by more than one internal line, then the new diagram contains some internal lines going from the new vertex back to itself. From the point of view of the present argument, such "daisy-buck loops" can be shrunk to the point, since they do not contain any δ-function. The procedure is then iterated until one obtains the simplest possible diagram with one vertex with all external lines attached to it and with the corresponding δ-function $\delta(\Sigma\, p_{\text{ext}})$. The final point is to realize that in Feynman diagrams momenta are oriented towards the vertices, while final state momenta are usually understood as flowing from the diagram, i.e. $\Sigma\, p_{\text{ext}} = p_f - p_i$.

After having identified the typical δ-function present in the T-matrix, we should face the annoying issue of the undefined square of this δ-function in $|T_{fi}|^2$. We will do this in a covardly manner, by an attempt to avoid the problem by the standard trick with the universe treated as a finite cube (length L) with the periodic boundary conditions. The allowed 3-momenta are $\vec{p} = \frac{2\pi}{L}\left(n_x, n_y, n_z\right)$ and the 3-momentum δ-function transfers to

$$\delta^3(\vec{p} - \vec{p}') = \int \frac{d^3x}{(2\pi)^3} e^{i(\vec{p}-\vec{p}').\vec{x}} \to \frac{V}{(2\pi)^3}\delta_{\vec{p}\vec{p}'} \stackrel{\text{def}}{=} \delta_V^3(\vec{p} - \vec{p}')$$

while the same trick performed with the time variable gives the analogous result $\delta(E - E') \to \frac{T}{2\pi}\delta_{EE'} \stackrel{\text{def}}{=} \delta_T(E - E')$. The calculation of $|T_{fi}|^2$ in the finite 4-dimensional cube presents no problem at all: $|T_{fi}|^2 = V^2T^2\delta_{\vec{p}_f\vec{p}_i}\delta_{E_fE_i}\left|M_{fi}\right|^2$. For reasons which become clear soon, we will write this as

$$|T_{fi}|^2 = VT\,(2\pi)^4\,\delta_{VT}^4(p_f - p_i)\left|M_{fi}\right|^2$$

where $\delta_{VT}^4 = \delta_V^3\delta_T$.

The finite volume normalization affects also the normalization of states, since $\langle\vec{p}|\vec{p}'\rangle = 2E_{\vec{p}}\,(2\pi)^3\,\delta^3\left(\vec{p} - \vec{p}'\right) \to 2E_{\vec{p}}V\delta_{\vec{P}\vec{P}'}$. The relation from the beginning of this paragraph between $|S_{fi}|^2$ (or $|T_{fi}|^2$ for $f \neq i$) and the corresponding probability therefore becomes: $\text{probability}_{\text{for } f\neq i} = |T_{fi}|^2\prod_{j=1}^n \frac{1}{2E_jV}$. Note that using the finite volume normalization, i.e. having discrete rather than continuous labeling of states, we should speak about probabilities rather than probability densities. Nevertheless, for the volume V big enough, this discrete distribution of states is very dense — one may call it quasi-continuous. In that case it is technically convenient

to work with the probability quasi-density, defined as the probability of decay into any state $|f\rangle = |\vec{p}_1, \ldots, \vec{p}_m\rangle$ within the phase-space element $d^3p_1 \ldots d^3p_m$. This is, of course, nothing else but the sum of the probabilities over all states within this phase-space element. If all the probabilities within the considered phase-space element were equal (the smaller the phase-space element is, the better is this assumption fulfilled) one could calculate the said sum simply by multiplying this probability by the number of states within the phase-space element. And this is exactly what is commonly used (the underlying reasoning is just a quasi-continuous version of the standard reasoning of integral calculus). And since the number of states within the interval d^3p is $\Delta n_x \Delta n_y \Delta n_z = V\, d^3p/\,(2\pi)^3$ one comes to the probability quasi-density (for $f \neq i$) being equal to $|T_{fi}|^2 \prod_{j=1}^{n} \frac{1}{2E_j V} \prod_{k=1}^{m} \frac{V\, d^3p}{(2\pi)^3}$. So for $f \neq i$ one has

$$\begin{matrix}\text{probability}\\ \text{quasidensity}\end{matrix} = VT\,(2\pi)^4\,\delta^4_{VT}(p_f - p_i)\,|M_{fi}|^2 \prod_{j=1}^{m} \frac{d^3p}{(2\pi)^3\,2E_j} \prod_{j=m+1}^{n} \frac{1}{2E_j V}.$$

Comparing this to the expressions for $d\Gamma$ and $d\sigma$ given in the Introductions (see p.18) we realize that we are getting really close to the final result. We just have to get rid of the awkward factors T and T/V in the probability quasi-densities for one and two initial particles respectively (and to find the relation between $1/E_A E_B$ and $[(p_A.p_B)^2 - m_A^2 m_B^2]^{-1/2}$ in case of the cross section). This step, however, is quite non-trivial. The point is that even if our result looks as if we are almost done, actually we are almost lost. Frankly speaking, the result is absurd: for the time T being long enough, the probability exceeds 1.

At this point we should critically reexamine our procedure and understand the source of the unexpected obscure factor T. Instead, we are going to follow the embarrassing tradition of QFT textbooks and use this evidently unreliable result for further reasoning, even if the word reasoning used for what follows is a clear euphemism[29]. The reason for this is quite simple: the present author is boldly unable to give a satisfactory exposition of these issues[30].

After having warned the reader about the unsoundness of what follows, we can proceed directly to the interpretation of the result for one-particle initial state in terms of the decay rate: From the linear time dependence of the probability density one can read out the probability density per unit time, this is equal to the time derivative of the probability density and for the exponential decay $\exp(-\Gamma t)$ this derivative taken at $t = 0$ is nothing else but the decay rate Γ (or $d\Gamma$ if we are interested only in decays with specific final states).

The previous statement is such a dense pack of lies that it would be hardly outmatched in an average election campaign. First of all, we did not get the linear time dependence, since T is not the "flowing time", but rather a single moment. Second, even if we could treat T as a variable, it definitely applies only to large times and, as far as we can see now, has absolutely nothing to say about infinitesimal times in the vicinity of $t = 0$. Third, if we took the linear time dependence seriously, then why to speak about exponential decay. Indeed, there is absolutely no indication of the exponential decay in our result.

[29] A fair comment on rates and cross sections is to be found in the Weinberg's book (p.134): *The proper way to approach these problems is by studying the way that experiments are actually done, using wave packets to represent particles localized far from each other before a collision, and then following the time history of these superpositions of multiparticle states. In what follows we will instead give a quick and easy derivation of the main results, actually more a mnemonic than a derivation, with the excuse that (as far as I know) no interesting open questions in physics hinge on getting the fine points right regarding these matters.*

[30] This incapability seems to be shared by virtually all authors of QFT textbooks (which perhaps brings some relief to any of them). There are many attempts, more or less related to each other, to introduce decay rates and cross sections. Some of them use finite volume, some use wave-packets (but do not closely follow the whole time evolution, as suggested by Weinberg's quotation), some combine the two approaches. And one feature is common to all of them: they leave much to be desired.

Nevertheless, for the reasons unclear at this point (they are discussed in the appendix **??**) the main conclusion of the above lamentable statement remains valid: decay rate is given by the probability quasi-density divided by T. After switching back to the infinite volume, which boils down to $\delta^4_{VT}(p_f - p_i) \to \delta^4(p_f - p_i)$, one obtains

$$d\Gamma = (2\pi)^4 \delta^4 (P_f - P_i) \frac{1}{2E_A} |M_{fi}|^2 \prod_{i=1}^{m} \frac{d^3 p_i}{(2\pi)^3 2E_i}$$

where E_A is the energy of the decaying particle.

Remark: *The exponential decay of unstable systems is a notoriously known matter, perhaps too familiar to realize how non-trivial issue it becomes in the quantum theory. Our primary understanding of the exponential decay is based on the fact that* $\exp(-\Gamma t)$ *is the solution of the simple differential equation* $dN/dt = -\Gamma N(t)$*, describing a population of individuals diminishing independently of a) each other b) the previous history. In quantum mechanics, however, the exponential decay should be an outcome of the completely different time evolution, namely the one described by the Schrödinger equation.*
Is it possible to have the exponential decay in the quantum mechanics, i.e. can one get $|\langle \psi_0 | \psi(t) \rangle|^2 = e^{-\Gamma t}$ *for* $|\psi(t)\rangle$ *being a solution of the Schrödinger equation with the initial condition given by* $|\psi_0\rangle$*? The answer is affirmative, e.g. one can easily convince him/herself that for the initial state* $|\psi_0\rangle = \sum c_\lambda |\varphi_\lambda\rangle$*, where* $H|\varphi_\lambda\rangle = E_\lambda |\varphi_\lambda\rangle$*, one obtains the exponential decay for* $\sum |c_\lambda|^2 \delta(E - E_\lambda) \stackrel{\text{def}}{=} p(E) = \frac{1}{2\pi} \frac{\Gamma}{(E-E_0)^2 + \Gamma^2/4}$ *(the so-called Breit-Wigner distribution of energy in the initial state).*
The Breit-Wigner distribution could nicely explain the exponential decay in quantum mechanics, were it not for the fact that this would require understanding of why this specific distribution is so typical for quantum systems. And this is very far from being obvious. Needless to say, the answer cannot be that the Breit-Wigner is typical because the exponential decay is typical, since this would immediately lead to a tautology. It would be nice to have a good understanding for the exponential decay in the quantum mechanics, but (unfortunately) we are not going to provide any.

For two particles in the initial state the reasoning is a bit more reasonable. The main trick is to use another set of the initial and final states, the one which enables semiclassical viewpoint, which further allows to give some sense to the suspicious factor T/V. We are speaking about well localized wave-packets with well defined momentum (to the extend allowed by the uncertainty principle). Let us call one particle the target, while the second one the beam. The target will be localized in all three space dimensions, the beam is localized just in one direction — the one of their relative momentum $\vec{p}_A - \vec{p}_B$. In the perpendicular directions the particle is completely unlocalized, the state corresponds to the strictly zero transverse momentum, in this way the particle simulates the transversely uniform beam. Let us note that if we want (as we do) to simulate a beam with a constant density independent of the extensiveness of the universe box, the beam particle state is to be normalized to L^2 rather than to 1.

Now in the finite-box-universe with the periodic boundary conditions the beam particle leaves the box from time to time, always simultaneously entering on the other side of the world. During the time T the scattering can therefore take place repeatedly. In the rest frame of the target particle the number of scatterings is $\frac{T}{L/v}$, where v is the velocity of the beam particle in this frame. The reasonable quantity (the cross-section[31]) in the target rest frame is therefore the probability quasi-density as obtained above, divided by the "repetition factor" vT/L and multiplied by the beam normalization factor L^2

$$d\sigma = (2\pi)^4 \delta^4_{VT}(p_f - p_i) \frac{1}{4E_A E_B v} |M_{fi}|^2 \prod_{j=1}^{m} \frac{d^3 p_j}{(2\pi)^3 2E_j}$$

Finally, one switches back to the infinite universe by $\delta^4_{VT}(p_f - p_i) \to \delta^4(p_f - p_i)$.

The energies E_A and E_B, as well as the beam particle velocity v, is understood in the target rest frame. To have a formula applicable in any frame, one should preferably find a Lorentz scalar, which in the target rest frame becomes equal to $E_A E_B v$. Such a scalar is provided by $[(p_A.p_B)^2 - m_A^2 m_B^2]^{1/2}$, since for $p_A = (m_A, \vec{0})$ it equals to $m_A(E_B^2 - m_B^2)^{1/2} = m_A |\vec{p}_B|$ and $v = |\vec{p}_B| / E_B$, and so we have come to the final result

$$d\sigma = (2\pi)^4 \delta^4 (P_f - P_i) \frac{|M_{fi}|^2}{4\sqrt{(p_A.p_B)^2 - m_A^2 m_B^2}} \prod_{j=1}^{m} \frac{d^3 p_j}{(2\pi)^3 2E_j}$$

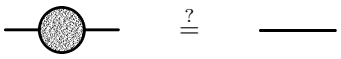
[31] Reacall that the cross-section is a (semi)classical notion, defined as follows. For a uniform beam of (semi)classical particles interacting with a uniformly distributed target particles, the number dn of beam particles scattered into an element of phase space dP_f is proportional to the flux j of the beam (density×velocity), the number N of the particles in target and dP_f itself: $dn \propto jN dP_f$. The coefficient of proportionality is the cross-section.

3.2 Standard approach

3.2.1 Consistency check of the naive approach

The naive approach to the quantum theory of interacting fields (particles) developed in the previous section was based on the natural assumption that time evolution of isolated particles is described by the free field Hamiltonian which is a part of the full Hamiltonian. This assumption, however, turns out to be incorrect. As a consequence of this incorrect assumption, calculations in the naive approach are plagued with inconsistencies.

To illustrate the problem, let us perform a small consistency check of the naive approach. We will consider the simplest problem in the theory of interacting fields, namely that of the propagation of the free particle. The transition amplitude from the quasi-localised state at one space-time point to a quasi-localised state at another space-time point is given by the two-point green function $g(x,y) = {}_{\text{out}}\langle 0|\, T\{\varphi_H(x)\varphi_H(y)\}\,|0\rangle_{\text{in}} = \lim_{T\to\infty}\langle 0|\, T\{U(T,-T)\varphi_I(x)\,\varphi_I(y)\}\,|0\rangle$. For the free field theory the green function is known explicitly $g_0(x,y) = d_F(x-y)$, which is represented by a simple line in the Feynman diagrams (we will call this simple line "the bare propagator"). For a theory of interacting fields the green function is only known as a perturbation series, which is represented by the sum of all two-legged Feynman diagrams (we will represent this sum by a two-legged diagram with a shaded circle and call it "the dressed propagator"). Now if the assumption of the naive approach is correct, then the dressed propagator should be equal to the bare one (naively an isolated particle is just a free particle, i.e. the propagation of an isolated particle in the full theory should be the same as in the "free part" of this theory). Our check of the naive approach is therefore a check of the following equation

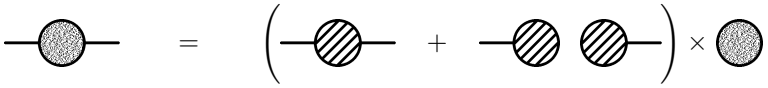


Before calculating any of the Feynman diagrams it is useful to elucidate the structure of the dressed propagator (this analysis will turn out to be very useful not only here, but even later when discussing renormalization of QFT). The first useful step is to find out whether the green function $g(x,y)$ is – just like the $g_0(x,y)$ – a function of only one variable, namely $x-y$. The answer seems to be trivially affirmative due to the translational symmetry, but some care is advisable. As to the spatial translations everything is OK, the invariance is guaranteed and the green function is indeed a function of $\vec{x}-\vec{y}$ rather than $\vec{x}$ and $\vec{y}$.[32] But as to the temporal translation, the question is whether our vacuum state is or is not an eigenstate of the full Hamiltonian. If $H|0\rangle = E_0|0\rangle$, then the vacuum is time-translation invariant (up to a phase) and the green function depends on $x-y$.[33] If, on the other hand, the vacuum is not an eigenstate of the full Hamiltonian, the green function is not guaranteed to enjoy temporal translation invariance and it may depend separately on x_0 and y_0.

[32] Using the explicit form of the translations for the field operators $\varphi_H(z) = e^{-iz.P}\varphi_H(0)e^{iz.P}$ and the definition of the in and out states ($|0\rangle_{\text{in}} = e^{-iTH}|0\rangle$ and ${}_{\text{out}}\langle 0| = \langle 0|\, e^{-iTH}$) one obtains for the two-point green function $g(x,y) = \lim_{T\to\infty}\vartheta(x_0-y_0)\,\langle 0|\, e^{-iTH}e^{-ix.P}\varphi_H(0)e^{i(x-y).P}\varphi_H(0)e^{iy.P}e^{-iTH}\,|0\rangle + x\leftrightarrow y$. Now one exploits the commutation relations $[H,P_\mu]=0$ and spatial translational invariance of the vacuum $P_i|0\rangle = 0$ to get $g(x,y) = \lim_{T\to\infty}\vartheta(x_0-y_0)\,\langle 0|\, e^{-i(T+x_0)H}\varphi_H(0)e^{i(x-y).P}\varphi_H(0)e^{-i(T-y_0)H}\,|0\rangle + x\leftrightarrow y$. So as to the spatial components, the green function $g(x,y)$ is indeed a function of $\vec{x}-\vec{y}$.

[33] For Hamiltonians in the form of a normal product with at least one annihilation operator (say, the Hamiltonian of particles with pair interactions) one clearly has $H|0\rangle = 0$. In such a case the two-point green function becomes $g(x,y) = \lim_{T\to\infty}\vartheta(x_0-y_0)\,\langle 0|\,\varphi_H(0)e^{i(x-y).P}\varphi_H(0)\,|0\rangle + x\leftrightarrow y$. This, however, is not guaranteed in general.

The next useful step in the preliminary analysis of the green function is to distinguish between connected and disconnected parts of diagrams. If we denote the sum of all connected parts by a diagonally hatched circle, then one gets[34]

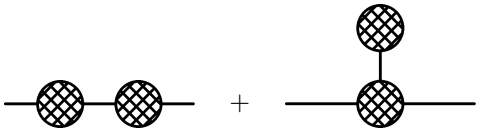


In this graphical equation, the no-leg diagram (the vacuum bubble) corresponds to the zero-point green function $\lim_{T\to\infty} \langle 0| U(T,-T) |0\rangle$ which is a constant (it simply does not have anything to depend on). The one-legged diagram (called tadpole) multiplied by the vacuum bubble corresponds to the one-point green function $g(z) = \lim_{T\to\infty} \langle 0| T\{U(T,-T)\varphi_I(z)\} |0\rangle$ which, due to the spatial translational symmetry[35] is just a function of z_0. If our vacuum was an eigenstate of the full Hamiltonian, then $g(z)$ would not depend even on z_0 and would be simply a constant. And since there is no constant term in the bare propagator, the tadpoles should better vanish in such a case. We did not rule out the possibility of non-vanishing tadpoles or the possibility of cancelation of the inappropriate contribution of the tadpoles with some contribution of the connected part of the dressed propagator, nevertheless it seems plausible to investigate at first the case of vanishing tadpoles and to comment on possible non-zero tadpoles afterwards. Our next task is therefore to understand the structure of the connected propagator assuming zero tadpole contributions.

