QUANTUM FIELD THEORY

Lecture notes by

Jan Ambjørn
and
Jens Lyng Petersen

$$\int \mathcal{D}[\Phi] e^{iS[\Phi]/\hbar}$$

$$\sum_{\{q_i\}} e^{-H[\{q_i\}]/kT}$$

University of Copenhagen
The Niels Bohr Institute
January 1998

Contents

Chapter 1

The Feynman Path Integral

1.1 The Idea

Modern elementary particle physics and field theory tend to describe quantum mechanical phenomena in either of two ways:(i) The traditional operator or Hilbert space formulation, or (ii) the path integral formulation invented by Feynman in the late 1940'ies (with earlier ideas in the same direction by Dirac).

Although this latter formulation led Feynman to "derive" the Feynman rules of Quantum Electrodynamics in an attractive, intuitive way, his path integral formulation of quantum theory enjoyed little popularity before the 1970'ies. However, in the early 70'ies this situation was rather radically changed for several reasons. First it became gradually clear that gauge field theories represented the right language within which to understand the "fundamental" forces of electromagnetism, weak interactions and strong interactions. For the last two, non-Abelian gauge theories were required, finally leading to the famous *Standard Model* based on the gauge group

$$SU(3)_C \otimes SU(2)_W \otimes U(1)_W$$

(C stands for "colour" and W for "weak"). And it turned out that the technical problems of quantizing non-Abelian gauge theories were handled far more efficiently and elegantly in the path integral formulation.

In a separate development it was becoming clear that there was a very deep similarity between statistical mechanics of a $(D+1)$-dimensional system on the one hand and a quantum theory of a closely related D-dimensional system (with the extra dimension being time) on the other. The relation involves "going to euclidean time", meaning considering purely imaginary values for the time-coordinate. This device had previously been much applied in connection with the practical evaluation of Feynman-diagrams involving loops – the so-called Wick-rotation – but now it took on a more significant role. It lead to two major improvements in people's understanding of field theories.

First it provided a very convenient new framework for mathematically rigorous studies. The idea was to consider the euclidean field theory "on a lattice", i.e. replacing continuous space-time by a discrete set of points, thereby providing what is known as an ultra violet cut-off in the theory. To recover the *continuum* field theory involves a very interesting process, whereby one first has to identify a 2nd order phase transition in the statistical system, at which point correlation lengths grow to "infinity" measured in *lattice units*. This phenomenon may then be reinterpreted in the continuum field theory language by regarding instead the lattice-distance as shrinking to zero while keeping the physical correlation length fixed in *physical units*.

This leads to the second benefit: An entirely new insight was provided on the process of *renormalization* of a quantum field theory. Previously to most people (including Dirac to his death!) renormalization represented a rather suspicious, dirty physicist's procedure for sweeping divergent (i.e. meaningless) loop-integrals under the rug. Now, thanks to work by Wilson in particular, the renormalization process could be treated in a way that was both much more physically and mathematically convincing, and at the same time it was related in an exciting way to critical phenomena in statistical mechanics.

Another major advantage of the path integral quantization is its intuitive appeal: one gets a kind of picture of how a quantum mechanical amplitude is built. It does not at all remove the mysteries of quantum theory, rather it highlights them, but one gets the added intuition.

In the coming several chapters we shall also see examples of how powerful the formalism is: many famous results can be obtained with remarkable ease in the scheme. At this point however, a word of warning should be sounded: The path integral formalism has a tendency to "seduce" one into formal manipulations which are not always justifiable in a more careful analysis. It should always be kept in mind that the manipulations we carry out are to be understood in a framework where a cut-off has been provided (for instance by means of a lattice). And the limit whereby the continuum is recovered has to be carefully analyzed. This is renormalization theory which will be treated only in the second part of the course. Despite this warning it remains very useful to present derivations in such a somewhat simplified manner. After one has studied renormalization theory, one should know in principle at what points care is especially needed.

Last but not least the euclidean path integral formulation has opened up for *computer simulations* which have become possible because the system one studies is equivalent to a statistical mechanical system and complex amplitudes are replaced by positive definite probability distributions, easy to simulate in computer work. At the moment this technique seems to be the most promising one whereby one can study non-perturbative phenomena such as quark confinement in QCD. But the scope of computer simulations is much wider than just that.

As a result of all these nice features, the situation around 1980 appeared to be that the path integral formulation was perhaps going to be the *only* remaining formulation of quantum theory in the future. But developments in the 1980'ies have countered that trend. In fact, one of the outgrowths of string theory has been the so-called two-dimensional conformal field theories, introduced by Belavin, Polyakov and Zamolodchikov. Such theories are required in the *path integral* description of quantum strings: the string traces out a 2-dimensional (one space - one time) surface, the string world-sheet, and things like the position of a point on the string may be regarded as a field variable on this 2-dimensional surface. For "asymptotic" strings, or strings representing fluctuations around a given "string vacuum" (representing space-time among other things — it is not the purpose here to explain string theory so we leave several notions hanging in the air) the ensuing field theory is invariant under conformal transformations of the world sheet coordinates: it is a conformal field theory. The same turns out to be true with the theories arising in the above mentioned limit of a critical 2-dimensional system.

An amazing fact which has been realized in the 1980'ies is that the set of these conformal field theories not only have a tremendously rich structure but also that in infinitely

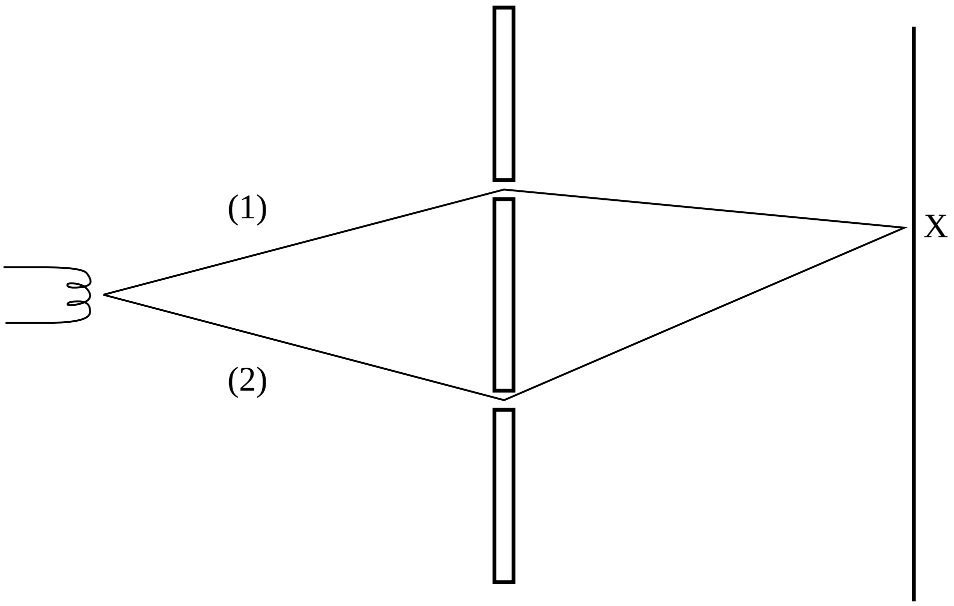


Figure 1.1: The famous double slit "experiment". The amplitude to find the particle at the point, X, is the sum of two sub-amplitudes, representing in turn the amplitudes for the particle to develop its history along the two paths, (1) and (2).

many (nearly all?) cases they can be solved *exactly*, a rare case in the past for a nontrivial theory. The key ingredient in the possibility of exact solution is the existence of an infinite dimensional algebraic structure representing conformal invariance (and generalizations thereof). It now turns out that such theories are best formulated in the *operator* formulation of quantum theory. In fact in most cases it is completely unknown what action would be needed as a starting point for a path integral formulation (see below). Nor is there any need to introduce a cut-off and carry out a renormalization procedure: one is dealing with the final renormalized theory *ab initio*, it is not intrinsically defined via a singular limiting procedure. (Nevertheless it should also be stressed that very interesting comparisons have been made in some of those cases where both formulations are possible.)

In short, the moral for the student of quantum field theory in the 1990'ies is that some knowledge of *both* the operator formulation *and* the path integral formulation is necessary. In this course we shall concentrate on the latter but derive it from the former and make correspondences between the two from time to time.

We finish these introductory remarks by a brief indication of the ideas that led Feynman to the path integral.

Let us consider the famous double slit experiment which illustrates some of the very peculiar features of quantum mechanics, fig. 1.1. As is well known, electrons from the source will form an interference pattern on the photographic plate when both slits are open, but not when only one is. We concentrate on the first case. The standard treatment in quantum mechanics starts from the *wave*-nature of the electron and uses Huygen's principle to work out the interference of elementary waves emanating from the two slits. This calculation produces an *amplitude*, the (absolute) square of which is the probability to find the electron at a certain point on the plate.

Feynman's observation was that we may calculate the very same amplitude using seemingly quite different notions, namely thinking of the electrons as *particles* rather than waves, but accepting that there is an amplitude somehow for the electron to go through slit 1 and a different amplitude for it to go through slit 2, given that it ends up at the same point on the photographic plate. These are the two paths in fig. 1. Thus we would build the total amplitude as

$$A(\text{total}) = A(\text{path } 1) + A(\text{path } 2)$$

Naturally the values of the amplitudes have to be the same as the values of the waves in the first calculation, and Feynman observed that this value could be interpreted as

$$A(\text{path}_i) = e^{\frac{i}{\hbar}S(\text{path}_i)}$$

Here S is the *action* of the system. It plays a fundamental role in the path integral formulation. It may be given as the time integral along the path of the lagrangian, $L(q_i, \dot{q}_i)$ with $\{q_i\}$ being a set of (generalized) coordinates and $\{\dot{q}_i\}$ the associated velocities. Then

$$S(\text{path}) = \int_{\text{path}} L(q_i(t), \dot{q}_i(t))dt$$

where the path is defined by the map:

$$t \mapsto q_i(t), \; t \in \text{some time interval}$$

Given one set of boundary conditions (in our case that the electron leaves the source and hits the plate at a given point), there is a *unique* classical path, namely the one which minimizes the action. It is given by that solution to the Euler-Lagrange *equations of motion*:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}$$

which has the right boundary conditions. For a simple non-relativistic particle of mass m, moving in a conserved force field described by a potential $V(q_i)$, the Lagrangian is given by

$$L(q_i, \dot{q}_i) = \frac{1}{2}m\sum_j \dot{q}_j^2 - V(q_i)$$

giving of course Newton's 2nd law

$$m\ddot{q}_i = -\partial_i V(q_j) \equiv F_i$$

($\partial_i \equiv \partial/\partial q_i$) but this is special for non-relativistic mechanics. In general we should just think of the action as some *Lorentz-invariant* functional of the path. Different functionals give rise to different theories. In practice we shall also require the existence of a lagrangian.

Feynman was intrigues by the fact that classically only an infinitesimal amount of the information stored in the action is physically relevant: only the path for which the action is extremal (usually minimal) is of interest.

But now we are led to conjecture (and prove in section 1.3) that a quantum mechanical amplitude may be represented as

$$A = \sum_{\text{all paths}} e^{\frac{i}{\hbar}S(\text{path})}$$

where the sum goes over physical as well as unphysical paths (classically speaking). They should all, however, conform to appropriate *boundary conditions*, specifying what amplitude more precisely we want to calculate. In general the amplitude will describe the transition between two states, and the boundary conditions should contain the same sort of information that is needed to specify the states. This will be the subject of some discussion in coming sections. Also we shall have to make more precise what "the sum over all paths" actually means. It is perhaps clear that such a strange object can only be given a meaning after the introduction of a "cut-off", and a subsequent limiting procedure. In field theory the problem is far more complicated still, by the fact that instead of a finite number of degrees of freedom, such as we have here – the q_i's –, we have an infinite number of degrees of freedom, namely one or more field values at each space point, $\vec{x}$: $q_i(t)$ is replaced by $\phi_r(\vec{x},t)$, so that the discrete index, i is replaced by the continuous "index", $\vec{x}$ (possibly in association with one or more discrete indices, here collected in the letter, r). Again we see the need for a cut-off. This of course is at the centre of renormalization in field theory.

We emphasize that with this path integral formulation of quantum theory, all objects involved are like "classical" ones: fields for example take ordinary numerical values, they are not operators. The *difference* from classical mechanics is that not only the classical path, satisfying the classical equations of motion, is included in the "sum over paths": all paths are. This provides a somewhat intuitive feel for what "quantum fluctuations " mean.

Another interesting intuitive qualitative consequence of Feynman's expression for the amplitude follows immediately:

Consider a situation where quantum mechanics plays a minor role. What does that mean? We see that it means that the "typical action" is much larger than the fundamental Planck action quantum, $\hbar$. In that case, namely we may expect that contributions in the sum over all paths have a tendency to wash out – cancel, because the phase $\exp[iS/\hbar]$ varies rapidly for different paths: two neighbouring paths will have action values differing by many units of $\hbar$. The only exception to this argument is for paths in the neighbourhood of the classical path: here neighbouring paths have essentially the same value of S since by definition the action is stationary (minimal) here. Hence the path integral should be dominated by the classical path in these "near classical" situations. In this sense classical mechanics can be said to be a consequence of the path integral in the appropriate limit. This should help us to accept the path integral formulation of quantum mechanics as a fundamental starting point.

1.2 The Action

1.2.1 Local Field Theories

In this course we shall deal with quantum field theories for which we can write down an action, so that we can get the path integral machinery to work. Further we shall restrict ourselves to *local* theories. Let us explain what this means, and provide some examples.

First we require the physical system, i.e. our field theory, to be specified by "(*very*) generalized coordinates", namely: a classical configuration of the system at a given time,

t, is specified by a *field*, a map

$$(t, \vec{x}) \mapsto \phi_r(t, \vec{x})$$

where the field ϕ_r takes values in a certain set, often denoted *the target space.* In the simplest case of a scalar field, this is just the set of real or complex numbers. For electromagnetism we would rather use the 4-vector potential, A^μ taking values in the (vector-) space of objects that transform as 4-vectors under Lorentz transformations. In that case the situation is further complicated by gauge-invariance which makes the precise definition of the target space somewhat more involved. Other complications will be met in later sections where we introduce *Grassmann*-numbers ("anticommuting numbers", see chapter 3.) appropriate for the description of fermionic fields, such as the electron field $\psi(x)$. Such a field is further characterized by belonging to the set of Dirac 4-spinors (in 4-dimensional space-time). For the gauge field theories of the standard model, the degrees of freedom, i.e. the generalized coordinates on the target space, are in a certain sense *geometrical*: things like the gauge potential $A^\mu(x)$ or the Fermi field $\psi(x)$ are to be compared to connections and vector fields, described by coordinates on an abstract differentiable manifold. Only the "geometrical properties" of the manifold are the real degrees of freedom. In practice this situation is dealt with as in geometry: one uses coordinate descriptions (here A^μ and ψ) but demands geometric (here physical) properties to be independent of coordinate transformations, here: gauge transformations. Complications arising from these effects will be treated in detail in the second part of this course.

Having identified the *configuration space*, namely the set of maps (fields) from space-time (Minkowski- or otherwise, see later) to the target space, for the kind of systems (field theories) we shall be dealing with, we now give an idea of the kind of actions we shall have in mind.

In general the action is a functional of a field: a given field development in space and time is mapped into one number, the action. It is given as the time integral over the *Lagrangian*:

$$S[\phi_r] = \int dt L[\phi_r(t, \cdot), \dot{\phi}_r(t, \cdot)]$$

where the lagrangian, L itself is a functional of the field configuration, one, however, that only depends on the field values at a given *time.* Thus a Lorentz-non-invariant, inertial-frame-dependent object, contrary to the action, which is fully invariant under Lorentz-transformations.

To the reader, who asks: Why? Well, invariance of the action is taken as an axiom. In practice we construct relativistically invariant theories by providing invariant action-functionals.

We furthermore restrict ourselves to *local* theories for which the lagrangian can be expressed as a local integral over the *lagrangian density*, $\mathcal{L}$. That in turn is taken to be an ordinary *function* of the field values and a finite number of field derivatives (in practice we shall restrict ourselves to the first order derivatives), all referred to a given space-time point x:

$$\mathcal{L} = \mathcal{L}(\phi_r(x), \partial_\mu \phi_r(x))$$

Then, for

$$(x) = (x^\mu) = (x^0, x^1, x^2, x^3) = (t, \vec{x})$$

we have

$$L[\phi_r(t,\cdot)] = \int d^3x \mathcal{L}(\phi_r(x), \partial_\mu \phi_r(x))$$

and the action takes the invariant form

$$S[\phi_r] = \int d^4x \mathcal{L}(\phi_r(x), \partial_\mu \phi_r(x))$$

showing that the lagrangian *density* is a convenient, invariant object, like the action itself, but contrary to the lagrangian.

Notice, that in units where $\hbar = c = 1$, the action is dimensionless (like $\hbar$) whereas $\mathcal{L}$ has dimension $L^{-d} = E^d$ where d is the dimension of space-time and L and E denote lengths and energies, respectively. It is often of interest to consider field theories in space-time dimensions other than 3+1. Most of the examples given below will make immediate sense for general d.

Our convention for Minkowskian metric etc. is as follows: upper (covariant) and lower (contravariant) Lorentz indices, usually denoted by small *greek* letters, are raised and lowered by the Minkowski-metric

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

So

$$\eta^\mu_\nu = \delta^\mu_\nu$$

We use the notation

$$\partial_\mu \equiv \partial/\partial x^\mu, \ \partial^\mu \equiv \partial/\partial x_\mu$$

$$k \cdot x \equiv k^\mu x_\mu = k_\mu x^\mu = \vec{k} \cdot \vec{x} - k^0 x^0$$

Hence, for the 4-momentum p of a particle of mass m we have

$$p^2 = -m^2$$

We now provide a sample of actions (or lagrangian densities, rather) for various field theories to illustrate the sort of things we have been describing:

1. *Scalar field theory*

   We take the field configuration to be specified by a number $R \in \mathsf{N}$ of real, scalar fields $\phi_r(x)$, $r = 1, ..., R$.

   We consider $\mathcal{L}$ being a sum of terms (summation on r implied):

   $$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_I + \mathcal{L}_J$$

   where

   $$\begin{aligned} \mathcal{L}_0 &= -\frac{1}{2}(\partial_\mu \phi_r(x) \partial^\mu \phi_r(x) + m_r^2 \phi_r(x) \phi_r(x)) \\ \mathcal{L}_I &= \sum_{n=3}^{N} \sum_{\{r_i\}} \lambda^{(n)}_{\{r_i\}} \phi_{r_1}(x) \cdot \phi_{r_2}(x) \cdot \ldots \cdot \phi_{r_n}(x) \\ \mathcal{L}_J &= J_r(x) \phi_r(x) \end{aligned} \qquad (1.1)$$

Here $\mathcal{L}_0$ describes a free non-interacting part quadratic in the fields (see later), $\mathcal{L}_I$ describes the self-interactions among the scalar fields of at least cubic power in the fields, and $\mathcal{L}_J$ is a source term, linear in the fields, describing in a generic way how the fields could couple to external sources, $J_r(x)$. In practice such source terms will allow a mathematically convenient way of producing generating functionals for field theory amplitudes: Greens functions, S-matrix elements, etc.

In 4-dimensional space-time (and higher) most of these theories turn out to be inconsistent as *quantum*-field theories unless

$$\lambda^{(n)}_{\{r_i\}} \equiv 0, \text{ for } n > 4$$

Only then will they be *renormalizable.* We refer to the second part of the course for discussion. A famous example for $R = 1$ is the $\lambda\phi^4$-theory:

$$\mathcal{L}_{\lambda\phi^4} = -\frac{1}{2}[\partial_\mu\phi\partial^\mu\phi + m^2\phi^2] + \frac{\lambda}{4!}\phi^4$$

In 2-dimensional space-time another famous example is the *Liouville* theory

$$\mathcal{L}_{\text{Liouville}} = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \mu^2 e^\phi$$

In 2 dimensions there is no restriction coming from renormalizability on the maximal power N above in the self-interactions.

2. *U(1)-gauge theory of Dirac Fermions: QED*

$$\mathcal{L}_{QED} = -\overline{\psi}(x)[\not{D} + m]\psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x)$$

Here $\psi(x)$ is a Dirac-spinor (Grassmann-number valued, see Chapter 3),

$$D_\mu \equiv \partial_\mu - ieA_\mu(x)$$

is a gauge-covariant derivative with $A_\mu(x)$ being the gauge potential and e the coupling constant to the field ψ. Finally

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$$

is the field strength tensor. The slash through the covariant derivative involves the notation

$$\not{a} \equiv a_\mu\gamma^\mu$$

where the γ^μ's are the Dirac gamma-matrices, satisfying the (anti-) commutation relations

$$\{\gamma_\mu, \gamma_\nu\} = 2\eta_{\mu\nu}I$$

with I the 4×4 unit matrix (in 4 dimensions).

$$\gamma_i = \gamma^i = \gamma_i^\dagger$$

and

$$\gamma_0 = -\gamma^0 = -\gamma_0^\dagger$$

in this metric. Notice that the lagrangian (density) may be written as a free Fermi part, bilinear in ψ, a free "photon" part, bilinear in A_μ's and an interaction part trilinear in fields. This split-up will be used for developing perturbation theory, however it breaks the gauge-invariance otherwise possessed by the theory. Also, to actually carry out the perturbation theory programme, additional "gauge-fixing" terms will have to be considered. However, none of that will concern us here.

3. *The Yukawa coupling*

   In the Standard Model, mass-terms for all fundamental fields are forbidden by gauge-invariance. However, by couplings to "Higgs"-fields with non-vanishing vacuum expectation values, field excitations nevertheless may *appear* massive. The coupling of Fermions to scalars is provided by so-called Yukawa terms of a form like

   $$\mathcal{L}_{\text{Yukawa}} = g_r \overline{\psi}(x)\psi(x)\phi_r(x)$$

   So that for a scalar field $\phi_r(x)$ constant in space and time, this indeed looks like a mass term. g_r is a coupling constant.

4. *Non-linear sigma model*

   In this example, the target-space of the field is a group-manifold, or more precisely, the set of matrices which form a certain *representation* of the group:

   $$x \mapsto g_{ij}(x)$$

   where i, j are matrix indices running from 1 to the dimension of the representation in question. A possible lagrangian with interesting group invariance properties is

   $$\mathcal{L}_{NLSM} = Tr[\partial_\mu g^{-1} \partial^\mu g]$$

5. *The Einstein-Hilbert action*

   In all of the above examples, space-time was assumed to be Minkowskian, so that the space-time metric is given by $\eta_{\mu\nu}$. In most cases, however, it is relatively straight forward, using the methods of general relativity, to generalize the actions into generally covariant forms, valid in an arbitrary gravitational background space-time, given by a metric tensor $g_{\mu\nu}(x)$. It is also well known that we may write down an elegant action for gravity itself, namely the Einstein-Hilbert action

   $$S_{EH} = \int d^4x \sqrt{-g(x)} R(x)$$

   where $g(x)$ denotes the determinant of the metric tensor and $R(x)$ the curvature scalar. It turns out, however, that this action does not lead to a renormalizable quantum field theory of gravity. That a theory is non-renormalizable, really means, that in principle many different consistent quantum theories could result in the same "low-energy" semi-classical theory. Each of these different theories contain their own *new physics.* Which one of the possibilities is the correct one, cannot be decided merely by the fact that it reproduces Einstein gravity. Thus quantum gravity must be dealt with in a profoundly different way, possibly via string theory. In this course we shall not consider quantum gravity.

These examples should provide some idea of the kind of situations we shall be interested in. I this first part of the course we shall try to be rather general and mostly not have any one particular field theory in mind, even though the general techniques will often be treated by way of an example field theory.

1.2.2 Functional Methods.

A large part of this course will be taken up by *perturbation theory.* This may be seen as a way of allowing in a systematic way for field configurations in the neighbourhood of the classical solution. Hence, even though we are concerned with *quantum* field theory, the classical solutions will play a crucial role.

A classical solution to a field theory is singled out by two requirements: (i) It corresponds to definite *boundary conditions* reflecting the physics of the situation, and (ii) it minimizes the action under the given set of boundary conditions. The second requirement leads to the Euler-Lagrange equations of motion, which we now consider. Let $\phi_r(x)$ denote a generic set of fields, and let $\mathcal{L} = \mathcal{L}(\phi_r, \partial_\mu\phi_r)$. We now work out a variation of the action corresponding to fixed boundary conditions:

$$\begin{aligned} \delta S(\phi_r) &= \int d^dx[\frac{\partial\mathcal{L}}{\partial\phi_r}\delta\phi_r(x) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\delta\partial_\mu\phi_r(x)] \\ &= \int d^dx[\frac{\partial\mathcal{L}}{\partial\phi_r} - \partial_\mu\{\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)}\}]\delta\phi_r(x) \end{aligned} \tag{1.2}$$

Here, in the second term we did the following

1. We used $\delta\partial_\mu\phi_r(x) = \partial_\mu\delta\phi_r(x)$ which follows from

$$\delta\partial_\mu\phi_r(x) = \partial_\mu[\phi_r(x) + \delta\phi_r(x)] - \partial_\mu\phi_r(x)$$

2. We performed a partial integration *for fixed boundary conditions.* This allowed us to discard the boundary terms in the partial integration (since $\delta\phi_r(x)$ vanishes on the "boundaries").

The above expression for the variation of S has to be valid for arbitrary field variations. This gives the EL-equations of motion:

$$\frac{\partial\mathcal{L}}{\partial\phi_r} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r)} = 0 \tag{1.3}$$

Now let us use the derivation of the equations of motion as an excuse to introduce some *functional techniques and notation.* In following sections this will prove very convenient. First let us introduce the *functional derivative.* To motivate the definition, consider first an ordinary set of variables $\{x_i\}$, $i = 1, 2, ...$ When these are independent we clearly have

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}$$

Actually we may use instead the name $\{\phi_i\}$ for the variables and write

$$\frac{\partial\phi_i}{\partial\phi_j} = \delta_{ij}$$

We may also think of a *fixed set* of the variables as a map from the integers, $i \in \mathsf{N}$ to the real or complex numbers:

$$i \mapsto \phi_i$$

Our *variables*, $\{\phi_i\}$, correspond to the *set* of all possible such maps.

Now we generalize this to maps from space-time points (x) in d dimensions to some target space. Such maps are our *fields.* And we define the functional derivative by the rule

$$\frac{\delta\phi_r(x)}{\delta\phi_s(y)} \equiv \delta^d(x-y)\delta_{rs}$$

This formalism in fact is really nothing but a particularly convenient way of doing variational calculus. As an example, consider the (highly unrealistic) case where the lagrangian density is independent of field derivatives. And consider one field only. Then for the action

$$S[\phi] = \int d^dx \mathcal{L}(\phi(x))$$

we may rederive the variation in the new language as

$$\begin{aligned} \frac{\delta S[\phi]}{\delta\phi(y)} &= \int d^dx \frac{\partial\mathcal{L}}{\partial\phi(x)}\frac{\delta\phi(x)}{\delta\phi(y)} \\ &= \int d^dx \frac{\partial\mathcal{L}}{\partial\phi(x)}\delta^d(x-y) \\ &= \frac{\partial\mathcal{L}}{\partial\phi(y)} \end{aligned} \tag{1.4}$$

We see that the idea of the functional derivative is nothing but a convenient way of doing variational calculus. Now, to actually get something analogous to the equations of motion, we must understand also the ∂_μ type terms. Therefore, quite generally let us consider linear operators on (classical) fields. These may be thought of as generalizing linear operators on *vectors*, (ϕ_i):

$$A : \phi_i \mapsto (A\phi)_i = \sum_j A_{ij}\phi_j$$

showing how the linear operator is realized as a matrix.

Analogously for fields $\phi(x)$ rather than finite dimensional vectors (ϕ_i), we may think of linear operators in terms of "functional matrices" or *integration kernels:*

$$A : \phi(x) \mapsto (A\phi)(x) = \int d^dy A(x,y)\phi(y)$$

Very often we shall use a similar idea also when the functional matrices are actually distributions. As a good example, consider the derivative operation, which indeed is linear:

$$\partial_\mu : \phi(x) \mapsto \partial_\mu\phi(x) = \int d^dy \frac{\partial}{\partial x^\mu}\delta^d(x-y)\phi(y) = -\int d^dy \frac{\partial}{\partial y^\mu}\delta^d(x-y)\phi(y)$$

One sees by partial integration (provided that is allowed without boundary terms!) that also the last expression gives the correct result. Thus ∂_μ has the integration kernel (functional matrix element)

$$\partial_\mu(x,y) = \frac{\partial}{\partial x^\mu}\delta^d(x-y) = -\frac{\partial}{\partial y^\mu}\delta^d(x-y)$$

Notice that when we differentiate a vector with respect to a vector-coordinate, the result is a *matrix*:

$$\frac{\partial}{\partial x_i}(Ax)_n = \frac{\partial}{\partial x_i}\sum_m A_{nm}x_m = A_{ni}$$

In a very compressed matrix notation we may sometimes write this as

$$\frac{\partial}{\partial x}Ax = A$$

Similar compressed notations are sometimes used in the functional case. For example we might write

$$\frac{\delta}{\delta\phi}\partial_\mu\phi = \partial_\mu$$

as a short hand of the expression with the kernel distribution. When in doubt one should always reintroduce the "generalized matrix elements" explicitly as above.

Now we may summarize the variational calculation leading to the equations of motion as

$$\frac{\delta}{\delta\phi_r(x)}S[\phi] = \frac{\partial\mathcal{L}}{\partial\phi_r(x)} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_r(x))}$$

And we may conclude that the classical field is the one, which (i) satisfies the appropriate boundary conditions, and (ii) for which the functional derivative of the action with respect to the field vanishes identically.

Finally we mention that we may construct the *hamiltonian density* as follows. First we find the *momentum density* conjugate to the field $\phi_r(x)$ as

$$\pi_r(x) = \frac{\partial\mathcal{L}}{\partial\dot{\phi}_r(x)} = \frac{\delta S[\phi]}{\delta\dot{\phi}_r(x)} \tag{1.5}$$

then the hamiltonian density is given by

$$\mathcal{H}(\phi_r(x), \partial_i\phi_r(x), \pi_r(x)) = \sum_s \pi_s(x)\dot{\phi}_s(x) - \mathcal{L} \tag{1.6}$$

where it is understood that when $\dot{\phi}_r(x)$ appears in what remains on the right hand side it should be reexpressed in terms of $\pi_r(x)$ using the definition.

1.2.3 Free Field Theory as a Sum of Harmonic Oscillators

In sections 4 and 5 we shall treat in detail the harmonic oscillator from various points of view. It may seem strange that one has to study a one dimensional non-relativistic mechanical system in order to learn about relativistic quantum field theory. However, as we shall now illustrate, a relativistic, *free* field (with an infinite number of degrees of freedom) is equivalent to an infinite set of harmonic oscillators. The intuitive physical reason is not surprising: we can set up harmonic plane waves in the field, ones for every choice of momentum (i.e. wave-) vector. This should motivate our efforts later on.

For simplicity we concentrate on a single scalar field with lagrangian (density)

$$\mathcal{L}(\phi) = -\frac{1}{2}(\partial_\mu\phi\partial^\mu\phi + m^2\phi^2) \tag{1.7}$$

It is trivial to work out the equations of motions and find the Klein-Gordon equation

$$(\partial_\mu \partial^\mu - m^2)\phi(x) = 0 \tag{1.8}$$

the most general solution of which may be expressed in Fourier modes or momentum space as

$$\phi(x) = \sum_{\vec{k}} [a(\vec{k})e^{ikx} + a^*(\vec{k})e^{-ikx}] \tag{1.9}$$

Here k^0 is given by the positive energy mass shell condition

$$k^2 = -m^2, \quad k^0 = +\sqrt{\vec{k}^2 + m^2}$$

and we use the notation

$$\begin{aligned} \sum_{\vec{k}} &\equiv \int \frac{d^3\vec{k}}{2k^0(2\pi)^3} \\ \delta_{\vec{k}\vec{k}'} &\equiv 2k^0(2\pi)^3\delta^3(\vec{k} - \vec{k}') \end{aligned} \tag{1.10}$$

in 3 space dimensions. Then of course

$$\sum_{\vec{k}} \delta_{\vec{k}\vec{k}'} = 1$$

In the operator formulation of quantum field theory, the variables, $a(\vec{k})$ and $a^*(\vec{k})$ become annihilation and creation operators respectively, $\hat{a}(\vec{k})$ and $\hat{a}^\dagger(\vec{k})$ corresponding to "single particle" field excitations of momentum, $\vec{k}$, energy k^0 and mass, m. Continuing with the classical case we find

$$\begin{aligned} \pi(x) &= \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)} = \dot{\phi}(x) \\ \mathcal{H}(\phi, \partial_i\phi, \pi) &= \frac{1}{2}[\pi^2(x) + \vec{\nabla}\phi(x)^2 + m^2\phi^2(x)] \end{aligned} \tag{1.11}$$

Let us find the free hamiltonian

$$\begin{aligned} H_0 &= \int d^3x \mathcal{H} \\ &= \frac{1}{2}\int d^3x \sum_{\vec{k},\vec{k}'} \Big\{ -(k^0k^{0\prime} + \vec{k}\cdot\vec{k}')\,[a(\vec{k})e^{ikx} - a^*(\vec{k})e^{-ikx}][a(\vec{k}')e^{ik'x} - a^*(\vec{k}')e^{-ik'x}] \\ &+ m^2[a(\vec{k})e^{ikx} + a^*(\vec{k})e^{-ikx}][a(\vec{k}')e^{ik'x} + a^*(\vec{k}')e^{-ik'x}]\ \Big\} \\ &= \frac{1}{2}\int d^3x \sum_{\vec{k},\vec{k}'} \{ \\ &\quad a(\vec{k})a(\vec{k}')e^{-it(k^0+k^{0\prime})}e^{i\vec{x}\cdot(\vec{k}+\vec{k}')}(m^2 - (k^0k^{0\prime} + \vec{k}\cdot\vec{k}')) \\ &+ a^*(\vec{k})a^*(\vec{k}')e^{it(k^0+k^{0\prime})}e^{-i\vec{x}\cdot(\vec{k}+\vec{k}')}(m^2 - (k^0k^{0\prime} + \vec{k}\cdot\vec{k}')) \\ &+ a(\vec{k})a^*(\vec{k}')e^{it(k^{0\prime}-k^0)}e^{i\vec{x}\cdot(\vec{k}-\vec{k}')}(m^2 + (k^0k^{0\prime} + \vec{k}\cdot\vec{k}')) \\ &+ a^*(\vec{k})a(\vec{k}')e^{it(k^0-k^{0\prime})}e^{i\vec{x}\cdot(\vec{k}'-\vec{k})}(m^2 + (k^0k^{0\prime} + \vec{k}\cdot\vec{k}'))\ \} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2}\sum_{\vec{k}} \frac{1}{2k^0}\left\{a(\vec{k})a(-\vec{k})e^{-2itk^0}(m^2-m^2)\right.\\
&+ a^*(\vec{k})a^*(-\vec{k})e^{2itk^0}(m^2-m^2)\\
&+ \left.(a(\vec{k})a^*(\vec{k})+a^*(\vec{k})a(\vec{k}))((k^0)^2-\vec{k}^2+(k^0)^2+\vec{k}^2)\right\}\\
&= \frac{1}{2}\sum_{\vec{k}} k^0(a(\vec{k})a^*(\vec{k})+a^*(\vec{k})a(\vec{k}))
\end{aligned} \tag{1.12}$$

Let us compare this with the well known 1-dimensional harmonic oscillator with mass $m = 1$ and angular frequency ω. The hamiltonian may be written

$$H_{HO}(p,q) = \frac{1}{2}(p^2+\omega^2 q^2)$$

with $p =$ momentum and $q =$ position. Let us change variables to

$$\begin{aligned}
a &= \frac{1}{\sqrt{2\omega}}(\omega q + ip) \qquad & a^* &= \frac{1}{\sqrt{2\omega}}(\omega q - ip)\\
p &= -i\sqrt{\frac{\omega}{2}}(a-a^*) \qquad & q &= \frac{1}{\sqrt{2\omega}}(a+a^*)
\end{aligned} \tag{1.13}$$

Then

$$H = \frac{1}{2}\omega(aa^*+a^*a)$$

Now the comparison is evident: The relativistic free scalar field represents a sum of independent harmonic oscillators, one for each value of the momentum 3-vector, $\vec{k}$.

Quantum mechanically, p and q become operators, $\hat{p}, \hat{q}$ with

$$[\hat{p},\hat{q}] = -i\hbar = -i$$

And a, a^* become the operators $\hat{a}, \hat{a}^\dagger$ with

$$[\hat{a},\hat{a}^\dagger] = 1$$

the usual annihilation and creation operators for oscillator phonon excitations.

As is well known, a more careful treatment of the quantum hamiltonian yields the result

$$\hat{H} = \frac{1}{2}\omega(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}) = \omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$$

The operator $\hat{a}^\dagger\hat{a}$ is the number operator , the eigenvalue of which is zero on the oscillator ground state, the "vacuum". The term $\frac{1}{2}\omega$ is a "zero point fluctuation energy". Similar remarks would apply to the quantum field case.

1.3 Quantum Mechanics

1.3.1 The Feynman Path Integral in Phase Space

We start slowly by considering a physical system with a single degree of freedom, position, q. Conjugate to that we have a single momentum variable, p. As an example we shall soon study the one dimensional harmonic oscillator.

Our first objective will be to derive a path integral expression for the *transition amplitude* defined as follows:

Suppose the particle is known to be at position q at time t, what is the quantum mechanical amplitude to find the particle at position q' at time $t' > t$? The answer is given by the transition amplitude

$$F(q',t';q,t) = \text{ transition amplitude}$$

First we establish the standard expression in terms of the hamiltonian of the system. Let $|q,t\rangle_H$ be the *Heisenberg*-state-vector describing the initial knowledge. Notice that it is independent of time as a Heisenberg state vector should be. The t-label is a specification: the particle is localized at the time value, t. For $\tau = t$ the state $|q,t\rangle_H$ is an eigenstate of the time dependent Heisenberg position operator $\hat{q}_H(\tau)$:

$$\hat{q}_H(\tau = t)|q,t\rangle_H = q|q,t\rangle_H$$

whereas $|q,t\rangle_H$ is not an eigenstate of $\hat{q}_H(t)$ for $\tau \neq t$. The time-independent Schrödinger position operator $\hat{q}_S$ is obtained by

$$\hat{q}_H(t) = e^{i\hat{H}t}\hat{q}_S e^{-i\hat{H}t}$$

where $\hat{H}$ is the quantum hamiltonian of the system. The state vector $|q\rangle$ defined by

$$|q\rangle = e^{-i\hat{H}t}|q,t\rangle_H$$

is independent of the time label and is an eigenstate of $\hat{q}_S$:

$$\begin{aligned} \hat{q}_S|q\rangle &= \hat{q}_S e^{-i\hat{H}t}|q,t\rangle_H = e^{-i\hat{H}t}\hat{q}_H(t)|q,t\rangle_H = qe^{-i\hat{H}t}|q,t\rangle_H \\ &= q|q\rangle \end{aligned} \tag{1.14}$$

Using the definition we now obtain the formulas

$$F(q',t';q,t) = {}_H\langle q',t'|q,t\rangle_H = \langle q'|\exp\{-i\hat{H}(t'-t)\}|q\rangle \tag{1.15}$$

This is the traditional quantum mechanical expression. We now derive a path integral expression for this.