The main tool for the analysis of the structure of the connected propagator is the so-called One Particle Irreducibile (OPI) diagram. It is a connected Feynman diagram, which cannot be divided into two disconnected ones by cutting just a single internal line. The sum of all OPI diagrams with given external legs is going to play a very important role in what follows, so we will use a specific symbol for it – the cross-hatched blob. So we have, on top of the full dressed propagator, three different graphical representations for three different sorts of propagators

the bare one the OPI one the connected one

and there is a very important relation between them. To reveal it, let us start investigating diagrams for which there exists precisely one internal leg, cutting of which would split the diagram into two disconnected parts. How does the sum of all such diagrams look like? First of all, the parts disconnected after the cut have to be OPI (were they not, there would be at least two lines, cutting of which would lead to splitting of the diagram). Once the cut is made, either each disconnected part contains one (original) external leg, or one disconnected part contains both external legs and the other one does not contain external legs. Consequently, the result must be



[34] In any two-leg diagram the external legs are either connected to each other or they are not. And any diagram comes, as always, multiplied by the sum of all diagrams with no external legs at all.

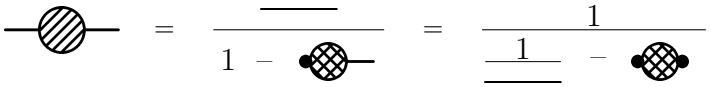
[35] The gymnastics goes exactly in the same steps as the one for the two-point green function above, leading to $g(z) = \lim_{T\to\infty} \langle 0| e^{-i(T+z_0)H} \varphi_H(0) e^{-i(T-z_0)H} |0\rangle$.

(if we assume vanishing tadpole diagrams, then the second part is missing).

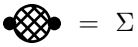
Now let us proceed to diagrams for which there exist precisely two internal legs, cutting of which would split the diagram into two disconnected parts. Using the arguments similar to those above, one obtains

Analogous results are obtained for diagrams with n internal legs which if cut would split the diagram. And since it is quite clear that the connected part of the dressed propagator is the sum of such classes of diagrams for $n = 0, \cdots, \infty$ (plus the bare propagator) one gets the announced important relation (without tadpole diagrams which are supposed to vanish)

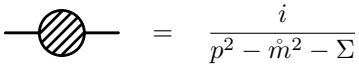
The RHS of the last equation is nothing but a geometric series, called *the Dyson series*. With the first term and the quotient  (the dot symbolizes a slot for the amputated leg) the summation is readily performed[36]



To move further it is convenient to consider the Fourier transforms of the connected and the bare propagators (i.e. the propagators in the p-representation). As to the bare one, we have seen that it is almost, even if not quite, equal to $\frac{i}{p^2-\mathring{m}^2}$. Using this expression and denoting[37]



one can finally write



This is as far as one can get without really calculating diagrams. The result looks encouraging. After some more or less plausible assumptions (vacuum as an eigenstate of the full Hamiltonian, vanishing tadpole diagrams, imperfect bare propagator) one obtained the structure of the dressed propagator very similar to the bare propagator. The last step is to demonstrate that $\Sigma = 0$ and we are done. But this last step is almost never achieved. The outcome of explicit Σ calculations (which we will perform later) is usually neither zero, nor small, as a rule it turns out to be huge. Our consistency check was just to pass, but it fails dramatically at the last step.

[36] The geometric series is convergent only for |quotient| < 1. Nevertheless, the function $\frac{1}{1-z}$ is an analytic function in the whole complex plane with an exception of a single point $z = 1$ (at which the function has a simple pole) and so it is the analytic continuation of the geometric series. For |quotient| ≥ 1 we will understand the sum as this analytic continuation.

[37] The quantity denoted by Σ is usually called self-energy. It is not a constant, it rather depends on the momenta of the (amputated) external legs. We will discuss this dependence later on.

To understand the relevance of the fail, as well as to find out a way of recovery, it is convenient (just for a moment and for purely didactic reasons) to consider a specific hypothetical situation of Σ being a constant. For a constant Σ we got the correct (free) propagator for an isolated particle, even if not with the expected mass. The squared mass of a freely propagating isolated particle did not come out to be $\mathring{m}^2$ but rather $m^2 = \mathring{m}^2 + \Sigma$.

What is a physical substance of the particle with the mass m? Is it a completely new particle present on top of the original paritcle or is it the original particle with the shifted mass? If it was a new particle present together with the original one, then one would expect the presence of the original particle to be manifest in the propagator. This, however, is not what we got – our assumption of a constant Σ lead us to the dressed propagator with just m and no sign of $\mathring{m}$ whatsoever. So the interpretation of the outcome in terms of the original particles with the shifted mass seems to be the only natural one.

The shift of the mass is in fact something to be expected. In relativistic theories the mass is related to the energy, so the shift in energies would pronounce itself also as a change of masses. What is surprising is the amount of the shift. If the interaction Hamiltonian is in a sense small, all the changes are anticipated to be small. In QFT, however, this expectation is not fulfilled.

The large shift in the mass imposes two questions. First, if an isolated particle propagates as something with the mass m (which we shall call the physical mass), then what is the meaning of the mass $\mathring{m}$ (which we shall call the bare mass)? And second, if the eigenstates of the free Hamiltonian H_0 are not the physical states of isolated particles, then how should one redefine the basic ingredients of the QFT (like the green functions or the s-matrix) so that they correspond to transition amplitudes for real processes with physical particles? The answer to the first question is called renormalization. The answer to the second question is nothing else but replacing the naive approach by the standard one.

The correct interpretation of the $\mathring{m}$ parameter is that it is just a parameter of the theory. One could naively consider it to be a mass of particles described by the theory, but such an interpretation turned out to be misleading. The process of pinning down the value of $\mathring{m}$ so that one obtains the right experimental value of the physical mass m is called mass renormalization. In our simple example it is straightforward (after the Σ is calculated, of course). One simply puts $\mathring{m} = \sqrt{m^2 - \Sigma}$. In general, the renormalization procedure is much more involved. We will learn about it later on.

If the real physical particles are those of the mass m, then the transition amplitudes needed for comparison with experimental data are, of course, the ones between the states of these real particles. In the case of the scattering this means that the proper transition amplitudes in the Heisenberg picture are

$$_{\text{out}}\langle \boldsymbol{p}_1, \ldots, \boldsymbol{p}_m | \boldsymbol{p}_{m+1}, \ldots, \boldsymbol{p}_n \rangle_{\text{in}}$$

where we have used boldface notation for momenta to distinguish these multiparticle states of the real physical particles with the mass m from the states of the original free particles with the mass $\mathring{m}$. This distinction is to be crucial for the standard approach.[38]

[38] Note also that we have used 4-vectors rather than 3-vectors for denoting particle states. The point is that for the same 3-momentum the states of particles with masses $\mathring{m}$ and m differ just it the (dependent) zeroth component of the 4-momentum. It is therefore reasonable to denote them $|p\rangle$ and $|\boldsymbol{p}\rangle$ respectively (rather than $|\vec{p}\rangle$ and $|\vec{\boldsymbol{p}}\rangle$).

The transition amplitude between states of the physical particles with the mass m, however, cannot be calculated with the Feynman diagrams developed so far. The point is that we do not know explicitly the physical states $|\boldsymbol{p}\rangle$ which are quite different from $|p\rangle = \sqrt{2\omega_{\vec{p}}}\, a^+_{\vec{p}}\, |0\rangle$. In other words, we do not know how to express $|\boldsymbol{p}\rangle$ in terms of $|p\rangle$. So after rewriting the transition amplitude in the interaction picture

$$\lim_{T\to\infty} {}_{\text{out},I}\langle \boldsymbol{p}_1, \ldots, \boldsymbol{p}_m |\, U_I\,(T,-T)\, |\boldsymbol{p}_{m+1}, \ldots, \boldsymbol{p}_n\rangle_{I,\text{in}}$$

we do not know how to write this as the vacuum expectation value of T-products of fields and creation/annihilation operators. But this was the crucial step leading to the Wick's theorem and the Feynman diagrams machinery. Without these tools we are stuck.

Our simplified assumptions of vanishing tadpoles and constant Σ forced us to realize that we need to rethink the derivation of the Feynman rules. The necessity of this rethinking, however, is not based on the said assumptions. Indeed, even without them the fail of the consistency check, when combined with a notoriously known experimental fact, leads us inevitably to the notion of physical particles with the mass different form the $\mathring{m}$.

The point is that we know one thing about isolated particles for sure: they behave like free particles. This is a simple and well known experimental fact. As a consequence, any relativistic quantum theory which has an ambition to describe the real world (or a toy world in the case of toy models), must provide this behavior for isolated particles (i.e. particles far away from each other as well as from anything else).

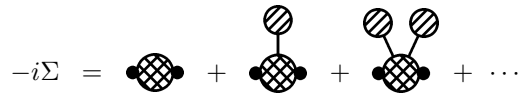
The naive approach assumed that isolated particles are described by the free part of the full Hamiltonian. This, however, turned out not to be true (by the fail of the consistency check). Isolated particles do not propagate as a free particles with the mass $\mathring{m}$. But they still behave like free ones, even if with different mass. Their dynamics is not dictated by the free part of the full Hamiltonian, but rather by some other free Hamiltonian.

The standard approach is based on identification of the role of this "other free Hamiltonian", on subsequent reformulation of definitions of the important transition amplitudes (the S-matrix elements and the Green functions) and finally on reformulating the Feynman rules for these redefined transition amplitudes. It will turn out that the new Feynman rules are identical to the Feynman rules of the naive approach with three welcome additions:

1. $i\varepsilon$ in the denominator of the propagator (in the p-representation)
2. no loops on external legs (and a well-defined factor $\sqrt{Z}$ instead of them)
3. no vacuum bubbles

These are precisely the only three missing ingredients of the naive Feynman rules (when compared to the first Introduction, i.e. Conclusions). So once we understand these three points we can assert that we have derived and understood the Feynman rules presented at the beginning of these lecture notes.

Remark: *So far, our discussion was based on several simplifying assumptions, e.g vanishing tadpoles or constant Σ. None of these assumptions was crucial for the main conclusion and they can be relaxed easily. We can take the sum of the tadpole diagrams to be non-zero. The non-vanishing tadpoles would modify the Dyson series (by additional terms), but it is easy to convince oneself that the sum is unchanged, provided one redefines the self energy to be*



This self energy is a function of external legs momenta (there are two external legs in the game, but due to the momentum δ-function in vertices their momenta have to be equal, so one has just one momentum). If one now denotes the contribution of the tadpole diagrams by t and the contribution of the vacuum bubbles by v, then the consistency check (equality of the dressed and the bare propagators) rewritten from pictures to expressions would require

$$\left(\frac{i}{p^2-\mathring{m}^2-\Sigma(p)}+t^2\right)\times v=\frac{i}{p^2-\mathring{m}^2}$$

However, explicit calculations of Σ, t and v never support this claim.

Remark: *We defined the mass renormalization as finding out the value of the unknown bare mass $\mathring{m}$ from the experimentally known physical mass m. In principle, however, four different combinations are possible*

	$\mathring{m}$	m
1.	*unknown*	*known*
2.	*known*	*unknown*
3.	*known*	*known*
4.	*unknown*	*unknown*

1. The first possibility is typical in QFT and our definition of the mass renormalization applies to this very case. Let us stress here that the basic purpose of renormalization is not to deal with infinities occurring in QFT calculations. Renormalization is just an ordinary fit of theoretical parameters by experiment and it would be necessary even if there were no infinities in the game.
2. The second possibility shows up, in a sense, in the solid state physics. For an electron within a crystal, the bare mass is given by the well-known electron mass (in the vacuum), while the physical mass is an effective electron mass within the crystal. Notice, that the table value $0.51MeV$ which has played a role of the physical mass m in the previous point, plays the role of the bare mass $\mathring{m}$ here. Note also that if the effective electron mass is experimentally accessible in some (even if indirect) way, then this case is actually an example of the third possibility.
3. The third possibility occurs almost, but not quite, in QED of pions. Pions come as neutral as well as charged particles. Were they true elementary particles blind to all interactions except of electromagnetic ones, one should have $\mathring{m}=m_{\pi^0}$ and $m=m_{\pi^\pm}$. However, pions are composite particles consisting of quarks and gluons, and they feel all the basic interactions. Due to this fact one knows experimentally the bare mass of neither $\pi^\pm$ nor π^0.
4. The fourth possibility is typical for QCD. The problem with QCD is that there are no free quarks and we therefore have no direct experimental information about their physical masses. What is understood by mass renormalization in this case evidently needs some further discussion.

3.2.2 Double role of the free Hamiltonian and the $i\epsilon$ trick

One thing we know about particles is this: if far enough from everything else, every particle behaves like a free one. This is a fundamental experimental fact and every realistic particle theory should be in accord with it. The naive approach to the QFT was based on three assumptions:

1. Dynamics of particles is given by a Hamiltonian which is the sum of two parts – the free Hamiltonian and the interaction Hamiltonian.
2. Dynamics of isolated particles (far away from each other) is given by the free Hamiltonian.
3. The free Hamiltonian in the point 2 is the same as the free Hamiltonian in the point 1.

The fail of our consistency check force us to relax at least one of these assumptions. The most natural way of doing this is to relax the point 3 and to replace it by the new assumption:

3'. The free Hamiltonian in the point 2 is the same as the complete Hamiltonian in the point 1.

The replacement of the assumption 3 by the assumption 3' is the crucial step in advancement from the naive approach to the standard one. It does not mean that the full Hamiltonian is equal to some free Hamiltonian in the whole Hilbert space. It only means that they coincide in some region of the Hilbert space spanned by states of particles far away from each other. This situation is typical for effective theories – for some phenomena a full theory is equivalent to (and can be replaced by) another theory, usually a much simpler one.

So in the standard approach we have the full Hamiltonian (expressed in terms of $a^+_{\vec{p}}$ and $a_{\vec{p}}$)

$$H = H_0 + H' \qquad H_0 = \int \frac{d^3p}{(2\pi)^3}\, \omega_{\vec{p}}\, a^+_{\vec{p}}\, a_{\vec{p}} \qquad \omega_{\vec{p}} = \sqrt{\vec{p}^{\,2} + \mathring{m}^2}$$

and the effective Hamiltonian (expressed in terms of effective operators $\boldsymbol{a}^+_{\vec{p}}$ and $\boldsymbol{a}_{\vec{p}}$)

$$H \approx H^{\text{eff}} \qquad H^{\text{eff}} = \int \frac{d^3p}{(2\pi)^3}\, \boldsymbol{\omega}_{\vec{p}}\, \boldsymbol{a}^+_{\vec{p}}\, \boldsymbol{a}_{\vec{p}} + \boldsymbol{E}_\Omega \qquad \boldsymbol{\omega}_{\vec{p}} = \sqrt{\vec{p}^{\,2} + m^2}$$

where the symbol $\approx$ stands for equality within some region of the Hilbert space and $\boldsymbol{E}_\Omega$ has to account for possible non-zero energy of the physical vacuum, which is defined as the lowest energy eigenstate of the full Hamiltonian. This state, which we shall denote by the symbol $|\Omega\rangle$, is supposed to be also the ground state (with the same energy) of the effective Hamiltonian (there is usually no proof of this assertion, but it is nevertheless commonly assumed to be true). So the physical vacuum of the standard approach is the state $|\Omega\rangle$ satisfying

$$\begin{aligned} H|\Omega\rangle &= \boldsymbol{E}_\Omega|\Omega\rangle \qquad \text{definition} \\ H^{\text{eff}}|\Omega\rangle &= \boldsymbol{E}_\Omega|\Omega\rangle \qquad \text{assumption} \end{aligned}$$

This state is furthermore supposed to be annihilated by all the effective annihilation operators

$$\boldsymbol{a}_{\vec{p}}\,|\Omega\rangle = 0$$

and therefore the Fock space for H^{eff} spanned the states obtained by subsequent application of the effective creation operators on the physical vacuum

$$\begin{aligned} &|\Omega\rangle \\ &|\boldsymbol{p}\,\rangle = \sqrt{2\boldsymbol{\omega}_{\vec{p}}}\,\boldsymbol{a}^+_{\vec{p}}\;|\Omega\rangle \\ &|\boldsymbol{p}, \boldsymbol{p}\,'\rangle = \sqrt{2\boldsymbol{\omega}_{\vec{p}}}\,\boldsymbol{a}^+_{\vec{p}}\,|\vec{\boldsymbol{p}}\,'\rangle \\ &\vdots \end{aligned}$$

Let us now consider a typical prediction of QFT, namely that of a non-stable particle life-time or a scattering cross-section. What are the initial states for a particle decay or for a scattering process? They are, just like in the naive approach, momentum eigenstates superpositions, which are well localized in both position and momentum space, far away from each other and, in the case of scattering, heading to the same small area of spacetime (where the actual scattering takes place). But in contrary to the naive approach, the momentum eigenstates are not the $|p\rangle$ ones but rather the $|\boldsymbol{p}\rangle$ ones. So for a particle decay the initial state should be a one-particle state well localized in both momentum and position

$$|i\rangle = \int d^3p\, f(\boldsymbol{p})\ |\boldsymbol{p}\rangle$$

with an appropriate function $f(\boldsymbol{p})$. The multi-particle initial state for a scattering process is

$$|i\rangle = \int d^3p_{m+1} \dots d^3p_n\ f_{m+1}(\boldsymbol{p}_{m+1}) \dots f_n(\boldsymbol{p}_n)\ \ |\boldsymbol{p}_{m+1}, \dots, \boldsymbol{p}_n\rangle$$

The same is true for the state in which we are interested in the final stage of the scattering, i.e.

$$|f\rangle = \int d^3p_1 \dots d^3p_m\ f_1(\boldsymbol{p}_1) \dots f_m(\boldsymbol{p}_m)\ \ |\boldsymbol{p}_1, \dots, \boldsymbol{p}_m\rangle$$

The genuine S-matrix of the standard approach is now defined as[39]

$$\begin{aligned}
S_{fi} &= \langle f|\, e^{-iH(t_f - t_i)}\, |i\rangle && \text{Schrödinger picture} \\
S_{fi} &= {}_{\mathrm{H,out}}\langle f|i\rangle_{\mathrm{H,in}} && \text{Heisenberg picture} \\
S_{fi} &= {}_{\mathrm{I,out}}\langle f|\, U_I(t_f, t_i)\, |i\rangle_{\mathrm{I,in}} && \text{interaction picture}
\end{aligned}$$

The problem with the S-matrix is that one cannot calculate it along the lines developed for the calculation of the naive approach s-matrix. The point is that the time evolution in the interaction picture (where the perturbative calculations are usually performed) is given in terms of the original creation/annihilation operators $a^+_{\vec{p}}$, $a_{\vec{p}}$ while the states are given in terms of the effective creation/annihilation operators $\boldsymbol{a}^+_{\vec{p}}$, $\boldsymbol{a}_{\vec{p}}$. And since we do not know explicit expressions of one set of these operators in terms of the others, we cannot use the Wick theorem to obtain the Feynman rules.

To proceed, one seemingly has to express the effective operators in terms of the original ones or vice versa. And this is by no means an easy task. So it is a kind of relief that there is a clever shortcut which enables us to avoid the explicit calculation. The shortcut is nothing else than considering a slightly imaginary time. Technically it renders down to a simple replacement $t \to t(1 - i\varepsilon)$. The effect of such replacement is best understood in the Schrödinger picture and then utilized in the interaction picture.

[39] As to the Green function in the standard approach, the natural guess for the generalization from the naive approach $g(x_1, \dots, x_n) = {}_{\mathrm{out}}\langle 0|T\{\varphi(x_1) \dots \varphi(x_n)\}|0\rangle_{\mathrm{in}}$ would be $G(x_1, \dots, x_n) = {}_{\mathrm{out}}\langle \Omega|T\{\boldsymbol{\varphi}(x_1) \dots \boldsymbol{\varphi}(x_n)\}|\Omega\rangle_{\mathrm{in}}$ where the effective field $\boldsymbol{\varphi}(x)$ is the standard superposition of the effective creation and annihilation operators $\boldsymbol{a}^+_{\vec{p}}$, $\boldsymbol{a}_{\vec{p}}$. This, however, is not the standard choice. The reason is that a different definition of the Green function will turn out to be much more useful.

the trick

The time evolution of states in the Schrödinger picture expressed in terms of the Hamiltonian normalized eigenstates $|\psi_i\rangle$ and eigenvalues E_i reads

$$|\psi(t)\rangle = e^{-iH(t-t_0)}\,|\psi(t_0)\rangle = \sum_n e^{-iE_n(t-t_0)}\,|\psi_n\rangle\,\langle\psi_n|\psi(t_0)\rangle$$

After the replacement $(t-t_0) \to (t-t_0)(1-i\varepsilon)$ the sum on the RHS is for large t dominated by the contribution with the lowest energy, i.e. it is given by $e^{-iE_0(t-t_0)}\,|\psi_0\rangle\,\langle\psi_0|\psi(t_0)\rangle$. This means that one can obtain the ground state of the system by evolving any state (with non-vanishing scalar product with the ground state) long enough in the slightly imaginary time

$$|\psi_0\rangle = \lim_{(t-t_0)\to\infty(1-i\varepsilon)} \frac{e^{-iH(t-t_0)}\,|\psi(t_0)\rangle}{e^{-iE_0(t-t_0)}\langle\psi_0|\psi(t_0)\rangle}$$

The first excited state can be obtained in a similar way, one just has to start with a state $|\psi(t_0)\rangle$ satisfying $\langle\psi_0|\psi(t_0)\rangle = 0$ and $\langle\psi_1|\psi(t_0)\rangle \neq 0$

$$|\psi_1\rangle = \lim_{(t-t_0)\to\infty(1-i\varepsilon)} \frac{e^{-iH(t-t_0)}\,|\psi(t_0)\rangle}{e^{-iE_1(t-t_0)}\langle\psi_1|\psi(t_0)\rangle}$$

Generalization to higher excited states is straightforward.

One can now use this trick to relate the in and out states in the standard and naive approaches. If e.g. $\langle\Omega|0\rangle \neq 0$, then

$$|\Omega\rangle = \lim_{T\to\infty(1-i\varepsilon)} \frac{1}{e^{-i\boldsymbol{E}_\Omega T}\langle\Omega|0\rangle}\, e^{-iHT}\,|0\rangle$$

And since $\langle\Omega|p\rangle = 0$ (for non-zero $\vec{p}$ this is the scalar product of two different eigenstates of the 3-momentum operator with different eigenvalues), a plausible assumption $\langle\boldsymbol{p}|p\rangle \neq 0$ leads to[40]

$$|\boldsymbol{p}\rangle = \lim_{T\to\infty(1-i\varepsilon)} \frac{(2\pi)^3\,2\boldsymbol{E}_{\vec{p}}}{e^{-i(\boldsymbol{E}_{\vec{p}}+\boldsymbol{E}_\Omega)T}\langle\boldsymbol{p}|p\rangle}\, e^{-iHT}\,|p\rangle$$

where the factor $(2\pi)^3\,2\boldsymbol{E}_{\vec{p}}$ (with $\boldsymbol{E}_{\vec{p}} = \boldsymbol{\omega}_{\vec{p}}$) comes from the relativistic normalization of one-particle states (which are not normalized to unity). Generalization to multi-particle states is again straightforward. And once we have the standard in- and out- states expressed in terms of the naive in- and out- states, we can express the standard S-matrix in terms of the naive s-matrix

$$S_{fi} = s_{fi}|_{T\to T(1-i\varepsilon)} \frac{1}{e^{-i\boldsymbol{E}_\Omega 2T}} \prod_{i=1}^{m} \frac{(2\pi)^3\,2\boldsymbol{E}_{\vec{p}_i}}{e^{-i\boldsymbol{E}_{\vec{p}_i}T}\langle p_i|\boldsymbol{p}_i\rangle} \prod_{j=m+1}^{n} \frac{(2\pi)^3\,2\boldsymbol{E}_{\vec{p}_j}}{e^{-i\boldsymbol{E}_{\vec{p}_j}T}\langle\boldsymbol{p}_j|p_j\rangle}$$

So to calculate the S-matrix one just needs to calculate the s-matrix, which can be done by use of the Feynman rules. The only catch is that one must learn how to do it in the "slightly imaginary time" and how to take care of the extra factors multiplying the s-matrix. But all this can be done in surprisingly easy way.

[40] We have assumed that one-particle state $|\boldsymbol{p}\rangle$ has the lowest energy among the states with the 3-momentum $\vec{p}$.

How should one perform the replacement $t \to t(1-i\varepsilon)$ in the Feynman rules? Well, the only place where time is present in the Feynman rules is the propagator in the x-representation

$$d_F(\xi) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\vec{p}}} \left(\vartheta(\xi^0)e^{-ip\xi} + \vartheta(-\xi^0)e^{ip\xi}\right)$$

and therefore the said replacement is to be done just in the naive propagator, which is promoted in this way to the standard Feynman propagator

$$D_F(\xi) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2\omega_{\vec{p}}} \left(\vartheta(\xi^0)e^{-ip\xi}e^{-p_0\xi_0\varepsilon} + \vartheta(-\xi^0)e^{ip\xi}e^{p_0\xi_0\varepsilon}\right)$$

The discussion of the naive propagator at the page 97 showed us that this can be written as[41]

$$D_F(\xi) = \int \frac{d^4p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\varepsilon} e^{-ip\xi}$$

so that the Fourier transform of the Feynman propagator (used in the p-representation rules) is

$$D_F(p) = \frac{i}{p^2 - m^2 + i\varepsilon}$$

The $i\varepsilon$ in the propagator is the p-representation incarnation of the slightly imaginary time trick.

diagram = diagram s oblečenými vonkajšími nohami
olečné nohy, čas T, akurát exponenciály

[41] $\frac{i}{p^2-m^2+i\varepsilon\omega_{\vec{p}}^2}$

Stav s najnižšou energiou môžeme dostať trikom

$$e^{-iHT}\left|\psi\right\rangle = \sum_n e^{-iE_nT}\left|\varphi_n\right\rangle\left\langle\varphi_n\right|\left.\psi\right\rangle$$

$$\left|\varphi_0\right\rangle = \lim_{T\to\infty(1-i\varepsilon)} \frac{e^{-iHT}\left|\psi\right\rangle}{e^{-i\boldsymbol{E}_0T}\left\langle\varphi_0\right|\left.\psi\right\rangle}$$

$$\left|\Omega\right\rangle_{\text{in}} = e^{-iHT}\left|\Omega\right\rangle = \lim_{T\to\infty(1-i\varepsilon)} \frac{e^{-iHT}\left|0\right\rangle}{e^{-i\boldsymbol{E}_\Omega T}\left\langle\Omega|0\right\rangle}$$

$$\left|\vec{\boldsymbol{p}}\right\rangle_{\text{in}} = e^{-iHT}\left|\vec{\boldsymbol{p}}\right\rangle = \lim_{T\to\infty(1-i\varepsilon)} \frac{(2\pi)^{3/2}\sqrt{2\boldsymbol{E}_{\vec{p}}}}{\left\langle\vec{\boldsymbol{p}}\right|\left.\vec{p}\right\rangle} \frac{e^{-iHT}}{e^{-i(\boldsymbol{E}_\Omega+\boldsymbol{E}_{\vec{p}})T}}\left|\vec{p}\right\rangle$$

$${}_{\text{out}}\!\left\langle\Omega\right| = \left\langle\Omega\right| e^{-iHT} = \lim_{T\to\infty(1-i\varepsilon)} \frac{\left\langle 0\right| e^{-iHT}}{e^{-i\boldsymbol{E}_\Omega T}\left\langle 0|\Omega\right\rangle}$$

$${}_{\text{out}}\!\left\langle\vec{\boldsymbol{p}}\right| = \left\langle\vec{\boldsymbol{p}}\right| e^{-iHT} = \lim_{T\to\infty(1-i\varepsilon)} \left\langle\vec{p}\right| \frac{e^{-iHT}}{e^{-i(\boldsymbol{E}_\Omega+\boldsymbol{E}_{\vec{p}})T}} \frac{(2\pi)^{3/2}\sqrt{2\boldsymbol{E}_{\vec{p}}}}{\left\langle\vec{p}\right|\left.\vec{\boldsymbol{p}}\right\rangle}$$

$$S_{fi} = {}_{\text{out}}\!\left\langle\vec{\boldsymbol{p}}_1,\ldots,\vec{\boldsymbol{p}}_m\right|\left.\vec{\boldsymbol{p}}_{m+1},\ldots,\vec{\boldsymbol{p}}_n\right\rangle_{\text{in}}$$

$$S_{fi} = \left.\frac{s_{fi}}{e^{-i\boldsymbol{E}2T}}\right|_{t\to t(1-i\epsilon)} \prod_{k=1}^{m} \frac{(2\pi)^{3/2}\sqrt{2\boldsymbol{E}_{\vec{p}_k}}}{\left\langle\vec{p}_k\right|\left.\vec{\boldsymbol{p}}_k\right\rangle} \prod_{l=m+1}^{n} \frac{(2\pi)^{3/2}\sqrt{2\boldsymbol{E}_{\vec{p}_l}}}{\left\langle\vec{\boldsymbol{p}}_l\right|\left.\vec{p}_l\right\rangle}$$

$$\boldsymbol{E} = \boldsymbol{E}_\Omega + \boldsymbol{E}_{\vec{p}_1} + \cdots + \boldsymbol{E}_{\vec{p}_m} = \boldsymbol{E}_\Omega + \boldsymbol{E}_{\vec{p}_{m+1}} + \cdots + \boldsymbol{E}_{\vec{p}_n}$$

- zmenu $t \to t(1-i\epsilon)$ zabezpečí i epsilon v menovateli propagátora
- vlnové balíky a krátkodosahovosť (exponenciálny pokles D(x) so vzdialenosťou) spôsobia, že okrem krátkeho času sa časový vývoj "deje na vonkajších nohách", časovej exponenciály máme teda šancu sa zbaviť odstránením vonkajších nôh
- čomu presne zodpovedá oblečená vonkajšia noha? Je to časový vývoj od -T (resp do T) do (resp. od) nejakého času, v ktorom sú už častice blízko pri sebe. Každá vonkajšia noha končí v nejakom vertexe, ktorý je už prepojený s aspoň jednou inou vonkajšou nohou a z toho vertexu je noha napojená na jedno $\varphi(x)$. Vchádzajúce a vychádzajúce oblečené nohy sú teda rovnaké, ako v amplitúdach

$${}_{\text{out}}\!\left\langle\Omega\right| T\left\{\varphi(x_{m+1})\ldots\varphi(x_n)\right\}\left|\vec{\boldsymbol{p}}_{m+1},\ldots,\vec{\boldsymbol{p}}_n\right\rangle_{\text{in}}$$

$${}_{\text{out}}\!\left\langle\vec{\boldsymbol{p}}_1,\ldots,\vec{\boldsymbol{p}}_m\right| T\left\{\varphi(x_1)\ldots\varphi(x_m)\right\}\left|\Omega\right\rangle_{\text{in}}$$

Prepíšem do interakčného obrazu a v časoch blízkych interakcii vložím jednotkový operátor z vl. stavov efektívneho hamiltoniánu. Dostanem presne príspevok od vonkajších nôh a

okrem toho príspevok od prechodu vákuum-vákuum (najprv od -T do τ okolo 0, potom od τ' okolo 0 do T), plus ešte

$$\langle\Omega|\, T\left\{\varphi(x_{m+1})\dots\varphi(x_n)\right\}|\vec{\boldsymbol{p}}_{m+1},\dots,\vec{\boldsymbol{p}}_n\,\rangle$$

$$\langle\vec{\boldsymbol{p}}_1,\dots,\vec{\boldsymbol{p}}_m\,|\,T\left\{\varphi(x_1)\dots\varphi(x_m)\right\}|\Omega\rangle$$

Čiže amputovaním oblečených vonkajších nôh (vydelením amplitúdy prvými dvomi výrazmi v tomto bode) sa zbavím všetkého, čo stálo pri s_{fi}, ale navyše dostanem v menovateli amplitúdu vákuum-vákuum (tej sa zbavím ľahko, stačí nepočítať vákuové bubliny) a členmi danými druhými dvomi výrazmi v tomto bode. Ak chcem mať povodný výraz, musím okrem amputácie oblačených vonkajších nôh a nepočítania vákuových bublín ešte násobiť celú vec týmito dvomi výrazmi.

- vlnové balíky:

$$\langle\Omega|\, T\left\{\varphi(x_{m+1})\dots\varphi(x_n)\right\}|\vec{\boldsymbol{p}}_{m+1},\dots,\vec{\boldsymbol{p}}_n\,\rangle = \langle\Omega|\,\varphi(x_{m+1})\,|\vec{\boldsymbol{p}}_{m+1}\,\rangle\dots\langle\Omega|\,\varphi(x_n)\,|\vec{\boldsymbol{p}}_n\,\rangle$$

$$\langle\vec{\boldsymbol{p}}_1,\dots,\vec{\boldsymbol{p}}_m\,|\,T\left\{\varphi(x_1)\dots\varphi(x_m)\right\}|\Omega\rangle = \langle\vec{\boldsymbol{p}}_1|\,\varphi(x_1)\,|\Omega\rangle\dots\langle\vec{\boldsymbol{p}}_m\,|\,\varphi(x_m)\,|\Omega\rangle$$

translácie:

$$\langle\Omega|\,\varphi(x)\,|\vec{\boldsymbol{p}}\rangle = \langle\Omega|\,e^{-ixP}\varphi(0)e^{ixP}\,|\vec{\boldsymbol{p}}\rangle = \langle\Omega|\,\varphi(0)\,|\vec{\boldsymbol{p}}\rangle\,e^{ix\boldsymbol{p}}$$

boosty:

$$\langle\Omega|\,\varphi(0)\,|\vec{\boldsymbol{p}}\rangle = \langle\Omega|\,U\,U^{-1}\varphi(0)U^{-1}\,U\,|\vec{\boldsymbol{p}}\rangle = \langle\Omega|\,\varphi(0)|\vec{\boldsymbol{0}}\,\rangle = \sqrt{Z}\times phasefactor$$

consistency check of the standard approach

Let us finish the formulation of the standard approach by performing the same consistency check which revealed the problems of the naive approach. To do so, one has to investigate the structure of the dressed propagator, defined as

$$G(x,y) = {}_{\text{out}}\langle\Omega|\, T\{\varphi_H(x)\varphi_H(y)\}\, |\Omega\rangle_{\text{in}}$$

The aim is to show that it corresponds to the propagation of a free particle with a mass m (in the naive approach the mass ought to have been $\mathring{m}$). One advances precisely as in the naive case (see the beginning of this chapter). It turns out that individual steps are even easier now, e.g. the same translation invariance arguments which lead to the dependence of the naive $g(x,y)$ on x_0, y_0 and $\vec{x}-\vec{y}$ now lead to the dependence of the standard $G(x,y)$ on just $x-y$ (because the physical vacuum $|\Omega\rangle$ is, unlike the perturbative vacuum $|0\rangle$, an eigenstate of the full Hamiltonian). Another simplification is that there is no need for discussion of the vacuum bubbles (there are none in the results of the standard approach). And the sum of the tadpole diagrams is, again due to the translational invariance of the physical vacuum, a constant (namely $\langle\Omega|\, \varphi_H(0)\, |\Omega\rangle$.