Let us subdivide the time interval from t to t' in $n+1$ small pieces of length

$$\epsilon = \frac{t'-t}{n+1}$$

with

$$\begin{aligned} t_0 &= t \\ t_1 &= t+\epsilon \\ &\vdots \\ t_i &= t+i\epsilon \\ &\vdots \\ t_{n+1} &= t' \end{aligned} \tag{1.16}$$

we now use the fact that for every value, t_i of the time label, the set of states

$$\{|q_i,t_i\rangle_H | q_i \in \mathsf{R}\}$$

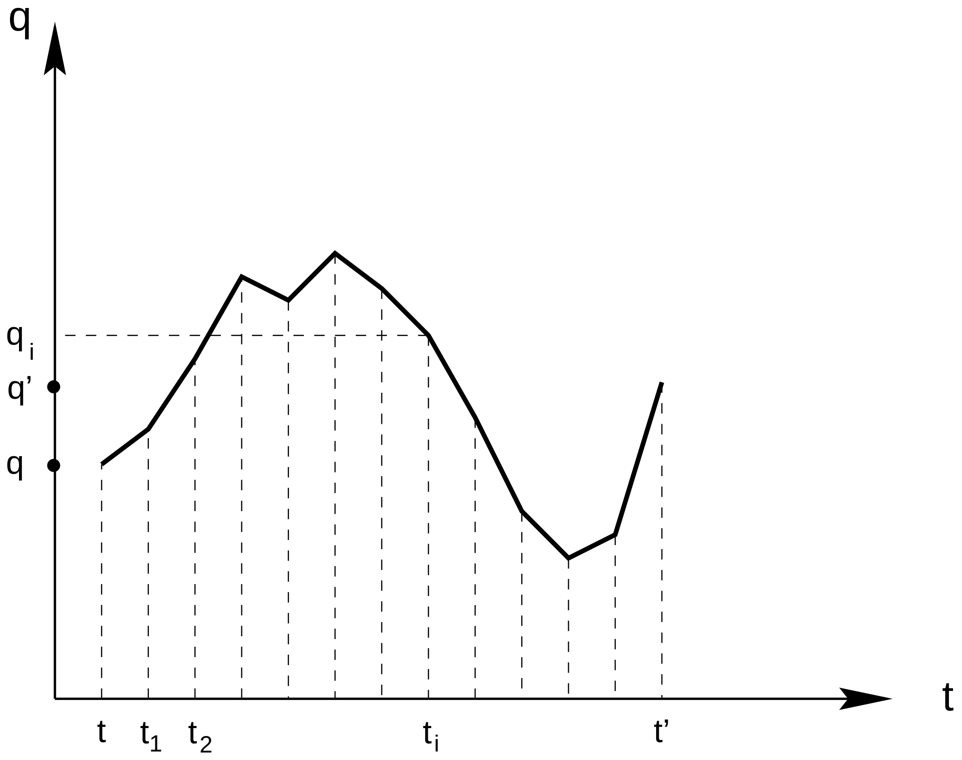


Figure 1.2: An “arbitrary” discretized path

form a complete set so that the unit operator may be expressed as

$$\int dq_i |q_i, t_i\rangle\langle q_i, t_i|$$

where for simplicity we dropped the Heisenberg label. Hence we may express the transition amplitude as

$$\begin{aligned} F(q', t'; q, t) &= \langle q', t'|q, t\rangle \\ &= \int dq_1 dq_2 \cdots dq_n \langle q', t'|q_n, t_n\rangle\langle q_n, t_n|q_{n-1}, t_{n-1}\rangle \cdots \\ &\quad \cdots |q_1, t_1\rangle\langle q_1, t_1|q, t\rangle \end{aligned} \tag{1.17}$$

For a fixed set of the integration variables

$$q_i \equiv q(t_i)$$

we may think of the “function” $q(t_i)$ as representing a certain path taken by the particle. Here it is inherently defined with a “regularization”: we have introduced a discrete set of times, t_i, but we may imagine that in the limit $n \to \infty$ these paths become “sufficiently close to” the actual possible physical paths for a particle not subject to classical equations of motion, cf. fig.1.2. In this limit we shall further allow ourselves the notation

$$\mathcal{D}q(t) \equiv \prod_{i=1}^{n} dq(t_i)$$

In this sense the above expression is already a path integral: We have what we shall call a functional integral over “all” paths.

We next want to relate the integrand to Feynman's expression e^{iS}. Rather than presenting the very shortest treatment let us be slightly more general than Feynman and first introduce yet another complete set of states, namely the momentum states, writing

$$\begin{aligned} {}_H\langle q(t_i), t_i|q(t_{i-1}), t_{i-1}\rangle_H &= \langle q(t_i)|e^{-i\epsilon\hat{H}}|q(t_{i-1})\rangle \\ &= \int \frac{dp(t_i)}{2\pi}\langle q(t_i)|p(t_i)\rangle\langle p(t_i)|e^{-i\epsilon\hat{H}}|q(t_{i-1})\rangle \end{aligned} \tag{1.18}$$

Notice the normalizations used:

$$\begin{aligned} |q\rangle &= \int \frac{dp}{2\pi}|p\rangle\langle p|q\rangle = \int \frac{dp}{2\pi}e^{-ipq}|p\rangle \\ |p\rangle &= \int dq|q\rangle\langle q|p\rangle = \int dq e^{ipq}|q\rangle \\ &= \int dq\int\frac{dp'}{2\pi}e^{iq(p-p')}|p'\rangle \\ &= \int\frac{dp'}{2\pi}2\pi\delta(p-p')|p'\rangle \\ &= |p\rangle \end{aligned} \tag{1.19}$$

showing the consistency.

Let

$$H = H(p, q)$$

be the *classical* hamiltonian expressed as a function of (think of a polynomium in) p and q. The *quantum* hamiltonian $\hat{H}(\hat{p}, \hat{q})$ is constructed by replacing the classical variables by their quantum equivalents. However, because $\hat{p}$ and $\hat{q}$ do not commute, there is an ambiguity if the dependencies on (p, q) is not separable. In most of our applications they will in fact be separable, the harmonic oscillator being our favorite example:

$$H_{HO}(p, q) = \frac{1}{2}(p^2 + \omega q^2)$$

In general we shall *define* the meaning of the quantum hamiltonian such that it is ordered in such a way that all $\hat{p}$'s stand to the left of all $\hat{q}$'s. In that case we may then (to first order in ϵ) express the matrix elements as

$$\begin{aligned} \langle p(t_i)|e^{-i\epsilon\hat{H}(\hat{p},\hat{q})}|q(t_{i-1})\rangle &= e^{-i\epsilon H(p(t_i),q(t_{i-1}))}\langle p(t_i)|q(t_{i-1}\rangle \\ &= \exp\{-i[p(t_i)q(t_{i-1}) + \epsilon H(p(t_i), q(t_{i-1}))]\} \end{aligned} \tag{1.20}$$

We may now express the transition amplitude as the following *path integral in phase space*

$$\begin{aligned} F(q', t'; q, t) &= \int \prod_{i=1}^{n+1}\frac{dp(t_i)}{2\pi}\prod_{j=1}^{n} dq(t_j) \\ & \quad \langle q(t_{n+1})|p(t_{n+1})\rangle\langle p(t_{n+1})|e^{-i\epsilon\hat{H}}|q(t_n)\rangle\cdots \\ & \quad \cdots\langle q(t_1)|p(t_1)\rangle\langle p(t_1)|e^{-i\epsilon\hat{H}}|q(t_0)\rangle \\ &= \int \mathcal{D}p(t)\mathcal{D}q(t)\exp\{i[p(t_{n+1})(q(t_{n+1}) - q(t_n)) + \cdots + p(t_1)(q(t_1) - q(t_0)) \\ &- \epsilon(H(p(t_{n+1}), q(t_n)) + \cdots + H(p(t_1), q(t_0)))]\} \\ &= \int \mathcal{D}p(t)\mathcal{D}q(t)\exp\{i\int_{t_0}^{t_{n+1}} dt[p(t)\dot{q}(t) - H(p(t), q(t))]\} \end{aligned} \tag{1.21}$$

Notice that we have defined

$$\mathcal{D}p(t) \equiv \prod_i \frac{dp(t_i)}{2\pi}$$

This expression is referred to as the path integral in phase space as opposed to the Feynman path integral in configuration space. The expression in the exponent looks just like the time integral over the lagrangian, i.e. as the action, however, the difference is that the lagrangian depends on $(q(t), \dot{q}(t))$, and since $\dot{q}(t)$ is the derivative of $q(t)$ it is a functional of a path in *configuration* space. In the expression now derived we have an integral over paths in *phase* space, the dimension of which is twice as large: $p(t)$ above is *not* linked to $q(t)$ via $\dot{q}(t)$ and the classical equations of motion.

However, we shall consider the very important special case where the *dependence on p is quadratic.* In this case we shall be able to carry out the functional integral over $p(t)$ rather trivially using gaussian integration. Whenever that is the case we do indeed end up with Feynman's path integral. Even in that case, however, it is sometimes best to start with the path integral in phase space whenever subtleties require special care.

Let us first pause to give several crucial formulas for gaussian integration. These provide the key to understanding the manipulations of the path integral formulation of perturbative quantum field theory.

Gaussian integration

1.

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \tag{1.22}$$

Proof:

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} dx e^{-x^2} \times \int_{-\infty}^{\infty} dy e^{-y^2} = \int\int_{-\infty}^{\infty} dx dy e^{-(x^2+y^2)} \\ &= \int_0^{2\pi} d\phi \int_0^{\infty} r dr e^{-r^2} = \pi \int_0^{\infty} dr^2 e^{-r^2} = \pi \end{aligned} \tag{1.23}$$

qed

2.

$$\int_{-\infty}^{\infty} dx e^{-\frac{1}{2}ax^2} = \sqrt{\frac{2\pi}{a}} \quad \text{for Re}a > 0 \tag{1.24}$$

Proof:

Change variable $\frac{1}{2}ax^2 \to y^2$; discuss the case Im$a \neq 0$.

3. Let $\{A_{ij}\}$ be a real symmetric $n \times n$ matrix with positive eigenvalues. Then

$$\int d^n x e^{-\frac{1}{2}x_i A_{ij} x_j} = (2\pi)^{n/2} (\det A)^{-\frac{1}{2}} \tag{1.25}$$

Proof:

(i) Trivial for A diagonal. (ii) Unaffected by orthogonal transformation: $d^n x$ and $\det A$ invariant under these. Fill in details.

4. Let $z_j = x_j + iy_j$ be n complex variables and let $\overline{z}$ denote the complex conjugate. Let A be hermitian with positive eigenvalues. Then
$$\int d^n z d^n \overline{z} e^{-\frac{1}{2}\overline{z}^T A z} = (2\pi)^n (\det A)^{-1} \tag{1.26}$$

   Proof:

   Exercise! Discuss how $d^n z d^n \overline{z}$ is normalized.Compare with 3.

5. Let x and b be real n-vectors and let A be positive (i.e. positive eigenvalues), real symmetric $n \times n$.
$$S(x) \equiv \frac{1}{2} x^T A x + b^T x \tag{1.27}$$

   (Superscript T denotes transposition). Let x_C be the critical (or "classical") value of the vector x for which (the "classical equations of motion")
$$\frac{\partial}{\partial x_i} S(x) = 0$$

   holds, i.e.
$$x_C = -A^{-1} b$$

   so that
$$S(x_C) = \frac{1}{2}(A^{-1}b)^T A (A^{-1}b) - b^T A^{-1} b = -\frac{1}{2} b^T A^{-1} b$$

   Notice that this is *minus* the contribution of the first ("free") term in eq.(1.27). Then
$$\begin{aligned} \int d^n x e^{-S(x)} &\equiv \int d^n x e^{-\frac{1}{2}x^T A x - b^T x} \\ &= (2\pi)^{n/2} (detA)^{-\frac{1}{2}} e^{-S(x_C)} \\ &= (2\pi)^{n/2} (detA)^{-\frac{1}{2}} e^{+\frac{1}{2} b^T A^{-1} b} \end{aligned} \tag{1.28}$$

   Proof:
$$\frac{1}{2} x^T A x + b^T x = \frac{1}{2}[x^T A x + 2b^T x] = \frac{1}{2}[(x - x_C)^T A (x - x_C) - b^T A^{-1} b]$$

   Now change variables from x to $x - x_C = x + A^{-1}b$, and show that the Jacobian of the transformation is 1.

Eq.(1.28) will be our fundamental rule of gaussian integration. The idea in the proof is referred to as "completion of the square". It is easily generalized to the complex case analogous to 4.

1.3.2 The Feynman Path Integral in Configuration Space

Let us now go back to the path integral expression in phase space, eq.(1.21), and consider the case where the variables p and q are separable and in particular where the dependence on p is *quadratic*:
$$H(p, q) = \frac{1}{2} p^2 + V(q)$$

Then the "path integral" over $p(t)$ in eq.(1.21) becomes

$$\int \mathcal{D}p \exp\{i\int_t^{t'} d\tau[p(\tau)\dot{q}(\tau) - \frac{1}{2}p^2(\tau)]\} \tag{1.29}$$

To apply the rules of gaussian integration just developed we see that we have to subdivide again the integration interval and think of $p(\tau_i)$ as playing the role of x_i. Also

$$\sum_{i,j=1}^{n} x_i A_{ij} x_j \to i\int_t^{t'} d\tau p(\tau)^2 \sim i\sum_{i=1}^{n} \Delta\tau p(\tau_i)p(\tau_i)$$

Then

$$A_{ij} \to i\Delta\tau\delta_{ij}$$

and similarly

$$(A^{-1})_{ij} = \frac{1}{i\Delta\tau}\delta_{ij} \ , \ b_i = -i\Delta\tau\dot{q}(\tau_i)$$

Therefore

$$\frac{1}{2}b^T A^{-1} b \to \frac{i}{2}\int_t^{t'} d\tau\dot{q}^2(\tau)$$

The normalization of the path integral involves the determinant of A which is given by the strange expression

$$(\Delta\tau)^n$$

and that becomes singular in the limit $n \to \infty$. This is a common feature of path integral discussions. Actually as we shall later see, such normalization constants are really of no interest as long as they do not depend on variables of physical significance. In our case the constant only depends on the fictitious cut-off $\Delta\tau$. So we lump these uninteresting factors into "the normalization constant" $\mathcal{N}$. We have finally derived the *Feynman path integral*

$$\begin{aligned} F(q',t';q,t) &= \mathcal{N}\int \mathcal{D}q \exp\{i\int_t^{t'} d\tau[\frac{1}{2}\dot{q}^2(\tau) - V(q)]\} \\ &= \mathcal{N}\int \mathcal{D}q \exp\{i\int_t^{t'} d\tau[L(q(\tau),\dot{q}(\tau))\} \\ &= \mathcal{N}\int \mathcal{D}q \exp\{iS[q]\} \end{aligned} \tag{1.30}$$

1.3.3 Correlation functions

Amplitudes of physical interest are often more directly related to the so-called correlation functions (a name borrowed from statistical mechanics, see later sections) or Greens functions (of which we shall meet several kinds, see later sections).

Let us start by considering the *one point function*

$$\langle q',t'|\hat{q}(\bar{t})|q,t\rangle$$

where we now omit the Heisenberg label, H, and where $\bar{t} \in [t,t']$. This is the unnormalized expectation value of the position at time $\bar{t}$, given the transition represented by the initial and final state. The corresponding *normalized* expectation value is obtained by dividing by the transition amplitude itself. When this is done the unwanted strange normalization constant $\mathcal{N}$ disappears.

A path integral expression for the one point function is easily obtained by subdividing the interval again so that $\bar{t}$ coincides with one of the division points, say t_k. The calculation is entirely identical to what we have just done, with the exception that the small transition from t_{k-1} to t_k contains the position operator. But that operator is now acting on an eigenstate, and we simply get a factor q_k in front of everything in the integrand. When this is replaced by the name $q(t_k)$ and we pass to continuum notation we obtain

$$\langle q', t'|\hat{q}(\bar{t})|q, t\rangle = \int \mathcal{D}q e^{iS[q]} q(\bar{t}) \tag{1.31}$$

It is equally simple to work out the result for the two point function

$$\langle q', t'|\hat{q}(\bar{t}_2)\hat{q}(\bar{t}_1)|q, t\rangle$$

provided

$$\bar{t}_1 < \bar{t}_2$$

Subdividing such that these time values coincide with t_{k_1}, t_{k_2} respectively, we see that we get again the same integrand but with factors

$$q_{k_2} q_{k_1} \sim q(t_{k_2}) q(t_{k_1})$$

Therefore the path integral automatical provides us with the *time ordered* correlator or Greens function:

$$\begin{aligned} & \int \mathcal{D}q e^{iS[q]} q(\bar{t}_1) q(\bar{t}_2) \\ = \quad & \langle q', t'|T\{\hat{q}(\bar{t}_2)\hat{q}(\bar{t}_1)\}|q, t\rangle \end{aligned} \tag{1.32}$$

where we have introduced Dyson's time-ordering symbol, T, ordering operators such that the operator with the earliest times stand farthest to the right.

Clearly in the general case we find for the *N-point function*

$$\begin{aligned} G_N(\bar{t}_1, \bar{t}_2, ..., \bar{t}_N) & \equiv \langle q', t'|T\{\hat{q}(\bar{t}_1)\hat{q}(\bar{t}_2)...\hat{q}(\bar{t}_N)\}|q, t\rangle \\ & = \int \mathcal{D}q e^{iS[q]} q(\bar{t}_1) q(\bar{t}_2)...q(\bar{t}_N) \end{aligned} \tag{1.33}$$

Notice that the Greens function depends on our choice of boundary conditions, specified by $\langle q', t'|$ and $|q, t\rangle$, however, we refrain from adding more labels to it by now. Notice also that inside the time-ordering sign T, the $\hat{q}(\bar{t}_i)$ *operators* commute just as if they were "classical" variables like the ones inside the path integral sign.

1.3.4 The Generating Functional, or, the Partition Function

When one deals with a sequence of numbers $\{a_i | i \in \mathsf{N}\}$ it is very often mathematically convenient to study this sequence instead by the *generating function*

$$F(z) \equiv \sum_{n \in \mathsf{N}} \frac{1}{n!} a_n z^n$$

from which the sequence is recovered by

$$a_n = \frac{d^n}{dz^n} F(z)|_{z=0}$$

Likewise for Greens functions, N-point functions, we shall find it convenient to introduce a generating functional. In field theory we shall introduce several, and we shall see how the relation to statistical mechanics suggests the name "partition function" for the one we consider now. Given an "arbitrary" external "current" or "driving force", $J(t)$, which however, we shall always assume vanishes identically in the sufficiently distant past and future, we define the partition function as a functional of that current as

$$Z[J] \equiv \sum_{N=0}^{\infty} \int dt_1 dt_2 ... dt_N \frac{i^N}{N!} G_N(t_1, t_2, ..., t_N) \times J(t_1) J(t_2) \cdot ... \cdot J(t_N) \tag{1.34}$$

So that the Greens functions themselves are recovered by functional derivatives:

$$G_N(t_1, t_2, ..., t_N) = \frac{\delta}{i\delta J(t_1)} \frac{\delta}{i\delta J(t_2)} \cdots \frac{\delta}{i\delta J(t_N)} Z[J]|_{J\equiv 0} \tag{1.35}$$

This is easily verified using

$$\frac{\delta J(t_i')}{\delta J(t_j)} = \delta(t_i' - t_j)$$

After the functional derivatives are carried out we are instructed to put the current to zero. The factors i and $N!$ are purely for convenience.

It is now easy to see that there is an elegant path integral expression for this generating functional. Indeed, consider the modified action, where we add to the lagrangian a driving force term given by J:

$$L(q(t), \dot{q}(t)) \rightarrow L(q(t), \dot{q}(t); J(t)) \equiv L(q(t), \dot{q}(t)) + J(t) q(t) \tag{1.36}$$

and correspondingly

$$S[q; J] \equiv \int dt L(q(t), \dot{q}(t); J(t))$$

Notice that in the simple case where

$$L(q(t), \dot{q}(t); J(t)) = \frac{1}{2}\dot{q}^2 - V(q(t)) + J(t) q(t) \tag{1.37}$$

the equation of motion gets modified into

$$\ddot{q} = -V'(q) + J$$

showing that the "current", $J(t)$ is indeed an external driving force in that case.

Theorem

$$Z[J] = \int \mathcal{D}q e^{iS[q;J]} = \int \mathcal{D}q e^{iS[q] + i\int dt J(t) q(t)} \tag{1.38}$$

Proof:

$$\begin{aligned} \frac{\delta}{i\delta J(\bar{t})} \int \mathcal{D}q e^{iS[q;J]} &= \int \mathcal{D}q e^{iS[q;J]} \frac{\delta}{i\delta J(\bar{t})} i \int dt J(t) q(t) \\ &= \int \mathcal{D}q e^{iS[q;J]} q(\bar{t}) \end{aligned} \tag{1.39}$$

Similarly

$$\frac{\delta^N}{i\delta J(\bar{t}_1)\cdots i\delta J(\bar{t}_N)}\int \mathcal{D}q e^{iS[q;J]} = \int \mathcal{D}q e^{iS[q;J]} q(\bar{t}_1)\cdots q(\bar{t}_N) \tag{1.40}$$

So putting $J \equiv 0$ in the end

$$\begin{aligned}\frac{\delta^N}{i\delta J(\bar{t}_1)\cdots i\delta J(\bar{t}_N)}\int \mathcal{D}q e^{iS[q;J]}|_{J\equiv 0} &= \int \mathcal{D}q e^{iS[q]} q(\bar{t}_1)\cdots q(\bar{t}_N) \\ &= G_N(\bar{t}_1, ..., \bar{t}_N) \qquad (1.41)\end{aligned}$$

Hence

$$Z[J] = \int \mathcal{D}q e^{iS[q;J]} \tag{1.42}$$

qed

The Greens functions we have considered here and indeed our entire path integral derivation, made use of one particular kind of boundary conditions, or, correspondingly, one particular kind of specification of initial and final states. However, there is nothing sacred in the choice we have presented. On the contrary, in practical applications it is usually *not* a good idea to specify position states, since these can only make sense at one instant of time. The same is true of the analogous states in the field theory case where we might try to impose particular *values* for the field variables.

A much more practical specification both for comparison with actual measurements and for further development of the theory, would be ones where we used instead eigenstates of the energy operator without driving force, in the distant past and future. This choice of boundary condition we want to analyze in quite some detail by studying in the next subsection the harmonic oscillator.

1.4 The Harmonic Oscillator

1.4.1 The Naive Treatment

The example of the harmonic oscillator with an external driving force contains essentially all the complications we need for field theory perturbation theory. The action is given by

$$S_{HO} = S[q; J] = \int dt\{\frac{1}{2}(\dot{q}^2 - \omega^2 q^2) + J(t)q(t)\} \tag{1.43}$$

In this case we shall be able to derive a closed expression for the path integral. Our ability to do that depends on the problem being gaussian: the degree of freedom, $q(t)$ appears quadratically and linearly only, in the action, so that gaussian integration can be performed. We would like to emphasize two points:

1. How simple it is to derive the result using gaussian integration.

2. Nevertheless there are some very important subtleties associated with the choice of boundary condition.

In order to first emphasize the first point, let us present a sloppy treatment. We make the usually unjustified assumption that we can perform a partial integration in the action *without boundary terms*, so that the action can be written as

$$S[q;J] = \int dt\{-\frac{1}{2}q(t)[\frac{d^2}{dt^2}+\omega^2]q(t)+J(t)q(t)\} \tag{1.44}$$

This shows that the path integral is gaussian:

$$\int \mathcal{D}q \exp\{-\frac{1}{2}q\Delta^{-1}q+iJq\} \tag{1.45}$$

where we have introduced "functional notation". The operator, Δ^{-1} is a linear operator on functions $q(t)$ defined by

$$\Delta^{-1}q(t) \equiv i[\frac{d^2}{dt^2}+\omega^2]q(t) \tag{1.46}$$

and of course

$$Jq \equiv \int dt J(t)q(t)$$

etc. The name Δ^{-1} derives of course from the fact, that the inverse (suitably defined!) is the interesting object, the *propagator*. Using blindly the rule of gaussian integration above, we then find

$$Z[J] = \int \mathcal{D}q e^{-\frac{1}{2}q\Delta^{-1}q+iJq} = \mathcal{N}' e^{-\frac{1}{2}J\Delta J} \tag{1.47}$$

where the normalization among other things involve

$$(\det\Delta)^{\frac{1}{2}}$$

As we shall see we do not need to know too much about this determinant, but we shall need to understand in particular the meaning of Δ, the inverse of Δ^{-1}. Naively it looks like we might write

$$\Delta(t,t') = i\int \frac{dp}{2\pi}\frac{e^{-ip(t-t')}}{p^2-\omega^2} \tag{1.48}$$

Then, namely, we can easily prove that (formally) $\Delta^{-1}\Delta = 1$. More precisely we get

$$\begin{aligned}
\Delta^{-1}\Delta(t,t') &= i[\frac{d^2}{dt^2}+\omega^2]i\int\frac{dp}{2\pi}\frac{e^{-ip(t-t')}}{p^2-\omega^2}\\
&= -\int\frac{dp}{2\pi}\frac{[\frac{d^2}{dt^2}+\omega^2]e^{-ip(t-t')}}{p^2-\omega^2}\\
&= -\int\frac{dp}{2\pi}\frac{[-p^2+\omega^2]e^{-ip(t-t')}}{p^2-\omega^2}\\
&= \int\frac{dp}{2\pi}e^{-ip(t-t')}\\
&= \delta(t-t')
\end{aligned} \tag{1.49}$$

This shows that we are on the right track. And

$$Z[J] = \mathcal{N}' e^{-\frac{1}{2}J\Delta J} \equiv \mathcal{N}' \exp\{-\frac{1}{2}\int dt\int dt' J(t)\Delta(t,t')J(t')\} \tag{1.50}$$

So far so good. Now let us discuss the problems, and where we have been too sloppy.

If we try to evaluate the propagator we meet the problem that the denominator in the integrand is singular at $p^2 = \omega^2$. Something is wrong, but what is going on? Let us think a little harder, and try to be more mathematical. We have defined the operator Δ^{-1} which is certainly a linear operator acting on some vector space of functions. Hence we think in analogy with a finite dimensional matrix. But is that operator (matrix) really invertible? The answer depends on the set of functions, or better: the vector space of functions, on which we mean to define it. If we take the space of *all* functions $q(t)$ (just with some smoothness properties), we shall certainly go wrong because the operator in question has "zero-modes". This is the analogous situation to a matrix which has eigenvectors with zero eigenvalues. Such a matrix of course cannot be inverted: it is singular. Indeed there does exists eigenfunctions $q_0(t)$ with zero eigenvalues:

$$[\frac{d^2}{dt^2} + \omega^2]q_0(t) = 0 \tag{1.51}$$

Furthermore, far from being pathological, these zero modes are exactly the classical solutions for the free oscillator! Very physical ones! So, if we try to consider the operator, Δ^{-1} on a function space *including such functions,* it cannot possibly be inverted.

What will happen, however, is that when we impose *boundary conditions* corresponding to the initial and final *state specifications,* we shall *restrict* the set of functions $q(t)$ over which we path integrate (namely by forcing them to conform to the boundary conditions) to be such, that exactly there are no zero modes to be integrated over.

Let us now discuss this in some detail, starting with the situation encountered in the last section where we calculated $\langle q_f, t_f | q_b, t_b \rangle$. We obtained this matrix element by performing the integration over all continuous path $q(t)$, $t \in [t_b, t_f]$ with the boundary conditions $q(t_b) = q_b$ and $q(t_f) = q_f$. It is easy to check that the (1.51) has solution $q_{cl}(t) \neq 0$ if q_b or q_f are different from zero (except for special values of $\omega(t_f - t_b)$). Let us write

$$q(t) = q_{cl}(t) + \tilde{q}(t), \tag{1.52}$$

where $\tilde{q}(t)$ satisfies the "vanishing" boundary condition:

$$\tilde{q}(t_b) = \tilde{q}(t_f) = 0. \tag{1.53}$$

If $q_{cl}(t) \neq 0$ the classical equation (1.51) has no solution which satisfy the boundary condition (1.53). Now comes a very common idea in the functional approach to quantum field theory: We may change variable in the path integral from q to $\tilde{q}$ since the difference is just a fixed path:

$$\int \mathcal{D}q \; e^{iS[q_{cl}+\tilde{q};J]} = \int \mathcal{D}\tilde{q} \; e^{iS[q_{cl}+\tilde{q};J]} \tag{1.54}$$

This step is entirely analogous to the variable change

$$\prod_i dx^i = \prod_i d(x_0^i + \tilde{x}^i) = \prod_i d\tilde{x}^i$$

in the case of finitely many variables. The analogy is

$$\begin{aligned} i &\to t \\ x^i &\to q(t) \\ x_0^i &\to q_{cl}(t) \\ \tilde{x}^i &\to \tilde{q}(t) \end{aligned} \tag{1.55}$$

If we use that $q_cl(t)$ satisfies (1.51) as well as the boundary conditions (1.53) it follows that

$$S[q_{cl}+\tilde{q};J)] = S[q_{cl};J] + S[\tilde{q};J]. \tag{1.56}$$

This result is only correct for an action $S(q;J)$ which is quadratic in q. It implies that we can write the path integral (1.54) as

$$Z(J) = e^{iS[q_{cl};J]}\int \mathcal{D}\tilde{q}\; e^{iS[\tilde{q};J]}. \tag{1.57}$$

Using this more detailed formulation of the boundary problem, the functional integral *is* well defined. We are on a finite interval and the boundary condition for $\tilde{q}(t)$ is a standard Sturm-Liouville boundary condition where (for generic values of ω and $t_f - t_b$) (1.51) has no solutions. Thus the Green function Δ_0, where the suffix zero refers to the boundary condition for $\tilde{q}$, exists (exercise: find it using the notes on Green functions). We can finally write

$$Z(J) = \mathcal{N}'\, e^{iS[q_{cl};J]}\, e^{-\frac{1}{2}J\Delta_0 J} = \mathcal{N}''\, e^{-\frac{1}{2}J\Delta_0 J + q_{cl}J}, \tag{1.58}$$

where the normalization factor among other things involve the determinant $\det\Delta_0$, and where the last equality is obtained by using eq. (1.51) and absorbing a J-independent boundary term coming from a partial integration of the kinetic term in $S[q_{cl};J]$ in the normalization constant.

Let us now consider the situation where we project on the ground state of the harmonic oscillator, i.e. we calculate

$$Z(J) \equiv \int dq(t_f)\int dq(t_b)\; \langle\Psi_0(q_f,t_f)|q_f,t_f\rangle\langle q_f,t_f|q_b,t_b\rangle\langle q_b,t_b|\Psi_0(q_b,t_b)\rangle, \tag{1.59}$$

where all vectors as usually are in the Heisenberg representation. Furthermore we will assume that $t_b \to -\infty$ and $t_f \to \infty$. The amplitude, which we (with a slight abuse of notation) still denote $Z(J)$, is in fact a very natural quantity. It is the amplitude that a harmonic oscillator, original in the ground state, remains there after an external force $J(t)$ is applied to the system. Also, as will be discussed later, it is precisely this amplitude which is natural to consider when we generalize from quantum mechanics to relativistic field theory.

We will now show that

1. The integration over q_b and q_f can be included in the path integral in a natural way.
2. By this inclusion we arrive at a differential operator (1.46) which *can* be inverted and where the corresponding Green function (the Feynman propagator) is characterized by the fact that only positive frequency oscillations $e^{-i\omega t}$ are propagated forward in time and only negative frequency oscillations $e^{i\omega t}$ are propagated backwards in time (outside the time interval where $J(t)$ is non-zero).

A priori it is surprising that folding of the wave functions $\Psi_0(q_f)$ and $\Psi_0(q_b)$ into the path integral for $\langle q_f,t_f|q_b,t_b\rangle$ leads to a well defined boundary problem for differential operator (1.46) since we are integrating over all q_f and all q_b. As will be clear, it is very important that we consider the projection on the *ground state*. Only in this case a simple picture emerges.

Let us use that the ground wave functional of the harmonic oscillator is

$$\Psi_0(q) \propto e^{-\frac{1}{2}\omega q^2}. \tag{1.60}$$

Hence (1.59) can be written as (with $t_f = \infty$, $t_b = -\infty$)

$$Z(J) = \int dq(\infty) \int dq(-\infty)\; e^{-\frac{1}{2}\omega[q^2(\infty)+q^2(-\infty)]} \int \mathcal{D}q(t)\; e^{\frac{i}{2}\int_{-\infty}^{\infty} dt\,(\dot{q}^2-\omega^2 q^2(t))}. \tag{1.61}$$

In this formula the functional integral does not include the endpoints $q(\pm\infty)$ (clearly we are a little sloppy and formal at this point). We use the following formula (which is valid for functions $q^2(t)$ which behave reasonable nice at $t = \pm\infty$; investigate how nice!):

$$q^2(\infty) + q^2(-\infty) = \lim_{\varepsilon\to 0} \varepsilon \int_{-\infty}^{\infty} dt\; e^{-\varepsilon|t|}\, q^2(t). \tag{1.62}$$

We can now write

$$Z(J) = \int \mathcal{D}q(t)\; e^{\frac{i}{2}\int_{-\infty}^{\infty} dt\,(\dot{q}^2-(\omega^2 - i\varepsilon e^{-\varepsilon|t|}\omega)q^2(t))}, \tag{1.63}$$

where the integration over the endpoints $q(\pm\infty)$ are included in the functional integral and where the limit $\varepsilon \to 0$ is understood[1]. Summarizing, effectively an infinitesimal imaginary term has been added to ω^2:

$$\omega^2 \to \omega^2 - i\varepsilon'. \tag{1.64}$$

With this slight modification we can now return to our heuristic derivation (1.45)-(1.49) of eq. (1.48) for the propagator. With the change (1.64) (the so-called Feynman prescription) we obtain

$$\Delta_F(t,t') = i \int \frac{dp}{2\pi} \frac{e^{-ip(t'-t)}}{p^2 - \omega^2 + i\epsilon}. \tag{1.65}$$

The Feynman prescription has displaced the poles encountered at $p = \pm\omega$ infinitesimally into the complex plane:

$$p = \pm(\omega - i\varepsilon') \quad \text{(where we assume } \omega > 0 \text{ and } \varepsilon' > 0) \tag{1.66}$$

and thus made the integral in (1.65) well defined. Further, for $t' \to \infty$ (in fact, for $t' - t > 0$), the (Fourier) exponential function falls off exponentially in the lower half plane as $\exp\{-|\mathrm{Im} p|(t'-t)\}$, so that we may close the integration contour in the *lower* half plane. Thus using the Cauchy theorem of contour integration, we pick up the pole at $p = \omega - i\epsilon'$ with residue

$$\frac{i}{2\pi}\frac{1}{2\omega} e^{-i\omega(t'-t)}$$

Hence we find the useful formula

$$\Delta_F(t',t) = \frac{1}{2\omega} e^{-i\omega|t'-t|} \tag{1.67}$$

---

[1] Let us at this point remark that it is important that we project on the ground-state and not a higher energy eigenstate. Such higher energy states will always have zero-points x_0 where $\Psi_n(x_0) = 0$ and cannot be written as an exponential, which was important for the above arguments.

Here we used that the contour integral is obtained as $2\pi i$ times the residue, times a factor -1 because the contour runs clockwise when closed in the lower half plane. A similar calculation applies when $t - t' > 0$.

Let us now study the solution to the differential equation

$$\Big[\frac{d^2}{dt^2} + \omega^2\Big] f(t) = J(t) \tag{1.68}$$

using the Feynman Green function $\Delta_F(t', t)$ given by eq. (1.67). One has

$$f(t) = \int dt' \; \Delta_F(t, t') \, J(t'). \tag{1.69}$$

Assume that the support of $J(t)$ is confined to the interval $[-T, T]$. Then we have

$$f(t) \;=\; A\, e^{-i\omega t} \qquad \text{for} \quad t > T, \tag{1.70}$$

$$f(t) \;=\; A^*\, e^{i\omega t} \qquad \text{for} \quad t < T, \tag{1.71}$$

where

$$A = \int_{-T}^{T} dt' \, \frac{e^{i\omega t'}}{2\omega} \, J(t'). \tag{1.72}$$

Hence we can say that from a mathematical point of view the Feynman prescription (1.64) corresponds to solving the differential equation (1.68) with the boundary conditions that only positive frequencies propagate forward in time and only negative frequencies backwards in time. This boundary condition allows no solution to the homogeneous equation (1.51), and there is thus no zero eigenvalue. Consequently the Green function exists (and is given by (1.67).

To summarize: we have obtain the following fundamental result for the ground-state to ground-state amplitude (1.59) for the harmonic oscillator

$$\boxed{Z(J) = \mathcal{N}'' e^{-\frac{1}{2} J \Delta_F J}}\,, \tag{1.73}$$

where $\mathcal{N}''$ is a normalization factor. For a given external source $J(t)$ acting on the harmonic oscillator, the probability that the oscillation will remain in the ground state is thus given by

$$\begin{aligned} \frac{|Z(J)|^2}{|Z(0)|^2} &= \exp\Big(-\mathrm{Im} \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dt' J(t) \Delta_F(t, t') J(t')\Big) \\ &= \exp\Big(-\mathrm{Im} \int_{-\infty}^{\infty} \frac{dp}{2\pi} \frac{|\tilde{J}(p)|^2}{p^2 - \omega^2 + i\varepsilon}\Big) = \exp\Big(-\frac{|\tilde{J}(\omega)|^2}{2\omega}\Big), \end{aligned} \tag{1.74}$$

where $\tilde{J}(p)$ denotes the Fourier transform of $J(t)$. Note that this calculation is actually simpler than the corresponding calculation using conventional quantum mechanics.

Let us finally remark on the determinant of Δ_F. We could think of that as something like the product of all the eigenvalues λ_i of the differential operator, i.e. the λ_i's satisfying

$$\Big[-\frac{d^2}{dt^2} - \omega^2\Big] f_i(t) = \lambda_i f_i(t) \tag{1.75}$$

with the given boundary conditions[2]. The trouble is that this product would diverge wildly. The fact however, is that even if we regularize the determinate (say by including only eigenvalues less than some cut-off (the Planck energy (?!)) for some physically relevant particle physics problem), it would still be quite uninteresting: we have seen that what we need, to obtain the Greens functions, are *derivatives* after the external current, but the determinant cannot depend on that external current, it only depends on ω and the cut-off. Hence it is merely one more irrelevant contribution to the overall normalization constant. For this reason we shall ignore it in the following. (Please note, however, that there are many more complicated physical situations in which the determinant precisely does depend on physically relevant parameters. In those cases one must understand the determinant).

1.4.2 Harmonic Phase Space

To be even more careful, and to consider in detail the rather particular kind of boundary condition, we shall be mostly interested in, we actually go back to the path integral in phase space, and we take this opportunity to introduce some powerful technique based on the idea of *harmonic coordinates* in phase space: "harmonic phase space". The harmonic oscillator furnishes a very simple example of that.

Consider the free classical oscillator

$$\begin{aligned} L &= \frac{1}{2}\dot{q}^2 - \frac{1}{2}\omega^2 q^2 = p\dot{q} - [\frac{1}{2}\dot{q}^2 + \frac{1}{2}\omega^2 q^2] \\ &= p\dot{q} - H(p,q) \end{aligned} \tag{1.76}$$

with

$$p = \frac{\partial L}{\partial \dot{q}} = \dot{q}; \quad \dot{p} = \frac{\partial L}{\partial q} = -\omega^2 q \tag{1.77}$$

The general solution to the equation of motion

$$\ddot{q} = -\omega^2 q$$

is on the form

$$\begin{aligned} q(t) &= \frac{1}{\sqrt{2\omega}}[a(t) + a^*(t)] \\ p(t) &= i\sqrt{\frac{\omega}{2}}[-a(t) + a^*(t)] \\ a(t) &= \frac{1}{\sqrt{2\omega}}[\omega q + ip] = a(0)e^{-i\omega t} \\ a^*(t) &= \frac{1}{\sqrt{2\omega}}[\omega q - ip] = a^*(0)e^{i\omega t} \\ H(a^*, a) &= \frac{1}{2}[\dot{q}^2 + \omega^2 q^2] = \omega a^* a \end{aligned} \tag{1.78}$$

---

[2] Note that this is strictly speaking only true if the operator is defined on a finite interval. In case $t_f, t_b = \pm\infty$ there exists only *generalized* eigenfunctions with do not satisfy the boundary conditions. In addition the spectrum corresponding to these generalized eigenfunctions is no longer discrete, but is labeled by a continuous index, e.g.

$$\lambda(p) = \pm\sqrt{p^2 - \omega^2}, \qquad |p| > \omega.$$

Effectively we have passed from (p, q) coordinates in phase space to the harmonic ones (a^*, a), and we see that for the free oscillator, the classical trajectory in harmonic phase space is merely a phase-circle for each variable, a or a^*.