Let us remark that the case of the real scalar field, corresponding to spin-less neutral particles, is the only case in which one can obtain a non-vanishing vacuum expectation value $\langle\Omega|\, \varphi_H(0)\, |\Omega\rangle$. The reason is a corollary[42] of the Schur's lemma stating that this matrix element vanishes unless $\varphi(0)$ shares the (assumed) trivial transformation properties of the physical vacuum.[43] This requires vanishing spin (rotations) as well as vanishing any quantum number like electric charge (internal symmetries). Therefore the sum of tadpole diagrams can be non-zero only for spin-less neutral particles. And even for these particles the vacuum expectation value of the corresponding field can be made zero by simple redefinition of the field $\varphi_H(x) \to \varphi_H(x) - \langle\Omega|\, \varphi_H(0)\, |\Omega\rangle$. As a consequence, it is usually possible to ignore the tadpole diagrams in QFT calculations.

For vanishing tadpoles one obtains (by following the same steps as for the naive approach)

$$G_2(p) = \frac{i}{p^2 - \mathring{m}^2 + i\varepsilon - \Sigma\left(p^2; \mathring{m}, \mathring{g}\right)}$$

where the dependence of the self-energy on the momentum squared[44] and parameters of the theory ($\mathring{m}$, $\mathring{g}$) was presented explicitly. Now for the consistency check to go through two conditions are sufficient to fulfill. First, the propagator should have precisely one pole in the p^2-variable. And second, it should decrease fast enough with p_0 going to infinity (so that one can use the trick with integration in the complex plane). Under these circumstances the dressed propagator would correspond to propagation of particle with the mass given by the position of the pole, which is exactly what is wanted. The question is, if these assumptions are supported by explicit calculations. The answer is affirmative.

[42] Statement (corollary of the Schur's lemma): Let the states $|\alpha\rangle$, $|\beta\rangle$ and the operators A_i transform according to irreducible representations R_α, R_β and R_A (of the group G) respectively. Let us decompose the direct product $R_\beta \otimes R_A$ to irreducible representation $R_\beta \otimes R_A = \oplus_\gamma R_\gamma$. If R_α is not present among R_γ, then $\langle\alpha|\, A_i\, |\beta\rangle = 0$.

[43] If the vacuum was not invariant with respect to the Lorentz transformations, one could measure which vacuum is present in one's inertial frame and to distinguish between different inertial frames by the result of this measurement. This, however, should not be possible in a relativistic theory.

[44] Both $G_2(p)$ and Σ depend on the single variable p^2 (and the parameters $\mathring{m}$, $\mathring{g}$). First of all, they depend on momenta of their external legs. Second, because of momentum δ-functions in vertices, the momenta of these external legs are equal to each other. And finally, both G_2 and Σ are Lorentz scalars (Feynman diagrams are scalars, G_2 is a sum of diagrams, Σ is a sum of diagrams with amputated legs, hence they are functions of the only scalar which can be build out of the vector p.

the spectral representation of the dressed propagator

Question: How to find Z? Answer: closer look on the structure of the dressed propagator The main trick is just an insertion of the unit operator, in a clever (and notoriously known) form, at an appropriate place. The clever form is

$$1 = |\Omega\rangle\langle\Omega| + \sum_N \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{N\vec{p}}}|N,\vec{p}\rangle\langle N,\vec{p}|$$

where $(2E_{N\vec{p}})^{-1/2}|N,\vec{p}\rangle$ are *the full Hamiltonian* eigenstates (an orthonormal basis). Recall that since the Hamiltonian commutes with 3-momentum, these eigenstates can be labelled by the overall 3-momentum (plus additional quantum number N). Moreover, the eigenstates with different $\vec{p}$ are related by a Lorentz boost and the eigenvalues satisfy $E_{N\vec{p}} = \sqrt{m_N^2 + \vec{p}^2}$, where m_N is the so-called rest energy, i.e. the energy of the state with the zero 3-momentum. Writing the ground state contribution $|\Omega\rangle\langle\Omega|$ outside the sum reflects the natural *assumption* that the physical vacuum is Lorentz invariant, i.e. the boosted vacuum is equal to the original one.

The appropriate place for the insertion of this unit operator in the dressed propagator is (not surprisingly) between the two fields. If the time ordering is ignored for a while, one gets immediately

$$\langle\Omega|\varphi_H(x)\varphi_H(y)|\Omega\rangle = |\langle\Omega|\varphi_H(0)|\Omega\rangle|^2 + \sum_N \int \frac{d^3p}{(2\pi)^3}\frac{e^{-ip(x-y)}}{2E_{N\vec{p}}}|\langle\Omega|\varphi_H(0)|N,\vec{p}\rangle|^2$$

Now one inserts another unit operator written as $U_{\vec{p}}^{-1}U_{\vec{p}}$, where $U_{\vec{p}}$ is the representation of the Lorentz boost transforming $\vec{p}$ to $\vec{0}$. Making use of the Lorentz invariance of both the nonperturbative vacuum $\langle\Omega|U_{\vec{p}}^{-1} = \langle\Omega|$ and the scalar field $U_{\vec{p}}\varphi_H(0)U_{\vec{p}}^{-1} = \varphi_H(0)$, one can get rid of the explicit $\vec{p}$-dependence in the matrix element[45]

$$\langle\Omega|\varphi_H(0)|N,\vec{p}\rangle = \langle\Omega|U_{\vec{p}}^{-1}U_{\vec{p}}\varphi_H(0)U_{\vec{p}}^{-1}U_{\vec{p}}|N,\vec{p}\rangle = \langle\Omega|\varphi_H(0)|N,\vec{0}\rangle$$

With the time ordering reinstalled, one can therefore write

$$\langle\Omega|T\{\varphi_H(x)\varphi_H(y)|\Omega\rangle = v_\varphi^2 + \sum_N \int \frac{d^4p}{(2\pi)^4}\frac{ie^{-ip(x-y)}}{p^2 - m_N^2 + i\varepsilon}\left|\langle\Omega|\varphi_H(0)|N,\vec{0}\rangle\right|^2$$

with $v_\varphi = |\langle\Omega|\varphi_H(0)|\Omega\rangle|$ (this constant usually vanishes, so we will drop it from now on). The dressed propagator can be written as a superposition of the bare propagators corresponding to masses m_N, with the weight given by $\sum_N |\langle\Omega|\varphi_H(0)|N,\vec{0}\rangle|^2$. This is usually written in the form (the Källén–Lehmann spectral representation)

$$G(x,y) = \int_0^\infty \frac{dM^2}{2\pi}G_0(x,y;M^2)\rho\left(M^2\right)$$

where the so-called spectral density function $\rho\left(\mu^2\right)$ is defined as

$$\rho\left(M^2\right) = \sum_N 2\pi\delta\left(M^2 - m_N^2\right)\left|\langle\Omega|\varphi_H(0)|N,\vec{0}\rangle\right|^2$$

[45] For higher spins one should take into account nontrivial transformation properties of the field components. But the main achievement, which is that one gets rid of the explicit 3-momentum dependence, remains unchanged.

Now what was the purpose of all these formal manipulations? Is there anything useful one can learn from the spectral representation of the dressed propagator? Frankly, without further assumptions about the spectral function, not that much. But once we adopt a set of plausible assumptions, the spectral representation will tell us quite some deal about the structure of the dressed propagator. It will shed some new light on the structure of $G(x,y)$ anticipated within the first look on the dressed propagator, which is not very important, but nevertheless pleasant. On top of that, it will enable us to make some nontrivial statements about the analytic properties of the dressed propagator. This again is not that important, although interesting. The real point is that virtually the same reasoning will be used afterwards in the case of the n-point Green function, and then all the work will really pay off.

The usefulness of the spectral representation of the dressed propagator is based on the following assumptions:
1. the spectrum of the Hamiltonian at $\vec{p}=0$ contains discrete and continuous parts, corresponding to one-particle and multi-particle states respectively[46]
2. the lowest energy eigenstate (its energy will be denoted by m) with the nonvanishing matrix element $\langle\Omega|\varphi_H(0)|N,\vec{0}\rangle$ is a 1-particle one[47]
3. the minimum energy of the continuous spectrum is somewhere around $2m$, and if there are any other 1-particle states with nonvanishing $\langle\Omega|\varphi_H(0)|N,\vec{0}\rangle$, their energy is also around or above this threshold[48]

Within these assumptions one gets

$$G(x,y) = Z\, G_0(x,y;m^2) + \int_{\sim 4m^2}^{\infty} \frac{dM^2}{2\pi} G_0(x,y;M^2)\rho\left(M^2\right)$$

where $Z = \left|\langle\Omega|\varphi_H(0)|N,\vec{0}\rangle\right|^2$. In the p-representation this corresponds to

$$\widetilde{G}(p) = \frac{iZ}{p^2-m^2+i\epsilon} + \int_{\sim 4m^2}^{\infty} \frac{dM^2}{2\pi} \frac{i\rho\left(M^2\right)}{p^2-M^2+i\epsilon}$$

As to the analycity, the third term is clearly analytic (one can differentiate with respect to the complex p^2) everywhere except of the straight line from $\sim 4m^2$-$i\varepsilon$ to ∞-$i\varepsilon$. On this line, the behavior of the integral depends on the structure of $\rho\left(M^2\right)$. If there are any additional 1-particle states with nonvanishing $\langle\Omega|\mathring{\varphi}(0)|N,\vec{0}\rangle$, their contribution to $\rho\left(M^2\right)$ is proportional to the δ-function, leading to additional poles in $\widetilde{G}^{\circ}(p)$. Otherwise $\rho\left(M^2\right)$ is supposed to be a decent function, in which case $\widetilde{G}^{\circ}(p)$ has a branch cut along the line[49].

[46] A multiparticle state can be characterized by an overall 3-momentum (which is zero in our case), relative 3-momenta (which are the continuous parameters not present in single-particle states) and discrete quantum numbers like mass, spin, charges etc. It is natural to assume that energy of a multiparticle state is a continuous function of the relative 3-momenta. And since for one-particle states there is no known continuous characteristics, except of 3-momentum which is here fixed to be zero, there is no room for continuous change of energy in this case. Note that adopting this philosophy, bound states of several particles are considered one-particle states. This may seem as a contradiction, but is not, it is just a matter of terminology.

[47] This assumption is not of the vital importance, and it can be relaxed easily. One may view it as a typical case and stay open-minded for more exotic possibilities (e.g. no 1-particle state with non-vanishing $\langle\Omega|\mathring{\varphi}(0)|N,\vec{0}\rangle$, which may occur in theories with confinement).

[48] This again is not very important and can be relaxed if needed. On the other hand, it is quite natural assumption, reflecting the intuitive expectation of two-particle states having mass approximately twice as big as one particle. While this is something to be expected if the interaction is small enough for perturbation theory to be useful, it cannot be taken for granted in every circumstances. Also here one better stays open-minded.

[49] Proof: The imaginary part of the integral exhibits a discontinuity along this line, which follows directly from the famous relation $\lim_{\varepsilon\to 0}\int dx \frac{f(x)}{x\pm i\varepsilon} = \mp i\pi f(0) + v.p.\int dx \frac{f(x)}{x}$

Chapter 4

Renormalization

4.1 Renormalization without infinities

4.1.1 Renormalization conditions

Let us denote the real part of the pole position of the dressed propagator as m^2, i.e. let us assume that the dressed propagator has a pole (in the variable p^2) sitting at $p^2 = m^2 + i\varepsilon$ The value of m is given by the equation

$$m^2 - \mathring{m}^2 - \Sigma(m^2; \mathring{m}, \mathring{g}) = 0$$

This equation will serve us as one of renormalization conditions.

Let us expand $\Sigma(p^2; \mathring{m}, \mathring{g})$ in the variable p^2 around m^2

$$\Sigma(p^2; \mathring{m}, \mathring{g}) = \Sigma(m^2; \mathring{m}, \mathring{g}) + \Sigma'(m^2; \mathring{m}, \mathring{g}) \left(p^2 - m^2\right) + o\left(p^2 - m^2\right)$$

where $\Sigma'(p^2; \mathring{m}, \mathring{g}) \equiv \frac{\partial}{\partial p^2}\Sigma(p^2; \mathring{m}, \mathring{g})$. The dressed propagator becomes

$$G_2(p^2) = \frac{1}{1 - \Sigma'(m^2; \mathring{m}, \mathring{g})} \frac{i}{p^2 - m^2 + i\varepsilon} \left(1 + \mathcal{O}\left(p^2 - m^2\right)\right)$$

where we have already use $m^2 = \mathring{m}^2 + \Sigma(m^2; \mathring{m}, \mathring{g})$. What we got is the second renormalization condition

$$Z^{-1} = 1 - \Sigma'(m^2; \mathring{m}, \mathring{g})$$

renormalization conditions:

$$\begin{aligned} m_i^2 &= \mathring{m}_i^2 + \Sigma_i(m^2; \mathring{m}_1, \mathring{m}_2, \ldots, \mathring{g}, \ldots) \\ Z_i^{-1} &= 1 - \Sigma_i'(m^2; \mathring{m}_1, \mathring{m}_2, \ldots, \mathring{g}, \ldots) \\ g_{n,R}\gamma_n(R) &= \mathring{g}_n \, \Gamma_n(R; \mathring{m}_1, \mathring{m}_2, \ldots, \mathring{g}, \ldots) \prod_{j=1}^{n} \sqrt{Z_j} \end{aligned}$$

where $|g_R|$ is from experiment: $|M_{fi}| = |g_R\gamma(R)|$.

4.1.2 Loop expansion

Renormalization conditions are usually solved by expanding relevant quantities in powers of an appropriate parameter. One can use expansion in powers of the coupling constant, but since this constant is one of the unknowns, this expansion does not look like the most suitable one. We will therefore stick to the loop expansion, since it gives no opportunity to mix the expansion parameter with other quantities. The loop expansion, however, is defined so far only in diagrammatical terms (by the number of independent loops in diagrams). Our first task is therefore to reformulate it as a power expansion. Fortunately, this is easy to achieve.

Let us consider a rescaled Lagrangian

$$\mathcal{L}(\lambda) = \lambda^{-1}\mathcal{L}$$

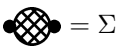
and find out how do the vertices, propagators and diagrams scale with λ. The vertices are given by the derivatives of the Lagrangian, the propagators are proportional to inverse vertices (two-legged ones) and the diagrams consist of the vertices and the propagators, i.e.

$$\begin{aligned} \text{vertex} &\to \lambda^{-1}\,\text{vertex} \\ \text{propagator} &\to \lambda\,\text{propagator} \\ \text{diagram} &\to \lambda^{I-V}\,\text{diagram} \end{aligned}$$

where I and V are numbers of internal lines and vertices respectively. Now according to Euler's theorem $I - V = L - 1$ (where L is a number of independent loops) one has

$$\text{diagram} \to \lambda^{L-1}\,\text{diagram}$$

i.e. λ times L-loop diagram scales like λ^L. In other words, the expansion in powers of λ is nothing else than the loop expansion.[1]

Exercise: *Rederive the renormalization conditions for* $\mathcal{L}(\lambda) = \lambda^{-1}\mathcal{L}$. *Show that they remain unchanged if one defines* $i\lambda$  $= \Sigma$ *and* $i\lambda$  $= \mathring{g}\,\Gamma$.

The loop expansions of the self-energy and the dressed vertex can be now written as

$$\Sigma(p^2; \mathring{m}, \mathring{g}) = \sum_{n=1}^{\infty} \lambda^n \Sigma_n(p^2; \mathring{m}, \mathring{g})$$

$$\Gamma(P; \mathring{m}, \mathring{g}) = \sum_{n=0}^{\infty} \lambda^n \Gamma_n(P; \mathring{m}, \mathring{g})$$

where Σ_n and Γ_n are sums of the corresponding n-loop diagrams.

Exercise: *Show that for the* φ^3*-theory:* $-i\Sigma_1 =$  $-i\Sigma_2 =$ 

$-i\Sigma_3 =$  *+ another five diagrams*

[1] Let us remark that in the path integral formulation of QFT the action (the integral of a Lagrangian density) appears divided by the Planck constant $\hbar$. So the role of our formal parameter λ can be played by the physical parameter $\hbar$. This possibility is masked in units where $\hbar = 1$, nevertheless it should be clear that the loop expansion can be understood as the expansion in powers of $\hbar$.

The loop expansion of the renormalization conditions for the φ^n-theory reads

$$\begin{aligned} m^2 &= \mathring{m}^2 + \sum_{n=1}^{\infty} \lambda^n \Sigma_n(m^2; \mathring{m}, \mathring{g}) \\ 1 &= Z\left(1 - \sum_{n=1}^{\infty} \lambda^n \Sigma'_n(m^2; \mathring{m}, \mathring{g})\right) \\ g_R &= \mathring{g}\ Z^{n/2}\ \sum_{n=0}^{\infty} \lambda^n \Gamma_n(R; \mathring{m}, \mathring{g}) \end{aligned}$$

where in the general case there should be $g_R\gamma(R)$ instead of g_R on the LHS of the third equation. The equations depend on the parameter λ, and so do the solutions. Expanding $\mathring{m}(\lambda)$, $Z(\lambda)$ and $\mathring{g}(\lambda)$ in powers of λ

$$\begin{aligned} \mathring{m}(\lambda) &= \sum_{n=0}^{\infty} \lambda^n \mathring{m}_n \\ Z(\lambda) &= \sum_{n=0}^{\infty} \lambda^n Z_n \\ \mathring{g}(\lambda) &= \sum_{n=0}^{\infty} \lambda^n \mathring{g}_n \end{aligned}$$

directly solves the equations order by order. At the zeroth order the result is

$$\begin{aligned} \mathring{m}_0 &= m \\ Z_0 &= 1 \\ \mathring{g}_0 &= g_R \end{aligned}$$

which means no renormalization at the tree level (as expected). At the first order in λ one gets (after realizing independence of m and g_R on λ)

$$\begin{aligned} \mathring{m}_1 &= -\frac{1}{2\mathring{m}_0}\Sigma_1(m^2; \mathring{m}_0, \mathring{g}_0) \\ Z_1 &= \Sigma'_1(m^2; \mathring{m}_0, \mathring{g}_0) \\ \mathring{g}_1 &= -\mathring{g}_0\left(\frac{n}{2}Z_1 + \Gamma_1(R; \mathring{m}_0, \mathring{g}_0)\right) \end{aligned}$$

At this point the reader is strongly recommended to work out the result at the second order in λ him/herself and to check it with what is given below only afterwards

$$\begin{aligned} \mathring{m}_2 &= -\frac{1}{2\mathring{m}_0}\left(\mathring{m}_1^2 + \mathring{m}_1\frac{\partial}{\partial\mathring{m}}\Sigma_1(m^2; \mathring{m}_0, \mathring{g}_0) + \mathring{g}_1\frac{\partial}{\partial\mathring{g}}\Sigma_1(m^2; \mathring{m}_0, \mathring{g}_0) + \Sigma_2(m^2; \mathring{m}_0, \mathring{g}_0)\right) \\ Z_2 &= Z_1\Sigma'_1(m^2; \mathring{m}_0, \mathring{g}_0) + \mathring{m}_1\frac{\partial}{\partial\mathring{m}}\Sigma'_1(m^2; \mathring{m}_0, \mathring{g}_0) + \mathring{g}_1\frac{\partial}{\partial\mathring{g}}\Sigma'_1(m^2; \mathring{m}_0, \mathring{g}_0) + \Sigma'_2(m^2; \mathring{m}_0, \mathring{g}_0) \\ \mathring{g}_2 &= \end{aligned}$$

.

Remark: *Renormalization is quite often performed using the so-called counterterm technique. The idea is to insist that the mass in the free Lagrangian is the physical one and to compensate for discrepancies by an appropriate term in the interaction Lagrangian. Technically it amounts to splitting the bare mass to two parts: the physical mass m and the counterterm δm*

$$\mathring{m} = m + \delta m$$

where m is used in the free Lagrangian and the term with δm is understood as a part of interaction Lagrangian (leading to a special two-legged vertex represented by a cross in Feynman diagrams). At any order of a perturbation theory the counterterm is chosen in such a way that it compensates the difference between $\mathring{m}$ and m at this order (which means that new counterterm vertices are added at every order of perturbation theory and they are treated as having a corresponding order, even if they do not add appropriate powers of the coupling constant or loop expansion). Relation between our approach and the mass counterterm is straightforward

$$\delta m = \mathring{m} - m = \sum_{n=1}^{\infty} \lambda^n \mathring{m}_n$$

The same philosophy can be used also for coupling constants. One insists that the tree level vertex (at some kinematical point) contributes in such a way that it saturates an experimental result. This again means nothing else but splitting

$$\mathring{g} = g_R + \delta g$$

where both parts remain in the interaction Lagrangian (leading to equivalent vertices) and the δg-term is to compensate (order by order) for discrepancies arising from the vertex with g. Again, it is straightforward to relate results of this counterterm technique to out approach

$$\delta g = \mathring{g} - g_R = \sum_{n=1}^{\infty} \lambda^n \mathring{g}_n$$

4.2 Regularization of infinities

Many loop integrals in perturbative QFT are divergent. This means that they are, frankly speaking, just nonsenses (even if fancy ones). However, if such nonsenses are combined in a right way, they may lead to a perfectly legitimate result. As an example let us take two limits: $\lim_{\Lambda\to\infty} \sin\Lambda$ and $\lim_{\Lambda\to\infty} \ln\Lambda$. They do not exist and neither does their ratio $\frac{\lim_{\Lambda\to\infty}\sin\Lambda}{\lim_{\Lambda\to\infty}\ln\Lambda}$ or any other function. Nevertheless, the limit of the ratio $\lim_{\Lambda\to\infty}\frac{\sin\Lambda}{\ln\Lambda}$ does exist (even if it vanishes). Regularization of loop integrals is nothing else but writing these integrals as particular limits and not taking the limits until the very end of calculations (when they are all taken simultaneously as one limit). If all such limits of relations between measurable physical quantities become well defined and finite, the theory is considered to be healthy enough (and is called renormalizable).

The most natural regularization of improper integrals is their very definition as a limit of integrals over increasing integration region: e.g. $\int_0^\infty = \lim_{\Lambda\to\infty}\int_0^\Lambda$. This regularization (in the radial spherical coordinate in the momentum space) is really used and it is called the cut-off regularization. It is, however, not the only possibility. Another quite natural one is to treat space-time as a lattice (rather than continuum). This will introduce an effective cut-off in the momentum space which is, however, not the same one as the cut-off in momentum spherical coordinates. So one has at least two natural regularizations, which one is the right one?

There is no such thing as the right regularization. Without regularization many QFTs do not make sense and regularization procedure can be defined in various ways. Some of them are more natural or advantageous, some are less, but none of them is "the one and only". Ok, but do they all provide the same results? No, they do not. One QFT with different regularizations can, in principle, provide different results. How come? The answer is perhaps not a surprising one: the regularization procedure is simply a part of definition of a QFT.

Dimensional regularization is based on a simple, even if strange idea of calculating the loop integrals in D dimensions instead of 4. The result is then understood as a function of a complex D and limit $D \to 4$ is taken.

4.2.1 Dimensional regularization

We will explain the general procedure using a specific example, namely the 1-loop contribution to the self-energy in the φ^3-theory

$$\Sigma_1 = i \quad = \frac{i}{2}\int \frac{d^Dk}{(2\pi)^D}(-i\mathring{g})\frac{i}{k^2-\mathring{m}^2+i\varepsilon}(-i\mathring{g})\frac{i}{(p-k)^2-\mathring{m}^2+i\varepsilon}$$

Calculations proceed in three typical steps called:
1. Feynman parametrization
2. Wick rotation
3. dimensional regularization

Feynman parametrization

In the first step the product of propagators is written in a specific form suitable for further processing. The aim is to rewrite the integrand (which is a function of the 4-vector k) as a function of the scalar k^2 (so that one can use spherical coordinates and integrate easily over the angles). The trick is based on a simple identity[2]

$$\frac{1}{ab} = \int_0^1 \frac{dz}{\left[az+b\,(1-z)\right]^2}$$

which is a special case of a more general identity[3]

$$\frac{1}{a_1\dots a_n} = (n-1)!\int_0^1 dz_1\dots dz_n \frac{\delta(z_1+\dots+z_n-1)}{[z_1a_1+\dots+z_na_n]^n}$$

By using of this identity one gets

$$\Sigma_1 = \frac{i\mathring{g}^2}{2}\int\frac{d^Dk}{(2\pi)^D}\int_0^1 dz \frac{1}{\left[(k^2-\mathring{m}^2+i\varepsilon)z+((p-k)^2-\mathring{m}^2+i\varepsilon)\,(1-z)\right]^2}$$

Completing the square in the denominator $\left[p^2(1-z)-2p.k(1-z)+k^2-\mathring{m}^2+i\varepsilon\right]^2$ by adding and subtracting $p^2\,(1-z)^2$ one obtains $[(p(1-z)-k)^2+p^2z(1-z)-\mathring{m}^2+i\varepsilon]^2$ and after the substitution $k\to k+p\,(1-z)$ the integral becomes

$$\Sigma_1 = \frac{i\mathring{g}^2}{2}\int\frac{d^Dk}{(2\pi)^D}\int_0^1 dz \frac{1}{[k^2-\Delta(p^2,z;\mathring{m})]^2}$$

where

$$\Delta(p^2,z;\mathring{m}) = \mathring{m}^2-p^2z(1-z)-i\varepsilon$$

The final move of this first step is interchange of order of integrations over k and z

$$\Sigma_1 = \frac{i\mathring{g}^2}{2}\int_0^1 dz\int\frac{d^Dk}{(2\pi)^D}\frac{1}{[k^2-\Delta(p^2,z;\mathring{m})]^2}$$

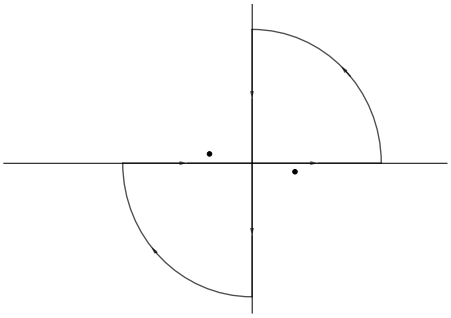
---

[2]Proof: $\frac{1}{ab}=\frac{1}{b-a}\left(\frac{1}{a}-\frac{1}{b}\right)=\frac{1}{b-a}\int_a^b\frac{dx}{x^2}$ and now one sets $x=az+b\,(1-z)$

[3]Proof (most simple with the so-called Schwinger parametrization $\frac{1}{a}=\int_0^\infty dx\,e^{-ax}$ which is valid for $\mathrm{Re}\,a>0$): $\frac{1}{a_1\dots a_n}=\int_0^\infty dx_1\dots dx_n\,e^{-\sum a_ix_i}$ where one inserts $\int_{-\epsilon}^\infty dX\,\delta(\sum x_i-X)=1$ and then one makes substitution $z_i=\frac{x_i}{X}$ to get $\frac{1}{a_1\dots a_n}=\int_0^1 dz_1\dots dz_n\int_{-\epsilon}^\infty dX\,X^n\,\frac{\delta(\sum z_i-1)}{X}\,e^{-X\sum_i a_iz_i}=\int_0^1 dz_1\dots dz_n\,\frac{\delta(1-\sum z_i)}{[a_1z_1+\dots a_nz_n]^n}\Gamma(n)$

Wick rotation

The purpose of the whole gymnastics in the first step was to get the integral in the form suitable for the use of spherical coordinates. The problem, however, is how to understand such coordinates in the Minkowski space. Wick rotation is the trick enabling us to rewrite the integral in Euclidean space. It concerns the integral over the component k_0. This integral is envisaged in the complex plane $k_0 + ik_4$ (the reason for the strange notation for the imaginary axis part will become clear shortly) where the integrand is an analytic function inside the contour displayed in the picture:



Due to the analyticity the integral over the whole contour is zero. The integral over the arcs vanishes as well, since the integrand decreases more rapidly than the inverse radius of arcs. The integral over the real axis can be therefore replaced by the integral over the imaginary axis

$$\int_{-\infty}^{\infty} dk^0 f(k^0) = \int_{-\infty}^{\infty} dik_4 \; f(ik_4)$$

This replacement is called the Wick rotation. For a loop integral it amounts to replacement

$$d^D k_{\text{Minkowski}} \; f(k^2) \to i \; d^D k_{\text{Euclid}} \; f(-k^2)$$

and once the integral is expressed in the 4-dimensional Euclidean space (with coordinates $k_1, \ldots, k_4$) one can use spherical coordinates happily. So after the Wick rotation our integral becomes

$$\Sigma_1 = -\frac{\mathring{g}^2}{2} \int_0^1 dz \int \frac{d^D k}{(2\pi)^D} \frac{1}{[-k^2 - \Delta(p^2, z; \mathring{m})]^2}$$

which can be written in the spherical coordinates $d^D k = k^{D-1} dk \; d\Omega_D$ as

$$\Sigma_1 = -\frac{\mathring{g}^2}{2} \int_0^1 dz \int \frac{d\Omega_D}{(2\pi)^D} \int_0^{\infty} \frac{k^{D-1} dk}{[k^2 + \Delta(p^2, z; \mathring{m})]^2}$$

Once the integral is written in this form, the integration over $d\Omega_D$ can be performed either as explicit integration over the angles, or using a neat trick based on the integral $\int_{-\infty}^{\infty} dx \; e^{-x^2} = \sqrt{\pi}$. In D dimensions this integral equals to [4] $\pi^{D/2}$ and in the spherical coordinates one obtains

$$\int_{-\infty}^{\infty} d^D x \; e^{-x^2} = \int d\Omega_D \int_0^{\infty} dx \; x^{D-1} e^{-x^2} \overset{y=x^2}{\longrightarrow} \int d\Omega_D \int_0^{\infty} \frac{dy}{2} \; y^{\frac{D}{2}-1} e^{-y} = \frac{\Gamma(D/2)}{2} \int d\Omega_D$$

and therefore

$$\int d\Omega_D = \frac{2\pi^{D/2}}{\Gamma(D/2)}$$

The last relation is going to play a crucial role in the dimensional regularization.

[4] $\int_{-\infty}^{\infty} d^D x \; e^{-x^2} = \int_{-\infty}^{\infty} dx_1 \ldots dx_D \; e^{-\sum_{i=1}^{D} x_i^2} = \prod_{i=1}^{D} \int_{-\infty}^{\infty} dx_i \; e^{-x_i^2} = \pi^{D/2}$

dimensional regularization

At this moment one sees clearly that for $D = 4$ the integral over dk is divergent. One can see it directly in the cut-off regularization, i.e. by replacing $\int_0^\infty dk \ldots$ by $\int_0^\Lambda dk \ldots$ and let $\Lambda \to \infty$ at the end of the day.[5] Another kind of regularization, however, is much more common in the contemporary theory of elementary particles.

Dimensional regularization is based on a simple observation that the integral $\int_0^\infty dk \frac{k^{D-1}}{(k^2+\Delta)^2}$ is convergent for any real $0 < D < 4$. The parameter D entered this integral as number of dimensions, so only the integer D seems to make sense, but the integral over dk is well defined also for non-integer D. We shall, therefore, consider a non-integer

$$D = 4 - 2\epsilon$$

and let $\epsilon \to 0$ at the end of the day. For such a D the integral over dk is calculated by a clever substitution $y = \frac{\Delta}{k^2+\Delta}$, i.e. $k^2 = \frac{\Delta}{y} - \Delta$ and $2kdk = -\Delta y^{-2} dy$

$$\int_0^\infty dk \frac{k^{D-1}}{[k^2+\Delta]^n} = \frac{1}{2}\Delta^{2-n-\epsilon} \int_0^1 dy\, y^{n-2} \left(\frac{1}{y} - 1\right)^{1-\epsilon} = \frac{1}{2}\Delta^{2-n-\epsilon} \int_0^1 dy\, y^{n-3+\epsilon} (1-y)^{1-\epsilon}$$

$$= \frac{1}{2}\Delta^{2-n-\epsilon} B(n-2+\epsilon, 2-\epsilon) = \frac{1}{2}\Delta^{2-n-\epsilon} \frac{\Gamma(n-2+\epsilon)\Gamma(2-\epsilon)}{\Gamma(n)}$$

$$\int_0^\infty dk \frac{k^{D-1}}{[k^2+\Delta]^2} = \frac{1}{2}\Delta^{-\epsilon}\Gamma(\epsilon)\Gamma(2-\epsilon)$$

Reminder[6]: $\Gamma(z) = \int_0^\infty dt\, t^{z-1}e^t$, $B(x,y) = \int_0^1 dt\, t^{x-1}(1-t)^{y-1}$ and $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

The tricky part of the dimensional regularization is not the dk integration but rather the $d\Omega_D$ integration. What is the meaning of Ω_D for non-integer D? This looks like a deep and potentially fatal question, but it is not. We do not need to have a geometric picture of non-integer dimensionality. We just need an expression which could serve as the integral $\int d\Omega_D$ for such a dimensionality. And this can be any decent function of D which is equal to the said integral $\int d\Omega_D$ for positive integer D. We have found such a function already and we can use it as a definition of the integral for non-integer D. So from now on

$$\int d\Omega_D \stackrel{\text{def}}{=} \frac{2\pi^{D/2}}{\Gamma(D/2)} = \frac{2\pi^{2-\epsilon}}{\Gamma(2-\epsilon)}$$

As to the dimensional regularization, one remark is still at place. When messing with space-time dimensionality, one should be careful enough not to mess with physical dimensionality as well. The point is that since d^4k is not dimensionless, one should not replace it by d^Dk but rather by something like $\mu^{4-D}d^Dk$ where μ is an arbitrary quantity with the dimension of momentum. So from now on the following replacement

$$d^4k \to \mu^{4-D}d^Dk$$

is going to be a part of our definition of the dimensional regularization.