In quantum mechanics the harmonic phase space variables are replaced by creation and annihilation operators $(\hat{a}^\dagger, \hat{a})$ satisfying the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = 1$$

and the Hilbert space of the oscillator is spanned by the basis vectors in Fock space, the vacuum, the singly excited, doubly excited,... etc. states:

$$|0\rangle,\ \hat{a}^\dagger|0\rangle,\ (\hat{a}^\dagger)^2|0\rangle,\ ...$$

However, we shall now introduce a different basis, more convenient for the present (as well as for many other) purpose, the *coherent states* that are eigenstates of the annihilation operator:

Let a be a real or complex number, then define the state

$$\begin{aligned} |a\rangle &\equiv e^{a\hat{a}^\dagger}|0\rangle \\ &\equiv \sum_{n=0}^{\infty} \frac{a^n}{n!}(\hat{a}^\dagger)^n|0\rangle \end{aligned} \tag{1.79}$$

One easily checks that

$$\hat{a}|a\rangle = [\hat{a}, e^{a\hat{a}^\dagger}]|0\rangle = a e^{a\hat{a}^\dagger}|0\rangle = a|a\rangle$$

(Notice the useful rule: $[\hat{a}, F(\hat{a}^\dagger)] = F'(\hat{a}^\dagger)$; exercise: prove that!).

Similarly we introduce bra-states: given a^* complex, define

$$\langle a^*| \equiv \langle 0|e^{\hat{a}a^*}$$

for which

$$\langle a^*|\hat{a}^\dagger = \langle a^*|a^*$$

showing that these are eigen (bra-) states of the creation operator. The normalization follows from

$$\langle a^*|a\rangle = \langle 0|e^{\hat{a}a^*}|a\rangle = e^{aa^*}\langle 0|a\rangle = e^{aa^*}\langle 0|0\rangle = e^{aa^*} \tag{1.80}$$

They also satisfy the completeness relation

$$I \equiv \int \frac{da da^*}{2\pi i} e^{-a^* a}|a\rangle\langle a^*| \tag{1.81}$$

Proof:

$$\begin{aligned} \langle b^*|I|b\rangle &= \int \frac{da da^*}{2\pi i} e^{-a^* a}\langle b^*|a\rangle\langle a^*|b\rangle \\ &= \int \frac{da da^*}{2\pi i} \exp\{-a^* a + b^* a + a^* b\} \\ &= \int \frac{da da^*}{2\pi i} \exp\{-[a^* - b^*][a - b]\} e^{+b^* b} \\ &= e^{b^* b} = \langle b^*|b\rangle \end{aligned} \tag{1.82}$$

where we used gaussian integration in the last step. This completes the proof of the completeness relation. Notice that we have not used that b^* should be the complex conjugate of b.

Let us now derive a path integral expression for the transition amplitude in a situation where the initial state at time t is a coherent state $|a,t\rangle$, and the final state at time t' is a coherent bra state $\langle a^*,t'|$. Further, let us immediately generalize the model to include a driving force:

$$H(a^*,a) \to H(a^*,a;t) \equiv \omega a^* a - \gamma(t)a^* - \overline{\gamma}(t)a \tag{1.83}$$

Notice that if we use instead the name

$$\gamma(t) = \frac{J(t)}{\sqrt{2\omega}}\ ,\ \overline{\gamma}(t) = \frac{\overline{J}(t)}{\sqrt{2\omega}}$$

then for $\overline{J} = J$, J is just again the external time-dependent driving force. (Proof: use eqs.(1.78)).

The matrix element of the quantum hamiltonian

$$\hat{H}(\hat{a}^\dagger,\hat{a};t) \equiv \omega\hat{a}^\dagger\hat{a} - \gamma(t)\hat{a}^\dagger - \overline{\gamma}(t)\hat{a} \tag{1.84}$$

(notice the ordering adopted, and notice that we have ignored the zero point energy which would merely contribute to the normalization constant) is given in the coherent state basis as

$$\langle a^*|\hat{H}(\hat{a}^\dagger,\hat{a};t)|b\rangle = H(a^*,b;t)\langle a^*|b\rangle = H(a^*,b;t)e^{a^*b} \tag{1.85}$$

We have now in the usual way for the transition amplitude *between coherent states*

$$\begin{aligned}
F(a^*,t';a,t) &\equiv \langle a^*,t'|a,t\rangle \\
&= \int\prod_i \frac{da(t_i)da^*(t_i)}{2\pi i}e^{-a^*(t_i)a(t_i)} \\
&\cdot\ \langle a^*(t')|e^{-i\epsilon\hat{H}}|a(t_n)\rangle\langle a^*(t_n)|e^{-i\epsilon\hat{H}}|a(t_{n-1})\rangle\langle\cdots \\
&\quad \cdots\rangle\langle a^*(t_1)|e^{-i\epsilon\hat{H}}|a(t)\rangle \\
&= \int\prod_i \frac{da(t_i)da^*(t_i)}{2\pi i} \\
&\cdot\ \exp\{a^*(t')a(t_n) - a^*(t_n)a(t_n) + a^*(t_n)a(t_{n-1}) - a^*(t_{n-1})a(t_{n-1}) + ... \\
&\quad ... - a^*(t_1)a(t_1) + a^*(t_1)a(t) \\
&-\ i\int_t^{t'} d\tau H(a^*(\tau),a(\tau);\tau)\} \\
&= \int\prod_i \frac{da(t_i)da^*(t_i)}{2\pi i} \\
&\quad \exp\{\int_t^{t'} d\tau[\dot{a}^*(\tau)a(\tau) - iH(a^*(\tau),a(\tau);\tau)] + a^*(t)a(t)\}
\end{aligned} \tag{1.86}$$

Here we have introduced the names $a^*(t') \equiv a^*$ and $a(t) \equiv a$. Also, in the last step we rearranged the exponent as

$$\begin{aligned}
&[a^*(t') - a^*(t_n)]a(t_n) + [a^*(t_n) - a^*(t_{n-1})]a(t_{n-1}) + ... \\
+\ &[a^*(t_2) - a^*(t_1)]a(t_1) + a^*(t_1)a(t) - i\int_t^{t'} d\tau H(a^*(\tau),a(\tau);\tau)
\end{aligned} \tag{1.87}$$

We see that we have expressed the transition amplitude as a path integral over paths in harmonic phase space:

$$F(a^*,t';a,t)=\int \mathcal{D}a\mathcal{D}a^*\exp\{i\int_t^{t'}d\tau[\frac{1}{i}\dot{a}^*(\tau)a(\tau)-H]+a^*(t)a(t)\} \tag{1.88}$$

where the first term in the exponential corresponds to the action part. We may notice that by a partial integration of the $\dot{a}^*a$ term in the integrand we may regain the symmetry with respect to the end points. The form in eq.(1.88) involves paths in phase space with the following boundary specification:

- At time t, $a(t)\equiv a$, $a^*(t)$ unspecified.
- At time t', $a(t')$ unspecified, $a^*(t')\equiv a^*$.

Notice that this kind of specification, where we only fix half of the phase space coordinates at each end point of the trajectory, makes sense both quantum mechanically and classically. Classically we have the additional possibility of specifying *all* phase space coordinates at the initial point say, and none at the final point, which is then a result of the dynamical evolution. That is impossible quantum mechanically, since the quantum variables do not commute.

The equations of motion corresponding to the above form are found by writing down the stationarity condition of the integral:

$$0=\int_t^{t'}d\tau\{\delta a[\dot{a}^*-i\frac{\partial H}{\partial a}]-\delta a^*[\dot{a}+i\frac{\partial H}{\partial a^*}]\}$$

or in our case

$$\dot{a}+i\omega a-i\gamma=0\ ,\ \dot{a}^*-i\omega a^*+i\overline{\gamma}=0 \tag{1.89}$$

The classical solutions with the right boundary conditions are immediately written down:

$$\begin{aligned} a(\tau) &= ae^{i\omega(t-\tau)}+i\int_t^{\tau}e^{i\omega(s-\tau)}\gamma(s)ds \\ a^*(\tau) &= a^*e^{i\omega(\tau-t')}+i\int_{\tau}^{t'}e^{i\omega(\tau-s)}\overline{\gamma}(s)ds \end{aligned} \tag{1.90}$$

The result of doing the gaussian path integral in phase space, eq.(1.88) is given by the critical value of the exponent, which we thus evaluate:

For the action part we write

$$\int_t^{t'}d\tau[\dot{a}^*a-iH]=\int_t^{t'}d\tau[\dot{a}^*a-i(\omega a^*a-\gamma a^*-\overline{\gamma}a]]=\int_t^{t'}d\tau[i\gamma(\tau)a^*(\tau)]$$

where we used the equation of motion

$$\dot{a}^*-i\omega a^*+i\overline{\gamma}=0$$

The critical value of the exponent in eq.(1.88) is thus

$$a^*(t)a+i\int_t^{t'}d\tau\gamma(\tau)a^*(\tau)$$

$$
\begin{aligned}
&= a^* a e^{i\omega(t-t')} + ia\int_t^{t'} ds e^{i\omega(t-s)}\overline{\gamma}(s) \\
&+ i\int_t^{t'} ds\gamma(s)[a^* e^{i\omega(s-t')} + i\int_s^{t'} ds' e^{i\omega(s-s')}\overline{\gamma}(s')] \\
&= a^* a e^{-i\omega(t'-t)} \\
&+ i\int_t^{t'} ds[ae^{i\omega(t-s)}\overline{\gamma}(s) + a^* e^{i\omega(s-t')}\gamma(s)] \\
&- \frac{1}{2}\int_t^{t'}\int_t^{t'} ds ds' \gamma(s)\overline{\gamma}(s') e^{-i\omega|s'-s|} \qquad (1.91)
\end{aligned}
$$

We see here the structure of eq.(1.58) coming in: there is a current-independent term corresponding to the classical action for zero external current; there is a first order contribution in currents and finally the most interesting second order contribution from which we may read off the propagator corresponding to the choice of boundary conditions we have taken.

Notice that in the very important special case where the initial and final states both correspond to the *vacuum* state of the system, i.e. to

$$a = 0 = a^*$$

we only have this second order term. Also, we shall be particularly interested in the limits

$$t \to -\infty, \ t' \to +\infty$$

This is the case that we shall mostly consider in this course: The vacuum transition amplitude between the infinite past and the infinite future, in the presence of an arbitrary external current. Precisely since this current is arbitrary, that is in fact enough to learn everything there is to learn about the system.

For this amplitude we may therefore write (for J real)

$$
\begin{aligned}
Z[J] &\equiv \langle 0, +\infty | 0, -\infty\rangle_J \langle 0|0\rangle_0 \\
&= \exp\{-\frac{1}{2} J\cdot\Delta_F\cdot J\}\langle 0|0\rangle_0 \qquad (1.92)
\end{aligned}
$$

where we read off

$$\Delta_F(s,s') = \frac{1}{2\omega} e^{-i\omega|s-s'|} \qquad (1.93)$$

and where we have written the normalization constant as the vacuum to vacuum transition amplitude for zero external current/driving force.

This completes our derivation of the expressions eqs.(1.65,1.67) for the Feynman propagator based on a careful analysis of boundary conditions. We see that the vacuum boundary conditions correspond to $q(t)$ in eq.(1.78) being given by the creation part for $t \to -\infty$ and the annihilation part for $t \to +\infty$. This agrees with eqs.(1.70,1.71).

1.5 The Euclidean Formulation

1.5.1 The Feynman-Kac Formula

Consider the transition amplitude, say in configuration space, in the case of a time independent hamiltonian. Let $\{|n\rangle\}$ be an orthonormal set of energy eigenstates with energy eigenvalues, $\{E_n\}$ where all $E_n > 0$ except the vacuum eigenvalue, which we take to be zero: $E_0 = 0$. Then we have

$$\begin{aligned} \langle q', t'|q, t\rangle &= \langle q'|\exp\{-i\hat{H}(t'-t)\}|q\rangle = \sum_{n,m}\langle q'|n\rangle\langle n|e^{-i\hat{H}(t'-t)}|m\rangle\langle m|q\rangle \\ &= \sum_n \langle q'|n\rangle\langle n|q\rangle e^{-iE_n(t'-t)} = \sum_n \psi_n(q')\psi_n^*(q)e^{-iE_n(t'-t)} \end{aligned} \tag{1.94}$$

where $\psi_n(q) = \langle q|n\rangle$ is the wave function in q-space for the state at energy level, n. This expression is an analytic function of the variable, $t' - t \equiv \Delta t$. Let us make the analytic continuation, often described as *the Wick rotation* or *rotation to euclidean time*:

$$\Delta t \to -i\beta \tag{1.95}$$

giving

$$\langle q', \beta|q, 0\rangle = \sum_n \psi_n(q')\psi_n^*(q)e^{-E_n\beta} \tag{1.96}$$

If we furthermore specialize to $q' = q$ and then integrate over q we find *the statistical partition function of the system at temperature*

$$kT \equiv \frac{1}{\beta} \tag{1.97}$$

($k =$ Boltzmann's constant). This result we now cast into a nice path integral expression known as the *Feynman-Kac formula* as follows:

We see that we are led to define paths as functions of "euclidean time" $t = -it_E$. Also, the condition above of putting $q' = q$ means that we consider *closed* paths:

$$q' = q(t_E = \beta) = q(t_E = 0) = q$$

and finally, the fact that we integrate in the end over the value

$$q = q(0) = q(\beta) = q'$$

means that we consider the path integral over *all closed paths of euclidean time-length*, $\beta = 1/kT$.

Let us now see what happens to the action integral under the rotation to euclidean time. Let us consider the simple 1-dimensional potential system

$$L[q, \dot{q}] = \frac{1}{2}(\frac{dq}{dt})^2 - V(q)$$

Then

$$\begin{aligned} iS[q] &= i\int_0^{-i\beta}(-idt_E)\{\frac{1}{2}\left(\frac{dq}{d(-it_E)}\right)^2 - V(q)\} \\ &= -S_E[q] \\ S_E[q] &\equiv \int dt_E\{\frac{1}{2}(\frac{dq}{dt_E})^2 + V(q)\} \end{aligned} \tag{1.98}$$

This "euclidean action" is strictly positive if the hamiltonian of the original system is. We then finally obtain for the thermodynamical partition function the Feynman-Kac formula

$$\begin{aligned}
Z(\beta) &= \mathrm{Tr}\{e^{-\beta\hat{H}}\} \\
&= \int \mathcal{D}q e^{-S_E[q]}|_{\text{closed paths of length } \beta} \\
S_E[q] &= \oint dt_E L_E[q,\dot{q}] \\
q : t_E &\mapsto q(t_E) \text{ such that } q(t_E+\beta) = q(t_E)
\end{aligned} \tag{1.99}$$

1.5.2 The Vacuum Functional

A very important special case obtains for zero temperature, in which case (i) the partition function is given entirely by the vacuum contribution, and (ii) our closed paths have infinite length: $\beta \to \infty$ for $kT \to 0$. This provides us with a new very powerful way to study our favorite objects: the vacuum functionals. In the presence of an external driving force the treatment of the subsection above does not quite go through, but precisely for the vacuum functional it does *provided* we assume that the external driving force vanishes identically outside a finite time-interval. The remaining infinite time is enough to project out the vacuum. (Exercise: fill in details!)

Then for $t = -it_E$ we find

$$i\int dt J(t)q(t) = i\int d(-it_E)J(-it_E)q(-it_E) \equiv \int dt_E J_E(t_E)q_E(t_E)$$

with an obvious notation. We shall dispense with the indices E on J and q, and write

$$iS[q;J] \to -S_E[q;J]$$

Thus for the harmonic oscillator

$$S_E[q;J] = \int dt_E\{\frac{1}{2}(\dot{q}^2+\omega^2q^2) - Jq\}$$

Thus we may express the generating functional - the vacuum functional in the precence of the driving force - as the following euclidean path integral

$$Z_E[J] = \int \mathcal{D}q \exp\{-\frac{1}{2}\int dt[(\frac{dq}{dt})^2+\omega^2q^2] + \int dt J(t)q(t)\}$$

Now in fact it is completely legitimate to carry out the partial integration without boundary terms: the paths are closed. Thus we get (cf. sect. 1.4.1)

$$\begin{aligned}
Z_E[J] &= \int \mathcal{D}q \exp\{-\frac{1}{2}\int dt q(t)[-\frac{d^2}{dt^2}+\omega^2]q(t) + \int dt J(t)q(t)\} \\
&= \mathcal{N}\exp\{\frac{1}{2}\int ds\int ds' J(s)\Delta_E(s,s')J(s')\} \\
\Delta_E(s,s') &= (-\frac{d^2}{ds^2}+\omega^2)^{-1}(s,s') \\
&= \int \frac{dE_E}{2\pi}\frac{e^{-iE_E(s-s')}}{E_E^2+\omega^2}
\end{aligned} \tag{1.100}$$

Contrary to the naive treatment given above in sect. 1.4.1, the present one in the euclidean case is quite correct.

Now let us compare this euclidean formulation with the real time formulation by "rotating back" to real time, i.e. to imaginary values of $(s - s')$:

$$t \equiv -is \text{ and } E_E s = Et \Rightarrow E_E = -iE \tag{1.101}$$

More precisely, we have a pole at $E_E = \pm i\omega$ and the rotation of the integration contour is valid as long as we do not cross the pole. Hence we should actually not rotate all the way, but rather

$$E_E \to e^{-i(\frac{1}{2}\pi - \epsilon)}E = -iE + \epsilon' = -i(E + i\epsilon')$$

for $E > 0$ (and something analogous for $E < 0$). Therefore (with s euclidean time and t real time)

$$\Delta_E(s = it) = -i\int_{-\infty}^{\infty} \frac{dE}{2\pi} \frac{e^{-iEt}}{-E^2 + \omega^2 - i\epsilon} = +i\Delta_F(t)$$

Also $dsds' = -dtdt'$ so that we have complete agreement with the result we found using the very detailed discussion of boundary conditions.

Clearly this euclidean formulation is very powerful. It allows us to derive the vacuum expectation values with great ease. Also it is mathematically better defined: a point we have not emphasized is that in the real time formulation, the "gaussian" integral is purely oscillatory so that strictly speaking its meaning is somewhat dubious. Instead the euclidean path integral is completely well defined. The relation to the real time formulation is obtained by analytically continuing back in the time variable. In so doing we may get arbitrarily close to singularities, however, when we start in the euclidean domain, we know from which side to approach them, as is shown by the appearance of the $i\epsilon$ here. Therefore it has become customary in many cases to actually work entirely in the euclidean formulation, knowing that whenever necessary we can always continue analytically to the physical domain. Indeed a puristic point of view is that this is the *only* way in which the real time Greens functions can be defined: as the above analytic continuation of the euclidean ones. Hence in this present course we shall mostly work in the euclidean formulation.

A further very nice observation is that in the euclidean formulation, the path integral exactly takes the form of a the statistical partition function of a fictitious associated physical system with *classical* hamiltonian equal to the positive definite euclidean action. The temperature here can be considered to be given by $\hbar$ (modulo normalization factors to render the dimensions right) when we re-insert that:

$$e^{-S_E} \to e^{-\frac{S_E}{\hbar}}$$

Notice, however, that the associated statistical system in the euclidean formulation, has a dimension of 1 higher than the quantum system in real time: In our case we started with a point particle (with a single degree of freedom, q), but the hamiltonian in the euclidean case takes the form of a 1-dimensional *integral* over some variable (euclidean time), showing that we have statistical mechanics of a one-dimensional object, a string-like system, the configurations of which are specified by a function, $q(t_E)$. Similarly, the quantum field theories with which we shall be concerned, and which live in 3-dimensional space (+ 1-dimensional time) will be related to 4-dimensional statistical systems by rotating to euclidean time.

The fact that in the euclidean formulation we have a path integral with a positive weight that may be given a classical statistical probability interpretation, rather than a quantum mechanical amplitude interpretation, is the starting point for numerical simulations of quantum field theories.

Chapter 2

Bosonic Field Theory

2.1 The Field Theory Transcription

We have been extremely careful in our discussion of the path integral formulation of a quantum system with a single degree of freedom. We shall now transcribe the results to systems with arbitrarily many, indeed infinitely many degrees of freedoms as if that did not cause any extra problems. The very essential extra problems which in fact do occur will be dealt with in connection with discussions of renormalization theory.

First, it really is trivial to generalize our previous results to the case of a system with an arbitrary, but finite number of degrees of freedom. We just have to path integrate over paths $q_i(t)$ for each value of i:

$$\prod_i \mathcal{D}q_i$$

Next we consider the field theory case. In order not to have too many indices let us use a notation pertaining to a single scalar field. The further generalization to fields with indices taking a finite number of values is quite straight forward and will mostly be treated by way of the examples we meet.

So our system, has a configuration space, characterized by "coordinates"

$$q_i(t) \rightarrow \phi_{\vec{x}}(t) \equiv \phi(\vec{x}, t) \equiv \phi(x)$$

The difficulty associated with renormalization is that we should really first consider discretized space, where the indices $\vec{x}$ live on some lattice, and then carefully consider taking the continuum limit. Alternatively some other regularization may be used. But we should start with a regularization. However, we ignore that complication for now.

So we consider the scalar field theory action (cf. sect. 1.2.1)

$$\begin{aligned} S[\phi] &= \int d^d x \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \\ \mathcal{L} &= -[\frac{1}{2}\partial_\mu \phi \partial^\mu \phi + \frac{1}{2} m^2 \phi^2 - V(\phi)] \end{aligned} \tag{2.1}$$

having in mind mostly dimension $d = 4$.

Thus the Minkowski Greens function or n-point function is

$$\begin{aligned} G_n(x_1, ..., x_n) &= \langle 0 | T\{\hat{\phi}(x_1)...\hat{\phi}(x_n)\} | 0 \rangle \\ &= \int \mathcal{D}\phi e^{iS[\phi]} \phi(x_1)...\phi(x_n) \end{aligned} \tag{2.2}$$

where $\hat{\phi}(x)$ denotes the quantum field operator and $\phi(x)$ the c-number field over which we path integrate. Similarly we shall express the euclidean Greens function in terms of the euclidean path integral over field developments, $\phi(\vec{x}, t)$, periodic in the euclidean time coordinate, t with infinite period in order to project the vacuum expectation value. The euclidean action for the scalar field theory above becomes

$$S_E[\phi] = \int d^dx [\frac{1}{2}\partial_\mu\phi\partial_\mu\phi + \frac{1}{2}m^2\phi^2 + V(\phi)] \tag{2.3}$$

where rotation to euclidean space means a change of sign on the square of the time derivative. In fact we use the following notation

$$\begin{aligned} t_M &\equiv x^0 = -x_0 \\ &= -it_E \\ t_E &\equiv x_4 \equiv x^4 = ix^0 \end{aligned} \tag{2.4}$$

Thus in the Minkowski case we use the metric

$$a_\mu b^\mu \equiv \vec{a} \cdot \vec{b} - a^0 b^0 \tag{2.5}$$

whereas in the euclidean case we do not distinguish between upper and lower indices, but write them wherever there is best space:

$$\begin{aligned} a_\mu b_\mu &\equiv a^\mu b^\mu \equiv \sum_{\mu=1}^{4} a_\mu b_\mu \\ &= a_\mu b_\nu \delta_{\mu\nu} \end{aligned} \tag{2.6}$$

Then the euclidean Greens function is given by the euclidean path integral

$$G_n(x_1, ..., x_n) = \int \mathcal{D}\phi e^{-S_E[\phi]}\phi(x_1)...\phi(x_n) \tag{2.7}$$

It is related to the Minkowski Greens function by analytic continuation in the time variables from real to imaginary euclidean time.

These Greens functions are conveniently collected in the generating functional or partition function, to borrow a word from statistical mechanics:

$$Z[J] \equiv \int \mathcal{D}\phi \exp\{-S_E[\phi] + J \cdot \phi\} \tag{2.8}$$

with

$$J \cdot \phi = \int d^dx J(x)\phi(x)$$

so that evidently

$$G_n(x_1, ..., x_n) = \frac{\delta}{\delta J(x_1)} \cdots \frac{\delta}{\delta J(x_n)} \int \mathcal{D}\phi \exp\{-S_E[\phi] + J \cdot \phi\}|_{J\equiv 0} \tag{2.9}$$

Then also

$$Z[J] = \sum_{n=0}^{\infty} \frac{1}{n!} \int \prod_{i=1}^{n} d^dx_i G_n(x_1, ..., x_n) J(x_1)...J(x_n) \tag{2.10}$$

2.1.1 On Greens Functions and Scattering Amplitudes

We now embark on the project of setting up perturbation theory for a general quantum field theory. We shall mostly study Greens functions in this course. However, let us begin by making a few remarks about how S-matrix elements could be constructed.

The first step in constructing perturbation theory is to make a split in the action

$$S[\phi] = S_0[\phi] + S_I[\phi] \tag{2.11}$$

where ϕ is a collective name for all the fields that might enter the theory. The first term is denoted the *free* part, and we imagine that we can solve the theory exactly if that was the only part. The second part is the perturbation which we would like to consider small in some sense and in which we perturb: make a systematic expansion, the convergence of which is usually more than doubtful, but the usefulness has nevertheless been demonstrated both in practice in many cases and in model mathematical analyses. Naturally, however, there are important cases where perturbation theory is generally believed insufficient. This is true of the phenomenon of quark confinement, but now we want to study perturbation theory.

We have seen how a free scalar field theory is equivalent to an infinite set of harmonic oscillators, one for each possible 3-momentum. Therefore, when we consider an S-matrix-element, corresponding to a scattering amplitude describing the transition between initial and final states specified by free particles in the infinite past and future, it is plausible that we could generalize our treatment of the harmonic oscillator and develop a path integral with suitable boundary conditions describing that situation. This indeed is possible.

In practice another method is often more convenient. Namely one starts with the Greens functions already introduced, and considers the Fourier-transforms or momentum space version:

$$\tilde{G}_n(p_1, ..., p_n) = \int d^dx_1 ... d^dx_n e^{ip_1x_1 + ... + ip_nx_n} G_n(x_1, ...x_n) \tag{2.12}$$

Actually translation invariance (in space and time)

$$G_n(x_1, ..., x_n) = G_n(x_1 - X, ..., x_n - X) \tag{2.13}$$

implies that this momentum space Greens function is proportional to a delta function expressing momentum conservation. Putting $X = x_1$ and changing integration variables $x_i \to x_i + X$, $i = 2, 3, ..., n$, we find

$$\begin{aligned} \tilde{G}_n(p_1, ..., p_n) &= \int d^dx_1 e^{ix_1(p_1 + ... + p_n)} \int d^dx_2 ... d^dx_n e^{i(x_2p_2 + ... + x_np_n)} G_n(0, x_2, ..., x_n) \\ &= (2\pi)^d \delta^d(p_1 + ... + p_n) G_n(p_1, ..., p_n) \end{aligned} \tag{2.14}$$

As we shall show later, it will happen that as a function of the momenta, these Greens functions develop a pole singularity on the mass shell at

$$p_i^2 = -m^2 \tag{2.15}$$

in addition to lots of other singularities. The S-matrix element pertains to scattering of particles all of which are on the mass shell. Therefore it is perhaps not surprising that the multiple *residue* at all the poles on these external lines (meaning in all the variables p_i^2) is rather simply related to the S-matrix element. For this reason our studies of the Greens functions will contain in principle all there is to know also about scattering amplitudes. For theoretical studies it is often preferable to stick to the very simple boundary conditions corresponding to looking at vacuum expectation values. Hence we concentrate on the Greens functions.

2.1.2 Dyson's Formula

Back to perturbation theory, let us denote by $|0\rangle$ the vacuum of the free theory. The (time-ordered in the Minkowski case) expectation value of any function of fields, $\mathcal{O}[\{\phi\}]$ (example: $\mathcal{O}[\{\phi\}] \equiv \phi(x_1)\phi(x_2)$) in the free theory is given by

$$\begin{aligned} & \langle 0|\mathcal{O}[\{\hat{\phi}\}]|0\rangle \\ = & \int \mathcal{D}\phi e^{-S_0[\phi]}\mathcal{O}[\{\phi\}] \end{aligned} \tag{2.16}$$

In particular the full path integral in the interacting theory including the interaction part S_I may be written

$$\begin{aligned} & \int \mathcal{D}\phi e^{-S_0[\phi]-S_I[\phi]} \\ = & \int \mathcal{D}\phi e^{-S_0[\phi]}e^{-S_I[\phi]} \\ = & \langle 0|e^{-S_I[\hat{\phi}]}|0\rangle \end{aligned} \tag{2.17}$$

This is essentially Dyson's formula. It was originally written down for the S-matrix, thus not for vacuum expectation values, and in Minkowski space, thus with the time-ordering sign (and $-S_I \to i(S_I)_{\text{Minkowski}}$).

For the Greens functions we get

$$\begin{aligned} G(x_1, ..., x_n) &= \langle 0|\hat{\phi}(x_1)...\hat{\phi}(x_n)e^{-S_I[\hat{\phi}]}|0\rangle \\ &= \int \mathcal{D}\phi e^{-S_0[\phi]}\phi(x_1)...\phi(x_n)e^{-S_I[\phi]} \end{aligned} \tag{2.18}$$

For the (euclidean) generating functional finally

$$\begin{aligned} Z[J] &= \langle 0|e^{-S_I[\hat{\phi}]}e^{\int d^dx J(x)\hat{\phi}(x)}|0\rangle \\ &= \int \mathcal{D}\phi e^{-S_0[\phi]+J\cdot\phi}e^{-S_I[\phi]} \end{aligned} \tag{2.19}$$

This is what we shall mostly mean by Dyson's formula. We see that the path integral formulation renders the derivation nearly trivial.

2.2 The Solution of Free Field Theory

Let us solve the euclidean free theory by gaussian integration. We consider

$$\begin{aligned} Z_0[J] &\equiv \int \mathcal{D}\phi e^{-S_0[\phi]+J\cdot\phi} \\ &= \int \mathcal{D}\phi \exp\{-\frac{1}{2}\int d^dx[\partial_\mu\phi\partial_\mu\phi + m^2\phi^2] + J\cdot\phi\} \end{aligned}$$

$$\begin{aligned}
&= \int \mathcal{D}\phi \exp\{-\frac{1}{2}\int d^dx\ \phi[-\partial_\mu\partial_\mu + m^2]\phi + J\cdot\phi\} \\
&= \int \mathcal{D}\phi \exp\{-\frac{1}{2}\phi\cdot\Delta^{-1}\cdot\phi + J\cdot\phi\} \\
&= \int \mathcal{D}\phi \exp\{-\frac{1}{2}[\phi - J\cdot\Delta]\Delta^{-1}[\phi - \Delta\cdot J] + \frac{1}{2}J\cdot\Delta\cdot J\} \\
&= e^{\frac{1}{2}J\cdot\Delta\cdot J}\int \mathcal{D}\phi' e^{-\frac{1}{2}\phi'\Delta^{-1}\phi'} \qquad (2.20)
\end{aligned}$$

where we refrained from using the standard rule of gaussian integration but just derived it instead, changing integration variable from ϕ to $\phi' \equiv \phi - \Delta\cdot J$. As in the case of the harmonic oscillator, the remaining path integral is uninteresting, since it does not depend on anything of physical significance, it is just the vacuum normalization. Thus

$$Z_0[J] = e^{\frac{1}{2}J\cdot\Delta\cdot J}\langle 0|0\rangle \qquad (2.21)$$

with

$$\langle 0|0\rangle = Z_0[0]$$

which we shall put to 1. We easily find the propagator in the euclidean case, just as for the harmonic oscillator. It is the well-defined inverse of the operator

$$\Delta^{-1} \equiv -\partial_\mu\partial_\mu + m^2$$

Hence

$$\Delta(x,x') = \int \frac{d^dp}{(2\pi)^d}\frac{e^{ip(x-x')}}{p^2+m^2} \qquad (2.22)$$

with a nice, non-singular integrand. This euclidean propagator is continued in the variable p^2 (in momentum space) to allow negative values for that in the Minkowski case. In analogy with the case of the harmonic oscillator, the effect is to substitute

$$p_E^2 + m^2 \to p_M^2 + m^2 - i\epsilon$$

where the Minkowski value, p_M^2 can be negative.

2.3 Wick's Theorem

Let us consider any function of field variables, $F[\{\phi\}]$. As an example consider

$$F[\{\phi\}] \equiv \phi^2(x_1)\phi(x_2)\phi^4(x_3)$$

We can immediately write down an expression for the vacuum expectation value of such a function in the free theory, using the free generating functional. Indeed for the example

$$\begin{aligned}
\langle 0|F[\{\hat\phi\}]|0\rangle &= \int \mathcal{D}\phi e^{-S_0}\phi^2(x_1)\phi(x_2)\phi^4(x_3) \\
&= \left(\frac{\delta}{\delta J(x_1)}\right)^2 \frac{\delta}{\delta J(x_2)}\left(\frac{\delta}{\delta J(x_3)}\right)^4 \int \mathcal{D}\phi e^{-S+J\phi}|_{J\equiv 0} \qquad (2.23)
\end{aligned}$$

This example should make it clear that we have the general formula

$$\langle 0|F[\{\hat\phi\}]|0\rangle_J = F[\{\frac{\delta}{\delta J}\}]Z_0[J]|_J \qquad (2.24)$$

We now use this rule to calculate the partition function of the full interacting theory in terms of the one of the free theory. We use Dyson's formula eq.(2.19). Hence

$$F[\{\phi\}] = \exp\{-\int d^d x V(\phi(x))\}$$

where $V(\phi(x))$ is the interaction lagrangian. We have Wick's theorem

$$\begin{aligned} Z[J] &= \exp\{-\int d^d x V(\frac{\delta}{\delta J(x)})\} Z_0[J] \\ &= \exp\{-\int d^d x V(\frac{\delta}{\delta J(x)})\} \exp\{\frac{1}{2} J \Delta J\} \end{aligned} \qquad (2.25)$$

Had it not been for several problems, this would have been a most wonderful formula. It seems we have provided an exact solution of the full interacting quantum field theory. The solution is given in terms of a very complicated differential operator acting on the rather simple generating functional of the free theory. In practice the differential operator is unmanageably complicated unless we can allow ourselves a power expansion of the exponent hitting $Z_0[J]$. This is perturbation theory. Even here, as we shall see, the expression leads to divergent integrals. The meaning of those can be made clear however, by rethinking everything through in the light of a regularization. Again this is renormalization, which we do not cover for the time being. Suffice it to say that the illnesses can be cured in renormalizable theories. For now we therefore go on and use Wick's theorem as a starting point for the derivation of Feynman rules.

However, let us first pause to consider some simple examples of how the theorem works.

2.3.1 Examples of the Use of Wick's Theorem

Let us start by considering the free theory. Then let us work out the one-point function in the presence of the external current:

$$\begin{aligned} G_1^{(0)}(x_1)_J &= \frac{\delta}{\delta J(x_1)} e^{\frac{1}{2} J \Delta J} \\ &= \Delta J(x_1) e^{\frac{1}{2} J \Delta J} \end{aligned} \qquad (2.26)$$

The first subscript on G tells us that it is a 1-point function, the second that it is evaluated in the presence of the external current, whereas the superscript indicates that we consider the free theory. Also, our abbreviated notation stands for

$$\Delta J(x_1) e^{\frac{1}{2} J \Delta J} = \int d^d x \Delta(x_1, x) J(x) \cdot \exp\{\frac{1}{2} \int d^d x d^d y J(x) \Delta(x, y) J(y)\}$$

Notice that the propagator eq.(2.22) is symmetric in its variables $\Delta(x, y) = \Delta(y, x)$. We finish the calculation of the one-point function by putting $J = 0$:

$$G_1^{(0)}(x_1) \equiv 0$$

Similarly every n-point function is easily seen to be zero when n is odd.

Figure 2.1: Propagator in x-space

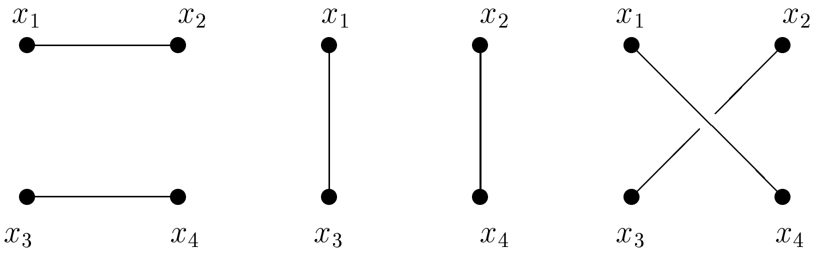


Figure 2.2: Feynman diagrams for the free 4-point function

We proceed to the 2-point-function:

$$\begin{aligned} G_2^{(0)}(x_1,x_2)_J &= \frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}e^{\frac{1}{2}J\Delta J} \\ &= \frac{\delta}{\delta J(x_1)}\{\Delta J(x_2)e^{\frac{1}{2}J\Delta J}\} \\ &= \{\Delta(x_1,x_2)+\Delta J(x_2)\Delta J(x_1)\}e^{\frac{1}{2}J\Delta J} \end{aligned} \tag{2.27}$$

Again we finish by putting $J=0$, giving

$$G_2^{(0)}(x_1,x_2)=\Delta(x_1,x_2) \tag{2.28}$$

This is an important result: The propagator is the free two-point function:

$$\Delta(x_1,x_2)=\langle 0|\hat{\phi}(x_1)\hat{\phi}(x_2)|0\rangle \tag{2.29}$$

We introduce for this object the "Feynman-diagram" fig. 2.1. The reader is invited to work out the 4-point function with the result

$$\begin{aligned} G_4^{(0)}(x_1,x_2,x_3,x_4) &= \prod_{i=1}^{4}\frac{\delta}{\delta J(x_i)}e^{\frac{1}{2}J\Delta J} \\ &= \Delta(x_1,x_2)\Delta(x_3,x_4)+\Delta(x_1,x_3)\Delta(x_2,x_4) \\ &+ \Delta(x_1,x_4)\Delta(x_3,x_2) \end{aligned} \tag{2.30}$$

with the Feynman diagram fig.2.2.

Clearly this result generalizes into the following Feynman rule for the free n-point function:

To calculate $G_n^{(0)}(x_1,...,x_n)$, for n even, mark n point on a piece of paper and name the points $x_1,...,x_n$. Construct all possible Feynman diagrams by connecting the points pairwise by propagators. For each diagram construct an amplitude which is the product

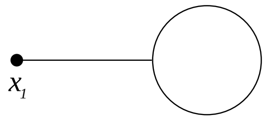


Figure 2.3: The "tadpole" diagram

of all these propagators, $\Delta(x_i, x_j)$ for the propagator linking points x_i, x_j. Then the free n-point function is given by the sum of amplitudes for individual Feynman diagrams.

Next let us consider a simple interaction term

$$V(\phi) = \frac{g}{3!}\phi^3$$

Actually this will give us a theory in which the hamiltonian will be unbounded from below, because the term ϕ^3 can become arbitrarily large, negative. However, the perturbation theory is well defined order by order in the coupling constant, and this example is just meant as a typical illustration. In QCD for instance there will be both triple-gluon interactions (analogous to the present case) and 4-gluon interactions, that will repair the hamiltonian.