[5] The result is logarithmicaly divergent: $\int_0^\Lambda dk \frac{k^3}{(k^2+\Delta)^2} = \frac{1}{2}\int_0^{\Lambda^2} dx \frac{x}{(x+\Delta)^2} = \frac{1}{2}\left[\ln\left(1+\frac{\Lambda^2}{\Delta}\right) - \frac{\Lambda^2}{\Lambda^2+\Delta}\right]$

[6] raz
dva
tri
štyri
päť
šesť

Now back to the self-energy. Taking all the pieces together one gets

$$\Sigma_1 = -\frac{\mathring{g}^2}{32\pi^2}\Gamma(\epsilon)(4\pi)^\epsilon \int_0^1 dz \left(\frac{\mu^2}{\Delta(p^2,z;\mathring{m})}\right)^\epsilon$$

The Γ-function is now evaluated using it's basic properties $\Gamma(z+1) = z\Gamma(z)$ and $\Gamma(n) = (n-1)!$ leading to $\epsilon\Gamma(\epsilon) = \Gamma(1+\epsilon) = \Gamma(1) + \epsilon\Gamma'(1) + \frac{1}{2}\epsilon^2\Gamma''(1) + \ldots$ from where one concludes

$$\Gamma(\epsilon) = \frac{1}{\epsilon} + \Gamma'(1) + \frac{\epsilon}{2}\Gamma''(1) + \ldots$$

The ϵ-powers are evaluated according to $x^\epsilon = e^{\ln x^\epsilon} = e^{\epsilon \ln x} = 1 + \epsilon \ln x + \ldots$, so that one obtains[7]

$$\Sigma_1 = -\frac{\mathring{g}^2}{32\pi^2}\left(\frac{1}{\epsilon} - \gamma + \ln 4\pi - \int_0^1 dz\, \ln\frac{\Delta(p^2,z;\mathring{m})}{\mu^2} + \mathcal{O}(\epsilon)\right)$$

As expected, this expression is divergent in the limit $\epsilon \to 0$. This is a typical feature of the dimensional regularization: infinities of divergent integrals pronounce themselves as poles in ϵ.

The last thing to be evaluated is the integral

$$\int_0^1 dz\, \ln\frac{\mathring{m}^2 - p^2 z(1-z) - i\varepsilon}{\mu^2} = \ln\frac{\mathring{m}^2}{\mu^2} + \int_0^1 dz\, \ln\left(1 - z(1-z)p^2/\mathring{m}^2 - i\varepsilon\right)$$

where the ε is not the one from the dimensional regularization, but the one from the propagator. In the renormalization conditions the self-energy $\Sigma(p^2;\mathring{m},\mathring{g})$ is needed around $p^2 = m^2$ and $\mathring{m} = m$. For such values the complex logarithm in the z-integration is of the type $\ln(x - i\varepsilon)$ with $x > 0$, for which $\lim_{\varepsilon\to 0}\ln(x - i\varepsilon) = \ln x$. So one can write

$$\Sigma_1 = -\frac{\mathring{g}^2}{32\pi^2}\left(\frac{1}{\epsilon} - \gamma + \ln 4\pi - \ln\frac{\mathring{m}^2}{\mu^2} - \int_0^1 dz\, \ln\left(1 - z(1-z)p^2/\mathring{m}^2\right) + \mathcal{O}(\epsilon)\right)$$

For $I(a) = \int_0^1 dz\ \ln(1 - z(1-z)a)$ *Mathematica* yields

$$I(a) = -2 - 2\frac{a-4}{\sqrt{(4-a)\,a}}\arctan\frac{\sqrt{a}}{\sqrt{4-a}}$$

and so we can finally write

$$\Sigma_1(m^2;m,g_R) = -\frac{g_R^2}{32\pi^2}\left(\frac{1}{\epsilon} - \gamma + \ln 4\pi - \ln\frac{m^2}{\mu^2} + 2 - \frac{\pi}{\sqrt{3}} + \mathcal{O}(\epsilon)\right)$$

and

$$\Sigma_1'(m^2;m,g_R) = -\frac{g_R^2}{32\pi^2 m^2}\left(1 - \frac{2\pi}{3\sqrt{3}} + \mathcal{O}(\epsilon)\right)$$

As to the renormalization conditions, this is all what is needed. However, if the self-energy loop is a part of a diagram contributing to some M_{fi}, one may need it for p^2 outside a vicinity of m^2 and in such a case the *Mathematica* result may turn out to be insufficient. We will return to this issue later on.

[7]Notation: $\Gamma'(1) = -\gamma$. Numerical value $\gamma = 0.5772\ldots$ Proposition (without proof): $\Gamma''(1) = \gamma^2 + \frac{\pi^2}{6}$.

Exercise: *Calculate the contribution of the tadpole diagram to the self-energy in the φ^3 theory. You should get* $\Sigma_{\text{tadpole}}(p^2;\mathring{m},\mathring{g}) = -\frac{\mathring{g}^2}{32\pi^2}(\frac{4\pi\mu^2}{\mathring{m}^2})^\epsilon\,\Gamma(\epsilon-1) = \frac{\mathring{g}^2}{32\pi^2}(\frac{1}{\epsilon}+1-\gamma+\ln\frac{4\pi\mu^2}{\mathring{m}^2}+\mathcal{O}(\epsilon))$

Remark: *The radial integral in Σ_{tadpole} is equal to $\int_0^\infty dk\frac{k^{D-1}}{k^2-\mathring{m}^2+i\varepsilon}$. In the cut-off regularization this integral is quadratically divergent, i.e. the divergence is proportional to Λ^2. This is evidently worse than the logarithmic divergence of the non-tadpole contribution. On the other hand, in the dimensional regularization the divergence is of order $1/\epsilon$, i.e. of the same order as the non-tadpole contribution. This is just a specific example of a general feature of the dimensional regularization: logarithmic divergences of the cut-off regularization becomes $1/\epsilon$ divergences in the dimensional regularization and higher than logarithmic divergences in the cut-off regularization come out finite in the dimensional regularization.*

Remark: *In the case of non-vanishing tadpole diagrams one can proceed in one of two ways. One can either perform the renormalization accounting for the contributions of the tadpole diagrams to the self-energy and the dressed vertex, or one can redefine the fields $\varphi(x)\to\varphi(x)-v$ so that the new fields will lead to vanishing tadpoles.*

Exercise: *Calculate the 1-loop contribution to the dressed vertex in the φ^3 theory:*

$$-i\mathring{g}\Gamma_1(p_1,p_2,p_3;\mathring{m},\mathring{g}) = \int \frac{\mu^{4-D}d^Dk}{(2\pi)^D}\frac{(-i\mathring{g})^3\,i^3}{((k+p_1)^2-\mathring{m}^2+i\varepsilon)((k-p_2)^2-\mathring{m}^2+i\varepsilon)(k^2-\mathring{m}^2+i\varepsilon)}$$

You should get

$$\Gamma_1 = 2\mathring{g}^2\int_0^1 dz_1\int_0^{1-z_1} dz_2\int\frac{\mu^{4-D}d^Dk}{(2\pi)^D}\frac{1}{(k^2+\Delta)^3} = \frac{\mathring{g}^2}{(4\pi)^2}(4\pi\mu^2)^\epsilon\,\Gamma(1+\epsilon)\int_0^1 dz_1\int_0^{1-z_1} dz_2\,\Delta^{-1}+\mathcal{O}(\epsilon)$$

where $\Delta = \mathring{m}^2 - z_1p_1^2 - z_2p_2^2 + (z_1p_1 - z_2p_2)^2 - i\varepsilon$

Remark: *Infrared divergences (why to keep $\epsilon \neq 0$ even in convergent integrals)*

$$\mathring{m} = m - \lambda \frac{1}{2m}\Sigma_1(m^2; m, g_R) = m + \lambda \frac{1}{2m}\frac{g_R^2}{32\pi^2}\left(\frac{1}{\epsilon} - \gamma + \ln 4\pi - \ln\frac{m^2}{\mu^2} + 2 - \frac{\pi}{\sqrt{3}}\right)$$

$$Z = 1 + \lambda\Sigma_1'(m^2; m, g_R) = 1 - \lambda\frac{g_R^2}{32\pi^2 m^2}\left(1 - \frac{2\pi}{3\sqrt{3}}\right)$$

$$\mathring{g} = g_R - \lambda g_R\left(\frac{3}{2}\Sigma_1'(m^2; m, g_R) + \Gamma_1(R; m, g_R)\right) = g_R - \lambda g_R\left(\frac{3g_R^2}{64\pi^2 m^2}\left(1 - \frac{2\pi}{3\sqrt{3}}\right) + \frac{g_R^2}{16\pi^2}C(R; m)\right)$$

where

$$C(R; m) = \int_0^1 \int_0^{1-z_1} \frac{dz_1 d\, z_2}{m^2 - z_1 p_1^2 - z_2 p_2^2 + (z_1 p_1 - z_2 p_2)^2 - i\varepsilon}$$

the z-integral

For $I(a) = \int_0^1 dz\ \ln(1 - z(1-z)a)$ *Mathematica* yields

$$I(a) = -2 - 2\frac{a-4}{\sqrt{(4-a)\,a}} \arctan \frac{\sqrt{a}}{\sqrt{4-a}}$$

For $0 < a < 4$ this is perfectly unambiguous result. For other values of a, however, some of the square roots are not single valued functions,they have cuts in the complex plane, and one has to be rather careful here. The subtleties involved for $a < 0$ and $a > 4$ are of great importance and cannot be ignored.

To really understand the result, it is perhaps worthwhile to calculate the integral by hand. For $a < 4$ the argument of the logarithm is positive, i.e. the logarithm is a single-valued real function. The integral can be easily taken by the partial integration[8], for any real a one has

$$\int_0^1 dz\ \ln|1 - z(1-z)a| = -\int_0^1 dz\ \frac{z\,(2z-1)\,a}{az^2 - az + 1}$$

and the RHS integral is easily evaluated[9]

$$\begin{aligned} &\text{for } a < 0 && I(a) = -2 - \sqrt{1-4/a}\,\ln\frac{\sqrt{1-4/a}+1}{\sqrt{1-4/a}-1} \\ &\text{for } 0 < a < 4 && I(a) = -2 + 2\sqrt{4/a-1}\,\arctan\sqrt{\tfrac{a}{4-a}} \end{aligned}$$

For $a > 4$, on the other hand, the argument $1 - z(1-z)a$ becomes negative for z between $\frac{1}{2}\left(1 \mp \sqrt{1-4/a}\right)$. The logarithm is not a single valued function in this interval. For $x \in \Re^-$ the logarithm develops a cut, where one has $\ln(x \pm i\epsilon) = \ln|x| \pm i\pi$. The imaginary part of the integral is the integral of the constant $-\pi$ over the interval of length $\sqrt{1-4/a}$. The real part is calculated using the same method as above

$$\text{for } a > 4 \qquad I(a) = -2 - \sqrt{1-4/a}\,\ln\frac{1+\sqrt{1-4/a}}{1-\sqrt{1-4/a}} - i\pi\sqrt{1-4/a}$$

---

[8] $\int dz\ \ln|1 - z\,(1-z)\,a| = z\ln|1 - z\,(1-z)\,a| - \int dz\ z\frac{(2z-1)a}{1-z(1-z)a}$ for any real a

[9] $\int dz \frac{z(2z-1)a}{az^2-az+1} = \int dz(2 + \frac{z-2/a}{z^2-z+1/a}) = 2z + \int dz' \frac{z'+1/2-2/a}{z'^2-1/4+1/a} = 2z + \frac{1}{2}\ln[z'^2 + \frac{4-a}{4a}] - 2\int dz'(1+\frac{4a}{4-a}z'^2)^{-1}$ where $z' = z - 1/2$ and the last integral $\int dz'(1+\frac{4a}{4-a}z'^2)^{-1}$ equals to $\sqrt{\frac{4-a}{4a}}\arctan\sqrt{\frac{4a}{4-a}}z'$ for $0 < a < 4$, while for $a < 0$ it equals to $\frac{1}{4}\sqrt{1-4/a}\ln\left|\frac{\sqrt{1-4/a}+2z}{\sqrt{1-4/a}-2z}\right|$

4.3 Vertices — the first look

4.3.1 the coupling constant renormalization

Our treatment of the loop effects on vertices is going to be closely related and simultaneously rather different from that of the propagators. Both aspects, the similarities and the distinctions, can be illustrated by the role played by some basic notions in the two cases.

In the preceding sections, the discussion was centered around the bare and dressed masses and propagators. In a close analogy, the bare and dressed coupling constants and vertices will be the main characters of the sections to follow. As to the bare objects, the analogy is still a very close one: both the bare masses and coupling constants are the parameters of the Lagrangian, and both the bare propagators and vertices are just what is given by the Feynman rules.

Now to the differences. While the dressed (physical) mass is a straightforward and unambiguous quantity characterizing a movement of a single particle[10], there is no such straightforward and unambiguous dressed (physical) coupling constant. If one really wants, as one usually does, to define a dressed coupling constant, one is invited to do so, but naturalness of such definition is a matter of taste, and it is far from being unambiguous. The lack of the straightforward and genuine definition adds some extra artificial flavor to the discussion of the coupling constant renormalization.

Fortunately, even if there is no genuine definition of the dressed coupling constant, there is one "almost natural". It concerns the dressed vertex, which sounds like a perfect analogy to the dressed propagator, but it is not. The point is that unlike the dressed propagator, which is the sum of all connected diagrams, the dressed vertex is the sum of all OPI (one particle irreducible) diagrams [11]. Depicting the sum of all connected and all OPI diagrams by the shaded and hatched blob respectively, one may consider three types of vertices

the bare one | the OPI one **the dressed one** | the noname one

to be compared with the corresponding three types of propagators[12].

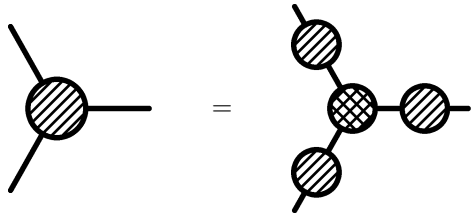
To understand why it is preferable to reserve the name "dressed vertex" to the hatched blob rather then to the shaded one, it is illustrative to express the latter in terms of the former. Once the role of the OPI diagrams is understood, we will proceed to the definition of the dressed coupling constant.

[10] This applies, of course, only if the particle can exist in a single-particle state, which is not the case e.g. for quarks and gluons. This is the reason why the treatment of quark masses in QCD looks more like treatment of coupling constants then that of ordinary masses.

[11] Of course, the diagrams with a given set of external legs are understood. These legs are usually of different types, corresponding to various particle species. We are nevertheless going to use only one type of leg in diagrams, at least in this bird-eye view discussion. This does not mean any loss of generality, the particle species can be distinguished by labels (which are anyway omitted in this section) rather than by typographical means.

[12] The dressed propagators were given by the sum of all connected diagrams. The sum of OPI diagrams have played only an auxiliary, even if very important, role in the calculation of the dressed propagator. For vertices, on the other hand, the sum of the connected diagrams does not play any particular role, which is why it did not deserve a specific name.

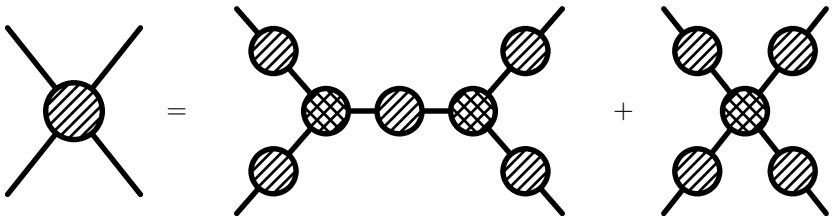
Let us start with the simplest case of three external legs. For sake of simplicity we will suppose vanishing contributions from tadpoles (see the discussion of the Dyson series in the preceding section), in which case one easily convinces him/herself that



Now if the process is experimentally accessible[13] for some combination of external momenta and other quantum numbers, then one can possibly extract some information about the value of the vertex in the real world. The connection between the experiment and the diagrams is provided by the S-matrix elements, which are given, as we already know, by the amputated Feynman diagrams. And since the propagators to be amputated are the dressed ones, the above diagrammatic equation is telling us that the sum of the amputated diagrams is nothing but the hatched blob.

At the tree level, the hatched blob is nothing but the bare vertex. So in this approximation the experiment tells us directly the value of the bare vertex. Beyond this approximation, the experiment tells the value of something else, namely the bare vertex plus the loop corrections. It seems therefore reasonable to call this corrected vertex the dressed one.

Going one step further, i.e. considering four external legs, one finds easily



where the first term on the RHS stands for the three similar contributions (with different pairs of external legs entering the first hatched blob).

Both terms on RHS resemble closely the tree diagrams. This again suggests the hatched blob to be considered the dressed partner of the bare vertex. Note, however, that the existence of the dressed vertex with a given set of external legs does not depend on the existence of the corresponding bare vertex. The above relations apply to the $\varphi^3+\varphi^4$-theory, with the bare 3- and 4-legs vertices corresponding to the dressed ones, as well as to the simple φ^3-theory, with no bare partner to the 4-legs dressed propagator.

If both the 3- and 4-legs processes were experimentally accessible, then one can determine the 3-legs dressed propagator from the 3-legs process, and then use its knowledge[14] to determine the 4-legs dressed propagator from the 4-legs process. One can, however, usually extract

[13] Processes with three external legs are not the typical ones in particle physics. Anyway, there are such processes, e.g. a decay of an unstable particle or the so-called scattering on a classical external field. Strictly speaking, they are just subprocesses of some more complex process, like creation of an unstable particle in some inelastic scattering and subsequent decay in the former case, or a scattering on some heavy particle in the latter case. Nevertheless, the approximation in which these subprocesses are treated on they own is usually a pretty safe one, so it makes a good sense to consider also the three-legs processes.

[14] An experimental information is, of course, available only for specific kinematics, i.e. for some combinations of particle momenta and other quantum numbers, but this is, as a rule, sufficient to determine the dressed vertex in full generality. The point is that the structure of the dressed vertex is in principle theoretically known (practically

an experimental information about both 3- and 4-legs dressed vertices, even if there is no experimentally accessible 3-legs process. The point is that the RHS of the relation for the 4-legs processes contains two terms with different analytic properties. The first one contains the pole corresponding to the dressed propagator (unless it is cancelled, due to some conspiracy, by the zero of the dressed vertex — which is usually not the case). The second one usually does not contain poles. So in the fitting procedure, one tries first to isolate the pole contribution, which enables to pin down the 3-legs dressed vertex, and then use the remainder to pin down the 4-legs one.