To zero'th order in the coupling constant, g, the Greens functions are just the free ones before. So let us look at the first order in g. Hence we replace

$$e^{-\int d^dx \frac{g}{3!}(\frac{\delta}{\delta J(x)})^3} \to -\int d^dx \frac{g}{3!}(\frac{\delta}{\delta J(x)})^3$$

So the order g^1 term in the generating functional is

$$\begin{aligned} & -\int d^dx \frac{g}{3!}(\frac{\delta}{\delta J(x)})^3 e^{\frac{1}{2}J\Delta J} \\ = & -\frac{g}{3!}\int d^dx (\frac{\delta}{\delta J(x)})^2 \{\Delta J(x) e^{\frac{1}{2}J\Delta J}\} \\ = & -\frac{g}{3!}\int d^dx \frac{\delta}{\delta J(x)}[\{\Delta(x,x) + (\Delta J(x))^2\} e^{\frac{1}{2}J\Delta J}] \\ = & -\frac{g}{3!}\int d^dx \{3\Delta(x,x)\Delta J(x) + (\Delta J(x))^3\} e^{\frac{1}{2}J\Delta J} \end{aligned} \tag{2.31}$$

We see that to this order, in fact we can only produce n-point functions with n odd (due to the odd nature of the interaction). Hence we may start by looking at the 1-point function, normally something one does not consider. Still we get a formal expression by taking one functional derivative with respect to $J(x_1)$ and then putting $J \equiv 0$ as usual:

$$G_1^{(1)}(x_1) = -\frac{1}{2}g\int d^dx \Delta(x,x)\Delta(x,x_1) \tag{2.32}$$

This expression may be represented by the Feynman diagram fig.2.3. Notice that the vertex with no name attached is integrated over. Looking at the formula for the propagator, we easily see that the expression $\Delta(x,x)$ diverges in space time dimensions greater than 1 (!) In most regularization schemes one can in fact show that such contributions can often be ignored. However, that is not always the case. For example the phenomenon of spontaneous symmetry breaking which is important in the Weinberg-Glashow-Salam

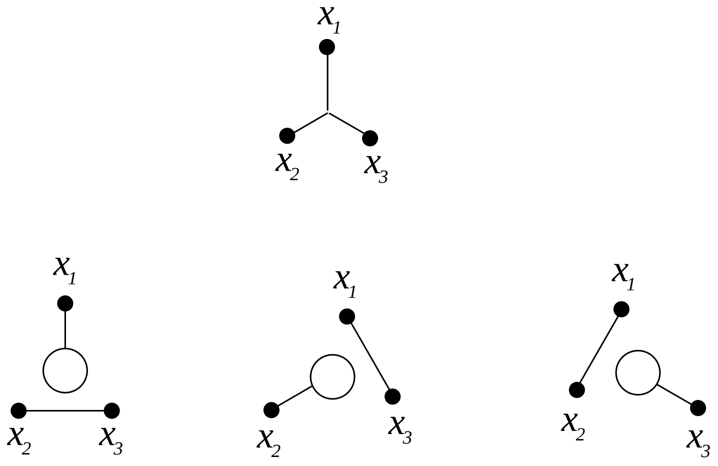


Figure 2.4: Feynman diagrams for the lowest order 3-point function in ϕ^3 theory

model, can be viewed in these terms. Our present concern, however, is only to study the general structure of the theory, so we shall leave these ill defined expressions standing.

Next we consider the 3-point function $G_3^{(1)}(x_1, x_2, x_3)$ to first order in g. To this purpose we only need the third order expression in the currents, so we may use the following approximation for the generating functional:

$$-\frac{g}{3!}\int d^dx\{3\Delta(x,x)\Delta J(x)\frac{1}{2}J\Delta J + (\Delta J(x))^3\}$$

It is now straight forward to act successively with $\frac{\delta}{\delta J(x_1)}$, $\frac{\delta}{\delta J(x_2)}$, $\frac{\delta}{\delta J(x_3)}$ and after putting $J \equiv 0$ obtain the result

$$\begin{aligned} & -g\int d^dx\{\Delta(x,x_1)\Delta(x,x_2)\Delta(x,x_3) \\ + & \frac{1}{2}\Delta(x,x)[\Delta(x,x_1)\Delta(x_2,x_3) \\ + & \Delta(x,x_2)\Delta(x_1,x_3) \\ + & \Delta(x,x_3)\Delta(x_2,x_1)]\} \end{aligned} \tag{2.33}$$

with the Feynman diagrams given in fig. 2.4. As a final example, consider the 0-point function to order g^2. To that order the generating functional is given by

$$\frac{1}{2!}\int d^dx\frac{g}{3!}(\frac{\delta}{\delta J(x)})^3\int d^dy\frac{g}{3!}(\frac{\delta}{\delta J(y)})^3 e^{\frac{1}{2}J\Delta J}$$

Now to get a 0-point function we should expand this expression to 0th order in the current. We have a total of 6 functional derivatives hitting Z_0. they will eat up 6 current factors. It follows that we must consider the term in Z_0 which is of order 6 in the current. That is the 3d order term in the expansion of the exponential:

$$\begin{aligned} Z_0 \rightarrow & \frac{1}{3!}(\frac{1}{2})^3\int d^dz_1d^dz_2d^dz_3d^dz_1'd^dz_2'd^dz_3' \\ \{ & J(z_1)\Delta(z_1,z_1')J(z_1') \\ \cdot & J(z_2)\Delta(z_2,z_2')J(z_2') \\ \cdot & J(z_3)\Delta(z_3,z_3')J(z_3')\} \end{aligned} \tag{2.34}$$

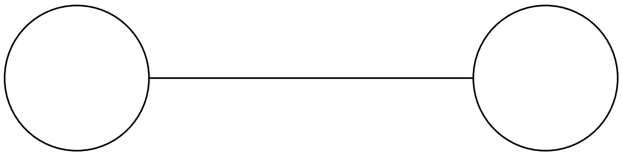


Figure 2.5: A second order vacuum bubble in ϕ^3 theory

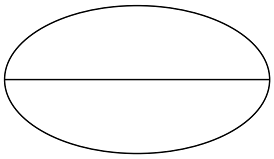


Figure 2.6: Another vacuum bubble to second order in ϕ^3 theory

When we let the 6 functional derivatives eat away at the J factors (meaning: when we carry out the differentiations), we see that we shall be able to produce two different kinds of terms:

1. It may happen that two functional derivatives with J having the *same* argument (say x), eat up the two J factors surrounding the *same* propagator (say $\Delta(z_1, z_1'))$ resulting in a term with a factor $\Delta(x,x)$ after integration over z_1, z_1'. By necessity, the remaining functional derivatives will finish by giving a term containing

   $$\frac{g^2}{2\cdot 3!}\int d^dx d^dy \Delta(x,x)\Delta(x,y)\Delta(y,y)$$

   corresponding to the Feynman diagram fig. 2.5. The reader is invited to work out the weight factor indicated. It comes about by considering all the different combinations giving the same contribution.

2. It may happen that we let the functional derivatives eat J factors in such a way, that all derivatives with J having the same argument are used on *different* propagators. This will give us a term

   $$\frac{g^2}{2\cdot 3!}\int d^dx d^dy \Delta^3(x,y)$$

   corresponding to the Feynman diagram fig. 2.6. Again the reader can try to figure out the combinatorics giving the overall factor.

Now hopefully the general idea is clear. It should also be clear that some general technique and insight would be very helpful. Hence the work involved in going from the Wick theorem level to the Feynman-rule level.

2.4 The Feynman Rules in x-Space

2.4.1 An Alternative Version of Wick's Theorem

First we recast Wick's theorem into a form, slightly more convenient for our purpose:

Wick's Theorem (second version)

$$Z[J]=e^{\frac{1}{2}\frac{\delta}{\delta\phi}\Delta\frac{\delta}{\delta\phi}}\{e^{-\int V[\phi]+J\phi}\}_{\phi\equiv 0} \tag{2.35}$$

Let us repeat the formula in detail:

$$\begin{aligned} Z[J] &= \exp\{\frac{1}{2}\int d^dx d^dy \Delta(x-y)\frac{\delta}{\delta\phi(x)}\frac{\delta}{\delta\phi(y)}\} \\ &\times \{\exp[\int d^dz(-V(\phi(z))+J(z)\phi(z))]\}_{\phi\equiv 0} \end{aligned} \tag{2.36}$$

The meaning is that the functional differential operator expressed by the first factor has to act on the second factor, and afterwards one should put $\phi\equiv 0$.

Eq.(2.35) is a direct consequence of Wick's theorem in the form eq.(2.25) and the following Lemma:

Lemma (Coleman)

Let F and G be functions of n-dimensional vectors $(x)=(x_1,...,x_n)$ or $(y)=(y_1,...,y_n)$:

$$F(x)=F(x_1,...,x_n),\quad G(y)=G(y_1,...,y_n)$$

Then

$$F(\frac{\partial}{\partial x})G(x)=G(\frac{\partial}{\partial y})\{F(y)e^{x\cdot y}\}_{y\equiv 0} \tag{2.37}$$

Proof:

By Fourier decomposition it is enough to prove the lemma for functions of the form

$$F(x)=e^{a\cdot x},\quad G(x)=e^{b\cdot x}$$

where $a\cdot x=a_1x_1+...+a_nx_n$ etc. Then consider the left hand side of (2.37):

$$F(\frac{\partial}{\partial x})G(x)=e^{a\cdot\frac{\partial}{\partial x}}e^{b\cdot x}=\sum_{m=0}^{\infty}\frac{1}{m!}(a\cdot\frac{\partial}{\partial x})^m e^{b\cdot x}$$

Now

$$\frac{\partial}{\partial x_i}e^{b\cdot x}=b_ie^{b\cdot x}$$

and

$$(\frac{\partial}{\partial x_i})^m e^{b\cdot x}=(b_i)^m e^{b\cdot x}$$

showing that the action of any *function of* $\partial/\partial x_i$ results in multiplication by the same *function of* b_i. Hence we get for the left hand side

$$e^{a\cdot\frac{\partial}{\partial x}}e^{b\cdot x}=e^{a\cdot b}e^{b\cdot x}=e^{b\cdot(a+x)}$$

Similarly the right hand side gives

$$\begin{aligned} & G(\frac{\partial}{\partial y})\{F(y)e^{x\cdot y}\}_{y\equiv 0}=e^{b\cdot\frac{\partial}{\partial y}}\{e^{a\cdot y}e^{x\cdot y}\}_{y\equiv 0} \\ =\ & e^{b\cdot(a+x)}e^{(a+x)\cdot y}|y\equiv 0 \\ =\ & e^{b\cdot(a+x)} \end{aligned} \tag{2.38}$$

This completes the proof of the lemma.

We now apply the lemma to the form of Wick's theorem, eq.(2.25) for $Z[J]$ with the following "translation": The vector (x) with components x_i is replaced by the "vector" J with components $J_x \equiv J(x)$. The vector (y) is likewise replaced with the trial field ϕ with components $\phi(x)$. Also

$$x \cdot y \to J \cdot \phi = \int d^d x J(x) \phi(x)$$

etc. For $F(x)$ we use

$$F[J] = e^{-\int d^d x V(J(x))}$$

and for $G(y)$ we use

$$G(\phi) = e^{\frac{1}{2}\phi\Delta\phi}$$

Then the left hand side is our old expression eq.(2.25) for Wick's theorem, and the right hand side is the new one eq.(2.35).

2.4.2 Statement of the Feynman Rules in x-Space

For definiteness let us consider a general interaction of the form

$$V(\phi) = g\phi^p \tag{2.39}$$

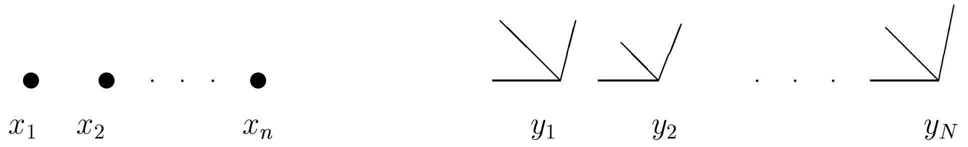
where p is a positive integer (usually greater than 2, since $p = 2$ corresponds to a mass term and could have been absorbed in the "free part" of the theory). It is then easy to see what happens if we have a sum of terms like that.

We now state the

Feynman rules in x-space for constructing an n-point Greens function in Nth order of perturbation theory in g

as the following 4 rules, later to be supplemented:

Rule I On a piece of paper, mark n dots and label them with the names of the arguments of the Greens function, $x_1, ..., x_n$. Further mark out N "vertices" like this



with p "legs" emanating from each, and mark the vertices $y_1, ..., y_N$.

Rule II Connect the dots to the legs of the vertices by simple lines. At this level the legs are treated as distinct. Draw a diagram for each way in which this may be done. These are called the Feynman diagrams in x-space.

Rule III For each diagram, D, construct it's *value* as follows:

$$F_D^{(N)}(x_1, ..., x_n) = \int d^d y_1 \cdots d^d y_N I_D^{(N)}(x_1, ..., x_n; y_1, ..., y_N) \tag{2.40}$$

$I_D^{(N)}$ is a product of factors:

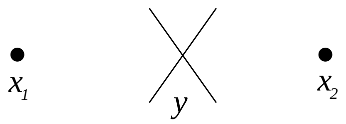
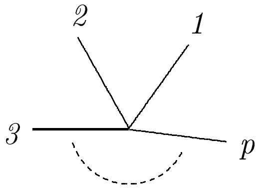


Figure 2.7: Beginning of the order g^1 Feynman diagrams for the 2-point function in ϕ^4 theory.

- For each vertex, put a vertex-factor

$-g =$



p legs

- For each line connecting dots and/or vertex-legs, put a propagator-factor

$\Delta(z-w) =$

where z and w are the labels of the dot and the vertex.

Rule IV

$$G^{(N)}(x_1, ..., x_n) = \frac{1}{N!}\sum_D F_D^{(N)}(x_1, ..., x_n) \tag{2.41}$$

Finally then

$$G(x_1, ..., x_n) = \sum_{N=0}^{\infty} G^{(N)}(x_1, ..., x_n) \tag{2.42}$$

These are in principle the complete rules of perturbation theory. *But* very many diagrams usually give exactly the same contribution. Often this allows a simplification in the rules, in particular when we consider the momentum space versions. Then quite often the factor $\frac{1}{N!}$ gets at least partly cancelled. We shall come back to all that. But it is useful to first prove the above 4 rules. Before we do that we consider an example.

Example of the Use of Feynman rules in x-space: The 2-point function to order g^1 for $p = 4$.

We start marking out the two dots, x_1, x_2 and the one vertex, y fig.2.7. Then we produce all possible diagrams. First the ones where x_1 is connected to x_2, i.e. we start with fig.2.8. and get figs. 2.9. These diagrams not only give the same contribution, they even give the same integrand. And we find

$$F_{D_1} = F_{D_2} = F_{D_3} = -g\int d^d y \Delta(x_1 - x_2)\Delta(y - y)\Delta(y - y)$$

Next we consider diagrams where x_1 is connected to one of the 4 legs from the y vertex. Choosing a leg, x_2 has to be connected to one of the remaining 3 legs, and the diagram

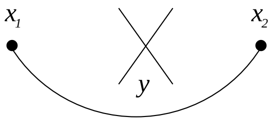


Figure 2.8: Beginning of the disconnected diagrams.

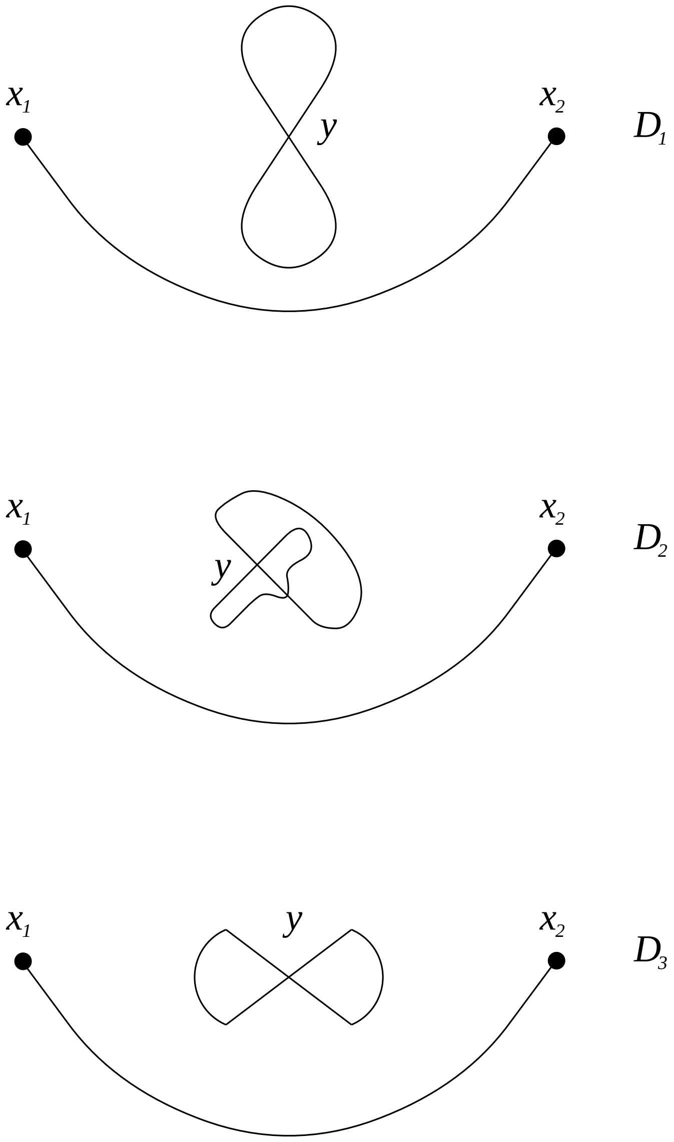


Figure 2.9: The 3 topologically equivalent disconnected diagrams.

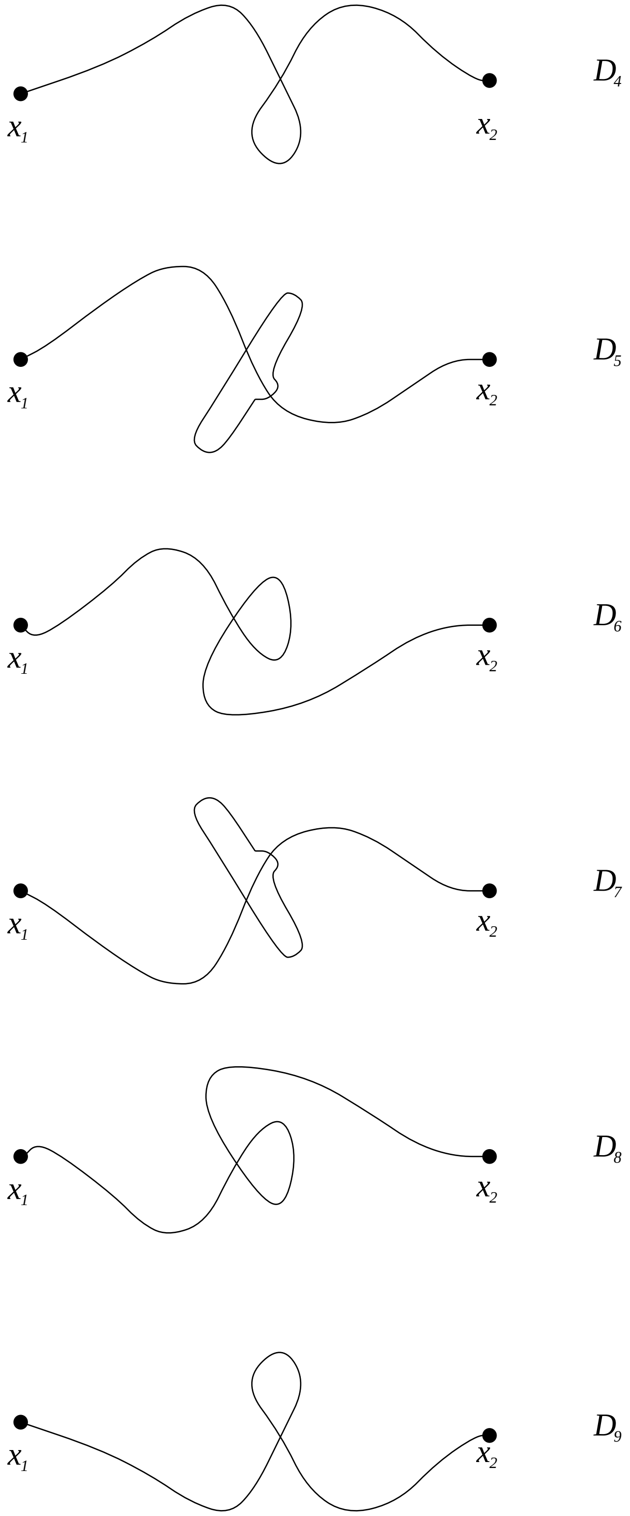


Figure 2.10: The first half of the topologically equivalent connected diagrams.

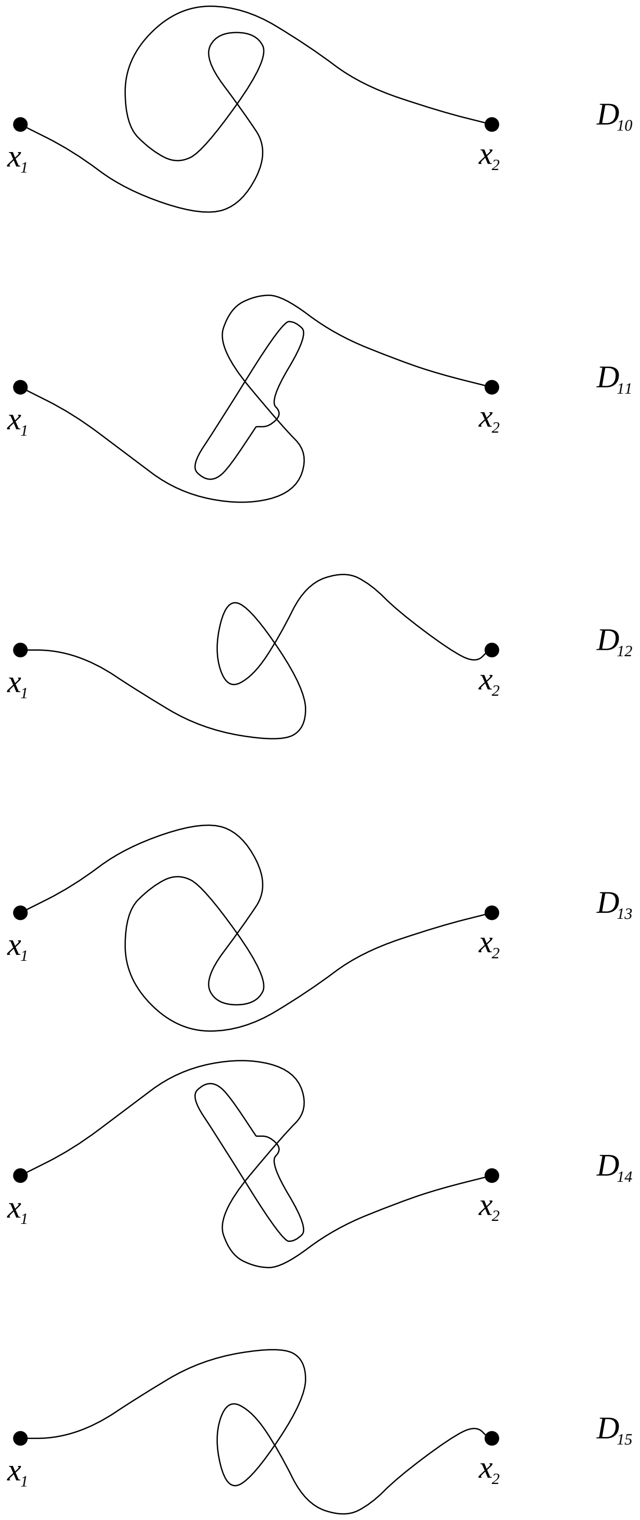


Figure 2.11: The second half of the topologically equivalent connected diagrams.

can be completed in just one way by connecting the last two legs: figs. 2.10,2.11. Again, evidently all these diagrams give exactly the same contribution, and we get

$$F_{D_4} = \cdots = F_{D_{15}} = -g \int d^d y \Delta(x_1 - y)\Delta(x_2 - y)\Delta(y - y)$$

This completely trivial example is already quite complicated to analyze using rules I–IV. Hence, we shall soon be very interested in deriving some simplifications. Nevertheless they form a good starting point.

2.4.3 Proof of the Feynman Rules

Using our second version of Wick's theorem and the definition of the Greens functions we get

$$\begin{aligned} G(x_1, ..., x_n) &= \frac{\delta^n Z[J]}{\delta J(x_1) \cdots \delta J(x_n)}|_{J=0} \\ &= e^{\frac{1}{2}\frac{\delta}{\delta\phi}\Delta\frac{\delta}{\delta\phi}}\{\phi(x_1)\cdot ... \cdot \phi(x_n) e^{-\int d^d x [V(\phi(x)) - J(x)\phi(x)]}\}_{\phi=0}|_{J=0} \quad (2.43)\end{aligned}$$

where we have to put first $J = 0$ and then $\phi = 0$ after the ϕ-derivatives have been carried out. But that shows we may drop the $J\phi$ term in the last exponential:

$$G(x_1, ..., x_n) = e^{\frac{1}{2}\frac{\delta}{\delta\phi}\Delta\frac{\delta}{\delta\phi}}\{\phi(x_1)\cdot ... \cdot \phi(x_n) e^{-\int d^d x V(\phi(x))}\}_{\phi=0} \quad (2.44)$$

We now want to consider the contribution from N'th order perturbation theory in g. It is obtained by the N'th term in the expansion of e^{-V}:

$$\begin{aligned} G^{(N)}(x_1, ..., x_n) &= e^{\frac{1}{2}\frac{\delta}{\delta\phi}\Delta\frac{\delta}{\delta\phi}} \\ &\times \{\phi(x_1)\cdot ... \cdot \phi(x_n)\frac{(-g)^N}{N!} \\ &\times \int d^d y_1 ... d^d y_N \phi^p(y_1)\cdot ... \cdot \phi^p(y_N)\}_{\phi\equiv 0} \quad (2.45)\end{aligned}$$

Now we begin to see some of the features of the rules emerging: the factors $\phi(x_1)...\phi(x_n)$ are responsible for the dots, labelled $x_1, ..., x_n$; the factors $\phi^p(y_1)...\phi^p(y_N)$ are responsible for the N p-leg vertices, labelled $y_1, ..., y_N$. We see that there is a factor $-g$ for each vertex and that we have a factor $1/N!$, as well as an integration over the position of the vertices.

But we still have to carry out the $\delta/\delta\phi$ differentiations. The curly bracket contains a total of

$$Q = n + p \cdot N$$

ϕ-factors. Since we put $\phi \equiv 0$ in the end, we need exactly the term with Q derivatives. We also see that the number of derivatives will always be even. Hence we can only get a contribution different from zero provided

$$Q = 2q$$

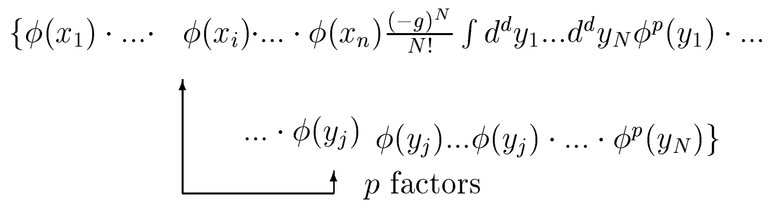
is even so that q is a positive integer. We then find

$$\begin{aligned} G^{(N)}(x_1,...,x_n) &= \frac{1}{q!}\frac{1}{2^q}\int d^dz_1d^dw_1\cdot...\cdot d^dz_qd^dw_q \\ & \quad \frac{\delta}{\delta\phi(z_1)}\Delta(z_1-w_1)\frac{\delta}{\delta\phi(w_1)}\cdot...\cdot\frac{\delta}{\delta\phi(z_q)}\Delta(z_q-w_q)\frac{\delta}{\delta\phi(w_q)} \\ &\times \quad \{\phi(x_1)\cdot...\cdot\phi(x_n)\frac{(-g)^N}{N!}\int d^dy_1...d^dy_N\phi^p(y_1)\cdot...\cdot\phi^p(y_N)\} \end{aligned} \tag{2.46}$$

Thanks to our second form of Wick's theorem, which we have been using, we are very close to the end. In fact the differential operators

$$\int d^dzd^dw\frac{\delta}{\delta\phi(z)}\Delta(z-w)\frac{\delta}{\delta\phi(w)}$$

exactly have the effect corresponding to connecting dots and legs by lines as in Rule II. For example, if $\delta/\delta\phi(w)$ acts on a factor $\phi(y)$, the result is $\delta^d(y-w)$ and the integral over w implies that we get $\Delta(z-w)$ replaced by $\Delta(z-y)$. Similarly with $\delta/\delta\phi(z)$. We see that we may keep track of the result by drawing a line between the two factors ϕ on which the $\delta/\delta\phi$'s acts. Thus, if $\delta/\delta\phi(w)$ acts on $\phi(x_i)$ and $\delta/\delta\phi(z)$ acts on the 2'nd factor $\phi(y_j)$, we draw the line in eq.2.47.



$$\{\phi(x_1)\cdot...\cdot\ \phi(x_i)\cdot...\cdot\phi(x_n)\tfrac{(-g)^N}{N!}\int d^dy_1...d^dy_N\phi^p(y_1)\cdot...$$

$$...\cdot\phi(y_j)\ \underbrace{\phi(y_j)...\phi(y_j)}_{p \text{ factors}}\cdot...\cdot\phi^p(y_N)\} \tag{2.47}$$

Such a line is called a Wick contraction, and we see that it gives the propagator factor $\Delta(x_i-y_j)$. Now there is a similar term when $\delta/\delta\phi(w)$ acts on the 2'nd factor $\phi(y_j)$ and $\delta/\delta\phi(z)$ acts on $\phi(x_i)$. This gives a factor 2 for each factor

$$\frac{\delta}{\delta\phi}\Delta\frac{\delta}{\delta\phi}$$

so that the factor $1/2^q$ gets cancelled. Likewise the contribution eq.(2.47) is obtained by letting *any* of the q factors with a propagator surrounded by two functional derivatives after ϕ act. Hence the factor $1/q!$ is also cancelled.

This completes the proof of rules I–IV.

We finish by the remark that a contribution identical to the one in eq.(2.47) is obtained by letting the Wick contraction line end on *any* of the factors $\phi(y_j)$. Hence we shall often get $p!$ identical contributions from fully contracted expressions where the Wick-contractions on the factors $\phi(y_j)...\phi(y_j)$ are permuted in all possible ways. This indicates that we might get a nicer notation if we write

$$g=\frac{\lambda_p}{p!}$$

Notice however, that the example in the previous subsection indicates that the situation is not always so simple. In figs.2.9 we only have 3 Wick contractions (D_1, D_2, D_3) instead of $4! = 24$. In figs.2.10,2.11 we have 12 $(D_4, ..., D_{15})$ instead of 24. We shall have to learn how to deal with this.

2.5 The Generating Functional for Connected Greens Functions

We have seen in the examples, that to a certain order in perturbation theory, very many Feynman diagrams are disconnected diagrams. It also follows from the form of the Feynman rules, that the corresponding contribution to the Greens function is simply the product of that of the sub diagrams of the disconnected diagram in question. For this reason it is enough to calculate all the connected diagrams, the disconnected ones being trivially formed.

Now it turns out that there is a very beautiful expression for the generating functional of connected diagrams. Indeed, the generating functional $Z[J]$ for the *full* Greens functions plays a role very analogous to the partition function in statistical mechanics, as we have discussed. In fact it *is* the partition function of a statistical system in d dimensions in the euclidean case. We now have the following:

Theorem

The generating functional, $-W[J]$, of connected Greens functions, is minus the free energy, $W[J]$, defined by:

$$Z[J] = e^{-W[J]} \tag{2.48}$$

Proof:

It is convenient to have the concept of n-point functions *in the presence of the external current,* J, i.e. without putting $J \equiv 0$. Such an n-point function is naturally defined as

$$G(x_1, ..., x_n)_J = \frac{\delta^n}{\delta J(x_1)...\delta J(x_n)} Z[J] \tag{2.49}$$

everything *without* putting $J \equiv 0$. At the very end we may still want to do that.

A further useful object, we shall come back to is the 1-point function in the presence of J. The lowest order contributions in a ϕ^4 theory is given by

$$\begin{aligned} G(x)_J &= \int d^d y \Delta(x-y) J(y) \\ &- g \int d^d z d^d y_1 d^d y_2 d^d y_3 \Delta(x-z)\Delta(y_1-z)\Delta(y_2-z)\Delta(y_3-z) J(y_1) J(y_2 J(y_3) \\ &+ ... \end{aligned} \tag{2.50}$$

corresponding to the Feynman diagrams of fig.2.12. Here and in future, we shall use the notation, that a cross indicates a current-factor. Further, vertices and crosses, that are not labelled imply that we integrate over the arguments. Let us also introduce the symbol

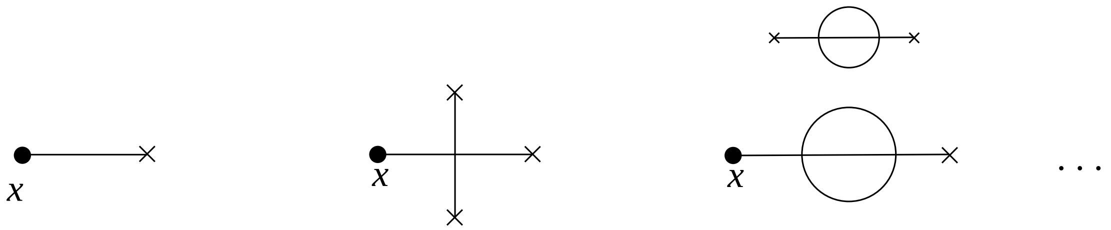


Figure 2.12: The first contributions to the 1-point function in ϕ^4 theory, in the presence of an external current (the crosses).

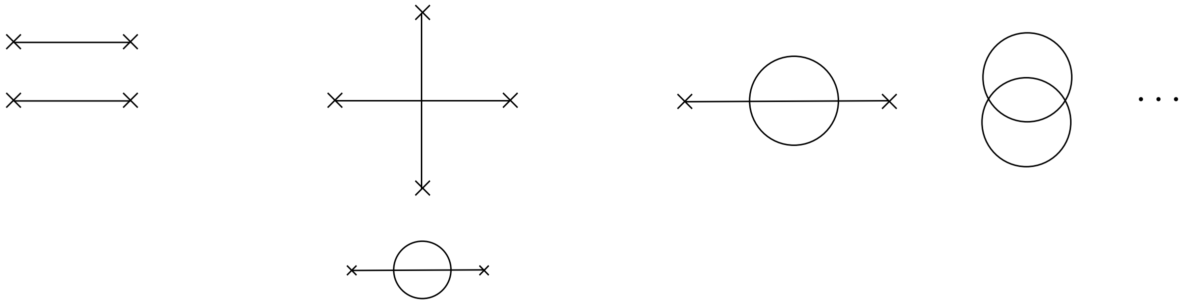
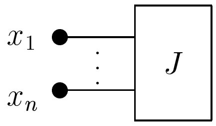
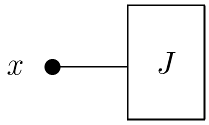


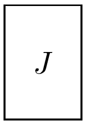
Figure 2.13: Some vacuum diagrams in the presence of an external current in ϕ^4 theory.



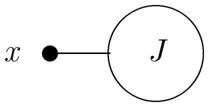
for the general, full n-point Greens function, in the presence of the current, J. In particular



is the full 1-point function, and

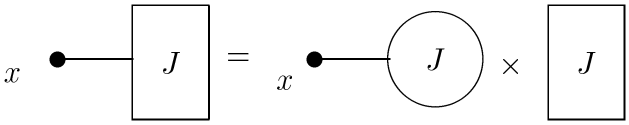


is the sum of all vacuum blobs, cf. fig.2.13. Similarly, let us use the symbol



for the sum of all *connected* diagrams for the 1-point function.

By looking at the diagrams and remembering how the Feynman rules work for disconnected pieces, it is immediately obvious, that the sum of all diagrams for the full 1-point function, is obtained by taking all the connected diagrams and multiplying them with all possible vacuum bubbles in all possible ways:



The left hand side is the sum of all Feynman diagrams with one external leg ending in the point, *x and in the presence of the external current,* i.e. with “external legs” ending in crosses representing factors of J at those points, which are then integrated over. The equation says that we may obtain this sum (which of course includes all sorts of disconnected components) by multiplying the sum og all *connected* one point diagrams with the sum of all vacuum bubbles, connected or disconnected (and possibly containing “external legs ending in crosses). Notice that this split up can work only for 1-point functions. For the 2-point function, we might have the two external legs belonging to disconnected pieces of the combined diagram. But the one external leg in the 1-point function neccessarily is attached to a diagram which is part of the sum of connected diagrams for the connected 1-point function.

All this may be expressed as

$$\frac{\delta Z[J]}{\delta J(x)} = -\frac{\delta W[J]}{\delta J(x)} Z[J]$$

(we have introduced the minus sign merely to get the free energy in the end). This is a (functional) differential equation, the solution of which is immediately written down as

$$Z[J] = \mathcal{N} e^{-W[J]}$$

where $\mathcal{N}$ is an integration constant, fixed by choosing a certain normalization for the vacuum bubbles. It clearly plays no role for the connected diagrams themselves, as long as it does not depend on anything.

This completes the proof of the theorem.

2.6 The Statistical Weight Factor

Because of the integrations implied in Feynman Rule III, it is clear that we have

Rule V Topologically equivalent diagrams give the same contribution.

Hence, in practice we only draw *topologically inequivalent* diagrams and then count the number of them. Also, using the coupling constant, $\lambda = \lambda_p$

$$g = \frac{\lambda_p}{p!}$$

we use the Feynman rules in the form

$\Delta(x-y) =$


$-\lambda =$

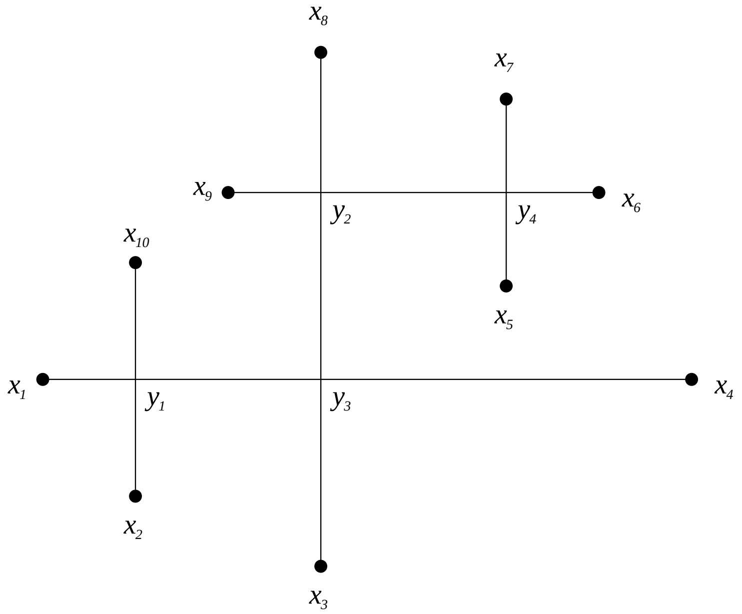



Figure 2.14: A rather general tree diagram in a ϕ^4 theory

Then we construct the diagram for each topologically inequivalent diagram, D and associate with that a weight factor w_D.

We now look into the construction of w_D.

First we prove that:

A tree diagram has weight factor, $w_{\mathrm{tree}} = 1$

To get the idea, consider some fairly general example in a ϕ^4 theory, fig.2.14. We have labelled also the vertices, $y_1, ..., y_4$ over which one integrates. Clearly we get the same contribution if these names are permuted. In general, at order N in λ there are N vertices, and $N!$ ways of labelling them, so that we understand how the factor $1/N!$ in our rule IV gets cancelled. Notice that in stating the rules I–IV we assumed that the position of dots and vertices from diagram to diagram was unchanged. Therefore, the diagram corresponding to interchanging say y_3 and y_2 in fig.2.14 would *not* be drawn like in fig.2.15, but rather like in fig.2.16 However, figs.2.15 and 2.16 are topologically equivalent.

Similarly, there are 4! ways of joining propagators to legs at a vertex, hence the factor 1/4! in the definition of λ gets cancelled. This completes the proof that $w_{\mathrm{tree}} = 1$.

Rather that giving a general rule for calculating the weight factor, w_D, we indicate how the calculation may be done in several simple examples. In more complicated cases, more powerful techniques are very helpful, for example, one may iterate the Schwinger-Dyson equation, to be introduced in a later part. In fig.2.17, we consider the vacuum blob to lowest order. Here $N = 1$, so we start by having (using λ rather than g)

$$\frac{1}{1!} \cdot (\frac{1}{4!})^1$$

To find the number of Wick contractions, consider the vertex with its 4 legs, fig.2.18. Leg 1 can be joined to another leg in 3 ways, after which the remaining two legs may be joined

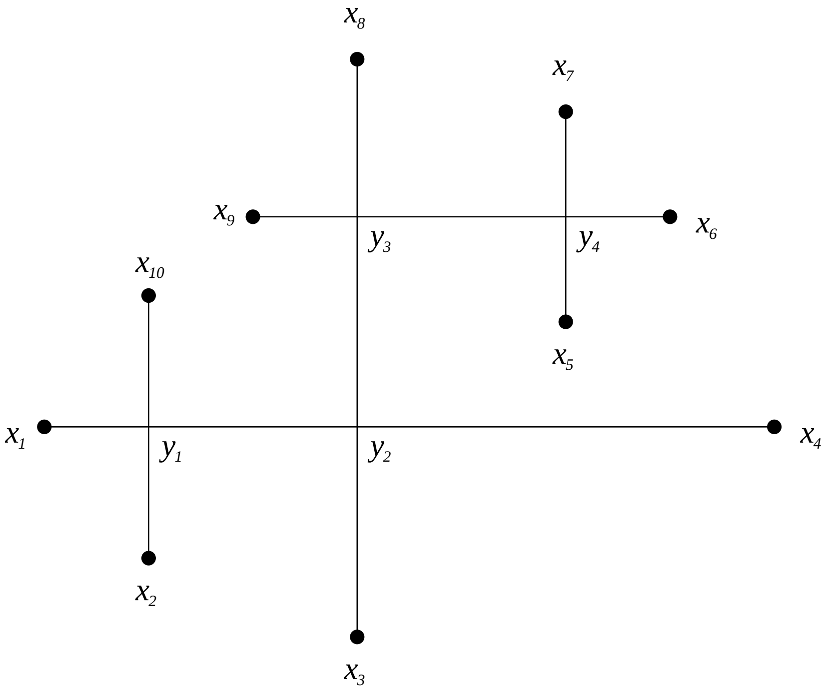


Figure 2.15: Interchanging y_3 and y_2 does not result in this figure!

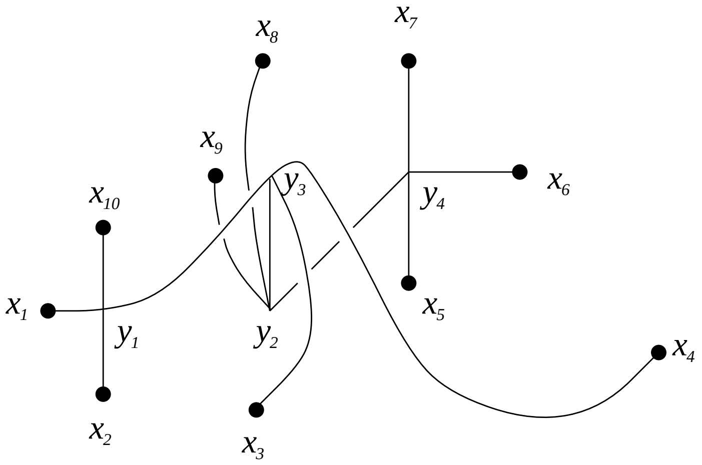


Figure 2.16: Interchanging y_3 and y_2 results in this figure!

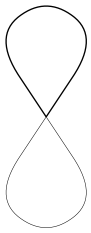


Figure 2.17: The lowest order vacuum diagram in a ϕ^4 theory.

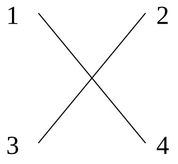


Figure 2.18: The ϕ^4 vertex with labelled legs.

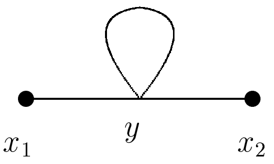


Figure 2.19: Lowest non-trivial order 2-point function.

in just 1 way. Hence

$$w_{D_1} = \frac{1}{1!} \cdot (\frac{1}{4!})^1 \cdot 3 = \frac{1}{8}$$

Next, we consider the 2-point function to order λ^1, fig.2.19. Again $N = 1$ and again we start with $\frac{1}{1!} \cdot (\frac{1}{4!})^1$. Starting the diagram as in fig.2.20, we see that x_1 may be connected to a y-leg in 4 ways. Then x_2 may be connected in 3 ways, and the remaining legs may be joined up in 1 way. Thus

$$w_{D_2} = \frac{1}{1!} \cdot (\frac{1}{4!})^1 \cdot 4 \cdot 3 = \frac{1}{2}$$

As a final example, consider the setting sun diagram fig.2.21. Here $N = 2$ and we start by $\frac{1}{2!}(\frac{1}{4!})^2$. We begin by looking at the non-contracted diagram (a) in fig.2.22. Now x_1 may be connected in 8 ways, giving, say (b) in fig.2.22. Then x_2 has to be connected to the *opposite* vertex, giving 4 possibilities like (c) in fig.2.22. Then leg a say, has to be connected to the other vertex in 1 of 3 ways, (d) in fig.2.22. Leg b say, has to be connected in 1 of 2 ways, (e) in fig.2.22. Finally the diagram may be finished in just 1 way, fig.2.23. This is topologically equivalent to fig.2.21, and we get

$$w_{\text{setting sun}} = \frac{1}{2!}(\frac{1}{4!})^2 \cdot 8 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = \frac{1}{6}$$

These examples should make the idea of the weight factor clear. In general the weight factor is different from 1 whenever the Feynman-diagram in question has some symmetries.

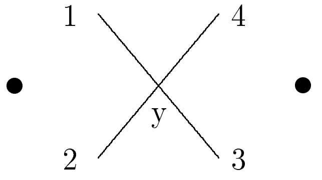


Figure 2.20: Start of the construction of the order 1, 2-point function.

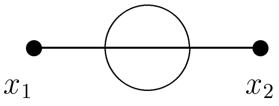


Figure 2.21: The setting sun diagram.

2.7 The Feynman Rules in Momentum Space

We have already introduced the Greens functions in (euclidean) momentum space:

$$\begin{aligned}\tilde{G}(p_1,...,p_n) &= \int \prod_i d^d x_i e^{i\sum_j x_j p_j} G(x_1,...,x_n) \\ &= (2\pi)^d \delta^d(p_1+...+p_n) G(p_1,...,p_n) \qquad (2.51)\end{aligned}$$

The advantage of working in momentum space is that all space integrations may then be carried out in favour of momentum integrations, and these last ones may be partly carried out because of translation invariance, implying momentum conservation at all vertices. The resulting expressions contain fewer remaining integrations, thus for tree diagrams there will be *no* remaining integrations.

Writing the propagator like

$$\Delta(x-y) = \int \frac{d^d p}{(2\pi)^d} \frac{e^{ip(x-y)}}{p^2+m^2} \qquad (2.52)$$

we see that in momentum space it will be natural to associate a momentum with every propagator. The direction of the momentum has to be chosen by convention. Letting it flow from x to y in the above expression, we have associated the exponent, e^{ipx} with the momentum *leaving* the vertex, x, and the exponential, e^{-ipy} with the momentum flowing *into* y. Following what happens at all vertices, it is obvious that the integration over the x-space position of a vertex, results in a momentum-space delta function expressing momentum conservation at that vertex. Afterwards the propagator-momenta may be trivially integrated over until no more delta functions are at hand. By then, we have a Feynman-diagram of the kind most often considered: each propagator has a momentum flowing in it, and momentum conservation has already been taken into account in order to minimize the number of free momentum integrations.

At the external lines, we see that the integration over x_i will constrain the propagator on the i'th external leg to have momentum, p_i, cf. the discussion in sect. 2.1.1.

It is of interest to count how many free momentum integrations remain. That number is equal to the number of independent *loops* in the Feynman diagram. Let us count it as follows:

Suppose we have a diagram with V vertices (i.e. V'th order in perturbation theory), I internal lines and E external lines (the number previously called n). Then to begin with we have I momentum variables, but they are constrained by V (d-dimensional) delta functions. However, these are not quite independent, since we have already assumed the external momenta are consistent with momentum conservation. Hence there are only $V-1$ (d-dimensional) constraints among the momenta, and we end up with

$$L = I - V + 1 \qquad (2.53)$$

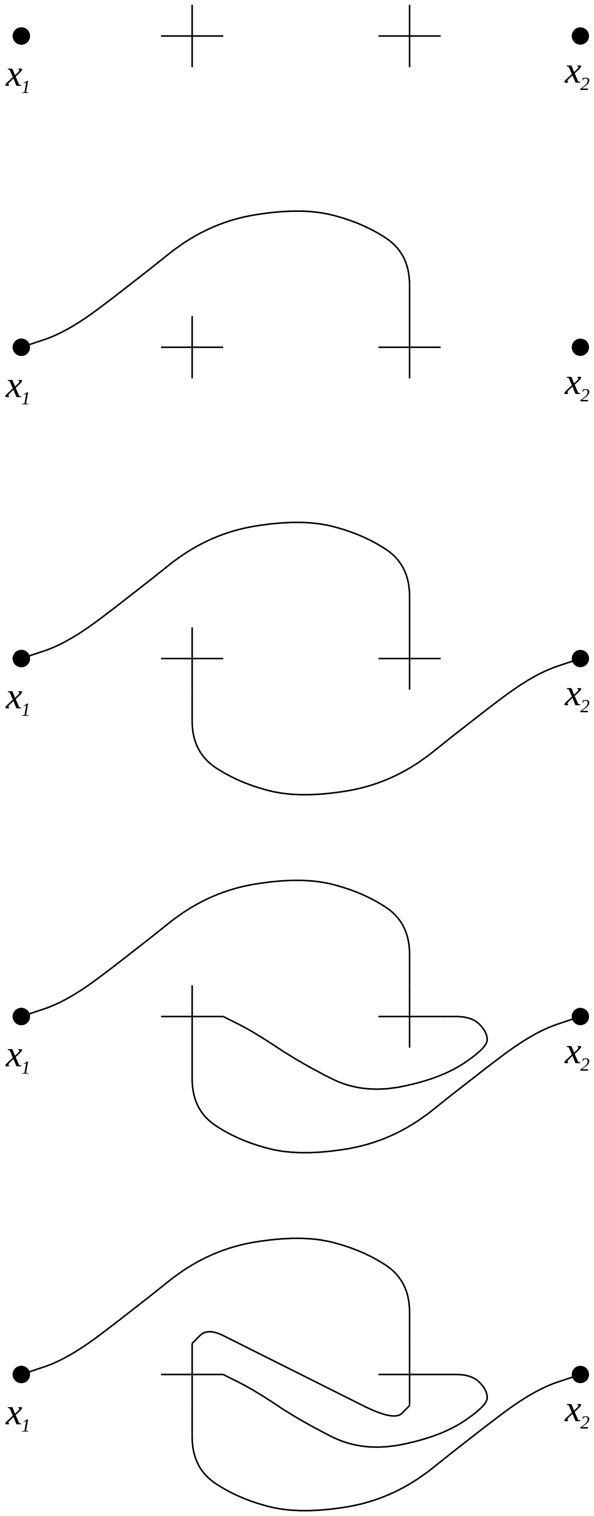


Figure 2.22: (*a*): The non-contracted setting sun diagram. (*b*): A once-contracted setting sun diagram. (*c*): A twice-contracted setting sun diagram. (*d*): A 3-contracted setting sun diagram. (*e*): A 4-contracted setting sun diagram.

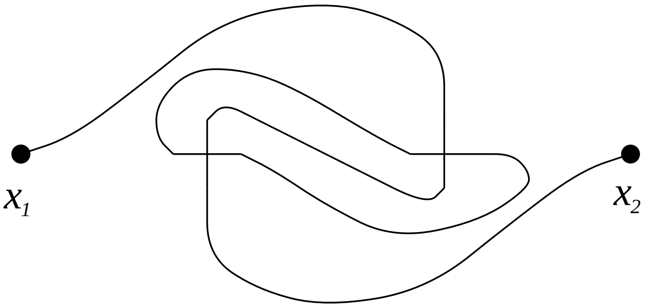


Figure 2.23: A version of the fully contracted setling sun diagram.

(d-dimensional) loop integrations. We may also introduce the order, p, of the vertices of the theory. If we imagine that we cut every internal propagator, we see that the total number of free legs become

$$E + 2I = pV \tag{2.54}$$

Hence

$$L = \frac{1}{2}(pV - E) - V + 1 = V(\frac{1}{2}p - 1) - \frac{1}{2}E \tag{2.55}$$

We may now summarize the **Feynman rules in momentum space**:

To construct the N'th order contribution to the momentum space connected Greens function, $G_C^{(N)}(p_1, ..., p_n)$ in perturbation theory:

1. Draw all connected, topologically inequivalent Feynman diagrams with n external lines and N vertices.

2. Label the external lines $p_1, ..., p_n$ with associated arrows indicating a flow direction of the momenta. Introduce a set of independent oriented loop-momenta, $\ell_1, ..., \ell_L$ and label all propagator-lines by their momenta q_j, calculated from the external momenta and the loop-momenta by momentum conservation at all vertices.

3. For each diagram, D, construct the contribution to the Greens function as

$$F_D(p_1, ..., p_n) = \int \frac{d^d\ell_1}{(2\pi)^d} \cdots \frac{d^d\ell_L}{(2\pi)^d} I_D(\{p_i\}, \{\ell_j\}) \tag{2.56}$$

4. Construct the integrand, I_D as a product of the following factors:

   - For each external line, p_i put

$$\frac{1}{p_i^2 + m^2}$$

  - for each internal line, q_j, put
$$\frac{1}{q_j^2+m^2}$$
where $q_j = q_j(\{p_i\},\{\ell_i\})$.
  - For each vertex, put a factor $-\lambda_p$.
  - calculate the weight factor w_D of the diagram and put that down as a factor.

5. Construct the total contribution to $G_C^{(N)}$ by summing over all Feynman diagrams.

2.7.1 The Most General Bosonic Field Theory

Let us briefly outline the strategy for how to deal with a general case with many bosonic fields, ϕ_r, where the index r can be a Lorentz-index, a flavour- or colour- group theory index, or some other index.

First we identify the "free theory" as the part bilinear in fields. After appropriate partial integrations the free action may be written as

$$S_0 = \int d^d x \sum_{r,s} \phi_r(x)\Delta^{-1}_{r,s}\phi_s(x) \tag{2.57}$$

where

$$\Delta^{-1}_{r,s}$$

is some differential operator. Inverting that operator gives the propagator. If that operator *cannot* be inverted, it is a sign that we are using a non-independent set of fields, and one has to think. In gauge theories in particular, a gauge-fixing has to be performed, possibly some "ghost-terms" have to be introduced, but *then* propagators may be found.

We may then straight forwardly derive a generalized version of Wick's theorem in both forms. In general there may be several different kinds of interactions coupling different fields together. A generic piece in the interaction part may be of the form

$$S_{r_1,\ldots,r_p} = \int d^d x \mathcal{A}_{r_1,\ldots,r_p}\phi_{r_1}(x)\cdots\phi_{r_p}(x) \tag{2.58}$$

(indices not summed over). The object $\mathcal{A}_{r_1,\ldots,r_p}$ may contain coupling constants, and may also contain *differential operators* acting on the fields. This interaction gives rise to a p-leg vertex coupling the fields, $\phi_{r_1},\ldots,\phi_{r_p}$ together. The corresponding vertex factor in the Feynman rules in momentum space with legs labelled by momenta $k_1,\ldots,k_p$, is obtained as follows:

$$\begin{aligned} &-\int d^d x_1\cdots d^d x_p e^{i(k_1x_1+\cdots+k_px_p)} \\ \times\ &\frac{\delta}{\delta\phi_{r_1}(x_1)}\cdots\frac{\delta}{\delta\phi_{r_p}(x_p)}S_{r_1,\ldots,r_p} \end{aligned} \tag{2.59}$$

This expression follows directly from the straight forward generalization of our treatment above, valid for a single field. Notice the minus sign. It comes because the path integral is weighted by e^{-S}. As a trivial example, consider the ϕ^4 theory

$$S_I = \frac{\lambda}{4!}\int d^d x\phi^4(x)$$

Using the rules we find the vertex

$$-\delta^d(k_1 + k_2 + k_3 + k_4) \cdot \lambda$$

The rule above will always produce the momentum conservation factor, but we know what the effect of that will be.

Chapter 3

Field Theory of Fermi Fields

3.1 Grassmann Numbers

3.1.1 Motivation

The aim of this chapter is to develop the path integral formalism of Fermi fields in close analogy with the case of bosonic fields. This turns out to be quite possible. However, there is a very important new point which has to be addressed: Because of Fermi statistics, it turns out that the "classical" Fermi fields over which we path integrate, cannot possibly take values in the set of complex numbers (or any modification of those, such as the set of complex d-vectors). Instead the "right numbers" for Fermi fields turn out to be so-called Grassmann numbers, the properties of which we shall first have to familiarize ourselves with. These are *anti-commuting* objects. We shall have to learn to develop concepts such as differentiation, and notably, integrations over Grassmann numbers.

To motivate the discussion, recall that our treatment of bosonic field theory was based on a close study of the ordinary harmonic oscillator, with creation and annihilation operators, $\hat{a}^\dagger$ and $\hat{a}$ satisfying the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = -[\hat{a}^\dagger, \hat{a}] = \hbar \tag{3.1}$$

where we have made Planck's constant explicit in order to study the classical limit, $\hbar \to 0$. In that limit, the operators will commute with one another, and hence we represent them by ordinary complex numbers. This in fact is what we do in the path integral. Another way of saying this is that in the derivation of the path integral we use complete sets of coherent states which are eigenstates of the operators with complex number valued eigenvalues, over which the path integral is performed.

Now that we want to deal with fermions, being excitations of Fermi fields, we have excitations satisfying Fermi statistics, including the Pauli exclusion principle. So our model is not the ordinary harmonic oscillator, but rather the "fermionic harmonic oscillator", defined by creation and annihilation operators, $\hat{b}^\dagger$ and $\hat{b}$ satisfying the *anti*-commutation relation

$$\{\hat{b}, \hat{b}^\dagger\} = \{\hat{b}^\dagger, \hat{b}\} = \hbar \tag{3.2}$$

But now we see the strange thing that in the classical limit $\hbar \to 0$, the operators *cannot* be represented by classical values b^* and b represented by complex numbers, since they have to *anti*-commute.

However, associated with b^* and b we may introduce a *Grassmann algebra.* It is a vector space over the field of complex numbers, and generated by b^*, b and 1. The usual axioms of associativity and distributivity etc. have to hold, but the algebra product satisfies the *anti*-commutativity property

$$\{b^*, b\} = \{b, b^*\} = \{b, b\} = \{b^*, b^*\} = 0 \tag{3.3}$$

In this simple example there are only 2 Grassmann generators. In the realistic field theory application, we shall have infinitely many, in general, where the b's are labelled by all the usual one-particle labels, like momentum and spin, and perhaps colour and flavour, and of course an index to indicate whether the excitation is a fermionic one or the corresponding anti-fermionic one. However, let us first study the case of just a finite number of distinct Grassmann generators.

We see that the algebra can be separated into the even and odd parts as follows: these may be expressed in terms of an even or odd product of the anti-commuting generators. Thus in our simple case, the odd part is the vector space spanned by b^* and b themselves, whereas the even part is spanned by $b^*b = -bb^*$ and 1. Notice that we cannot have an odd part with 3 Grassmann numbers, since in this very simple case there are only 2 different ones. Thus

$$bb^*b = -bbb^* = 0$$

since of course

$$bb = b^*b^* = 0$$

as a consequence of the anti-commutation relations. This may be regarded as the (formal) classical expression of Pauli-statistics.

The reader may wonder what the physical significance of the "classical" concept of Grassmann fields really is. For bosonic fields, such as the electromagnetic field, we know there are physical situations, where the true quantum field exhibits properties arbitrarily close to a classical field. But that depends on there being a large number of elementary excitations in the same quantum state. That precisely is impossible for fermions. Therefore the idea of "classical" Fermi fields is probably not backed by similarly significant physical situations. For us the device of Grassmann numbers may be considered a purely technical one which allows the path integral formalism to work for these as well. It is a fact though that this technical development has been extremely useful. It is crucial in discussions of supersymmetry and in the (also technical) discussion of gauge-fixing and ghosts.

3.1.2 Elementary Definitions

Consider a Grassmann algebra generated by the set of single generators, $\{x_i\}$ satisfying

$$\{x_i, x_j\} = 0$$

and 1, for which $[x_i, 1] = 0$. An arbitrary function of these, $F(x_i)$ may be expanded in a power series

$$F(x_i) = F^0 + \sum_i F_i^1 x_i + \sum_{i<j} F_{ij}^2 x_i x_j + \sum_{i<j<k} F_{ijk}^3 x_i x_j x_k + \ldots \tag{3.4}$$

Several comments are in order:

1. If the function values are complex numbers, or more generally, belong to the even part of the Grassmann algebra, then the coefficients above, F^n, are complex numbers for n even, and Grassmann numbers for n odd. Hence we shall usually allow ourselves such extra Grassmann coefficients to also be part of the algebra, even though they are not at the moment considered part of the *variables.*
2. The power series expansion is truncated if there is only a finite number, N of generators $\{x_i\}$. We cannot have terms of order higher than N, because in such a term at least one generator will occur twice making the term zero, because we can anticommute one copy of that generator through the others until it hits the first copy, giving zero.
3. Since $x_i x_j = -x_j x_i$ the coefficient, F^2_{ij} is taken antisymmetric in its indices, etc.

Differentiation is defined as follows:

$$\begin{aligned} \frac{\partial}{\partial x_i} x_j &= \delta_{ij} \\ \frac{\partial}{\partial x_i} x_j &= -x_j \frac{\partial}{\partial x_i} \quad \text{for} j \neq i \end{aligned} \tag{3.5}$$

Thus

$$\begin{aligned} \frac{\partial}{\partial x_i} x_1 x_2 x_3 ... x_n &= \delta_{i1} x_2 x_3 ... x_n - \delta_{i2} x_1 x_3 ... x_n + \delta_{i3} x_1 x_2 x_4 ... x_n \\ &+ \; ... + (-)^{n-1} \delta_{in} x_1 x_2 ... x_{n-1} \end{aligned} \tag{3.6}$$

The best way to think about differentiation, is to imagine that we anti-commute the variable to be differentiated after, all the way to the left, and then remove it. In the process a sign is picked up that we have to worry about.

If both f and g are bosonic functions of the Grassmann variables, meaning they belong to the even part of the algebra, we have

$$\frac{\partial}{\partial x_i}(f(x)g(x)) = [\frac{\partial}{\partial x_i} f(x)]g(x) + f(x)[\frac{\partial}{\partial x_i} g(x)] \tag{3.7}$$

If on the other hand f is fermionic meaning it belongs to the odd part of the algebra, we have

$$\frac{\partial}{\partial x_i}(f(x)g(x)) = [\frac{\partial}{\partial x_i} f(x)]g(x) - f(x)[\frac{\partial}{\partial x_i} g(x)] \tag{3.8}$$

As a relevant example, consider $\{x_i\}$ and $\{y_j\}$ being Grassmann generators. We can then form the exponential

$$e^{\sum_i x_i y_i} = 1 + \sum_i x_i y_i + \frac{1}{2} \sum_{ij} x_i y_i x_j y_j + ... \tag{3.9}$$

where the sum on the right hand side breaks off at the power corresponding to the number of different generators. We clearly get

$$\begin{aligned} \frac{\partial}{\partial x_k} e^{\sum_i x_i y_i} &= y_k e^{\sum_i x_i y_i} \\ \frac{\partial}{\partial y_k} e^{\sum_i x_i y_i} &= -x_k e^{\sum_i x_i y_i} \end{aligned} \tag{3.10}$$

We now come to the crucial and rather strange definition of **integration over Grassmann numbers**. As will be clear, the usual intuition that integration is like a *Riemann sum* is quite useless here. It is better to think of integration as a *linear operation.* Also it does not make sense to think of definite integrals, between finite limits. Everything is more analogous to

$$\int_{-\infty}^{\infty} dx$$

We first define integration for 1 variable. Since the most general function of one Grassmann variable, x, is a linear function, it is enough to provide the definitions

$$\begin{aligned} \int dx &= 0 \\ \int dx x &= 1 \end{aligned} \tag{3.11}$$

For functions of several variables, the rule follows from

$$\begin{aligned} \{dx_i, dx_j\} &= 0 \\ \{x_i, dx_j\} &= 0 \text{ for } i \neq j \end{aligned} \tag{3.12}$$

so that for example

$$\int dx_1 dx_2 x_1 x_2 = -\int dx_1 x_1 \int dx_2 x_2 = -1$$

whereas

$$\int dx_1 dx_2 x_2 x_1 = +\int dx_1 x_1 \int dx_2 x_2 = +1$$

One good thing about this definition, is that it gives a *translational* invariant definition:

$$\int dx f(x+a) = \int dx \{f^0 + f^1(x+a)\} = \int dx f^1 x = \int dx f(x) (= \pm f^1)$$

(depending on whether f^1 is bosonic or fermionic).

We have the following strange consequences of these definition:

1. **Delta function**

$$\delta(x) = x \tag{3.13}$$

Proof: Consider the bosonic function $f(x) = f^0 + f^1 x$ where f^1 is fermionic.

$$\begin{aligned} \int dx \delta(x-y) f(x) &= \int dx (x-y)(f^0 + f^1 x) = \int dx [x f^0 - y f^1 x] \\ &= f^0 + y \int dx f^1 x = f^0 - y f^1 = f^0 + f^1 y = f(y) \end{aligned} \tag{3.14}$$

qed. Thus $\int dx \delta(x) = 1 = -\int \delta(x) dx$.

2. **Differentiation is the *same* thing as integration**.

Consider the example:

$$\int dx_2 x_1 x_2 x_3 = -\int dx_2 x_2 x_1 x_3 = -x_1 x_3$$

and

$$\frac{\partial}{\partial x_2} x_1 x_2 x_3 = -\frac{\partial}{\partial x_2} x_2 x_1 x_3 = -x_1 x_3$$

This should convince the reader.

3. **Change of variable I.**

   For x, y Grassmann variables, consider the 1-dimensional change of variable

   $$y = ax, \ \ a \in \mathsf{C}$$

   Then

   $$1 = \int dxx = \int dyy = a \int dyx$$

   It follows that we must have

   $$y = ax \Rightarrow dy = a^{-1} dx \tag{3.15}$$

   This "upside-down" behaviour is the main point of Grassmann integration.

4. **Change of variable II.**

   Consider the case of several variables and the linear shift, $A_{ij} \in \mathsf{C}$:

   $$x_i = A_{ij} y_j$$

   (sum over j implied). Then using the definition of a determinant it is easy to see that

   $$\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} \sum_{j=1}^{n} A_{ij} y_j = det(A) \prod_{j=1}^{n} y_j$$

   Here we implied that $\prod_{i=1}^{n} x_i = x_1 x_2 ... x_n$. Let us also denote

   $$d^n x \equiv dx_n dx_{n-1} ... dx_2 dx_1$$

   Then we deduce

   $$1 = \int d^n x \prod_{i=1}^{n} x_i = \int d^n y \prod_{i=1}^{n} y_i = (det(A))^{-1} \int d^n y \prod_{i=1}^{n} x_i \tag{3.16}$$

   This implies

   $$x = Ay \Rightarrow d^n x = (det(A))^{-1} d^n y \tag{3.17}$$

   So the "Jacobian" is upside-down.

5. **Fourier expansion.**

   We have seen that the delta function is of the form

   $$\delta^n(x) = x_1 x_2 ... x_n$$

   Apart from a sign this is obtained by the Fourier formula

   $$\delta^n(x) = \int d^n p e^{\sum x_i p_i}$$

   In fact, expanding the exponential, we see that we only get a contribution from the term

   $$\frac{1}{n!} [\sum_i x_i p_i]^n$$

since we can only get a non-zero result from terms of the form $\pm x_1 x_2 ... x_n$. In fact, the above term gives us

$$\frac{1}{n!} n!(x_1 p_1)(x_2 p_2)...(x_n p_n) + \text{contributions giving zero in the integral}$$

since factors $(x_i p_i)$ and $(x_j p_j)$ commute because also the p's are fermionic. Hence, apart from a sign (which the reader can work out) we get

$$x_1 x_2 ... x_n p_1 p_2 ... p_n$$

From this follows, that we have Fourier's integral theorem:

$$\begin{aligned} F(p) &\equiv \int d^n x e^{x \cdot p} f(x) \Rightarrow \\ f(y) &= \epsilon \int d^n p e^{-y \cdot p} F(p) \end{aligned} \tag{3.18}$$

where again ϵ is a sign that we may or may not have to worry about.

3.1.3 Gaussian Integration I. Real Case

We have seen that the trick to develop perturbation theory is to understand gaussian integration. For a gaussian integral over ordinary (bosonic) variables we get some (irrelevant) π factors and an inverse determinant of the bilinear form in question. Better: for integration over *real* variables we get the inverse of the *square root* of the determinant, whereas for integration over complex variables we the get the inverse of the determinant itself. Even though in our applications the determinant did not play an important role, it actually does in many field theory applications. What we shall show now is that very similar results hold for "gaussian" Grassmann integration, however the determinants and their square roots appear "upside-down" compared to the bosonic case.

Theorem (real gaussian Grassmann integration)

Let $\{A_{ij}\}$ be a real, $(n \times n)$, *anti*-symmetric matrix such that the symmetric matrix A^2 has *negative*, non-zero eigenvalues only. Then n is necessarily even, $n = 2m$ and if $x_1, ..., x_n$ are real Grassmann variables

$$\int d^n x e^{x^T \mathsf{A} x} = 2^m \sqrt{det \mathsf{A}} \tag{3.19}$$

Proof:

First consider the simple case $n = 2, m = 1$. Then A is on the form

$$\mathsf{A} = \begin{pmatrix} 0 & \lambda \\ -\lambda & 0 \end{pmatrix}$$

and

$$det \mathsf{A} = \lambda^2 > 0$$

and

$$\mathsf{A}^2 = \begin{pmatrix} -\lambda^2 & 0 \\ 0 & -\lambda^2 \end{pmatrix}$$

has negative eigenvalues only. We may work out

$$x^T \mathsf{A} x = \lambda(x_1 x_2 - x_2 x_1) = 2\lambda x_1 x_2$$

$$\exp\{x^T \mathsf{A} x\} = \exp\{2\lambda x_1 x_2\} = 1 + 2\lambda x_1 x_2 + 0$$

since $(x_1 x_2)^2 = 0$. Then

$$\int d^2 x e^{x^T \mathsf{A} x} = \int dx_2 dx_1 [1 + 2\lambda x_1 x_2] = 2\lambda$$

So the result holds in this simple case. Notice that the *sign* of such expressions are very convention dependent. Similarly the sign in front of $x^T \mathsf{A} x$ is not very significant. In general we shall understand many of our formulas to be valid modulo signs only. In cases where this sign is important one should go over the conventions carefully.

We now introduce the

Lemma

There exists a regular, real orthogonal matrix B such that

$$\mathsf{B}^T \mathsf{A} \mathsf{B} = \begin{pmatrix} 0 & \lambda_1 & & & & & & \\ -\lambda_1 & 0 & & & & & & \\ & & 0 & \lambda_2 & & & & \\ & & -\lambda_2 & 0 & & & & \\ & & & & \cdot & \cdot & & \\ & & & & \cdot & \cdot & & \\ & & & & & & 0 & \lambda_m \\ & & & & & & -\lambda_m & 0 \end{pmatrix} \tag{3.20}$$

Here it is understood that elements not indicated are zero.

Proof of Lemma:

We want to establish the existence of a basis in n-dimensional space

$$\{\vec{e}_1, \vec{e}_{-1}, ..., \vec{e}_m, \vec{e}_{-m}\}$$

such that

$$\begin{aligned} \mathsf{A}\vec{e}_k &= \lambda_k \vec{e}_{-k} \\ \mathsf{A}\vec{e}_{-k} &= -\lambda_k \vec{e}_k, \quad k = 1, ..., m \end{aligned} \tag{3.21}$$

It follows that

$$\mathsf{A}^2 \vec{e}_k = \mathsf{A}(\lambda_k \vec{e}_{-k}) = -\lambda_k^2 \vec{e}_k$$

in other words, all the $\vec{e}_k$'s are eigenvectors of A^2. Also, notice that for any vector we have

$$v^T \mathsf{A}^2 v = v_i A_{ij} A_{jk} v_k = -(A_{ji} v_i)(A_{jk} v_k) < 0$$

when no eigenvalue is zero. So clearly, if $\vec{e}$ is an eigenvector of A^2 the eigenvalue may be denoted $-\lambda^2$. So we take a basis of eigenvectors of A. We now show that any eigenvalue must be (at least) two fold degenerate. In fact, let

$$\vec{e}\,' \equiv \frac{1}{\lambda} \mathsf{A} \vec{e}$$

Then

$$\begin{aligned} \mathsf{A}\vec{e} &= \lambda\vec{e}\,' \\ \mathsf{A}\vec{e}\,' &= \frac{1}{\lambda}\mathsf{A}^2\vec{e} = -\lambda\vec{e} \end{aligned} \tag{3.22}$$

This shows that both $\vec{e}$ and $\vec{e}\,'$ have the same eigen value of A^2. Also, they must be linearly independent, because otherwise $\vec{e}\,' = f\vec{e}$ and

$$\mathsf{A}\vec{e} = \lambda f\vec{e}$$

so that A would have the eigen vector $\vec{e}$. But that is impossible in a *real* vector space since then we could write

$$\mathsf{A}^2\vec{e} = -\lambda^2\vec{e} = f^2\lambda^2\vec{e}$$

making f imaginary.

So $\vec{e}$ and $\vec{e}\,'$ span a 2-dimensional subspace of eigenvectors of A^2 with the same eigen value. If we do the same thing for all eigenvalues of A we have found a basis in which A takes the form given in the Lemma. To finish we must show, that this basis is orthogonal. But that is true since

$$\vec{e}\cdot\vec{e}\,' = \frac{1}{\lambda}\vec{e}\mathsf{A}\vec{e} = 0$$

since A is antisymmetric. This finishes the proof of the Lemma.

To finish the proof of the Theorem, write

$$\begin{aligned} \int d^n x e^{x^T\mathsf{A}x} &= \int d^n x \exp\{(\mathsf{B}^T x)^T \mathsf{B}^T\mathsf{A}\mathsf{B}(\mathsf{B}^T x)\} \\ &= \int d^n x \exp\{x'^T \mathsf{B}^T\mathsf{A}\mathsf{B}x'\} \end{aligned} \tag{3.23}$$

where we put

$$x' = \mathsf{B}^T x$$

Now we have learned that we may do a linear change in the integration variable at the cost of a Jacobian that appears upside down. But in our case the Jacobian is 1, since the transformation is orthogonal. So

$$d^n x' = d^n x$$

This completes the proof of the Theorem.

It is instructive to see the consistency between the two upside down results: for the gaussian integration and for the Jacobian. In fact, consider a non-orthogonal variable change

$$x' = \mathsf{C}x$$

Then

$$\begin{aligned} 2^m\sqrt{det\mathsf{A}} &= \int d^n x \exp\{x^T\mathsf{A}x\} = \int d^n x' \exp\{x'^T\mathsf{A}x'\} \\ &= \int d^n x' \exp\{x^T[\mathsf{C}^T\mathsf{A}\mathsf{C}]x\} = \frac{1}{det\mathsf{C}}\int d^n x \exp\{x^T[\mathsf{C}^T\mathsf{A}\mathsf{C}]x\} \\ &= \frac{1}{det\mathsf{C}} 2^m\sqrt{det[\mathsf{C}^T\mathsf{A}\mathsf{C}]} = 2^m\sqrt{det\mathsf{A}} \end{aligned} \tag{3.24}$$

which is consistent.

3.1.4 Gaussian Integration II. Complex Case

This case is the one we shall mostly use. It is of interest to provide an independent proof of it.

Theorem (complex gaussian Grassmann integration)

Let x_i, y_i be two independent sets of Grassmann integration variables. Let A be anti-symmetric $n \times n$. Then

$$\int d^n x d^n y \exp\{y^T \mathsf{A} x\} = det\mathsf{A} \tag{3.25}$$

(Notice that we do not really define the concept of complex conjugation for Grassmann numbers here, we merely consider the y_i independent of the x_i without writing $y_i = x_i^*$. Sometimes ones does use that notation, however, see at the end of this subsection.)

Proof:

Consider directly the expansion of the integrand:

$$\exp\{y^T \mathsf{A} x\} = \sum_{N=0}^{\infty} \frac{1}{N!}(y_{i_1} A_{i_1 j_1} x_{j_1}) \cdot ... \cdot (y_{i_N} A_{i_N j_N} x_{j_N})$$

sums over repeated indices implied. The rules of Grassmann integration means that we shall get a non-vanishing contribution only when each x-variable and each y variable appear exactly once. Hence, only the term $N = n$ gives a contribution. Throwing away terms with more than one factor of each variable, we get

$$\frac{1}{n!}\sum_P \sum_Q (y_{P(1)} A_{P(1)Q(1)} x_{Q(1)}) \cdot ... \cdot (y_{P(n)} A_{P(n)Q(n)} x_{Q(n)})$$

where P and Q run over all permutations of $\{1, ..., n\}$. Each of the parentheses is bosonic: Grassmann-even, and may be permuted at will. For any permutation, P, we choose to re-order the parentheses so that we get

$$\frac{1}{n!}\sum_P \sum_Q (y_1 A_{1Q(P^{-1}(1))} x_{Q(P^{-1}(1))} \cdot ... \cdot (y_n A_{nQ(P^{-1}(n))} x_{Q(P^{-1}(n))})$$

Now, for a fixed permutation, P, $Q \circ P^{-1}$ will run once over all permutations when Q does. Hence, we obtain

$$\begin{aligned} & \frac{1}{n!}\sum_P \sum_{Q'} (y_1 A_{1Q'(1)} x_{Q'(1)} \cdot ... \cdot (y_n A_{nQ'(n)} x_{Q'(n)}) \\ = \; & \sum_Q (y_1 A_{1Q(1)} x_{Q(1)}) \cdot ... \cdot (y_n A_{nQ(n)} x_{Q(n)}) \\ = \; & \epsilon(y_1 y_2 ... y_n) \sum_Q A_{1Q(1)} x_{Q(1)} ... A_{nQ(n)} x_{Q(n)} \\ = \; & \epsilon(y_1 y_2 ... y_n)(x_1 x_2 ... x_n) \sum_Q \epsilon_Q A_{1Q(1)} ... A_{nQ(n)} \\ = \; & \epsilon(y_1 ... y_n)(x_1 ... x_n) det\mathsf{A} \end{aligned} \tag{3.26}$$

where ϵ is a Q-independent sign and ϵ_Q is the sign of the permutation, Q. This immediately proves the Theorem. (In fact our treatment here includes with more details the argument leading to the form of the Jacobian above).

It is instructive to see how the real and complex gaussian Grassmann integrations fit together. Hence let us write

$$x_i = a_i + ib_i; \quad y_i \equiv x_i^* = a_i - ib_i$$

Then

$$\begin{aligned} y^T \mathsf{A} x &= (a^T - ib^T)\mathsf{A}(a + ib) \\ &= a^T \mathsf{A} a + b^T \mathsf{A} b + i(a^T \mathsf{A} b - b^T \mathsf{A} a) \\ &= a^T \mathsf{A} a + b^T \mathsf{A} b \end{aligned} \tag{3.27}$$

since the last term vanishes by the anti symmetry of A:

$$a_i A_{ij} b_j - b_j A_{ji} a_i = a_i A_{ij} b_j + b_j A_{ij} a_i = 0$$

Now using the rule for changing variables we find

$$d^n x d^n y = J d^n a d^n b$$

It follows that from the rules of real gaussian Grassmann integration we should get the result

$$J 2^m (\sqrt{det \mathsf{A}})^2$$

for the integral. This is consistent since one may work out that

$$J = 2^{-m}$$

3.2 The Fermionic Oscillator

To set up the path integral formulation for Fermi fields, we follow as closely as possible our treatment for bose fields, using the formalism based on Grassmann numbers. Hence we start by considering the *fermionic oscillator*. Our treatment, however will be more brief than in the bosonic case, to which the reader is referred for details concerning the underlying philosophy.