The next step, concerning diagrams with five external legs, would introduce the 5-legs dressed vertices, etc. The RHS is always the sum of the so-called the skeleton diagrams, i.e. the tree diagrams with the bare propagators, external legs and vertices replaced by the dressed ones. We are not going to discuss this further, since for our purposes the relevant thing is a pair of a bare vertex (which usually has 3 or 4 legs) and the corresponding dressed one.

Let us take the above discussion as enough motivation for giving the name "dressed vertex" to the sum of the OPI diagrams. Our next task is to understand how is the dressed vertex used to define the dressed coupling constant. For the bare quantities, the relation between the vertex and coupling constant is

$$\text{the bare vertex} = -i\overset{\circ}{g}\,\gamma\,(P)$$

where $-i$ is the standard perturbation theory factor, $\overset{\circ}{g}$ is the bare coupling constant and $\gamma\,(P)$ is an explicitly known function of momenta and other quantum numbers (collectively denoted as P) characterizing the legs of the vertex. The function $\gamma\,(P)$ can be read out directly form the Lagrangian: for φ^3- and φ^4-theories the $\gamma\,(P)$ is not explicitly present ($\gamma\,(P) \equiv 1$), for the spinor quantum electrodynamics $\gamma\,(P)$ stands for Dirac matrices ($\gamma\,(P) \equiv \gamma^{\mu}_{\text{Dirac}}$), for the scalar electrodynamics $\gamma\,(P)$ stands for the sum of momenta ($\gamma\,(P) \equiv p_1^{\mu} + p_2^{\mu}$), etc.

The dressed vertex is some specific function of P, which is completely determined by the Lagrangian, although usually not easy to calculate. We shall denote this function, divided by $-i\overset{\circ}{g}$ for formal reasons, by $\Gamma\,(P)$

$$\text{the dressed vertex} = -i\overset{\circ}{g}\,\Gamma\,(P)$$

Now, were the functions $\gamma\,(P)$ and $\Gamma\,(P)$ just rescaled versions of each another, one would have a perfectly natural definition of the dressed coupling constant: it would be sufficient to express the dressed vertex simply as $-ig\,\gamma\,(P)$. Needless to say, life is not so easy. The function $\Gamma\,(P)$ may be, and as a rule indeed is, a bit different from $\gamma\,(P)$, leaving us with no straightforward definition of the dressed coupling constant.

it is known only approximately, but this does not matter here) and everything what the experiment has to pin down is usually just a single number.

Anyway, if one really likes the idea of writing the dressed vertex as $-ig\,\gamma\,(P)$, one is allowed to do so. The only problem is that with a given value of g, this is probably going to work only for one specific choice of P. For other P's one would need other values of g. Clearly, this idea looks like a foolish one. Nevertheless, it is precisely the way how the dressed coupling constant is defined.

The point is that one chooses arbitrary point P, which we shall call the renormalization point and denote by R, and only at this point one demands

$$-i\overset{\circ}{g}\,\Gamma\,(R) = -ig_R\,\gamma\,(R)$$

The subscript in g_R is frequently understood only as an abbreviation of the word "renormalized", which is often used as an alternative to the word "dressed", but for us it stands for the renormalization point and reminds us, that g_R is the dressed coupling constant defined at the point R. The particular choice of R is usually decided by the experimental convenience. From the measured value of the suitable cross section or decay rate, one extracts the value of the corresponding amplitude (squared), and from there one obtains the experimental (absolute) value of g_R.

As an example let us consider the φ^4-theory.[15] Let us suppose we have measured the cross-section at some kinematical point R. From the well-known relation between the cross section and the matrix element, which for four identical masses sounds (see e.g. Peskin–Schroeder, chapter 4) $d\sigma/d\Omega = |M_{fi}|^2/64\pi^2 E_{\mathrm{cm}}^2$ one learns the value of $|M_{fi}|^2$ at R. For the φ^4-theory, the 3-legs dressed vertex is zero (why?), so the amplitude is just the 4-legs dressed vertex multiplied by the wave-function renormalization factor $\sqrt{Z}$ for every external leg: $M_{fi} = -i\overset{\circ}{g}\Gamma\,(R)\,Z^2$, which at this very point is equivalent to $M_{fi} = -ig_R Z^2$ (recall that for the φ^4-theory $\gamma\,(R) \equiv 1$). We can therefore express the value of the dressed or renormalized coupling constant in terms of the measured cross-section and the calculated wave-function renormalization constant as[16]

$$g_R = 8\pi E_{\mathrm{cm}} Z^{-2} \sqrt{\left.\frac{d\sigma}{d\Omega}\right|_R}$$

Once the experimental value of g_R is known, one can use the theoretical result for g_R, which would be some function of R and the parameters of the Lagrangian, to pin down the value of $\overset{\circ}{g}$.

By the coupling constant renormalization one usually understands the proper choice of $\overset{\circ}{g}$, leading to the measured value of g_R.

This statement definitely needs some qualifications, if for nothing else then at least due to the fact, that it is $|g_R|$, rather then g_R itself, which is extracted from the experiment. Such qualifications are the subject matter of the subsequent set of remarks.

[15] The φ^3-theory would be a more complex example. The reason is that the decay is kinematically forbidden (at least for nonzero mass), so one has to extract experimental information about 3-legs dressed vertex from the 4-legs process.

[16] Of course, in this way we have pinned down only the absolute value of the dressed coupling constant. This is taken into account by allowing both signs in the square-root of the cross-section.

Remark: *The experimental information about g_R concerns its absolute value, not the phase. One should therefore find $\overset{\circ}{g}$ as a function of the unknown phase. Then one can use the fact that the bare coupling constant is usually real, which may already select the unique $\overset{\circ}{g}$ from the generally complex function $\overset{\circ}{g}$. If this condition is insufficient to determine the $\overset{\circ}{g}$ unambiguously, one has to consider all the real values of $\overset{\circ}{g}$ and check, which one is consistent with some additional experimental information.*

Remark: *As should be clear from what was said up to now, the renormalization is nothing but the experimental fit of the basic theoretical parameters. So one can use, in principle, any set of experimental numbers and look for the best fit in the space of these parameters. Renormalization is not bound to use physical (dressed) masses and coupling constants in this fitting procedure. Nevertheless, these quantities are in a sense natural, widely used, and perhaps also preferable from the point of view of error propagation.*

Remark: *Let us stress that the definition of the renormalized coupling constant is just a convention. If, for any reason, one would like to use another convention, one is allowed to do so. And even if it is not obvious, there are such reasons, mostly formal and technical ones. It is e.g. very popular to modify the definition of the dressed coupling constant g_R to*

$$-i\overset{\circ}{g}\,\Gamma\left(R\right) = -i\left[g_R + \Delta\left(R\right)\right]\gamma\left(R\right)$$

where $\Delta\left(R\right)$ is some explicitly given function. The particular choice of $\Delta\left(R\right)$ is called the choice of the renormalization scheme. From this point of view, our original definition of g_R was just a specific choice of the renormalization scheme, corresponding to $\Delta\left(R\right) = 0$. Once this viewpoint is adopted, the renormalized coupling constant depends on the choice of the renormalization point R, and on top of that on the choice of the renormalization scheme.

Remark: *Just like the mass and the wave-function renormalization, also the coupling constant renormalization can be formulated in the counterterm language. This time such formulation is not based on reshuffling terms between the free and the interaction Lagrangian, but rather on some specific splitting of one term in the interaction Lagrangian into two terms. The splitting of the coupling constant of the n-leg vertex is defined by*

$$\delta_g = Z^{n/2}\overset{\circ}{g} - g_R$$

(To understand this prescription consider e.g. $\overset{\circ}{g}\varphi^4 = \overset{\circ}{g}Z^2\varphi_r^4 = g_R\varphi_r^4 + \delta_g\varphi_r^4$.) The δ_g clearly depends on R, even if it is usually not emphasized explicitly.

Remark: *Many authors use the symbol Z not only for the wave-function renormalization constant, but also for the mass and coupling constant renormalization. If so, one replaces our Z by Z_φ, and defines the new quantities Z_g, Z_m*

$$\overset{\circ}{g} = Z_g g_R \qquad\qquad \overset{\circ}{m} = Z_m m$$

The Z_g clearly depends on R, even if it is usually not emphasized explicitly.

4.3.2 the running coupling constant

So far, the g_R was treated like a value at some fixed point R. However, since the point R is arbitrary, one may treat g_R as a function of R. In other words, we may use the definition of the renormalized coupling constant not only at the specific point R, but rather in full generality as $-i\overset{\circ}{g}\Gamma(P) = -ig_P\gamma(P)$. This is perfectly legal and often practical. The new quantity g_P is almost, but not quite, what is called the running coupling constant.

The benefits of using g_P are evident: once g_P is known, one has to consider only skeleton diagrams in any process, which reduces the amount of the required work significantly. In any realistic example, of course, the g_P is known only approximately, but this is sufficient for approximate calculations of the same order. The approximate knowledge of g_P introduces, however, a conceptually new aspect in the game.

The point is that g_P depends on the bare coupling constant, and the value of this parameter is pinned down by comparison to some experiment at some specific point R. Now, were the experiments precise and did we know g_P exactly, then we would get the same value of the bare coupling constant using any kinematical point P in the role of the renormalization point R. For approximately known g_P this is true no more and different renormalization points would give different values of $\overset{\circ}{g}$. Using the value of $\overset{\circ}{g}$ pinned down at the point R, the g_P would agree with the experiment completely at $P{=}R$, the agreement is expected to be quite good also for P in the vicinity of R, but with the increasing distance between P and R the approximate g_P becomes less and less reliable.

Surprisingly enough, there is a way of how to improve the reliability of an approximate result far away from the renormalization point. In the discussion of this improvement we shall follow the common habit of considering just a one-dimensional change of the renormalization point, namely the change in the overall energy scale μ. By the change of the energetic scale one means writing all momenta as multiples of some scale μ and then variation of this μ. In the other words, one writes $R = \mu_R N$, and then one studies the changes $\mu_R \to \mu$, i.e. the changes form R to the specific $P = \mu N$ (rather than a general change to an arbitrary P). Hence, the P-dependence of the dressed coupling constant g_P is reduced to the μ-dependence and it is this $g(\mu)$, which is called the running coupling constant. The generalization to g_P with an arbitrary P is straightforward (as we shall demonstrate) but almost never used[17].

The basic idea of the improvement of the approximately known $g(\mu)$ is to trust it only in the neighborhood of the renormalization point, but still to go far away from it. The trick is to change the renormalization point on the way. The tool is the so-called renormalization group equation (RGE).

The RGE is derived as follows. One starts from the dressed coupling constant as the function

[17]To understand why the change in μ is usually more important than any other change, one has first of all to realize that the unreliability of the approximate result usually becomes an issue only if P is really far away from R. The change of the overall scale μ can be arbitrary large, so this is can be a "dangerous change", which is as a rule confirmed by calculations. This does not hold for any change in P. To see the reason, at least to some extent, let us write individual momenta as $p_i = (E_i, \vec{p}_i) = |\vec{p}_i|\left(E'_i, \vec{n}_i\right) = \mu q_i \left(E'_i, \vec{n}_i\right)$, where $q_i = |\vec{p}_i|/\mu$ and $E'_i = E_i/|\vec{p}_i|$. Clearly, one cannot get very far by changing $\vec{n}_i$ (unless 2π is regarded a big number, but we are not going to be that faint-hearted). The parameter E'_i looks more dangerous, but in fact it is almost fixed by the value of μq_i. For external legs the reason is trivial: the momentum p_i is on-shell, i.e. $E'^2_i = 1 + m_i^2/\mu^2 q_i^2$. For internal legs the argument is much weaker, but still of some convincing power: internal legs are penalized by the propagator which is inversely proportional to the distance of E'^2_i from the on-shell value, i.e. going with E'_i far away from the on-shell value suppresses the amplitude significantly, making the process more difficult to measure, decreasing the demand for theoretical predictions in this direction. Finally one can vary q_i without limits. However, unless there are either theoretical or experimental indications of significant changes leading to experimentally accesible consequences, the demand/supply law explains why the changes in q_i are not a matter of general interest.

of the renormalization point and the bare coupling constant

$$g = G\left(\mu, \overset{\circ}{g}, \ldots\right)$$

where $G = \overset{\circ}{g}\,\Gamma(\mu, \overset{\circ}{g}, \ldots)\gamma^{-1}(\mu, \ldots)$ and ellipses stand for the scale independent part N of the renormalization point, as well as for other possible parameters like masses etc. The bare coupling constant can be expressed as a function of the dressed one

$$\overset{\circ}{g} = G^{-1}(\mu, g, \ldots)$$

where the inverse function[18] G^{-1} is understood at the fixed values of μ and all other parameters collected in the ellipses. The RGE is obtained by taking the derivative, with respect to μ, of the original relation and subsequent substitution of the inverse function into the result. For some formal reasons the whole equation is yet multiplied by μ, leading to the differential equation – the RGE[19]

$$\mu\frac{dg}{d\mu} = \beta(\mu, g) \qquad\qquad \beta(\mu, g) = \mu\left.\frac{\partial G(\mu, \overset{\circ}{g})}{\partial\mu}\right|_{\overset{\circ}{g}=G^{-1}(\mu,g)}$$

Note that we have refrained, in accord with the common habit, from writing the other variables (collected in the ellipses above) which are considered fixed.

What are the benefits of introducing this differential equation? If one knew the β-function and did not know the function G itself, one could find G, i.e. the $g(\mu)$ by solving the RGE. Otherwise it is of no use, at least as it stands. The starting function G is clearly a solution of the resulting differential equation, and for a well-behaved β-function, the solution is unique for a given initial value. So the whole outcome would be the function from which one had started.

At this point, it looks like if the RGE is useful only if one knows the β-function without knowing the G-function in advance. In QFT, however, this is never the case. One always obtains the β-function as a derivative of the corresponding G-function. The RGE is, nevertheless, a useful tool. It helps us, as already mentioned, to improve an approximately known G-function.

Let us consider G as a limit of a progression of approximate functions $G_{a,n}$, where the subscript a stands for *approximate*, and n stands for the (highest) order in the power expansion in some auxiliary parameter λ (usually the loop-expansion parameter). One can write the above differential equation for the function $g_{a,n}(\mu)$ and this would be the useless procedure, leading to the function $g_{a,n}(\mu)$ itself. Actually, in case of the approximate function the above equation (with g replaced by $g_{a,n}$) is not exactly what is called the RGE.

The RGE for an approximate function is obtained by spoiling the original equation a bit, namely by expanding the function $\beta_{a,n}$ in powers of λ and taking the result only up-to the n-th order. Let us emphasize that $\beta_{a,n}$ contains also the higher orders in λ, since it is the n-th order function with another n-th order function $\overset{\circ}{g} = G^{-1}(\mu, g)$ plugged in. The spoiled equation

$$\mu\frac{dg_{b,n}}{d\mu} = \beta_n(\mu, g_{b,n})$$

$$\beta_n(\mu, g_{b,n}) = \mu\left(\left.\frac{\partial G_{a,n}(\mu, \overset{\circ}{g})}{\partial\mu}\right|_{\overset{\circ}{g}=G^{-1}_{a,n}(\mu,g_{b,n})}\right)_{\text{up-to the } n\text{-th order only}}$$

[18] For a G non-monotonous in the variable $\overset{\circ}{g}$, one takes just one branch of the multi-valued function $G^{-1}(g)$ in what follows.

[19] The name is perhaps a bit misleading, since not much of the group theory is involved. Actually, nothing beyond the simple statement that for any line one can consider the group of translations along the line.

is called the RGE in case of approximate functions, and its solution can be quite different from the starting approximation. The subscript b now stands for *better*, because even if it is a solution of the spoiled equation, it is considered to be an improvement of the original approximation.

The reason why the $g_{b,n}$ is preferable to the $g_{a,n}$ is that whatever we are interested in, it is safer to treat it only up to the order to which the calculations were performed. Beyond this order, corrections are not sufficiently under control. And since in the RGE we are dealing with the relation of the function and its first derivative, we should take this relation seriously only up-to the order to which we have the situation under control.

As we shall see in the section on the loop expansion, the renormalization conditions for the coupling constant guarantee compensation of any potentially large corrections in G by the appropriate corrections in $\overset{\circ}{g}$. At the renormalization point all these corrections just vanish, in the vicinity of the renormalization point they are therefore expected to be tiny. But this is achieved only up to the order to which G was calculated. Any attempt to work beyond this order without complete calculation, would probably involve a work with possibly large corrections, which are perhaps just a fake, since they are going to vanish after the complete procedure at the higher order is performed.

To summarize, the perturbation theory provides us with a progression of functions $g_{a,n}(\mu)$ converging to the true running coupling constant $g(\mu)$. The RGE enables us to replace this progression by another one, of functions $g_{b,n}(\mu)$, having the same limit, but approaching it in a perhaps more decent way.

Let us emphasize that when calculating the β-function from an approximately known G-function, one has to work strictly to the order to which the G was calculated. From presentations in many textbooks a reader can get an impression that this is possible, but not mandatory. He or she may then consider the fact that the higher orders in the β-function are neglected, to be just a matter of convenience, rather than a matter of principle. Such a view is, however, incorrect. Did we not neglect the higher orders, we would reproduce the function we had started from, and no RGE improvement of the perturbation theory would be achieved.