Consider the fermionic oscillator with creation and annihilation operators $\hat{b}^\dagger$ and $\hat{b}$ satisfying ($\hbar = 1$)

$$\{\hat{b}, \hat{b}\} = 0 = \{\hat{b}^\dagger, \hat{b}^\dagger\}, \quad \{\hat{b}, \hat{b}^\dagger\} = 1$$

The Fock space is simply two-dimensional (rather than infinite dimensional as in the case of the ordinary oscillator), and is spanned by

$$|0\rangle, \quad \hat{b}^\dagger |0\rangle$$

($\hat{b}^\dagger \hat{b}^\dagger |0\rangle = 0$ by "Pauli statistics", $\{\hat{b}^\dagger, \hat{b}^\dagger\} = 0$).

Let b be a Grassmann number. Then

$$|b\rangle \equiv e^{\hat{b}^\dagger b} |0\rangle = (1 + \hat{b}^\dagger b)|0\rangle = (1 - b\hat{b}^\dagger)|0\rangle$$

is a coherent state with a grassmannian eigenvalue, b, for $\hat{b}$:

$$\hat{b}|b\rangle = (\hat{b} + b\hat{b}\hat{b}^\dagger)|0\rangle = b|0\rangle = b(1 - b\hat{b}^\dagger)|0\rangle = b|b\rangle$$

Similarly

$$\langle b^*| = \langle 0|e^{b^*\hat{b}}$$

is a coherent state with eigenvalue b^* of $\hat{b}^\dagger$. The normalization is

$$\langle b^*|b\rangle = e^{b^*b}$$

and the completeness relation is

$$I = \int db^* db |b\rangle\langle b^*| e^{-b^*b}$$

since sandwiching that between two arbitrary coherent states, gives back the normalization.

A very useful property of the coherent state technique for fermions is that these states are all "Grassmann-even": If we consider the vacuum state to have Fermion number 0 and the once excited state, $\hat{b}^\dagger|0\rangle$ to have Fermion number 1, then these have opposite Grassmann parity. However, in the coherent state the latter occurs with a Grassmann number coefficient. For this reason there is no question that a Grassmann number coefficient commutes with the coherent state. For states with odd Fermion numbers however, one has to consider those to anticommute with the Grassmann number coefficients, which is a nuisance.

The bosonic oscillator had a hamiltonian

$$\hat{H}_B = \frac{1}{2}\omega\{\hat{a}^\dagger, \hat{a}\} = \omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$$

Similarly, the fermionic oscillator is taken with the hamiltonian

$$\hat{H}_F = \frac{1}{2}\omega[\hat{b}^\dagger, \hat{b}] = \omega(\hat{b}^\dagger\hat{b} - \frac{1}{2})$$

The "zero-point" energies have opposite signs as is well known, but that in fact will not concern us here.

We modify the hamiltonian by adding grassmannian current terms, getting for the classical and quantum hamiltonian (and dropping the Fermi-index, F)

$$\begin{aligned} H(b^*, b) \to H(b^*, b; t) &= \omega b^* b - b^*\eta(t) - \overline{\eta}(t) b \\ \hat{H}(\hat{b}^\dagger, \hat{b}) \to \hat{H}(\hat{b}^\dagger, \hat{b}; t) &= \omega \hat{b}^\dagger\hat{b} - \hat{b}^\dagger\eta(t) - \overline{\eta}(t)\hat{b} \end{aligned} \tag{3.28}$$

Then

$$\langle b^*|\hat{H}(\hat{b}^\dagger, \hat{b}; t)|b\rangle = H(b^*, b; t)e^{b^*b}$$

We may then derive the path integral for the transition amplitude between coherent states in complete analogy with subsection 1.4.2, but paying particular attention to the order of Grassmann numbers and associated signs:

$$\begin{aligned} F(b^*, t'; b, t) &\equiv \langle b^*, t'|b, t\rangle \\ &= \int \prod_i db^*(t_i) db(t_i) e^{-b^*(t_i)b(t_i)} \\ &\cdot \ \langle b^*(t')|e^{-i\epsilon\hat{H}}|b(t_n)\rangle\langle b^*(t_n)|e^{-i\epsilon\hat{H}}|b(t_{n-1})\rangle\langle\cdots \end{aligned}$$

$$
\begin{aligned}
&\cdots\rangle\langle b^*(t_1)|e^{-i\epsilon\hat{H}}|b(t)\rangle \\
= &\int\prod_i db^*(t_i)db(t_i) \\
\cdot\ &\exp\{b^*(t')b(t_n)-b^*(t_n)b(t_n)+b^*(t_n)b(t_{n-1})-b^*(t_{n-1})b(t_{n-1})+... \\
&...-b^*(t_1)b(t_1)+b^*(t_1)b(t) \\
- &i\int_t^{t'}d\tau H(b^*(\tau),b(\tau);\tau)\} \\
= &\int\prod_i db^*(t_i)db(t_i) \\
&\exp\{\int_t^{t'}d\tau[\dot{b}^*(\tau)b(\tau)-iH(b^*(\tau),b(\tau);\tau)]+b^*(t)b(t)\} \\
= &\int\mathcal{D}b^*\mathcal{D}b\exp\{i\int_t^{t'}d\tau[-i\dot{b}^*(\tau)b(\tau)-H]+b^*(t)b(t)\} \qquad (3.29)
\end{aligned}
$$

Stationarity leads to the equations of motion:

$$
\begin{aligned}
\dot{b}^*-i\omega b^*+i\overline{\eta} &= 0 \\
\dot{b}+i\omega b-i\eta &= 0 \qquad (3.30)
\end{aligned}
$$

with the solutions

$$
\begin{aligned}
b(\tau) &= be^{i\omega(t-\tau)}+i\int_t^{\tau}e^{i\omega(s-\tau)}\eta(s)ds \\
b^*(\tau) &= b^*e^{i\omega(\tau-t')}+i\int_{\tau}^{t'}e^{i\omega(\tau-s)}\overline{\eta}(s)ds \qquad (3.31)
\end{aligned}
$$

The path integral evaluates as usual to the exponential of the critical value times an irrelevant determinant. As for the bosonic oscillator, we get for the action part of the exponential, evaluated on the classical path

$$
\begin{aligned}
\int_t^{t'}d\tau[\dot{b}^*(\tau)b(\tau)-iH(b^*(\tau),b(\tau);\tau)] &= \int_t^{t'}d\tau[\dot{b}^*b-i(\omega b^*b-b^*\eta-\overline{\eta}b)] \\
=\int_t^{t'}d\tau[(\dot{b}^*-i\omega b^*+i\overline{\eta})b+ib^*\eta] &= i\int_t^{t'}d\tau b^*(\tau)\eta(\tau) \qquad (3.32)
\end{aligned}
$$

This yields for the transition amplitude

$$
\begin{aligned}
F(b^*,t';b,t) &= \exp\{b^*be^{i\omega(t'-t)}+i\int_t^{t'}ds[\overline{\eta}(s)be^{i\omega(t-s)}+b^*\eta(s)e^{i\omega(s-t')}] \\
&- \int_t^{t'}d\tau\int_{\tau}^{t'}dse^{i\omega(\tau-s)}\overline{\eta}(s)\eta(\tau)\} \qquad (3.33)
\end{aligned}
$$

and the vacuum expectation value (the case $b=b^*=0$) for $t\to-\infty$, $t'\to+\infty$, in other words, the partition function, becomes

$$
Z[\overline{\eta},\eta]=\langle 0|0\rangle\exp\{-\int_{-\infty}^{+\infty}d\tau\int_{\tau}^{+\infty}dse^{i\omega(\tau-s)}\overline{\eta}(s)\eta(\tau)\} \qquad (3.34)
$$

So far things are really very similar to the bosonic case. Now comes a difference. In the bosonic case we could form a position, q, and a momentum, p, from the creation and annihilation operators. And we could build a lagrangian with p removed in favour of a $\dot{q}$ (historically one went the other way round). this $\dot{q}$ then appears *quadratically* in the lagrangian (the velocity squared). but such a possibility does not exist for the fermionic oscillator. Indeed the square of any Grassmann variable would be zero. In fact, we cannot really leave the present path integral in *phase space* in the fermionic case: we cannot go to configuration space. Likewise, we cannot attempt to put $\overline{\eta} = \eta$ like we did in the bosonic case (the common value of $\overline{\gamma}$ and γ got related to a real current J). We clearly see from the result that this would give zero in the exponential.

However, let us rewrite our path integral in phase space in a way that makes the transition to field theory more suggestive. In fact let us write

$$\begin{aligned} b &\rightarrow \psi \\ b^* &\rightarrow \overline{\psi} \end{aligned} \tag{3.35}$$

Also, in the action part we may write

$$\int_{-\infty}^{+\infty} dt \partial_t \overline{\psi}(t) \psi(t) = - \int_{-\infty}^{+\infty} dt \overline{\psi}(t) \partial_t \psi(t)$$

whenever we have vanishing boundary conditions corresponding to the vacuum expectation value. With this notation we get for the partition function

$$Z[\overline{\eta}, \eta] = \int \mathcal{D}\overline{\psi}\mathcal{D}\psi \exp\{i \int dt [i\overline{\psi}\partial_t \psi - \omega \overline{\psi}\psi + \overline{\psi}\eta + \overline{\eta}\psi]\}$$

Let us evaluate this by completing the square as usual (using "functional notation")

$$\begin{aligned} Z[\overline{\eta}, \eta] &= \int \mathcal{D}\overline{\psi}\mathcal{D}\psi \exp\{i \int dt [\overline{\psi}(i\partial_t - \omega)\psi + \overline{\psi}\eta + \overline{\eta}\psi]\} \\ &= \int \mathcal{D}\overline{\psi}\mathcal{D}\psi \exp\{i(\overline{\psi} + \overline{\eta}(i\partial_t - \omega)^{-1}) \cdot (i\partial_t - \omega) \cdot (\psi + (i\partial_t - \omega)^{-1}\eta)\} \\ &\cdot \ \exp\{-i\overline{\eta} \cdot (i\partial_t - \omega)^{-1} \cdot \eta\} \\ &= Z[0,0] \exp\{-i \int d\tau ds \overline{\eta}(s) D_F(s,\tau)\eta(\tau)\} \end{aligned} \tag{3.36}$$

Using eq.(3.34), we get the fermion propagator in the form

$$\begin{aligned} D_F(s,\tau) &\equiv (i\partial_t - \omega)^{-1}(s,\tau) \\ &= -i\theta(s-\tau)e^{-i\omega(s-\tau)} \\ &= \int \frac{dE}{2\pi} \frac{e^{-iE(s-\tau)}}{E - \omega + i\epsilon} \end{aligned} \tag{3.37}$$

We may check that

$$(i\partial_s - \omega)D_F(s,\tau) = \int \frac{dE}{2\pi}(E - \omega)\frac{e^{-iE(s-\tau)}}{E - \omega + i\epsilon} = \delta(s - \tau)$$

We now want to "go to euclidean time" following the idea of sect. 1.5. Putting $t \rightarrow -it_E$, we find

$$\begin{aligned} Z_E[\overline{\eta}, \eta] &= \int \mathcal{D}\overline{\psi}\mathcal{D}\psi \exp\{i \int d(-it_E)[\overline{\psi}(i\frac{\partial}{\partial(-it_E)} - \omega)\psi + \overline{\psi}\eta + \overline{\eta}\psi]\} \\ &= \int \mathcal{D}\overline{\psi}\mathcal{D}\psi \exp\{-(\overline{\psi} - \overline{\eta}(\partial_{t_E} + \omega)^{-1}) \cdot (\partial_{t_E} + \omega) \cdot (\psi - (\partial_{t_E} + \omega)^{-1}\eta)\} \\ &\cdot \ \exp\{\overline{\eta} \cdot (\partial_{t_E} + \omega)^{-1} \cdot \eta\} \\ &= Z[0,0]_E \exp\{\int d\tau ds \overline{\eta}(s) D(s,\tau)\eta(\tau)\} \end{aligned} \tag{3.38}$$

where the euclidean propagator is

$$D(s,\tau) \equiv (\partial_{t_E} + \omega)^{-1} = i \int \frac{dE}{2\pi} \frac{e^{-iE(s-\tau)}}{E + i\omega} \tag{3.39}$$

Indeed

$$(\partial_s + \omega) D(s,\tau) = i \int \frac{dE}{2\pi} (-iE + \omega) \frac{e^{-E(s-\tau)}}{E + i\omega} = \delta(s - \tau)$$

We may pass back to the Minkowski propagator as usual by analytically continuing $s \to is_M$ and $\tau \to i\tau_M$ (M for Minkowski), and at the same time rotate the integration contour by "not quite $-\frac{\pi}{2}$" so as to avoid crossing the pole at $E_E = -i\omega$. Thus we write E_E in the integral representation of the euclidean propagator and we rotate the integration contour as $E_E = (-i + \epsilon)E$. This gives

$$\begin{aligned} D(is_M, i\tau_M) &= i \int \frac{d(-iE)}{2\pi} \frac{e^{-i(-iE)(+i)(s_M - \tau_M)}}{(-i+\epsilon)E + i\omega} \\ &= i \int \frac{dE}{2\pi} \frac{e^{-iE(s_M - \tau_M)}}{E - \omega + i\epsilon E} \\ &= i \int \frac{dE}{2\pi} \frac{e^{-iE(s_M - \tau_M)}}{E - \omega + i\epsilon'} \\ &= iD_F(s_M, \tau_M) \end{aligned} \tag{3.40}$$

Here we used that the pole occurs for $E \approx \omega > 0$.

3.3 Fermi Fields

3.3.1 Free Fermi Fields

In sec. 1.5 we gave our notation for Fermi fields in Minkowski space. The free lagrangian density is

$$\mathcal{L}_F^M = -\overline{\psi}(\partial\!\!\!/ + m)\psi \tag{3.41}$$

Now the field operator contains an oscillator contribution from each wave vector. Not just the annihilation part, but also (for Dirac fermions) a creation part for anti-fermions:

$$\hat{\psi}(x) = \sum_{\vec{p},s} [\hat{b}_s(\vec{p}) u_s(\vec{p}) e^{ipx} + \hat{d}_s^\dagger(\vec{p}) v_s(\vec{p}) e^{-ipx}]$$

with $p_0 \equiv \sqrt{\vec{p}^2 + m^2}$. The spin index, s runs over the values $\pm\frac{1}{2}$, and the spinors $u_s(\vec{p})$ and $v_s(\vec{p})$ are fermion and anti-fermion solutions to the Dirac-equation. They have 4 Dirac indices which we do not write explicitly.

In the euclidean case we simply have (cf. the oscillator discussion),

$$\mathcal{L}_F^e = \overline{\psi}(\partial\!\!\!/ + m)\psi \tag{3.42}$$

the only difference being that (i) $\overline{\psi}$ and ψ have to be regarded as independent degrees of freedom (rather than being related by hermitian conjugation and γ_0), and (ii) the gamma-matrices are all hermitian:

$$\gamma_\mu = \gamma^\mu = \gamma_\mu^\dagger$$

and the anti-commutation relations are

$$\{\gamma_\mu, \gamma_\nu\} = 2I\delta_{\mu\nu} \tag{3.43}$$

I being the 4×4 unit matrix (in the 4-dimensional case). We then get for the euclidean path integral for the free fermion problem

$$\begin{aligned}
Z_F^0[\overline{\eta}, \eta] &= \int \mathcal{D}\overline{\psi}\mathcal{D}\psi \exp\{-\int d^4x(\overline{\psi}(\not\partial + m)\psi - \overline{\eta}\psi - \overline{\psi}\eta)\} \\
&= \int \mathcal{D}\overline{\psi}\mathcal{D}\psi \exp\{-(\overline{\psi} - \overline{\eta}(\not\partial + m)^{-1})(\not\partial + m)(\psi - (\not\partial + m)^{-1}\eta)\} \\
&\times \exp\{\overline{\eta}(\not\partial + m)^{-1}\eta\} \\
&= Z[0,0]\exp\{\overline{\eta}(\not\partial + m)^{-1}\eta\}
\end{aligned} \tag{3.44}$$

with the usual functional notation. In complete analogy with the oscillator case we get for the euclidean propagator

$$S_F(x,y) = i\int \frac{d^4p}{(2\pi)^4}\frac{e^{-ip(x-y)}}{\not p + im} \tag{3.45}$$

Notice that $1/(\not p + im)$ means the inverse of the 4×4 matrix, $(p_\mu\gamma_\mu + imI)$. On checks that

$$(\not\partial + m)_x i\int \frac{d^4p}{(2\pi)^4}\frac{e^{-ip(x-y)}}{\not p + im} = i\int \frac{d^4p}{(2\pi)^4}(-i\not p + m)\frac{e^{-ip(x-y)}}{\not p + im} = \delta^4(x-y) \tag{3.46}$$

The continuation to the Minkowski space may be performed just as for the oscillator case. The fermionic propagator is related to the bosonic one. In fact notice that

$$\begin{aligned}
(\not\partial + m)(\not\partial - m) &= \partial^2 - m^2 = -\Delta^{-1} \\
(\not p + im)(\not p - im) &= p^2 + m^2
\end{aligned} \tag{3.47}$$

Hence we would expect that

$$(\not\partial - m)^{-1}(\not\partial + m)^{-1} = -\Delta$$

or

$$(\not\partial - m)_x\Delta(x,y) = -S_F(x,y) \tag{3.48}$$

One easily verifies from the Fourier decomposition formulas that this is indeed correct.

Just as in the bosonic case, one can easily establish the relation between vacuum expectation values of operators, on the one hand, and path integral expectation values on the other. In the Minkowski case the relation is

$$\begin{aligned}
&\langle 0|T\{\hat\psi_{a_1}(x_1)\ldots\hat\psi_{a_N}(x_N)\hat{\overline{\psi}}_{b_1}(y_1)\ldots\hat{\overline{\psi}}_{b_M}(y_M)\}|0\rangle \\
= &\int \mathcal{D}\overline{\psi}\mathcal{D}\psi e^{iS_F^M}\{\psi_{a_1}(x_1)\ldots\psi_{a_N}(x_N)\overline{\psi}_{b_1}(y_1)\ldots\overline{\psi}_{b_M}(y_M)\}
\end{aligned} \tag{3.49}$$

Here we have written a time-ordering operator in the vacuum expectation value. It is analogous to the one in the bosonic case, except that the operators *anti*-commute under the time ordering sign, rather than commute as in the bosonic case. This of course agrees with the path integral expression. For completeness we have also indicated the Dirac-indices here in this correlator of N fermion operators and M anti-fermion operators , as it is loosely refereed to. (Of course, for vanishing external currents, this will vanish unless $N = M$.) The proof is a trivial generalization of the bosonic proof. The fact that Fermi-operators anti-commute, causes a certain amount of head-aches connected with rather trivial but tedious book keeping of signs. A useful technique for handling the situation, consists in multiplying each fermion operator with a constant Grassmann variable, indeed with a Grassmann spinor:

$$\psi_a(x) \to \overline{\zeta}\psi(x)$$

This combination behaves just like a commuting bosonic field. At the end of the calculation, one may then remove all the unwanted constant Grassmann numbers, and only think about the sign problems at that point.

Thus in the euclidean case we may write:

$$\begin{aligned} & \int \mathcal{D}\overline{\psi}\mathcal{D}\psi e^{-S^e_F}\{\overline{\zeta}_1\psi(x_1)\ldots\overline{\zeta}_N\psi(x_N)\overline{\psi}(y_1)\zeta_1\ldots\overline{\psi}(y_M)\zeta_M\} \\ = & \ \overline{\zeta}_1\frac{\delta}{\delta\overline{\eta}(x_1)}\ldots\overline{\zeta}_N\frac{\delta}{\delta\overline{\eta}(x_N)}\frac{\delta}{\delta\eta(y_1)}\zeta_1\ldots\frac{\delta}{\delta\eta(y_M)}\zeta_M Z[\overline{\eta},\eta]|_{\eta\equiv\overline{\eta}\equiv 0} \end{aligned} \tag{3.50}$$

Here for example

$$\overline{\zeta}_1\frac{\delta}{\delta\overline{\eta}(x_1)} \equiv \sum_{a=1}^{4}(\overline{\zeta}_1)_a\frac{\delta}{\delta(\overline{\eta})_a(x_1)}$$

3.3.2 Wick's Theorem for Fermi Fields

As an example of an interacting field theory involving Fermi fields, consider the euclidean QED case:

$$\begin{aligned} Z_{QED}[\overline{\eta},\eta,J] \equiv & \int \mathcal{D}\overline{\psi}\mathcal{D}\psi\mathcal{D}A_\mu \\ & \exp\{-\int d^4x[\overline{\psi}(\not{D}+m)\psi-\overline{\eta}\psi-\overline{\psi}\eta-J_\mu A_\mu-\frac{1}{4}F_{\mu\nu}F_{\mu\nu}]\} \\ = & \int \mathcal{D}\overline{\psi}\mathcal{D}\psi\mathcal{D}A_\mu \\ & \exp\{-\int d^4x[\overline{\psi}(\not{\partial}+m)\psi-\overline{\eta}\psi-\overline{\psi}\eta]\} \\ \cdot & \exp\{ie\int d^4x\overline{\psi}\not{A}\psi\}\exp\{\int d^4xJ\cdot A+\frac{1}{4}\int d^4F_{\mu\nu}F_{\mu\nu}\} \end{aligned} \tag{3.51}$$

Define

$$Z^0_A[J] \equiv \int \mathcal{D}A_\mu \exp\{\frac{1}{4}\int d^4xF_{\mu\nu}F_{\mu\nu}+\int d^4xJ\cdot A\} \tag{3.52}$$

Then we may write

$$\begin{aligned} Z_{QED}[\overline{\eta},\eta,J] = & \exp\{ie\int d^4x(-\frac{\delta}{\delta\eta(x)})\gamma_\mu\frac{\delta}{\delta J_\mu(x)}\frac{\delta}{\delta\overline{\eta}(x)}\} \\ & Z^0_F[\overline{\eta},\eta]Z^0_A[J] \end{aligned} \tag{3.53}$$

where Z^0_F was worked out in eq.(3.44).

Further, for a quite general interaction involving the Fermi field and the A_μ field (for purpose of illustration: not even necessarily gauge invariant),

$$S_I(\overline{\psi}, \psi, A_\mu)$$

we would get for the corresponding partition function

$$Z_I[\overline{\eta}, \eta, J] = \exp\{-S_I(-\frac{\delta}{\delta\eta}, \frac{\delta}{\delta\overline{\eta}}, \frac{\delta}{\delta J})\} Z_F^0[\overline{\eta}, \eta] Z_A^0[J] \tag{3.54}$$

And that is the (first version of) Wick's theorem, cf. sect. 2.3.

We now go to the second version of Wick's theorem, in analogy with the bosonic discussion in sect. 2.4. We first write down the:

Fermionic version of Coleman's Lemma

$$F(-\frac{\delta}{\delta\eta}, \frac{\delta}{\delta\overline{\eta}}) Z[\overline{\eta}, \eta] = Z[-\frac{\delta}{\delta\psi}, \frac{\delta}{\delta\overline{\psi}}](F(\overline{\psi}, \psi) e^{\overline{\psi}\eta + \overline{\eta}\psi})|_{\overline{\psi}\equiv\psi\equiv 0} \tag{3.55}$$

The proof is entirely analogous to the bosonic case and is left as an exercise for the reader.

If now we suppress the dependence on any other field than the fermionic field which we currently consider, we obtain the following:

Second version of Wick's theorem

$$\exp\{-S_I(-\frac{\delta}{\delta\eta}, \frac{\delta}{\delta\overline{\eta}})\} e^{\overline{\eta}\cdot S_F \eta} = \exp\{-\frac{\delta}{\delta\psi} S_F \frac{\delta}{\delta\overline{\psi}}\} \exp\{-S_I(\overline{\psi}, \psi) + \overline{\psi}\eta + \overline{\eta}\psi\}|_{\overline{\psi}\equiv\psi\equiv 0} \tag{3.56}$$

In the case of QED we are not yet ready to treat the photon field the same way. The reason is that because of gauge invariance, the photon propagator is an object that does not make sense until a gauge has been fixed. But the whole problem of gauge fixing will only be dealt with later on. However, once that has been achieved, basically there is a completely similar result. There is a photon propagator and a "free" photon partition function. (It depends on gauge-fixing terms in addition to the ones we have written down. Also, in the case of a non-Abelian gauge theory, we have a similar situation, but then the gauge fixing further implies the introduction of certain "ghost-fields").

3.3.3 Feynman Rules for Fermi Fields

Again everything is extremely similar to the bosonic case of sect. 2.4. Hence we mostly give an example and point out one or two important places where the characteristic sign-consideration of fermion problems come into play.

Consider the very simple Yukawa interaction

$$\mathcal{L}_Y = g\overline{\psi}\psi\phi \tag{3.57}$$

Then the generating functional for the full Greens functions may be written

$$Z[\overline{\eta}, \eta, J] = e^{\frac{1}{2}\frac{\delta}{\delta\phi}\Delta\frac{\delta}{\delta\phi}} e^{-\frac{\delta}{\delta\psi} S_F \frac{\delta}{\delta\overline{\psi}}} \exp\{-g\int d^4x \overline{\psi}\psi\phi + \int d^4x[\overline{\psi}\eta + \overline{\eta}\psi + J\phi]\}|_{\psi=\overline{\psi}=\phi=0} \tag{3.58}$$

Consider first the Fermion two-point function to lowest order:

$$-\frac{\delta}{\delta\psi} S_F \frac{\delta}{\delta\overline{\psi}} \{\psi(x)\overline{\psi}(y)\}|_{\psi=\overline{\psi}=\phi=0}$$

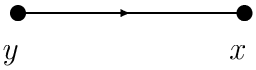


Figure 3.1: The fermionic two point function. The arrow is taken to go from the argument of a $\overline{\psi}$ to the argument of a ψ, in other words, we think of the fermionic (as opposed to the anti-fermionic) part in our convention.

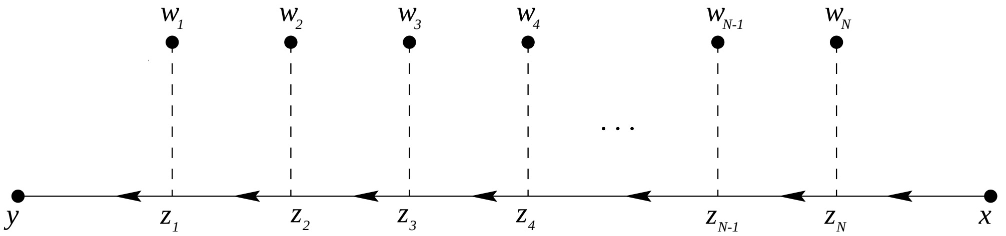


Figure 3.2: The tree diagram for a fermion line emitting N bosons

This is nothing but the fermionic propagator, in analogy with the bosonic result:

$$\langle 0|\hat{\psi}(x)\hat{\overline{\psi}}(y)|0\rangle = S_F(x,y)$$

Notice this is a matrix equation. Next consider the connected tree-diagram for a fermionic two-point function and a bosonic N-point function fig.3.2. We are led to work out

$$\begin{aligned}
& \frac{1}{N!}\prod_{i=1}^{N}\frac{1}{2}\int d^4u_i d^4v_i\frac{\delta}{\delta\phi(u_i)}\Delta(u_i,v_i)\frac{\delta}{\delta\phi(v_i)} \\
\cdot \quad & \frac{1}{(N+1)!}\prod_{j=1}^{N+1}\int d^4x_j d^4y_j(-\frac{\delta}{\delta\psi(x_j)}S_F(x_j,y_j)\frac{\delta}{\delta\overline{\psi}(y_j)}) \\
\{\tfrac{1}{N!} \quad & \prod_{l=1}^{N}\int d^4z_l(-g\overline{\psi}(z_l)\psi(z_l)\phi(z_l))\prod_{k=1}^{N}\phi(w_k)\psi(x)\overline{\psi}(y)\} \\
= \quad & \frac{1}{(N+1)!}\prod_{j=1}^{N+1}\int d^4x_j d^4y_j(-\frac{\delta}{\delta\psi(x_j)}S_F(x_j,y_j)\frac{\delta}{\delta\overline{\psi}(y_j)}) \\
\{ \quad & \psi(x)\frac{1}{N!}\prod_{l=1}^{N}\int d^4z_l(\overline{\psi}(z_l)\psi(z_l))\prod_{k=1}^{N}(-g\Delta(z_l,w_k))\overline{\psi}(y)\} \\
= \quad & \int\prod_{l=1}^{N}d^4z_l S_F(x,z_N)S_F(z_N,z_{N-1})\ldots S_F(z_1,y) \\
\cdot \quad & (-g)\Delta(w_1,z_1)(-g)\Delta(w_2,z_2)\ldots(-g)\Delta(w_N,z_N)
\end{aligned} \qquad (3.59)$$

Here we made use of the fact that

$$-\frac{\delta}{\delta\psi}S_F\frac{\delta}{\delta\overline{\psi}}$$

gives $+S_F$ when it hits a $\psi\overline{\psi}$ with the ψ to the *left* of the $\overline{\psi}$. Also we could move the $\psi(x)$ all the way to the left without picking up any signs since it was moved passed an even number of Grassmann fields. The reader may verify that the statistical weight factor

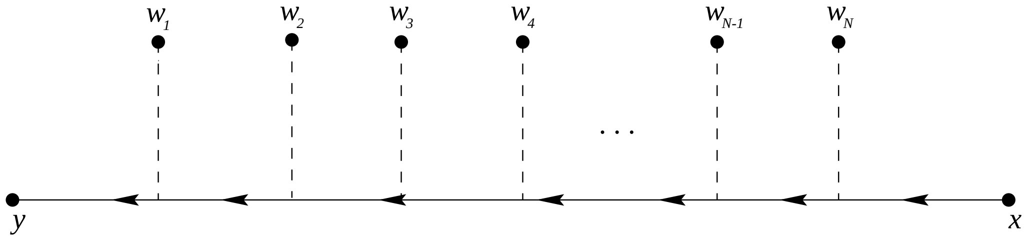


Figure 3.3: The reversed tree diagram for a fermion line emitting N bosons, i.e. the correlator involves $\overline{\psi}(x)\psi(y) = -\psi(y)\overline{\psi}(x)$.

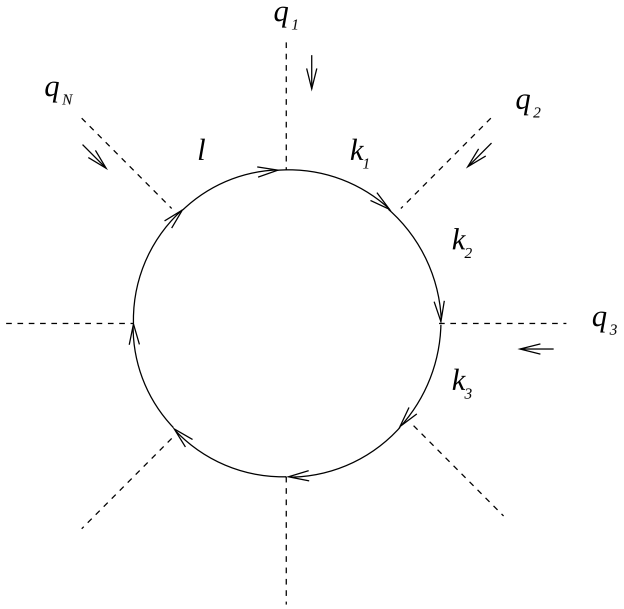


Figure 3.4: A fermion loop emitting N bosons

comes out to be 1 just as for purely bosonic tree-diagrams. We observe the Feynman rules, that the fermionic propagator is what we expect. In matrix notation the fermion propagators are written down from left to right when we make our way along the fermion line *against* the direction of the arrow. The vertex in this case is merely

$$-g$$

the minus sign coming from the fact that in the euclidean case we consider, we are dealing with the exponential of *minus* the (inter-)action. It is obvious, that if instead we considered the diagram fig.3.3 where it is understood that the fermionic operators are taken in the opposite order, then the result is the same *except for a sign change.* In Minkowski space, we would think of the time-ordered product of operators, and in fig.3.2 we would have a situation where the operator $\overline{\psi}(y)$ sometime in the past created a *fermion* that gets destroyed later at x by $\psi(x)$. Hence we think of a fermion line passing from the initial to the final state. In the reversed figure, correspondingly we would have an *anti-fermion* passing form the initial to the final state. We see that this costs a change of sign.

Finally we consider the *loop*-diagram fig.3.4 where the calculation is entirely analogous except we do not have the factors $\psi(x)\overline{\psi}(y)$. We see that in order to always produce the situation where a ψ stands to the left of a $\overline{\psi}$ we must move the last factor $\psi(z_N)$ all the way to the left. But that means anticommuting through $2(N-1)+1$ Grassmann fields. Thus

A fermion loop gives an extra factor -1

To understand how to pass to Feynman rules in momentum space, consider the Fourier transformed of the diagram in fig.3.2. As in sect. 2.1, we consider first (cf. eq.(3.59))

$$\begin{aligned}
\tilde{G}(p,p';q_1,\ldots,q_N) &= \int d^4x d^4y d^4w_1\ldots d^4w_N e^{-i(px+p'y+q_1w_1+\ldots+q_Nw_N)} \\
&\times \int d^4z_1\ldots d^4z_N i\int \frac{d^4k_{N+1}}{(2\pi)^4}\frac{e^{-ik_{N+1}(x-z_N)}}{\not{k}_{N+1}+im} i\int \frac{d^4k_N}{(2\pi)^4}\frac{e^{-k_N(z_N-z_{N-1})}}{\not{k}_N+im} \\
&\cdot \ \ldots i\int \frac{d^4k_1}{(2\pi)^4}\frac{e^{-ik_1(z_1-y)}}{\not{k}_1+im}(-g)\int \frac{d^4r_1}{(2\pi)^4}\frac{e^{ir_1(w_1-z_1)}}{r_1^2+M^2} \\
&\cdot \ \ldots(-g)\int \frac{d^4r_N}{(2\pi)^4}\frac{e^{ir_N(w_N-z_N)}}{r_N^2+M^2}
\end{aligned} \tag{3.60}$$

Here we denoted the boson mass by M to make it distinct from the fermion mass, m. To extract the usual delta function of 4-momentum conservation, we change variables (using translation invariance):

$$\begin{aligned}
z_l &\to x+z_l \\
y &\to x+y \\
w_l &\to x+w_l
\end{aligned} \tag{3.61}$$

giving

$$\begin{aligned}
\tilde{G}(p,p';q_1,\ldots,q_N) &= \int d^4x d^4y d^4w_1\ldots d^4w_N e^{-ix(p+p'+q_1+\ldots+q_N)}e^{-i(p'y+q_1w_1+\ldots+q_Nw_N)} \\
&\times \int d^4z_1\ldots d^4z_N i\int \frac{d^4k_{N+1}}{(2\pi)^4}\frac{e^{ik_{N+1}z_N}}{\not{k}_{N+1}+im} i\int \frac{d^4k_N}{(2\pi)^4}\frac{e^{-k_N(z_N-z_{N-1})}}{\not{k}_N+im} \\
&\cdot \ \ldots i\int \frac{d^4k_1}{(2\pi)^4}\frac{e^{-ik_1(z_1-y)}}{\not{k}_1+im}(-g)\int \frac{d^4r_1}{(2\pi)^4}\frac{e^{ir_1(w_1-z_1)}}{r_1^2+M^2} \\
&\cdot \ \ldots(-g)\int \frac{d^4r_N}{(2\pi)^4}\frac{e^{ir_N(w_N-z_N)}}{r_N^2+M^2}
\end{aligned} \tag{3.62}$$

We now perform the integrations over the various points in coordinate space successively with the following results:

$$\begin{aligned}
x &: (2\pi)^4\delta^4(p+p'+q_1+\ldots+q_N) \\
y &: (2\pi)^4\delta^4(-p'+k_1) \\
w_1 &: (2\pi)^4\delta^4(-q_1+r_1) \\
\vdots &\quad \vdots \\
w_N &: (2\pi)^4\delta^4(-q_N+r_N) \\
z_N &: (2\pi)^4\delta^4(-k_N+k_{N+1}-r_N) \\
\vdots &\quad \vdots \\
z_1 &: (2\pi)^4\delta^4(-k_1+k_2-r_1)
\end{aligned} \tag{3.63}$$

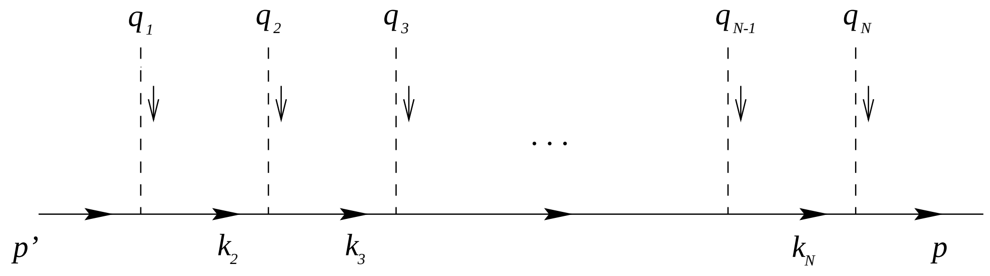


Figure 3.5: The same tree diagram as before, ready for evaluation in momentum space. The k_i's are fixed by momentum conservation.

We see that we produce energy-momentum conservation delta functions at each vertex as usual and that we may use fig.3.5 to write down the result for the "reduced" Greens function in momentum space (the one that has it's overall momentum conservation delta function removed) directly as:

$$\begin{aligned} G(p,p';q_1,\ldots,q_N) &= \frac{i}{\not p + im}\frac{i}{\not k_N + im}\cdots\frac{i}{\not k_2 + im}\frac{i}{\not p' + im} \\ &\cdot\ (-g)^N\frac{1}{q_1^2+M^2}\cdot\ldots\cdot\frac{1}{q_N^2+M^2} \end{aligned} \tag{3.64}$$

Notice that this is a matrix in Dirac indices, just like the Greens function we start from:

$$\langle 0|\hat\psi(x)\hat{\bar\psi}(y)\hat\phi(w_1)\ldots\hat\phi(w_N)|0\rangle$$

And notice that we work our way against the fermion-arrow. Hence we deduce the following Feynman-rules in momentum space:

Bosonic propagator $= 1/(q^2+M^2) =$

Fermionic propagator $= i/(\not k + im) =$

Fermion-Boson vertex $= -g =$

As a final example, consider the evaluation of the loop diagram fig.3.4 in momentum space. We can now directly write down:

$$\begin{aligned} G^{1-loop}(q_1,\ldots,q_N) &= -(-g)^N\prod_{i=1}^{N}\frac{1}{q_i^2+M^2} \\ &\times\ \int\frac{d^4l}{(2\pi)^4}Tr\{\frac{i}{\not l + im}\frac{i}{\not k_{N-1}+im}\cdots\frac{i}{\not k_1 + im}\} \end{aligned} \tag{3.65}$$

where

$$l \ = \ k_0 = k_N$$

$$\begin{aligned}
k_1 &= l + q_1 \\
&\vdots \\
k_i &= k_{i-1} + q_i \\
&\vdots
\end{aligned} \tag{3.66}$$

From these examples we see that we may generalize and obtain the Feynman rules for a completely general interaction part as follows:

Consider a piece in the interaction, $S_I[\overline{\psi}, \psi, \phi_1, \ldots, \phi_n]$, involving the fermion fields, ψ and $\overline{\psi}$ as well as bosonic fields, $\phi_1, \ldots \phi_n$. We have seen already how to find the propagator. The vertex corresponding to this interaction piece is

$$\begin{aligned}
&\int d^4xd^4yd^4z_1 \ldots d^4z_n e^{-i(px+p'y+q_1z_1+\ldots+q_nz_n)} \\
&\frac{\delta}{\delta\phi_1(z_1)} \cdots \frac{\delta}{\delta\phi_n(z_n)} (-\frac{\delta}{\delta\overline{\psi}(y)} \frac{\delta}{\delta\psi(x)}) \{-S_I[\overline{\psi}, \psi, \phi_1, \ldots, \phi_n]\}
\end{aligned} \tag{3.67}$$

Notice that this will produce a matrix in Dirac-space: $\frac{\delta}{\delta\overline{\psi}(y)}$ is a column whereas $\frac{\delta}{\delta\psi(x)}$ is a row.

Let us consider how the rule works on the two examples we have considered. For the simple Yukawa interaction, we find

$$\begin{aligned}
&\int d^4xd^4yd^4ze^{-i(px+p'y+qz)} \frac{\delta}{\delta\phi(z)} (-\frac{\delta}{\delta\overline{\psi}_a(y)} \frac{\delta}{\delta\psi_b(x)}) \{-g \int d^4w \overline{\psi}_c(w)\psi_c(w)\phi(w)\} \\
= \;& -g\delta_{ab}(2\pi)^4\delta^4(p + p' + q)
\end{aligned} \tag{3.68}$$

in accordance with what we would expect from the examples worked out. Similarly from the QED interaction we get

$$\begin{aligned}
&\int d^4x_1d^4x_2d^4x_3e^{-i(p_1x_1+p_2x_2+p_3x_3)} \frac{\delta}{\delta A_\mu(x_3)} (-\frac{\delta}{\delta\overline{\psi}_a(x_2)} \frac{\delta}{\delta\psi_b(x_1)}) \\
&\{ie \int d^4x \overline{\psi}_c(x)A_\nu(x)(\gamma_\nu)_{cd}\psi_d(x)\} \\
= \;& (2\pi)^4\delta^4(p_1 + p_2 + p_3)ie(\gamma_\mu)_{ab}
\end{aligned} \tag{3.69}$$

The technique indicated here will turn out to be particularly useful for deriving Feynman rules for interactions involving derivatives, such as is the case of non-Abelian theories.

Chapter 4

Formal Developments

4.1 The Effective Action

4.1.1 The Classical Field

Let us consider a rather general euclidean quantum field theory, described by the action

$$S[\phi]$$

where "ϕ" may be considered a general *set* of fields, but we want to "pollute" the notation as little as possible, so we shall refer to the space time point, x, as well as any other labels on the field collectively by an index:

$$\phi_i$$

Thus, for a gluon field, $A^a_\mu(x)$ with Lorentz index, μ, colour index, a, and space time point, $x = (t, \vec{x})$, we let i stand for

$$i \equiv \{x, a, \mu\}$$

Similarly we shall imply sums over repeated indices, i, and mean those to include if necessary any *integration* over space time points.

In previous sections we have seen that

$$Z[J] = \int \mathcal{D}\phi e^{-S[\phi]+J_i\phi_i}$$

is the generating functional for the "full" Greens functions. It is very analogous to the *partition function* in statistical mechanics. Likewise we have seen that $W[J]$ defined by

$$e^{-W[J]} = Z[J]$$

is the generating functional of the *connected* Greens functions. It is very analogous to the *free energy* in statistical mechanics.

We now want to introduce yet another generating functional which turns out to be very convenient in many cases. It is the *Effective Action.* It will be introduced as the generating functional of the *one particle irreducible* Greens functions, or 1PI Greens functions for short. These are essentially (see later) the sums of diagrams which are not only connected, but which also have the property that they cannot be separated in two by cutting a single

Figure 4.1: Examples of one particle reducible (a) and one particle irreducible (1PI) (b) Feynman diagrams

propagator line. Fig.4.1 provides examples.

It is probably rather unclear at the moment why such a strange definition should lead to anything particularly useful. However, we shall soon see some very nice properties of the effective action (justifying the name). It turns out to be the most natural object to work with in renormalization theory. The effctive action is very analogous to the *thermodynamic potential* in statistical mechanics.

We begin by introducing the *classical field*, $\Phi_i^{Cl}[J]$, for a given external current, J_i. By this we mean the field which would be observed in an actual classical experiment. It is precisely the quantum theoretical expectation value of the field, properly normalizied:

$$\begin{aligned}
\Phi^{Cl}(x;J) &\equiv \Phi_i^{Cl}[J] \equiv \langle 0|\hat{\phi}_i|0\rangle_J / \langle 0|0\rangle_J \\
&= \int \mathcal{D}\phi e^{-S[\phi]+J\cdot\phi}\phi_i / Z[J] \\
&= \frac{1}{Z[J]}\frac{\delta}{\delta J_i}Z[J] \\
&= \frac{\delta}{\delta J_i}(-W[J]) \\
&= G_1^c(x;J) \qquad (4.1)
\end{aligned}$$

where we sometimes like to write (x) instead of i. We see that this “classical field” has several interpretations:

1. It is the normalized, quantum theoretical expectation value of the field variable, something which usually vanishes unless the external current is non-vanishing. Classically this is very familiar. The electromagnetic field, for example, vanishes unless there are some external charges or magnets around. These are indeed described by external “currents”. From given external currents we may uniquely calculate the *classical*, “classical” field. The same is true in the quantum problem in principle. But the calculation is much more difficult in general. For an electromagnetic field it will not be enough to use the classical Maxwell equations of motions. There will be “quantum corrections” that Maxwell did not know about. The object Φ^{Cl} is what an actual measurement would find on the average.

2. We see that this classical field has the interpretation of being the *connected one-point function in the presence of the external currents.*

In accord with our previous notation we introduce the picture in fig.4.2 for the classical field. As the simplest possible example, let us consider the free bosonic field theory,

$$S_0[\phi] = \frac{1}{2}\int d^dx[\partial_\mu\phi\partial_\mu\phi + m^2\phi^2] = \frac{1}{2}\phi_i\Delta_{ij}^{-1}\phi_j$$

where

$$\Delta_{ij}^{-1}\phi_j \equiv [-\partial_\mu\partial_\mu + m^2]\phi(x)$$

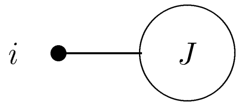


Figure 4.2: Pictorial representation of the connected one-point function: the classical field.

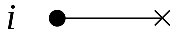


Figure 4.3: Pictorial representation of the classical field. A dot represents a fixed point in space time, whereas a cross represents a point at which a factor J is encountered and which is integrated over.

for $i \equiv (x)$. Adding an external current, the action is modified into

$$S_0[\phi] - J \cdot \phi = \frac{1}{2}\phi_i \Delta_{ij}^{-1} \phi_j - J_k \phi_k$$

and the classical equations of motion give

$$\Delta_{ij}^{-1} \phi_j - J_i = 0$$

or

$$\phi_i = \Delta_{ij} J_j$$

(Beware that $\Delta_{ij}^{-1} \neq 1/\Delta_{ij}$. The meaning is that it is the inverse *matrix* or *integration kernel* of Δ_{ij} so that $\Delta_{ik}\Delta_{kj}^{-1} = \delta_{ij}$, sum over k implied.) This is very well known. Thus, in elctrostatics, if the external "current" corresponds to a point charge at the origin, we may think of ϕ as representing the electrostatic potential. The classical equation of motion says that the laplacian acting on it is a delta function (the point charge). The solution is Coulomb's law, which is the "propagator" of the laplacian.

In sect. 2.2 we saw that the generating functional in this very simple, free case, is given by

$$-W_0[J] = \frac{1}{2} J_i \Delta_{ij} J_j$$

Thus we get for the classical field, $\Phi_i^{Cl,(0)}$, in this free case

$$\Phi_i^{Cl,(0)} = \frac{\delta}{\delta J_i}(-W_0[J]) = \Delta_{ij} J_j$$

We see that in that case there are no quantum corrections: The free fields are unable to feel the effect of quantum fluctuations of one another. We represent the Feynman diagram for the classical field as in fig.4.3. Similarly fig.4.4 represents several Feynman diagram for the classical field (the one point connected Greens function in the presence of the external current) for the case of a $\lambda\phi^4$ theory.

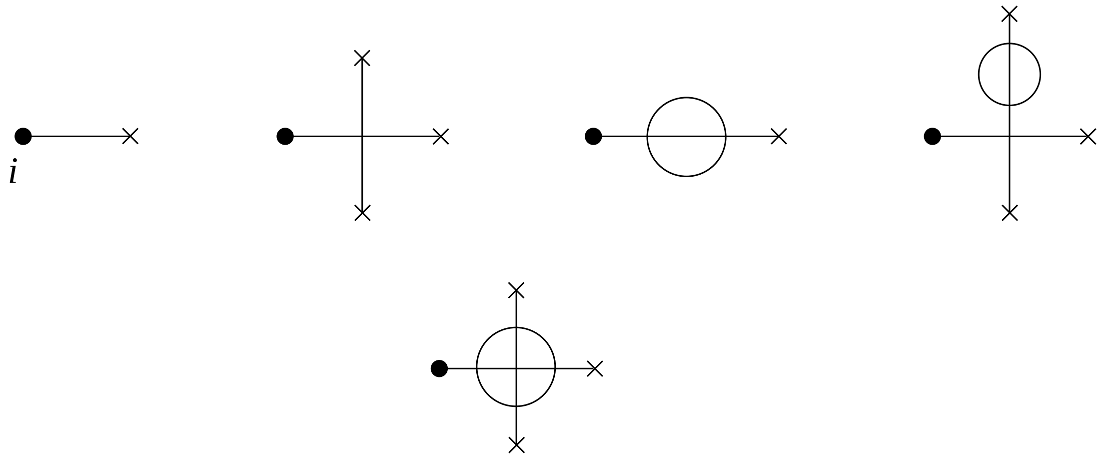


Figure 4.4: Several Feynman diagrams for the classical field in the $\lambda\phi^4$ theory.

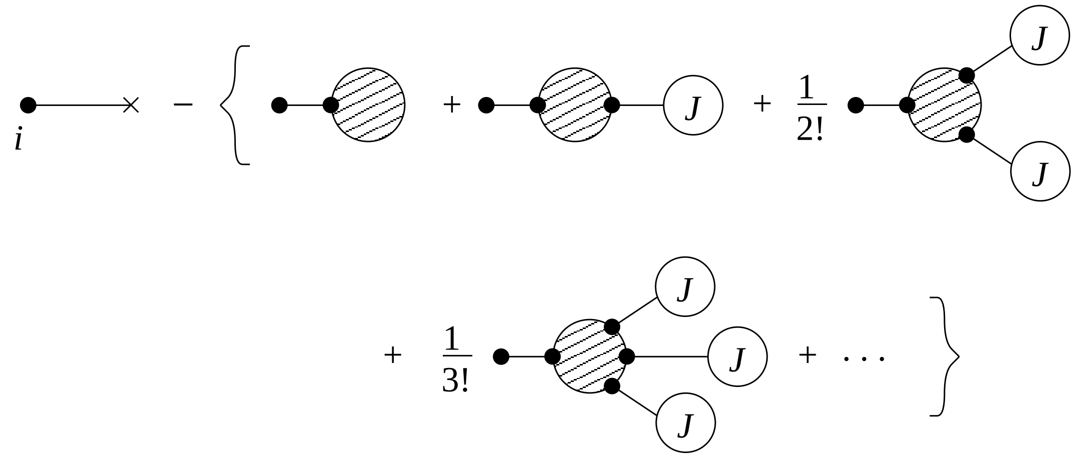


Figure 4.5: The formal expansion of the classical field, based on a classification of Feynman diagrams. The shaded blobs with dots rather than legs are our pictorial representation for the 1PI N-point functions.

4.1.2 The One Particle Irreducible (1PI) Greens Functions

We now introduce a classification of all the Feynman diagrams for the classical field based on 1PI N-point functions. The idea is given in fig.4.5. It may be described as follows: starting at the point, x (or, i), at which we think of evaluating the classical field, we will always find a propagator leaving that point. The simplest possibility is that the propagator ends at the external current and gives us the standard classical *free* field. However, there are many other possibilities. The next simplest thing to happen is that we find the propagator ending in a fully 1PI diagram: the diagram cannot be separated in two by cutting any one line at all (other than the external propagator starting at x). The sum of all these possibilities define the 1PI irreducible one point function, Γ_i or $\Gamma(x)$. Notice that our pictorial representation for that is a shaded blob *with no legs coming out.* Instead of legs we just put black dots. It is sometimes referred to as the *amputated* Greens function: legs have been removed.

The next possibility is that the first propagator ends in a diagram which is in fact one particle *re*-ducible, but which has the property that the first part of it takes the form of a 1PI piece with exactly two legs emanating from it. One leg is just our very first propagator, the other one may hook on to any new Feynman diagram at all. The sum of all these possibilities gives rise to the third graph in fig.4.5. It defines the two point 1PI Greens function and it is built with the original one point connected Greens function (i.e. the classical field) as a factor: It represents all the possibilities left over for the second leg to end in. For reasons that will become clear in a little while, we reserve the name Γ_{ij} for a slightly modified version of the 1PI two point function. The final result may be summarized in the expansion

$$\Phi_i^{Cl} = \Delta_{ij}\{J_j - (\Gamma_j + \Pi_{jk}\Phi_k^{Cl} + \frac{1}{2}\Gamma_{jkl}\Phi_k^{Cl}\Phi_l^{Cl}... + \frac{1}{(n-1)!}\Gamma_{ji_1\cdots i_n}\Phi_{i_1}^{Cl}\cdots\Phi_{i_n}^{Cl} + ...)\} \tag{4.2}$$

Notice that we have introduced a normalization convention for each term based on some combinatorics that turns out to be useful. Also the sign on the Γ's is conventional.

Finally we are able to introduce the definition of something which will almost, but not quite, be the effective action. It is, however, exactly the generating functional for 1PI Greens functions:

$$\hat{\Gamma}[\Phi^{Cl}] \equiv \Gamma_i\Phi_i^{Cl} + \frac{1}{2}\Pi_{ij}\Phi_i^{Cl}\Phi_j^{Cl} + \frac{1}{3!}\Gamma_{ijk}\Phi_i^{Cl}\Phi_j^{Cl}\Phi_k^{Cl} + ... \tag{4.3}$$

so that

$$\Gamma_{i_1 i_2\cdots i_n} = \frac{\delta}{\delta\Phi_{i_1}^{Cl}}\cdots\frac{\delta}{\delta\Phi_{i_n}^{Cl}}\hat{\Gamma}[\Phi^{Cl}]|_{\Phi^{Cl}\equiv 0} \tag{4.4}$$

except for the two point function where we use the name Π_{ij} rather than Γ_{ij}. Comparing eqs.(4.2) and (4.3), we see that we have

$$-\frac{\delta W}{\delta J_i} \equiv \Phi_i^{Cl} = \Delta_{ij}(J_j - \frac{\delta\hat{\Gamma}}{\delta\Phi_j^{Cl}}) \tag{4.5}$$

or

$$\frac{\delta\hat{\Gamma}}{\delta\Phi_i^{Cl}} + \Delta_{ij}^{-1}\Phi_j^{Cl} = J_i$$

or

$$\frac{\delta}{\delta \Phi_i^{Cl}}[\hat{\Gamma} + \frac{1}{2}\Phi_l^{Cl}\Delta_{lj}^{-1}\Phi_j^{Cl}] = J_i \tag{4.6}$$

This means that we feel very tempted to define

$$\Gamma[\Phi^{Cl}] \equiv \hat{\Gamma}[\Phi^{Cl}] + \frac{1}{2}\Phi_l^{Cl}\Delta_{lj}^{-1}\Phi_j^{Cl} \tag{4.7}$$

and that is the object we call the **effective action**. One reason for the name is that the effective action satisfies an equation very similar to the one satisfied by the *classical* action. Indeed, we see that eq.(4.6) becomes

$$\frac{\delta \Gamma}{\delta \Phi_i^{Cl}} = J_i \tag{4.8}$$

which is very similar to the classical equations of motion: $0 = \delta S/\delta\phi_i - J_i$, or

$$\frac{\delta S}{\delta \phi_i} = J_i \tag{4.9}$$

From the way we have introduced it, it is clear that the effective action is (almost) the generating functional of the 1PI Greens functions. The only difference is with the two point function. In fact we may write

$$\Gamma[\Phi^{Cl}] = \sum_{N=1}^{\infty} \frac{1}{N!}\Gamma_{i_1\cdots i_N}\Phi_{i_1}^{Cl}\cdots\Phi_{i_N}^{Cl} \tag{4.10}$$

and where from the above, the name Γ_{ij} now is seen to mean exactly

$$\Gamma_{ij} \equiv \Pi_{ij} + \Delta_{ij}^{-1} \tag{4.11}$$

It has become customary sometimes to refer loosely to Γ_{ij} as the 2 particle 1PI Greens function.

Notice, that for a given quantum field theory, the classical field is a unique functional of the external current: given J, Φ^{Cl} is in principle defined. When we talk about $W[J]$ as the generating functional of connected Greens functions, we think of it as a functional of this current. But when we talk about $\Gamma[\Phi^{Cl}]$ as the generating functional of the 1PI Greens functions, we think of it as a functional of Φ^{Cl}. However, in principle we may imagine that we may solve one in terms of the other as we please.

We now show that the following elegant path integral representation for the effective action is true:

Theorem (effective action)

$$e^{-\Gamma[\Phi^{Cl}]+J_i\Phi_i^{Cl}} = \int \mathcal{D}\phi e^{-S[\phi]+J_i\phi_i} \tag{4.12}$$

Clearly this theorem is yet another reason for the name: *effective* action. Notice that in the litterature, various conventions are given for which phase to choose for the $J \cdot \phi$ term. However, we should allways have the same phases on the left and the right in the formula above.

Proof of Theorem: The theorem is equivalent to the statement

$$W[J] = \Gamma[\Phi^{Cl}] - J \cdot \Phi^{Cl}$$

or that the effective action is the Legendre transform of the "Free energy" (as with the thermodynamic potential in statistical mechanics):

$$\Gamma[\Phi^{Cl}] = W[J] + J_k \Phi_k^{Cl} \tag{4.13}$$

This expression is correct because it implies that the effective action satisfies the correct quantum version of the "equation of motion". Indeed take a functional derivative with respect to Φ_i^{Cl} and obtain from the right hand side:

$$\frac{\delta W}{\delta J_k}\frac{\delta J_k}{\delta \Phi_i^{Cl}} + \frac{\delta J_k}{\delta \Phi_i^{Cl}}\Phi_k^{Cl} + J_i = J_i$$

since

$$-\frac{\delta W}{\delta J_k} = \Phi_k^{Cl}$$

Hence Γ given by eq.(4.13) satisfies the crucial eq.(4.8). It follows that eq.(4.13) must hold up to an additive contribution *independent* of J or Φ^{Cl}. such contributions do not interest us at present.

4.1.3 The Two Point Function

The 1PI two point function is Π_{ij}, not Γ_{ij}. But so what is Γ_{ij}? We now show that this Greens function nevertheless has a very nice interpretation: it is the *inverse*, full (connected) propagator, or two point function. To show this, recall that the connected one point function in the presence of the external current is

$$G_i^c[J] = -\frac{\delta W[J]}{\delta J_i} = \Phi_i^{Cl}[J]$$

so that the two point function in the presence of the external current becomes (our notation here is for commuting, bosonic fields; for anti commuting, fermionic fields there are a few extra minus signs which the reader can easily fix)

$$G_{ij}^c[J] = -\frac{\delta}{\delta J_j}\frac{\delta}{\delta J_i}W[J] = \frac{\delta \Phi_i^{Cl}}{\delta J_j}$$

Now use eq.(4.2) to obtain

$$\frac{\delta \Phi_i^{Cl}}{\delta J_j} = \Delta_{ij} - \Delta_{ik}\Pi_{kl}\frac{\delta \Phi_l^{Cl}}{\delta J_j} - ...$$

where the remaining terms vanish when $J \equiv \Phi^{Cl} \equiv 0$. Thus

$$G_{ij}^c = \Delta_{ij} - \Delta_{ik}\Pi_{kl}G_{lj}^c$$

or

$$(\delta_{ik} + \Delta_{il}\Pi_{lk})G_{kj}^c = \Delta_{ij}$$

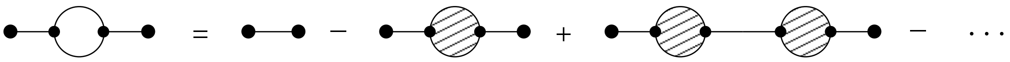


Figure 4.6: Graphical representation of the the connected two point function in terms of sums of graphs of higher and higher one particle reducibility

or

$$(\Delta^{-1}_{mk} + \Pi_{mk})G^c_{kj} = \delta_{mj}$$

or, using eq.(4.11),

$$\Gamma_{mk}G^c_{kj} = \delta_{mj} \tag{4.14}$$

This is what we set out to proove. Notice that we may write one of the above equations as the following matrix equation

$$(1 + \Delta\Pi)G^c = \Delta$$

or

$$G^c = (1 + \Delta\Pi)^{-1}\Delta = (1 - \Delta\Pi + \Delta\Pi\Delta\Pi - \Delta\Pi\Delta\Pi\Delta\Pi + ...)\Delta$$

(the standard proof of the identity $(1 - x)^{-1} = 1 + x + x^2 + x^3 + ...$ is easily seen to hold equally well if x is an operator) or

$$G^c_{ij} = \Delta_{ij} - \Delta_{il}\Pi_{lk}\Delta_{kj} + \Delta_{il}\Pi_{lk}\Delta_{km}\Pi_{mn}\Delta_{nj} - ... \tag{4.15}$$

This equation has the graphical representation shown in fig. 4.6. Both have an obvious interpretation: The first term is the free field contribution. The next term is the sum of all Feynman diagrams which are 1PI apart from the two external legs. Then comes terms of higher and higher one particle reucibility. Apart from the signs we could have written this expression down from the very beginning just by thinking in terms of such a one particle reducibility classification.

4.1.4 The Classical Action as the Generating Functional of Tree Diagrams

From the theorem, eq.(4.12) we see that in the classical limit when quantum fluctuations in the path integral are unimportant so that the effective quantum average is the same as the classical expression, the effective action is the same thing as the classical action. We have also seen that the effective action is a generating functional of diagrams. We now want to see that in the same sense, the classical action is the generating functional of *tree* diagrams. Indeed it is important to realize that the tree diagram approximation to quantum field theory *is the same thing as the classical approximation.* Actually that statement should be clarified. Many of the successes of quantum electro dynmics and other parts of the standard model are well approximated by tree diagrams. Does that really mean that we merely do classical physics and not quantum physics? No! we do use quantum theory heavily to interpret the result of our calculations in terms of quantum amplitudes with all the standard mysteries of interference and probablility interpretations. *But* realy, the tree diagram approximation does not teach us *anything* about the particular quantum aspects of the *field theory.* For that we have to go to loop diagrams.

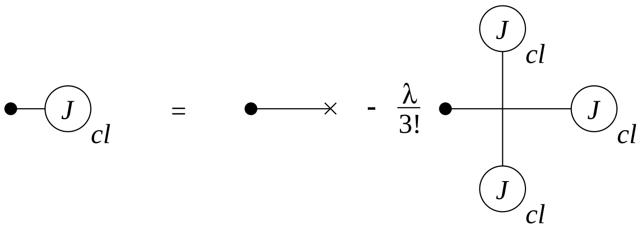


Figure 4.7: Self consistency condition for the solution of the classical field equations for $\lambda\phi^4$ theory.

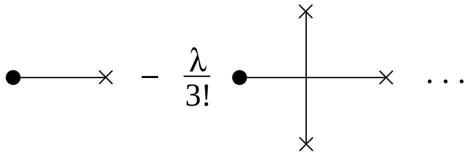


Figure 4.8: The classical field to first order in λ generates the 1- and 4-point tree diagrams.

We now examine this role of the classical action in a simple example: $\lambda\phi^4$ theory. The action is

$$S[\phi] = \frac{1}{2}\phi_i \Delta^{-1}_{ij}\phi_j + \frac{\lambda}{4!}\sum_i \phi_i^4$$

and the sums on indices are really integrations over space time. So the equation of motion in the presence of an external current, $\delta S/\delta\phi_i = J_i$, becomes

$$\Delta^{-1}_{ij}\phi_j + \frac{\lambda}{3!}\phi_i^3 = J_i$$

or

$$\phi_i = \Delta_{ij}J_j - \frac{\lambda}{3!}\Delta_{ij}\phi_j^3$$

an equation which we may display pictorially as in fig.4.7. We may think of solving this rather complicated classical equation in perturbation theory as a function of the self coupling constant, λ. Trivially, for $\lambda = 0$ we have the free field solution. A first order approximation consists in introducing that (order λ^0 solution) in the right hand side to obtain the first order solution fig.4.8. More systematicall, we may replace any classical field-"balloon" by the consistency equation. The first step gives fig. 4.9. This process may evidently be continued, and we see how indeed all possible tree diagrams get generated in the process.

4.2 The Dyson-Schwinger Equation

In the theory described by the action, $S[\phi]$ and extended by an external current, J for each field component,

$$S[\phi] \to S[\phi] - \phi \cdot J$$

the classical equations of motion take the form

$$\frac{\delta S[\phi]}{\delta\phi_i} - J_i = 0 \tag{4.16}$$

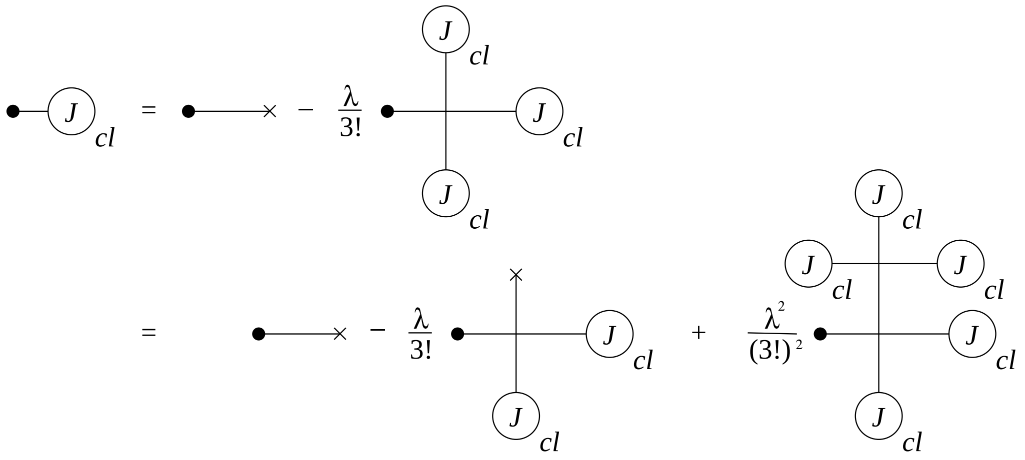


Figure 4.9: The top "balloon" (classical field) in the last part of the figure representing the equation of motion, has been replaced by the entire equation.

Ehrenfest's theorem states that these equations hold "on the average" in the quantum theory. Let us derive this result from the path integral. It gives us essentially what is usually called *the Dyson-Schwinger equations.*

To se the point consider first a simple one-dimensional function, $f(x)$, with the derivative integrable and $f(\pm\infty) = 0$. Hence we have

$$\int_{-\infty}^{+\infty} dx \frac{d}{dx} f(x) = 0$$

Similarly, for an arbitrary field component at any point, ϕ_i we get

$$0 = \int \mathcal{D}\phi \frac{\delta}{\delta\phi_i} e^{-S[\phi]+J\cdot\phi} \tag{4.17}$$

carrying out the differentiation we get

$$\begin{aligned} 0 &= \int \mathcal{D}\phi \{-\frac{\delta S[\phi]}{\delta\phi_i} + J_i\} e^{-S[\phi]+J\cdot\phi} \\ &= \{-\frac{\delta S[\phi]}{\delta\phi_i}|_{\phi\equiv\delta/\delta J} + J_i\} Z[J] \end{aligned} \tag{4.18}$$

clearly the first line is precisely the expression of Ehrenfest's theorem, whereas the second line will be our first version of the Dyson-Schwinger equation. To get that we used that the expectaion value of any function at all of the field may be obtained as usual by hitting the generating functional with that very function where ϕ has been replaced by $\delta/\delta J$.

Historically the equation was not obtained using this very fast path integral derivation. Dyson obtained it by analysing Feynman diagrams. Indeed the equation has a simple diagrammatic interpretation. That however differs in detail from one theory to the other.

Let us separate out a free part of the action and write quite generally,

$$S[\phi] = \sum_{ij} \frac{1}{2} \phi_i \Delta_{ij}^{-1} \phi_j + S_I[\phi]$$

so that

$$\frac{\delta S[\phi]}{\delta\phi_i} = \sum_j \Delta^{-1}_{ij}\phi_j + \frac{\delta S_I[\phi]}{\delta\phi_i}$$

and

$$\sum_i \Delta_{li}\frac{\delta S[\phi]}{\delta\phi_i} = \phi_l + \sum_i \Delta_{li}\frac{\delta S_I[\phi]}{\delta\phi_i}$$

This allows us to obtain the Dyson-Schwinger equation in the following form

$$\frac{\delta}{\delta J_l}Z[J] = \sum_i \Delta_{li}J_iZ[J] - \sum_i \Delta_{li}\frac{\delta S_I[\phi]}{\delta\phi_i}|_{\phi\equiv\delta/\delta J}Z[J] \tag{4.19}$$

It is difficult to proceed in a simple way without considering an example.

Hence, consider

$$S_I[\phi] = \frac{g_3}{3!}\sum_i \phi_i^3 + \frac{g_4}{4!}\sum_i \phi_i^4 \tag{4.20}$$

QCD will be a slight generalization of this: there we shall meet triple-gluon and quadruple-gluon couplings. In this theory we find

$$\frac{\delta S_I[\phi]}{\delta\phi_i} = \frac{g_3}{2}\phi_i^2 + \frac{g_4}{3!}\phi_i^3$$

Thus the Dyson-Schwinger equation becomes

$$\frac{\delta}{\delta J_l}Z[J] = \sum_i \Delta_{li}J_iZ[J] - \sum_i \Delta_{li}\{\frac{g_3}{2!}\frac{\delta^2}{\delta J_i\delta J_i} + \frac{g_4}{3!}\frac{\delta^3}{\delta J_i\delta J_i\delta J_i}\}Z[J] \tag{4.21}$$

(sum over i implied). We may easily convert that to an identity between Greens functions. Indeed let us (i) put $l = i_1$, (ii) then take functional derivatives on both sides with respect to $J_{i_2}, \cdots, J_{i_n}$, and (iii) finally put $J \equiv 0$ to obtain

$$\begin{aligned} G^{(n)}_{i_1i_2\cdots i_n} &= \Delta_{i_1i_2}G^{(n-2)}_{i_3i_4\cdots i_n} + \Delta_{i_1i_3}G^{(n-2)}_{i_2i_4\cdots i_n} + \cdots + \Delta_{i_1i_n}G^{(n-2)}_{i_2i_3\cdots i_{n-1}} \\ &- \sum_i \Delta_{i_1i}\{\frac{g_3}{2}G^{(n+1)}_{iii_2\cdots i_n} + \frac{g_4}{3!}G^{(n+2)}_{iiii_2\cdots i_n}\} \end{aligned} \tag{4.22}$$

This version of the Dyson-Schwinger equation has a very obvious diagrammatic interpretation shown in fig.4.10. It works as follows: Working from a particular line, say the one ending at i_1, we may follow that line, and one of several things may happen: (i) we may find that the line continues without ever meeting any vertex and finally ends in one of the remaining exterior points, $i_2, \cdots, i_n$; (ii) it may happen that following the line we do find a vertex, and that the first vertex we meet is a ϕ^3 vertex; (iii) finally it may happen that following the line, the first vertex we meet is a ϕ^4 vertex. We see how the point, labelled i, in the equation corresponds to the vertex being integrated over in the diagrams. The minus sign in the equation is part of the Feynman rules, whereas we have put explicitly the combinatoric factors from the equation.

One useful application of the Dyson-Schwinger equation consists in using it as a tool for deriving the combinatorial weight factor in Feynman diagrams. Indeed, the entire perturbation expansion may conveniently be generated by iterating again and again the Dyson-Schwinger equation. When several loops are involved, this can be considered the

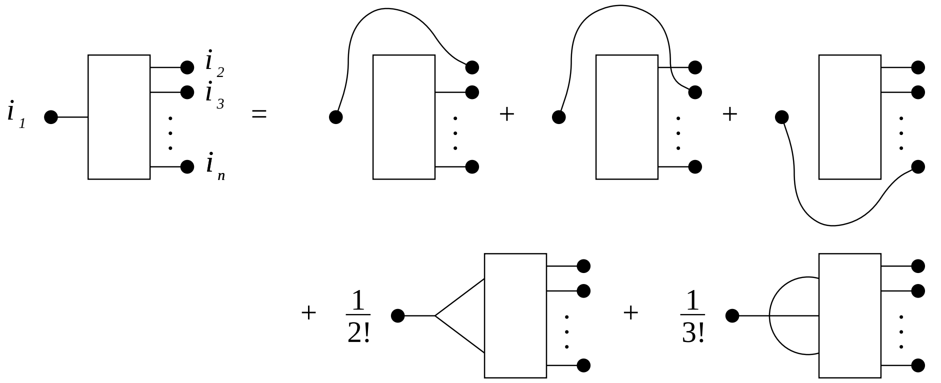


Figure 4.10: The diagrammatic version of the Dyson-Schwinger equation for a simple combined ϕ^3 and ϕ^4 theory.

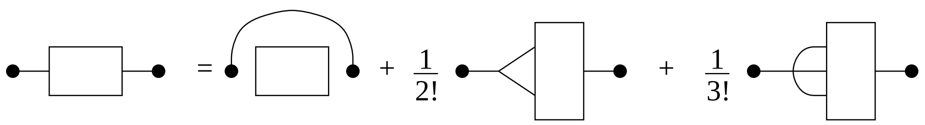


Figure 4.11: The Dyson-Schwinger equation for the two-point function (in a ϕ^3 plus ϕ^4 theory).

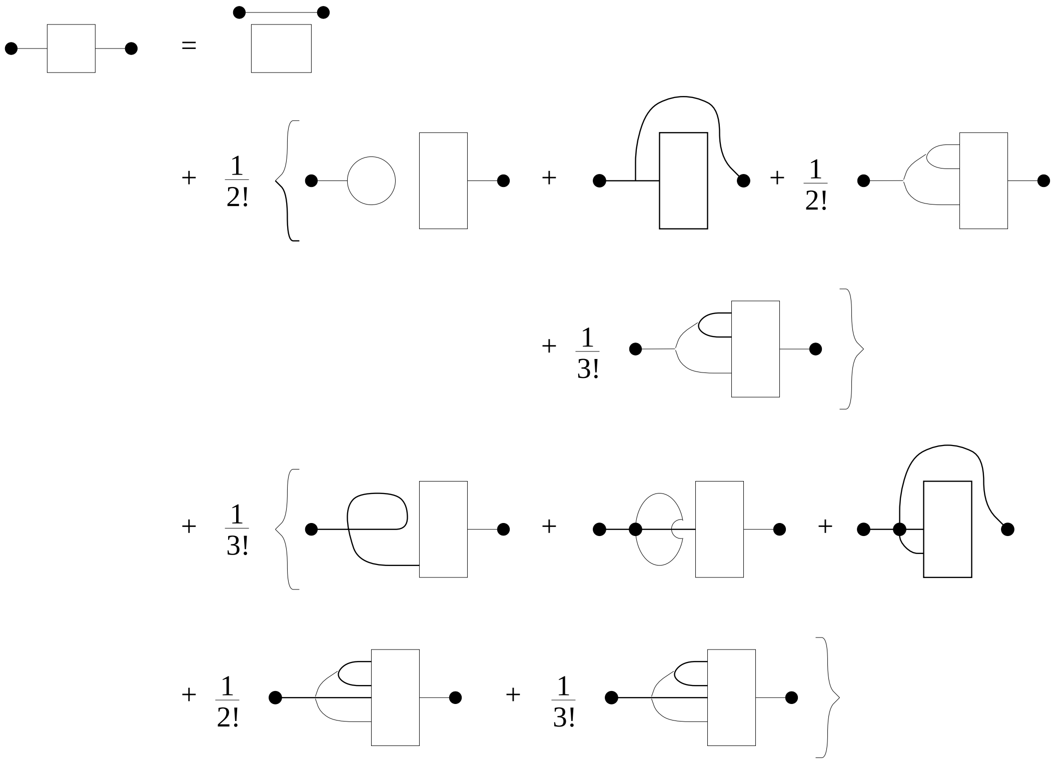


Figure 4.12: The once iterated version of the DS equation for the two point function (in the ϕ^3 plus ϕ^4 theory).

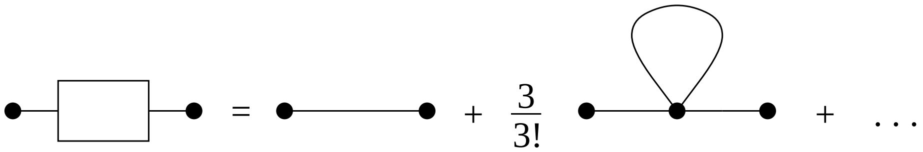


Figure 4.13: The two point function to order $g_3^0 g_4^1$. The weight factor is $\frac{1}{2}$.

most reliable and systematic way of calculating the weight factors. As an example consider the DS equation for the two point function, fig. 4.11.

We then treat the top line in the last two graphs in fig. 4.11 using the DS equations for the three-point and the four-point functions. The result is shown in fig. 4.12. Suppose we are interested only in the order $g_3^0 g_4^1$ contribution. This we get from the first three graphs in the last parenthesis (of which the first two are identical), and the result is given in fig. 4.13. We see in particular that the weight factor is $\frac{1}{2}$ a result already obtained previously in sect. 2.6 using more primitive techniques.

It is probably clear that we can produce yet other versions of the Dyson-Schwinger equation by using the generating functionals for connected diagrams or even 1PI diagrams. Since we know how these generating functionals are related to $Z[J]$, this is in principle straight forward even though the equations can become quite complicated. We shall not pursue this further here.

4.3 Symmetries and Ward Identities

4.3.1 Symmetry in classical mechanics

Consider a classical, mechanical system with a finite number of degrees of freedom, $\{q_i\}$, and described by the lagrangian

$$L(q_i, \dot{q}_i)$$

so that the classical equations of motion, the Euler-Lagrange equations, are

$$\dot{p}_i = \frac{\partial L}{\partial q_i} \tag{4.23}$$

where the conjugate momentum, p_i is defined by

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$$

Next suppose the lagrangian has an *invariance* or a *symmetry*. As an example, consider a non-relativistic point particle (with mass $m = 1$) moving in a plane in a potential centered at the origin and being rotationally invariant. In this case we have two degrees of freedom, $(q_1, q_2) = (x, y)$, and the potential, $V(|q|)$ is a function only of the norm, $|q|^2 = q_1^2 + q_2^2$. The lagrangian is

$$L(q_i, \dot{q}_i) = \frac{1}{2}(\dot{q}_1^2 + \dot{q}_2^2) - V(|q|)$$

It is evidently invariant under a rotation by an angle, ϵ:

$$\begin{aligned} q_1 &\rightarrow q_1 \cos\epsilon + q_2 \sin\epsilon \\ q_2 &\rightarrow -q_1 \sin\epsilon + q_2 \cos\epsilon \end{aligned} \tag{4.24}$$

For ϵ infinitesimal this becomes

$$\begin{aligned} \delta q_1 &= q_2\epsilon \\ \delta q_2 &= -q_1\epsilon \end{aligned} \tag{4.25}$$

The invariance of the lagrangian is made manifest by going to polar coordinates

$$\begin{aligned} q_1 &= r\cos\theta \\ q_2 &= r\sin\theta \end{aligned} \tag{4.26}$$

in which

$$L(r,\theta,\dot r,\dot\theta) = \frac{1}{2}(\dot r^2 + r^2\dot\theta^2) - V(r)$$

is manifestly independent of θ: θ is a *cyclic* variable.