Remark: *The RGE, even if it is not obvious at the first sight, is usually a relatively simple differential equation. The β-function is, as a rule, a polynomial in $g_{b,n}$ (due to the fact that the G-function is a polynomial in the bare coupling constant) and frequently it is μ-independent (due to the fact that the G-function uses to be proportional to $\ln\mu$).*

4.4 Vertices — the closer look

Unlike in case of propagators, the second look on vertices is devoted not to deepen, but rather to broaden the material of the first look. The key question of this section is: how does it come that in QED measured coupling constants of various particles are identical, in spite of the fact that loop corrections seem to be very different.

4.4.1 the Ward identities

The plan is to discuss the Ward identities on the level of Feynman diagrams and to answer the question at this level.

4.4.2 the relation between coupling constants and charges

The plan is to discuss the fact that the word electric charge is used for both the coupling constant and the quantity conserved due to the $U(1)$ invariance of QED. The Ward identities are rederived in language of matrix elements of conserved currents.

4.5 Vertices — the loop expansion

4.5.1 the renormalization conditions

In spite of the fact that we have discussed them separately, the mass, the wave function and the coupling constant renormalizations are to be performed simultaneously. One writes down the renormalization conditions discussed earlier together with the new renormalization condition between $\overset{\circ}{g}$ and g_R (the latter expressed in terms of the experimental value of the cross section and power of Z), to get the set of coupled equations

$$\begin{aligned}
\overset{\circ}{m}{}^2 &= m^2 - \Sigma(m^2; \overset{\circ}{m}{}^2, \overset{\circ}{g}) \\
Z^{-1} &= 1 - \Sigma'(m^2; \overset{\circ}{m}{}^2, \overset{\circ}{g}) \\
\overset{\circ}{g} &= g_R \, \gamma(R) \, \Gamma^{-1}(R; \overset{\circ}{m}{}^2, \overset{\circ}{g})
\end{aligned}$$

for the unknown Z, $\overset{\circ}{m}$ and $\overset{\circ}{g}$.

The whole thing can be, of course, formulated also in the counterterm version, where the renormalization conditions assumes the form

$$\begin{aligned}
0 &= \Sigma(m^2, g_R; \delta_{m^2}, \delta_Z, \delta_g) \\
0 &= \Sigma'(m^2, g_R; \delta_{m^2}, \delta_Z, \delta_g) \\
\delta_g &= -g_R + g_R \, (1+\delta_Z)^{n/2} \, \gamma(R) \, \Gamma^{-1}(R; m^2, g_R; \delta_{m^2}, \delta_Z, \delta_g)
\end{aligned}$$

4.5.2 the loop expansion — an example

The plan is to extend the analysis of the analogous paragraph dealing with propagators. The example is going to be the elastic scattering $\varphi\varphi \to \varphi\varphi$ within the φ^3-theory up to two loops, discussed in the same spirit as in the quoted paragraph. The RGE is also going to be discussed in the analogous way.

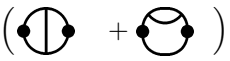
4.5.3 Loop expansion

Renormalization conditions are easily solved order by order in the loop expansion. Actually, at a given order the conditions are not merely the equations to be solved, but rather the explicit solutions themselves (provided the solutions at lower orders are known). Let us illustrate it on the φ^3-theory ($\mathcal{L}_{\rm int} = -\frac{1}{3!}\mathring{g}\varphi^3$). The loop expansion of the self-energy is

$$\Sigma(m^2;\mathring{m}^2) = \sum_{n=0}^{\infty} \lambda^n \Sigma_n(m^2;\mathring{m}^2)$$

$\Sigma_0(m^2;\mathring{m}^2)$ no contribution

$\Sigma_1(m^2;\mathring{m}^2)$ i 

$\Sigma_2(m^2;\mathring{m}^2)$ $i\Big($  $\Big)$

$\Sigma_3(m^2;\mathring{m}^2)$ $i\Big($ 

$+$ another five diagrams $)$

The parameter $\mathring{m}^2$, however, is also a function of the λ-parameter[20]. We will treat it as a power series, and Z is also going to be treated in this way

$$\mathring{m}^2(\lambda) = \textstyle\sum_{n=0}^{\infty} \lambda^n \mathring{m}^2{}_n \qquad\qquad \mathring{m}^2{}_0 = m^2$$

$$Z(\lambda) = \textstyle\sum_{n=0}^{\infty} \lambda^n Z_n \qquad\qquad Z_0 = 1$$

The lowest order coefficients are fixed by the fact that at the tree level the bare mass is equal to the physical one and Z is equal to 1. Expanding both sides of the renormalization conditions in λ one obtains for the first two corrections

$$\mathring{m}^2{}_1 = \Sigma_1(m^2;m^2) \qquad\qquad \mathring{m}^2{}_2 = \Sigma_2(m^2;m^2) + \Sigma_1^{(1)}(m^2;m^2)\mathring{m}^2{}_1$$

$$Z_1 = \Sigma_1'(m^2;m^2) \qquad\qquad Z_2 = \Sigma_2'(m^2;m^2) + \Sigma_1'^{(1)}(m^2;m^2)\mathring{m}^2{}_1 + Z_1^2$$

where $\Sigma_i^{(n)}(p^2;\mathring{m}^2) \equiv \partial^n \Sigma_i(p^2;\mathring{m}^2)/\partial(\mathring{m}^2)^n$, i.e. (n) stands for the derivative with respect to the second variable of the Σ (recall that the prime stands for the derivative with respect to to the first variable p^2).

The reader is invited to derive him- or herself the result at the next order.[21]

[20] The same is true also for the parameter $\mathring{g}(\lambda)$, which is not given explicitly in our formulae. At this stage, however, we are not going to power expand it (this is postponed to the discussion of the coupling constant renormalization), we are rather going to leave it untouched yet, which is perfectly legal, although not sufficient for any complete calculation at a given order.

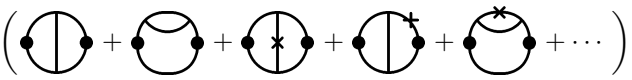
[21] He/she may want to check the result with $\mathring{m}^2{}_3 = \Sigma_3 + \Sigma_2^{(1)}\mathring{m}^2{}_1 + \Sigma_1^{(1)}\mathring{m}^2{}_2 + \frac{1}{2}\Sigma_1^{(2)}(\mathring{m}^2{}_1)^2$ and $Z_3 = \Sigma_3' + \Sigma_2'^{(1)}\mathring{m}^2{}_1 + \Sigma_1'^{(1)}\mathring{m}^2{}_2 + \frac{1}{2}\Sigma_1'^{(2)}(\mathring{m}^2{}_1)^2 + 2Z_1Z_2 - Z_1^3$, where the Sigmas are taken at $p^2 = m^2$ and $\mathring{m}^2 = m^2$, i.e. $\Sigma_3 \equiv \Sigma_3(m^2;m^2)$ etc.

The same procedure carried out within the counterterm approach:

$$\Sigma(m^2;\delta_{m^2},\delta_Z)=\sum_{n=0}^{\infty}\lambda^n\Sigma_n(m^2;\delta_{m^2},\delta_Z)$$

$\Sigma_0(m^2;\mathring{m}^2)$ $\quad i$ 

$\Sigma_1(m^2;\mathring{m}^2)$ $\quad i\Big($ $\Big)$

$\Sigma_2(m^2;\mathring{m}^2)$ $\quad i\Big($  $\Big)$

$\Sigma_3(m^2;\mathring{m}^2)$ acknowledgment: the author wants to thank the reader for filling up the diagrams for Σ_3

The cross represents the counterterm vertex (for $\mathcal{L}_{\rm ct}=\frac{1}{2}\delta_Z\partial_\mu\varphi\partial^\mu\varphi-\frac{1}{2}\delta_m\varphi^2$ the factor corresponding to the cross is $i(\delta_Z p^2-\delta_m)$). Every internal line can be decorated by any number of counterterm vertices. All such insertions can be summed up by the technique used in the derivation of the Dyson series (one simply sums up the geometric series), but there would be no much sense in doing this. The point is that what we are after is the power series in λ, and the counterterm expansions starts at $n=1$ (at the tree level counterterms vanish)

$$\delta_{m^2}(\lambda)=\textstyle\sum_{n=1}^{\infty}\lambda^n\delta_{m^2,n}$$

$$\delta_Z(\lambda)=\textstyle\sum_{n=1}^{\infty}\lambda^n\delta_{Z,n}$$

So one always needs only a first few counterterm insertions when working at the given order. It is therefore natural, and convenient for the bookkeeping purposes, to account only for such multiple counterterm insertions which could enter the results at the given order. Expanding both sides of the renormalization conditions in λ one obtains for the first two corrections

$$\begin{aligned}\delta_{m^2,1}&=-\Sigma_1+m^2\Sigma_1'\\ \delta_{Z,1}&=\Sigma_1'\end{aligned}$$

$$\begin{aligned}m^2\delta_{Z,2}-\delta_{m^2,2}&=\Sigma_2+\delta_{m^2,1}\Sigma_1^{(1,0)}+\delta_{Z,1}\Sigma_1^{(0,1)}\\ &\quad+\frac{1}{2}\delta_{m^2,1}^2\Sigma_0^{(2,0)}+\frac{1}{2}\delta_{Z,1}^2\Sigma_0^{(0,2)}+\delta_{m^2,1}\delta_{Z,1}\Sigma_0^{(1,1)}\\ \delta_{Z,2}&=\Sigma_2'+\delta_{m^2,1}\Sigma_1'^{(1,0)}+\delta_{Z,1}\Sigma_1'^{(0,1)}\\ &\quad+\frac{1}{2}\delta_{m^2,1}^2\Sigma_0'^{(2,0)}+\frac{1}{2}\delta_{Z,1}^2\Sigma_0'^{(0,2)}+\delta_{m^2,1}\delta_{Z,1}\Sigma_0'^{(1,1)}\end{aligned}$$

where $\Sigma_i^{(m,n)}\equiv\partial^{(m+n)}\Sigma_i/\partial(\delta_{m^2})^m\partial(\delta_Z)^n$, $\Sigma_i\equiv\Sigma_i(m^2;0,0)$ and we have utilized the explicitly known lowest order result $\Sigma_0(m^2;\delta_{m^2},\delta_Z)=-\delta_Z p^2+\delta_{m^2}$. Derivation of the next correction is a straightforward, even if tedious.

Exercise: *Play the same game as above for* ***a)*** φ^4*-theory* ***b)*** *QED.*

In practice, the self-energy is usually calculated only up-to a given order. To which order should one determine the bare mass and the wave-function renormalization constant (both being functions of λ) if the functions Σ_n are known only for $n \leq N$? There are two naturally looking possibilities: one can either solve the approximate N-th order renormalization conditions exactly, or one may expand the unknowns only up-to the N-th order

$$\overset{\circ}{m}{}^2(\lambda) = \sum_{n=0}^{N} \lambda^n \overset{\circ}{m}{}^2{}_n \qquad\qquad Z(\lambda) = \sum_{n=0}^{N} \lambda^n Z_n$$

and solve the renormalization conditions for $\overset{\circ}{m}{}^2{}_n$ and Z_n up-to the $n = N$.

The first possibility may look, at least at the first sight, as the preferable one. One may claim that even if working with approximate equations, one does not introduce additional error by solving them only approximately. However, this kind of reasoning is misleading.

The point is that solving the approximate renormalization conditions exactly corresponds to plugging-in the complete expansions for the unknowns into the incomplete expansions of the renormalization conditions. Of course, one can do that and one can determine the unknowns order by order up-to $n = \infty$ in this way, but needless to say, this would not be very consistent. One would demand, so to speak, too much from the poor N-th order renormalization conditions.

The essential difference between $\overset{\circ}{m}{}^2{}_n$ and Z_n for $n \leq N$ and for $n > N$, as determined from N-th order renormalization conditions, is that those for $n \leq N$ remain unchanged after the higher corrections in renormalization conditions are accounted for, while those for $n > N$ may change significantly. In other words, the values of the coefficients with $n \leq N$ ($n > N$), determined from N-th order renormalization conditions, are exact (unreliable) respectively.

Now for decent (quickly convergent) perturbation series, all this is hardly worth a discussion — the difference between the two approaches is of "the higher order", and as such it should by negligible anyway. But perhaps already in the second to the best of all worlds one can encounter rather slowly convergent perturbation series with numerically quite important (not to mention infinite) higher order corrections. In such case one may be tempted to estimate these higher corrections by not working strictly up-to the given order. Unfortunately, the more important, the less reliable, such estimations seem to be. The reason is that these estimations are nothing but some of the contributions to the complete higher order corrections. And if these partial contributions are sizeable, we have usually no reason to expect the rest to be much smaller.

The moral is that to be on the safe side, i.e. to make statements which will survive even after higher corrections are systematically calculated, one should work strictly only up-to a given order. This would mean not only to determine the bare mass and the wave-function renormalization constant up-to the N-th order, but furthermore to use them in the following way:

- calculate any quantity to the N-th order
- plug in $\overset{\circ}{m}{}^2(\lambda)$ and $Z(\lambda)$ to the N-th order
- expand in λ and trust the result only up-to the N-th order

Of course, the use of other prescription, which is equivalent to some way of estimating of higher order corrections, is not prohibited. The confidence in such corrections, however, is questionable.

Part II

Quantum Electrodynamics (Photons and Electrons)

Part III

The Standard Model (Leptons, Quarks, etc.)

Chapter 5

Particle Physics before the SM

5.1 From nuclear physics to particle physics

Everybody knows since the early childhood that atoms consist of electrons and nuclei. It may come later as a kind of surprise to learn that electrons were discovered more than a decade before atoms and that their discovery in fact meant that atoms of chemists are not atoms (they are not indivisible). Atomic nuclei were observed almost simultaneously with the atoms themselves (existence of atoms was proved in 1909 by the Perrin's confirmation of the Einstein's prediction for the Brownian motion; the atomic nucleus was discovered in 1911 by Rutherford). The first nucleus discovered by Rutherford was that of gold – it was observed by means of the alpha particles, i.e. of the helium nuclei (at that time, however, Rutherford was not aware that the alpha particles are the helium nuclei, even if he knew that they are doubly ionized helium atoms).

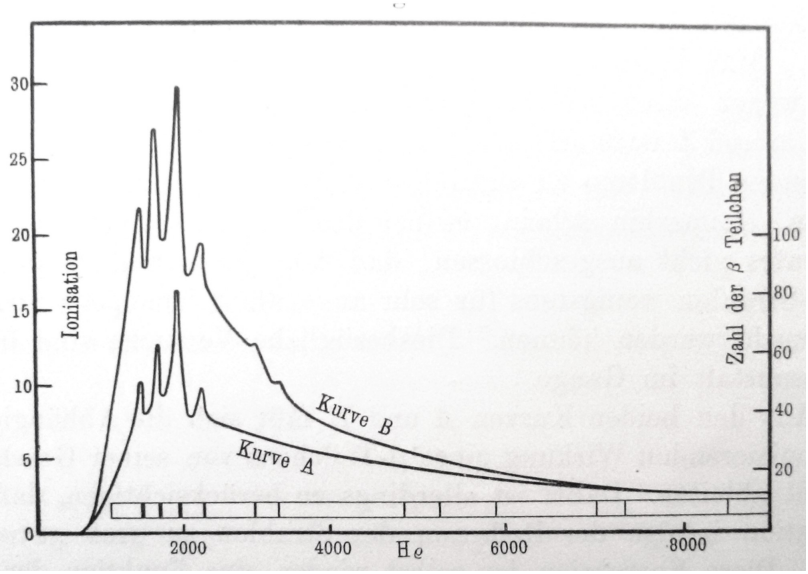
The physics of atoms, i.e. the physics of electrically charged electrons and nuclei became the quantum mechanics (with the electromagnetic field treated classically) and the quantum electrodynamics (with the electromagnetic field treated as a quantum one). These theories were extremely successful not only in explaining atomic spectra, but also in explaining the chemical bond, the periodic table and various properties of materials. Application of the quantum laws in the realm of nuclear physics was not that straightforward. The reason was that at the time of the advent of quantum mechanics neither the constituents of atomic nuclei, nor their interactions were known. The main facts about atomic nuclei known at the early twenties, were these:

- the masses of nuclei are (almost) integer multiples of the mass of the hydrogen nucleus[1]
- the charges of nuclei are integer multiples of the charge of the hydrogen nucleus[2]
- radioactivity (including the emission of electrons) is perhaps a nuclear phenomenon[3]

The most economic (even if wrong) model with these properties is the following one: an atomic nucleus with the relative atomic weight A (rounded to integer) and the atomic number Z consists of A protons and $A-Z$ electrons. In the β-decay one of the electrons is emitted from the nucleus. Models with nuclear electrons had couple of serious problems, the most prominent one being that of nitrogen nucleus spin. This was measured to be one, but for ${}^{14}_{7}N$ with 14 protons and 7 electrons it should be half-integer. A better model of nucleus and its β-decay became possible only after invention of neutrino and discovery of neutron.

[1] The term proton was coined by Rutherford in 1920.
[2] These multiples are different from the ones mentioned in the previous point.
[3] Since it is not affected by any phase or chemical changes, which do affect the orbital electrons.

5.1.1 Neutrino, neutron and Fermi's theory of the β-radioactivity



5.1.2 Pion

1935 Yukawa

he speculated about scalar potential field (0-th component of a 4-vector), we shall take scalar as the simplest possibility, but we should be open-minded as for the other possibiities (pseudoscalar, vector, ...) situation is similar to Fermi theory

predictions: mass of the pion, charge of the pion

1947 $\pi^{\pm}$ Powel, cosmic rays, emulsion

1950 π^0

1951-53 spin of $\pi^{\pm}$

spin, parity,

isospin

5.1.3 Delta

cyclotron

5.2 From the first family to the second one

5.2.1 Muon

5.2.2 Kaon

5.2.3 Lambda, Sigma, Ksi

1947 kaon 1950 lambda

5.2.4 rho, omega, eta

1961 bubble chamber, computers

5.3 From hadrons to quarks