Evidently, whenever we have an invariance, we may always find such a cyclic variable (or several). If q_i is cyclic, it follows from the equations of motion, eq.(4.23), that the conjugate momentum is conserved:

$$\dot p_i = 0$$

In our simple example, θ is cyclic:

$$\frac{\partial L}{\partial\theta} = 0$$

and

$$p_\theta \equiv \frac{\partial L}{\partial\dot\theta} = r^2\dot\theta$$

is the angular momentum perpendicular to the plane of the motion. We know it is conserved.

The purpose of this section is to see how this fundamental relation between symmetries and conserved quantities appear, first in classical field theory in the form of the *Nöther Theorem*, and then in quantum field theory in the form of *Ward identities.*

4.3.2 Nöther's Theorem

As an example of an invariance in field theory, consider a bosonic field theory depending on N scalar fields, ϕ^i, $i = 1, 2, \cdots, N$. Consider the (euclidean) lagrangian density

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_0 + \mathcal{L}_I \\ \mathcal{L}_0 &= \frac{1}{2}(\partial_\mu\phi^i\partial_\mu\phi^i + m^2\phi^i\phi^i) \\ \mathcal{L}_I &= V(|\phi|), \qquad |\phi|^2 \equiv \phi^i\phi^i \end{aligned} \tag{4.27}$$

where sums over i (and of course over the (euclidean) Lorentz index, μ) is implied. This lagrangian density is clearly invariant under the following group of transformations: Let A_{ij} be an orthogonal $N\times N$ matrix, in other words,

$$A \in O(N)$$

where $O(N)$ is the standard name for the group of orthogonal N-dimensional transformations. Consider the transformation on the fileds

$$\phi^i(x) \to (\phi^i(x))' \equiv \sum_j A_{ij}\phi^j(x) \tag{4.28}$$

It is clear that the lagrangian density is invariant under this transformation. Indeed, if u^i and v^i are N-dimensional vectors, then the scalar product

$$u \cdot v \equiv u^i v^i$$

is invariant under an orthogonal transformation (we do not attach any special meaning here to whether indices are upstairs or downstairs).

The invariance just considered here is called a *global* invariance, since the transformation matrix, A_{ij} is independent of the space-time point, x. For a *local* transformation, $A_{ij}(x)$, the lagrangian would *not* be invariant because of the derivatives. (One can extend the theory to include a *gauge-field* so that there actually is a local *gauge-invariance.* We do not consider this now).

It is convenient to consider the infinitesimal tranformation performed by an orthogonal matrix, A, which is close to the identity, here the $N \times N$ unit matrix, I. Then we may write

$$A_{ij} = \delta_{ij} + i\alpha_{ij} \tag{4.29}$$

where α_{ij} is an infinitesimal matrix. It is easily verified that orthogonality of A:

$$\sum_j A_{ij} A_{lj} = \delta_{il}$$

is equivalent to antisymmetry of α_{ij}:

$$\alpha_{ij} = -\alpha_{ji}$$

If A_{ij} is real, then α_{ij} is imaginary because of the i we have put. The reason is just that then also α_{ij} is *hermitean*, which is sometimes convenient.

Let us consider the set of all $N \times N$ antisymmetric matrices. They clearly form a vector space, and one may count that there are $\frac{1}{2}(N^2 - N) = \frac{1}{2}N(N-1)$ independent ones: the dimension of the vector space of antisymmetric matrices is this large. Let

$$T^a_{ij}, \quad a = 1, 2, \cdots, \frac{1}{2}N(N-1), \quad i, j = 1, 2, \cdots, N$$

be a basis for them. Then any infinitesimal antisymmetric matrix, α_{ij}, may be written

$$\alpha_{ij} = \sum_a \epsilon^a T^a_{ij} \tag{4.30}$$

where ϵ^a, $a = 1, 2, \cdots, \frac{1}{2}N(N-1)$ are infinitesmial numbers.

We may summarize what we have found by saying that the lagrangian density we consider is invariant under the global infinitesimal transformation

$$\delta\phi^i(x) = i \sum_j \alpha_{ij} \phi^j(x) = i \sum_a \epsilon^a \sum_j T^a_{ij} \phi^j(x) \tag{4.31}$$

The antisymmetric $N \times N$ matrices T^a are referred to as the *generators of the Lie-algebra* for $O(N)$.

We now leave the example an abstract the idea. Thus we consider any lagrangian density

$$\mathcal{L}(\phi^i, \partial_\mu \phi^i)$$

depending on a set of fields, $\{\phi^i\}$. And we assume that the lagrangian density is invariant under the set of transformations of the form eq.(4.31), but we do not want to specify in detail how many values the index, a, can assume or what Lie-algebra the generators T^a belong to. All we require is that the lagrangian density is invariant under this global tranformation, whether or not the fields $\phi^i(x)$ satisfy the classical equations of motion.

Under a general field variation, the lagrangian density is changed into

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi^i(x)}\delta\phi^i(x) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\delta\partial_\mu\phi^i(x) \\ &= [\frac{\partial\mathcal{L}}{\partial\phi^i(x)} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}]\delta\phi^i(x) \\ &+ \partial_\mu[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\delta\phi^i(x)] \end{aligned} \tag{4.32}$$

where we used as usual that

$$\delta\partial_\mu\phi^i(x) = \partial_\mu\delta\phi^i(x)$$

The above expression for the variation of the lagrangian density is completely general. We have not used that the variations should be of the form eq.(4.31), we have not used the equations of motion, we have not used any supposed invariance of the lagrangian density, and we have not assumed anything about the boundary conditions of the variations. Now we show how this general expression can be used in several quite different ways:

1. It may be used to derive the equations of motion. For this we assume the "arbitrary" variations, $\delta\phi^i(x)$, "vanish at infinity" so that the last term in $\delta\mathcal{L}$, the "total derivative term", integrates to zero in the variation of the action, S.

2. Next, suppose, $\phi^i(x)$ in fact *satisfies* the equtions of motion, and take in that case $\delta\phi^i(x)$ *not* to be an arbitrary variation, but to be given by the symmetry transformation, eq.(4.31). Then the first bracket in eq.(4.32) vanishes by the equations of motion, and we find

$$\begin{aligned}(\delta\mathcal{L})^{\text{Classical}}_{\text{symmetry}} &= i\partial_\mu[\sum_{i,j}\frac{\delta\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\alpha_{ij}\phi^j(x)] \\ &= i\sum_a\epsilon^a\partial_\mu[\sum_{ij}\frac{\delta\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}T^a_{ij}\phi^j(x)] \\ &= 0 \end{aligned} \tag{4.33}$$

where the last zero occurs becaus of our basic assumtion that the lagrangian density is invariant under the global transformation. This result may now be formulated as

Nöther's Theorem:

If a lagrangian density is invariant under the field transformations given by eq.(4.31), then we define the *Nöther currents*

$$j^a_\mu(x) \equiv \sum_{ij}\frac{\delta\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}T^a_{ij}\phi^j(x) \tag{4.34}$$

and when *the classical equations of motion are satisfied* the Nöther current is conserved:

$$\partial_\mu j^a_\mu(x) = 0 \tag{4.35}$$

Corrollary

The *Nöther charges* defined by

$$Q^a(t) \equiv \int d^3x j_0^a(t,\vec{x}) \tag{4.36}$$

are independent of time, i.e. they are *constants of the motion.*

Proof:

$$\begin{aligned} \dot{Q}^a(t) &= \int d^3x \partial_0 j_0^a(t,\vec{x}) \\ &= -\int d^3x \partial_i j_i^a(t,\vec{x}) \\ &= 0 \end{aligned} \tag{4.37}$$

In the second equality sign we used conservation of the current, and in the last one we used Gauss's theorem for currents "vanishng at infinity".

3. This time we do *not* assume that $\phi^i(x)$ satisfies the classical equations of motion, but we *do* take $\delta\phi^i(x)$ to be given by eq.(4.31) and we *do* assume the lagrangian density is invariant under the transformation for any field configuration (not only for the fields satisfying the classical equations of motion). From eq.(4.32) we then find

$$0 = \delta\mathcal{L} = i\sum_{ij}\alpha_{ij}\{[\frac{\partial\mathcal{L}}{\partial\phi^i(x)} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}]\phi^j(x) + \partial_\mu[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\phi^j(x)]\} \tag{4.38}$$

4. Finally consider the *local* version of the transformation eq.(4.31):

$$\delta\phi^i(x) = i\sum_j \alpha_{ij}(x)\phi^j(x) = i\sum_a \epsilon^a(x)\sum_j T^a_{ij}\phi^j(x) \tag{4.39}$$

Now we cannot assume that $\delta\mathcal{L}$ vanishes, but we still have eq.(4.32). The difference is that now we also get a term involving $\partial_\mu\alpha_{ij}$. In fact we find

$$\begin{aligned} \delta\mathcal{L} &= i\sum_{ij}\alpha_{ij}(x)\{[\frac{\partial\mathcal{L}}{\partial\phi^i(x)} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}]\phi^j(x) + \partial_\mu[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\phi^j(x)]\} \\ &+ i\sum_{ij}\partial_\mu\alpha_{ij}(x)[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\phi^j(x)] \end{aligned} \tag{4.40}$$

Here the term proprtional to $\alpha_{ij}(x)$ itself vanishes because of eq.(4.38) and we get

$$\begin{aligned} \delta S &= \int d^4\delta\mathcal{L} \\ &= i\int d^4x\sum_{ij}\partial_\mu\alpha_{ij}[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^i(x))}\phi^j(x)] \\ &= i\sum_a\int d^4x\partial_\mu\epsilon^a(x)j^a_\mu(x) \\ &= -i\sum_a\int d^4x\epsilon^a(x)\partial_\mu j^a_\mu(x) \end{aligned} \tag{4.41}$$

where we assumed in the last step that our arbitrary functions, $\epsilon^a(x)$ "vanish at infinity", and where we have introduced the Nöther currents, eq.(4.34).

This formula is a general result. It is valid whether or not we evaluate it on fields satisfying the classical equations of motion, so we shall use it in the next section where we give the quantum treatment. *If* in particular we use it for the classical fields, it represents another derivation of Nöther's Theorem: since δS vahishes on classical fileds, we deduce that $\partial_\mu j^a_\mu(x)$ on the classical fields because $\epsilon^a(x)$ are arbitrary functions.

From eq.(4.41) we find the following formula

$$\partial_\mu j^a_\mu(x) = i\frac{\delta S}{\delta\epsilon^a(x)} = -\frac{\delta S}{\phi^i(x)}T^a_{ij}\phi^j(x) \tag{4.42}$$

Example

Consider a field theory with an $O(2)$ symmetry, i.e. characterized by two field components, $\phi_1(x), \phi_2(x)$. Often in that case one prefers to introduce complex combinations:

$$\phi(x) \equiv \frac{1}{\sqrt{2}}(\phi_1(x) + i\phi_2(x)) \quad \phi^*(x) \equiv \frac{1}{\sqrt{2}}(\phi_1(x) - i\phi_2(x))$$

in terms of which the free part of the lagrangian becomes

$$\mathcal{L}_0 = \partial_\mu\phi^*\partial_\mu\phi + m^2\phi^*\phi$$

An infinitesimal roataion of the 2-vector (ϕ_1, ϕ_2) by an angle ϵ becomes a phase rotation on ϕ and the opposite phase roatation on ϕ^*:

$$\begin{aligned}\delta\phi &= i\epsilon\phi \\ \delta\phi^* &= -i\epsilon\phi^*\end{aligned} \tag{4.43}$$

Referring to the general formula, eq.(4.34), we still have two fields and we have one value of a with $T^a = +i$ for ϕ and $T^a = -i$ for ϕ^*. Also

$$\begin{aligned}\frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi)} &= \partial_\mu\phi^* \\ \frac{\delta\mathcal{L}}{\delta(\partial_\mu\phi^*)} &= \partial_\mu\phi\end{aligned} \tag{4.44}$$

So finally:

$$j_\mu(x) = i(\partial_\mu\phi^*(x)\phi(x) - \partial_\mu\phi(x)\phi^*(x)) \tag{4.45}$$

This current has an important physical significance: it is the electromagnetic current to which the electromagnetic field couples as a result of $U(1)$ gauge invariance.

4.3.3 Ward identities

Conservation of the Nöther current (and existence of the conserved Nöther charges) depended on the equations of motion being satisfied. However, in the path integral we integrate over all fields most of which do not satisfy the classical equations of motion. So do we have an analogous result? Well we may expect in fact that the *expectation value* of the Nöther current is conserved, since we have seen under our discussion of the Dyson-Schwinger equations that the classical equations of motion do hold as quantum averages. Also, invariance of the lagrangian density under the symmetry transformation does not require the classical equations of motion so it looks like we are in buissness. In fact, we shall derive Nöther's Theorem for a quantum average, but we shall emphasize one more important point which has to be fullfilled.

Let

$$\begin{aligned} \phi^i(x)' &\equiv \phi^i(x) + \delta\phi^i(x) \\ &= \phi^i(x) + i\alpha_{ij}(x)\phi^j(x) \end{aligned} \tag{4.46}$$

represent a *local* change of variable in the path integral. It is evident that we have

$$\int \mathcal{D}(\phi^i)' e^{-S[\phi']} = \int \mathcal{D}(\phi^i) e^{-S[\phi]}$$

since the result cannot depend on the name of the integration variable.

We now make the crucial *assumption* that the Jacobian to change variable from $\mathcal{D}(\phi^i)'$ to $\mathcal{D}(\phi^i)$ is 1. This assumption is nontrivial since we are dealing with complicated functional integrals the meaning of which should really be carefully examined as a result of introducing a cut-off and then considering some limiting procedure. In some cases, one can proove that indeed the Jacobian really is 1. But in other cases this is not true and one runs into the phenomenon of *quantum anomalies*. These occur in other words, when it looks like the theory has an invariance because the action has that invariance, but when it turns out that the path integral measure does *not* have that invariance. In this chapter, however, we shall not discuss anomalies, so we proceed assuming we are dealing with a situation where the integration measure is also invariant:

$$\mathcal{D}(\phi^i)' = \mathcal{D}\phi^i$$

We then find

$$\begin{aligned} 0 &= \int \mathcal{D}\phi^i [e^{-S[\phi']} - e^{-S[\phi]}] \\ &= \int \mathcal{D}\phi^i \delta e^{-S[\phi]} = -\int \mathcal{D}\phi^i \delta S[\phi] e^{-S[\phi]} \\ &= i\int d^4x \epsilon^a(x) \int \mathcal{D}\phi^i e^{-S[\phi]} \partial_\mu j^a_\mu(x) \end{aligned} \tag{4.47}$$

From this it follows that

$$\int \mathcal{D}\phi^i e^{-S[\phi]} \partial_\mu j^a_\mu(x) = 0 \tag{4.48}$$

or indeed that the Nöther current is conserved in the sense of a quantum average.

We may consider something a little more general, namely an arbitrary quantum field theoretic "correlator", an amplitude which is the quantum expectation value of some operators. Let us denote this operator generically as $A(\{\phi^i\})$ and write for the change under a local symmetry transformation:

$$\delta A(\{\phi^i\}) = \int d^4x \frac{\delta A(\{\phi^i\})}{\delta \epsilon^a(x)} \epsilon^a(x)$$

Then in a completely analogous way we find

$$\begin{aligned} 0 &= \int \mathcal{D}\phi^i \delta\{e^{-S[\phi]}A(\{\phi^i\})\} \\ &= i\int d^4x\epsilon^a(x)\int \mathcal{D}\phi^i e^{-S[\phi]}\{\partial_\mu j^a(x) - i\frac{\delta A(\{\phi^i\})}{\delta\epsilon^a(x)}\} \end{aligned} \tag{4.49}$$

This may also be written

$$\int \mathcal{D}\phi^i e^{-S[\phi]}\partial_\mu j^a(x) = i\int \mathcal{D}\phi^i e^{-S[\phi]}\frac{\delta A(\{\phi^i\})}{\delta\epsilon^a(x)} \tag{4.50}$$

which is a common way of expressing the *Ward identity.*

Ward identities may appear in many other forms. Let us just give one example. Using a generating functional with a "current term",

$$\int d^4x J_i(x)\phi^i(x) \equiv J\cdot\phi$$

(where the external current should not be confused with the Nöther current), we may apply the same idea and write

$$\begin{aligned} 0 &= \int \mathcal{D}\phi^i \frac{\delta}{\delta\epsilon^a(x)}e^{-S+J\cdot\phi} \\ &= i\int \mathcal{D}\phi^i\{-\frac{\delta S}{\delta\phi^i(x)}T^a_{ij}\phi^j(x) + J_iT^a_{ij}\phi^j(x)\}e^{-S+J\cdot\phi} \\ &= iT^a_{ij}\{-\frac{\delta S}{\delta\phi^i(x)}\left[\frac{\delta}{\delta J}\right] + J_i(x)\}\frac{\delta}{\delta J_j(x)}Z[J] \end{aligned} \tag{4.51}$$

Here we used that the quantum expectation value of any function of the ϕ^i's may be obtained as usual by hitting the generating functional by the same funciton of the derivatives after the external currents.

By further taking functional derivatives after the current, and finally putting the current equal to zero, we generate an infinity of Ward identities between Green functions.

Ward identities play a fundamental role in discussions of symmetry in quantum field theory, in connection with discussions of spontaneous symmetry breaking, in connection with proofs of keeping gauge invariance during renormalization etc.

Chapter 5

Regularization and Renormalization

5.1 Feynman diagrams in momentum space

We have already seen how the diagrammatic structure of perturbation theory leads to the evaluation of momentum integral when we want to calculate connected Green functions.

Let us recall the rules (section 2.7 in Bosonic Field Theory) for the connected momentum space Green functions for a scalar theory with the following lagrangian (in d-dimensional Euclidean space):

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2} m^2 \phi^2 + \sum_{n=3}^{N} \frac{\lambda_n}{n!} \phi^n \tag{5.1}$$

For a given connected diagram consisting of E external lines, I internal lines and V vertices we should associate

(A): to each external line i an external momentum p_i and a propagator

$$\frac{1}{p_i^2 + m^2}, \tag{5.2}$$

(B): to each internal line j an internal momentum q_j and a propagator

$$\frac{1}{q_j^2 + m^2}, \tag{5.3}$$

(C): to each vertex v_n of order n a factor

$$-\lambda_n (2\pi)^d \, \delta^d(\sum_{i \ni v} \varepsilon_{iv} p_i + \sum_{j \ni v} \varepsilon_{jv} q_j) \tag{5.4}$$

where the sign convention is that $\varepsilon_{jv} = 1$ if the momentum arrow of line j points away from vertex v and $\varepsilon_{jv} = -1$ if it point towards vertex v.

The diagram D will give a contribution $G_D(p_1, \ldots, p_E)$ to the connected Green function $G_C(p_1, \ldots, p_E)$ which can be written as

$$G_D(p_1, \ldots, p_E) = w_D \prod_{k=1}^{V} (-\lambda_{n_k}) \prod_{i=1}^{E} \frac{1}{p_i^2 + m^2} \cdot \tilde{F}_D(p_1, \ldots, p_E) \tag{5.5}$$

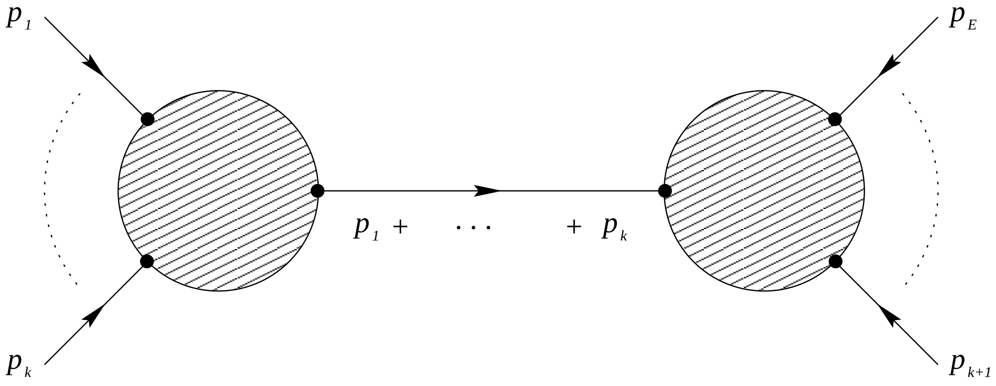


Figure 5.1: Splitting of a diagram in 1PI components

where w_D is a statistical weight factor (see section 2.7 of Bosonic Field Theory), n_k denotes the order of vertex v_k, λ_{n_k} the corresponding coupling constant. $\tilde{F}_D$ is given by

$$\tilde{F}_D(p_1, \ldots, p_E) = \int \prod_{j=1}^{I} \frac{dq_j}{(2\pi)^d} \frac{1}{q_j^2 + m^2} \quad \prod_{v=1}^{V} (2\pi)^d \delta^{(d)} \left(\sum_{i \ni v} \varepsilon_{iv} p_i + \sum_{j \ni v} \varepsilon_{vj} q_j \right) \tag{5.6}$$

We will now introduce the concept of a *one-particle irreducible* (1PI) diagram . It is a diagram which cannot be separated in two disconnected parts by cutting a single line. Such diagrams are the natural objects to analyze since the integral (5.6) splits in independent integrals for each 1PI component, the reason being that lines separating the 1PI components by momentum conservation at each vertex can be expressed entirely in terms of the external momenta. This is illustrated in fig.1.

The connected Green functions in Fourier space are then trivially built from the general diagrammatic rules considered earlier by taking products of propagators and the contributions from the 1PI diagrams. These we will denote $\tilde{I}_D(p_1, \ldots, p_E)$ and we will further extract the allover momentum conservation

$$\begin{aligned} \tilde{I}_D(p_1, \ldots, p_E) &= (2\pi)^d \; \delta^{(d)}(p_1 + \cdots + p_E) \; I_D(p_1, \ldots, p_E) && (5.7) \\ &= \int \prod_{j=1}^{I} \frac{d^d q_j}{(2\pi)^d (q_j^2 + m^2)} \prod_{v=1}^{V} (2\pi)^d \delta^{(d)} \left(\sum_{i \ni v} \varepsilon_{vi} p_i + \sum_{j \ni v} \varepsilon_{vj} q_j \right) && (5.8) \end{aligned}$$

As was shown in detail in chapter 4, it is often convenient to use instead of the connected Green functions G_C the 1PI Green functions $\tilde{\Gamma}$ and Γ, the only difference between $\tilde{\Gamma}$ and Γ being a δ-function for allover momentum conservation. The 1PI Green function Γ will be the sum of Γ_D's from the various diagrams D and each of the diagrams gives a contribution which can be written much like (5.5):

$$\Gamma_D(p_1, \ldots, p_E) = w_D \prod_{k=1}^{V} (-\lambda_{n_k}) \; \cdot \; I_D(p_1, \ldots, p_E) \tag{5.9}$$

In the next sections we shall mainly be discussing the integral $I_D(p_1, \ldots, p_E)$ since we will try to understand the structure of the divergences which are present when we perform the momentum integrals. Only after that we will turn to renormalization of the theory (and try to get rid of the divergences in a systematic way) and then the natural objects to consider will be the full 1PI Green functions $\Gamma(p_1,\ldots,p_E)$ and their diagrammatic expansion in terms of $\Gamma_D(p_1,\ldots,p_E)$.

We now have to perform the momentum integration for the integral I_D. V of these can be performed by elimination of the δ-functions. As already remarked in chapter 2 this leave us with L integration variables, where L is

$$L = I - V + 1. \tag{5.10}$$

The "1" in (5.10) is the total momentum conservation factored out in (5.7). Only V-1 of the δ-functions reduce the number of q_j integrations. L can be viewed as the number of *independent* loops in the diagram. Let us for a general 1PI graph denote by $q_\alpha, \alpha = 1,\ldots,L$ a set of internal integration momenta which survives the elimination of δ-functions in (5.7). The q_j's will then be linear combinations of the q_α's and the external momenta p_i:

$$q_j = \sum_{\alpha=1}^{L} c_{j\alpha}\, q_\alpha + \sum_{i=1}^{E} c_{ji}\, p_i \tag{5.11}$$

To find the relation (5.11) one has to eliminate the δ-functions explicitly for each graph. It is possible to develop a general formalism based on notation from electric current theory (q_j's are the currents), but it is not necessary for the lowest order graphs we consider here. With this notation I_D can be written as follows

$$I_D(p_1,\ldots p_E; m) = \int \prod_{\alpha=1}^{L} \frac{d^d q_\alpha}{(2\pi)^d} \prod_{j=1}^{I} \frac{1}{q_j^2 + m^2} \tag{5.12}$$

where it is understood that $q_j = q_j(p_i, q_\alpha)$ according to (5.11).

We can now finally address the question of convergence of the integral (5.12). It follows immediately from the form of the integral that *if* the integral is convergent it is a homogeneous function of degree $\omega(D)$ in p_i and m:

$$\begin{aligned} I_D(tp_i; tm) &= t^{\omega(D)}\, I_D(p_i, m) && (5.13)\\ \omega(D) &= dL - 2I. && (5.14) \end{aligned}$$

$\omega(D)$ is called the *superficial degree of divergence of the diagram D.*

Theorem 1: $\omega(D) < 0$ is necessary for convergence of I_D.

Proof:

From the inequality between arithmetic and geometric means we have

$$\prod_{i=1}^{I}(q_i^2 + m^2) \le \left(\sum_{i=1}^{I}(q_i^2 + m^2)\right)^I$$

We now use the representation (5.11). If q_α is sufficiently large (say $\ge \Lambda(p_i, m)$ where Λ might depend on m and p_i) we have:

$$\sum_{i=1}^{I}(q_i^2 + m^2) = \sum_{i=1}^{I} \left[(\sum_\alpha c_{i\alpha} q_\alpha + \sum_j c_{ij} p_j)^2 + m^2 \right] \le \mathrm{C} \sum_{\alpha=1}^{L} q_\alpha^2$$

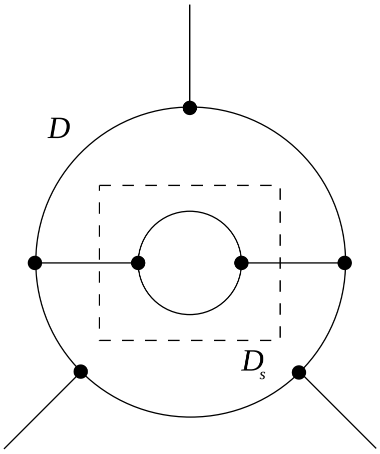


Figure 5.2: $\omega(D) = -6$, however $\omega(D_S) = 0$.

where C is some new constant. This leads to the following estimate:

$$I_D > \frac{1}{C^I}\int_{\sum_\alpha q_\alpha^2 > \Lambda^2} \prod_{\alpha=1}^{L} \frac{d^d q_\alpha}{(2\pi)^d} \frac{1}{(\sum_{\alpha=1}^{L} q_\alpha^2)^I} > \frac{1}{\tilde{C}}\int_{r>\Lambda} \frac{r^{dL-1}dr}{r^{2I}}$$

where again $\tilde{C}$ is some constant. This completes the proof.

It should be clear that the above consideration is nothing but simple power counting. In general $\omega(D) < 0$ is not sufficient for convergence. Subdiagrams "D_S" can have $\omega(D_S) > 0$ and they will lead to divergence of I_D. This follows heuristically by keeping the momenta coming from the rest of the integration fixed and doing the subdiagram integration first. The situation is illustrated in fig.2. If we consider the situation in four dimensions ($d = 4$) we have:

$$L_D = 3, \quad I_D = 9, \quad \omega(D) = d\,L_D - 2I_D = -6$$

However, the shown subdiagram has

$$L_{D_S} = 1, \quad I_{D_S} = 2, \quad \omega(D_S) = d\,L_{D_S} - 2I_{D_S} = 0$$

and is logarithmic divergent.

Definition: By a 1PI subdiagram D_S of D we understand a set of vertices and *all* lines joining them, such that D_S makes up an 1PI diagram.

Some of the internal lines in D can play the role as external lines for D_S as illustrated in fig.2 where D_S consists of two vertices and the four lines joining them. Two of these lines are external for D_S. Let $\mathcal{F}(D)$ denote the set of all 1PI subdiagrams of D (notice that $D \in \mathcal{F}(D)$).

Theorem 2: $\omega(D_S) < 0 \; \forall D_S \in \mathcal{F}(D) \Leftrightarrow I_D(p_1, \ldots, p_E)$ is an absolute convergent integral.

The proof will not be given here. The following should be noted: The "$\Rightarrow$" part of the theorem is true for any field theory. However, if we have spin 1/2 and spin 1 particles,

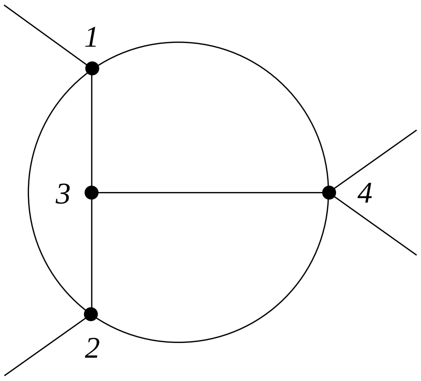


Figure 5.3: $n_1 = 4,\ n_2 = 4,\ n_3 = 3,\ n_4 = 5.$

$\omega(D) < 0$ is not necessary for convergence because the propagators are no longer positive definite functions. There might be cancellation between different terms in the numerators of these propagators.

5.2 Definition of renormalizable theories

We want to express $\omega(D)$ in a more transparent way. Let n_v denote the order of a vertex in a diagram D (see Fig. 3 for an example).

Since each internal line connects to two vertices and each external line to one vertex we have for any diagram:

$$2I + E = \sum_{v=1}^{V} n_v \tag{5.15}$$

If we combine this with the relation (5.10): $L = I - V + 1$ we get

$$\omega(D) \equiv dL - 2I = d - \frac{d-2}{2}E + \sum_{v=1}^{V}(\frac{d-2}{2}n_v - d) \tag{5.16}$$

Let us define ω_v as:

$$\omega_v = (d-2)n_v/2 \tag{5.17}$$

The interpretation of ω_v is as follows: If a vertex of order n_v appears in the diagram D we must have a term $\lambda_{n_v}\phi^{n_v}$ in the lagrangian:

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + \cdots + \frac{\lambda_{n_v}}{n!}\phi^{n_v} + \ldots \tag{5.18}$$

As $S \equiv \int d^dx\mathcal{L}(\phi)$ is dimensionless (if $\hbar = 1$) the canonical dimension of ϕ is (from the kinetic term)

$$[\phi] = (d-2)/2 \tag{5.19}$$

and we have

$$[\phi^{n_v}] = (d-2)n_v/2 = \omega_v \ ; \quad [\lambda_{n_v}] = d - \omega_v \tag{5.20}$$

and $\omega(D)$ can now be written as:

$$\begin{aligned}
\omega(D) &= d - \frac{d-2}{2}E + \sum_{v=1}^{V}(\omega_v - d) \\
&= d - \frac{d-2}{2}E - \sum_{v=1}^{V}[\lambda_v]
\end{aligned} \tag{5.21}$$

It is natural to *classify* the 1PI diagrams corresponding to a certain lagrangian according to the number of external lines, since these are precisely what associate them with the connected Green functions. In order to limit the *classes* of D where $\omega(D) \geq 0$ we *must* have

$$\boxed{\omega_v \leq d \quad \text{or} \quad [\lambda_v] \geq 0}. \tag{5.22}$$

If not, we can make arbitrary complicated diagrams with a given number of external lines E, but where $-\sum_v [\lambda_v]$ and therefore $\omega(D)$ is large positive.

This leads to the following classification of scalar field theories:

(A): *Non-renormalizable theories:* Contain at least one coupling constant with $[\lambda_v] < 0$. To all Green functions are associated superficially divergent diagrams.

(B): *Renormalizable theories:* Some $[\lambda_v] = 0$, the rest have $[\lambda_v] > 0$. Only a finite number of 1PI Green functions have superficially divergent diagrams (But the diagrams themselves can be arbitrary complicated).

(C): *Super-renormalizable theories:* All $[\lambda_v] > 0$. The superficial divergence of diagrams corresponding to a given class of 1PI diagrams (i.e E fixed) decreases with the number of vertices. Only a finite number of *diagrams* are superficially divergent.

The definition extends in a trivial way to theories which involve also spinor fields and vector fields. We leave the generalization as an exercise for the ambitious reader.

We shall see that it is possible to make sense of the renormalizable and super-renormalizable theories, even if there are infinities in the perturbation expansion. The infinities can in these cases be absorbed in a redefinition of the various coupling constants, as we shall discuss later. The important point is that only a finite number of Green functions have superficial divergences in these two cases, and the redefinition of the coupling constants will be sufficient to remove in a systematic way all divergences from the Green functions. If the theory is non-renormalizable there will superficial divergences associated with Green functions with an arbitrary number of external lines, and a redefinition of existing coupling constants in the lagrangian will not be sufficient to remove the divergences. If one should remove the superficial divergence of a n-point function one would have to introduce a new coupling constant and the corresponding new interaction $\lambda_n \phi^n / n!$ in the lagrangian (the details of this will be clear later). In this way one ends up with arbitrary polynomial interactions and the original theory is modified in a drastic and ill defined manner. It should be emphasized that this inability to make sense of non-renormalizable theories is linked to the perturbative expansion. In principle one could imagine the possibility of defining non-renormalizable theories in a non-perturbative way. We will discuss this later, and only note here that at present no interesting non-renormalizable theories have yet been defined.

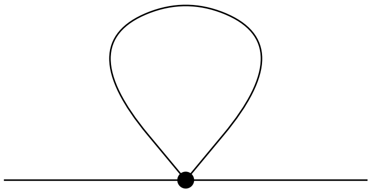


Figure 5.4: A tadpole diagram.

Example 1: Classification of renormalizable scalar field theories.

We are interested in finding the highest power N of ϕ which makes the lagrangian

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2}m^2\ \phi^2 + \sum_{n=3}^{N} \frac{\lambda_n}{n!}\phi^n \tag{5.23}$$

renormalizable, as a function of the dimension d of space-time. From (5.20) we have

$$[\lambda_n] \geq 0 \Rightarrow n \leq \frac{2d}{d-2} \tag{5.24}$$

d=2: All n are allowed: All polynomial interactions are (super-) renormalizable. One can even define renormalizable non-polynomial interactions.

d=3: $n \leq 6$. ϕ^3, ϕ^4, ϕ^5 are super-renormalizable and ϕ^6 is renormalizable.

d=4: $n \leq 4$. ϕ^3, ϕ^4 are allowed. ϕ^3 is super-renormalizable and ϕ^4 is renormalizable.

d=5: $n \leq 3$. Only ϕ^3 interactions are allowed. The interaction is super-renormalizable.

d=6: $n \leq 3$. The ϕ^3 interaction is renormalizable.

d>6: No interactions are allowed. Only a gaussian, i.e. a free field theory satisfies (5.24).

From the example it follows that only for $d \leq 4$ can we have renormalizable (or super renormalizable) scalar theories which have a chance of being non trivial for $d \leq 4$. From the point of view of perturbation theory ϕ^3 is renormalizable in $d = 6$ and we will sometimes use it for the purpose of illustration. However, clearly a ϕ^3 theory alone makes non sense in the functional integral, since ϕ^3 is not bounded from below.

Example 2: Superficially divergent 1PI diagrams in a ϕ^4 theory.

d=2: From (5.20) and (5.21) we have:

$$\omega_v = 0, \quad [\lambda_4] = 2, \quad \omega(D) = 2 - 2V$$

and only for $V = 1$ will we have a superficially divergent diagram ($\omega(D) \geq 0$) as shown in the fig. 4 (a *tadpole* diagram) Of course there will be many other diagrams where the tadpole diagram appears as a subdiagram as illustrated in fig.5 These diagrams are divergent, but have no superficial divergence, and once the divergence is removed from the tadpole diagram they will be convergent.

d=3: From (5.20) and (5.21) we have:

$$\omega_v = 2, \qquad [\lambda_4] = 1.$$

$$\omega(D) = 3 - E/2 - \sum_{v=1}^{V} = 3 - V - E/2$$

This leave us with only two 1PI superficial divergent diagrams corresponding to $V = 1, E = 2$ and $V = 2, E = 2$ as shown in fig.6.

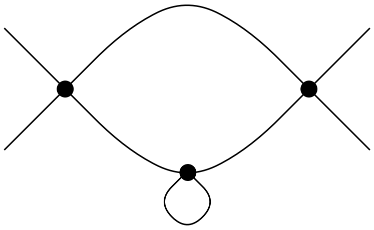


Figure 5.5: A diagram with a tadpole subdiagram.

Figure 5.6: The divergent ϕ^4 diagrams in $d = 3$.

d=4: From (5.20) and (5.21) we have:

$$\omega_v = 4, \qquad [\lambda_4] = 0, \qquad \omega(D) = 4 - E.$$

The four-dimensional theory is renormalizable, but not super renormalizable as the above examples in $d = 2, 3$. This means that infinite many 1PI diagrams are superficially divergent, in fact this divergence only depends on E. This is shown in fig.7 where a "blob" symbolizes an arbitrary complicated 1PI graph. Only diagrams with $E = 2$ and $E = 4$ can occur.

5.3 Regularization

As seen in the last sections we will in general encounter divergent integrals in our perturbative expansions. The volume of the momentum space is in a certain sense too large. The integrands (i.e. the product of propagators) are falling off at infinity, but too slowly to ensure convergence of the integrals. The philosophy for dealing with these divergences is as follows:

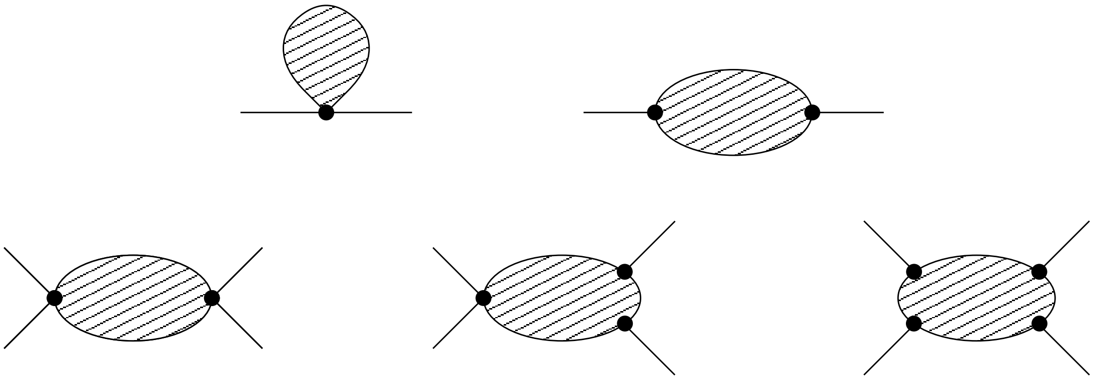


Figure 5.7: The structure of superficially divergent ϕ^4 diagrams in $d = 4$.

(A) : Regularize the divergent integrals in some way, that is: Introduce some kind of cut off Λ which makes them finite. This allows us to calculate the integrals in a meaningful way even if $\omega(D) > 0$, but in general our answer will now depend on Λ.

(B) : Show that the parts of the integrals that are divergent when the cut off is removed again can be absorbed in a redefinition of the coupling constants of the lagrangian.

In this section we discuss point (A), in the next we turn to point (B). The most naive choice of regularization is simply to introduce an upper integration limit: $|q| < \Lambda$ in all momentum integrals. This is more or less equivalent to putting the system on a lattice since it cuts off more or less any distance $x < 1/\Lambda$, and like a lattice approach the cut-off breaks explicitly euclidean invariance. We will discuss the lattice regularization in detail later, and here concentrate on regularizations which preserve euclidean invariance. Such regularizations are needed if we want to perform perturbation expansions "by hand", simply because the calculations otherwise become too cumbersome. However, it should be stressed that in principle any "sensible" regularization should be allowed. The usual way to regularize the integrals, while keeping euclidean invariance, is by modifying the momentum propagator by hand:

$$\frac{1}{q^2+m^2} \rightarrow \frac{1}{q^2+m^2+q^4/\Lambda^2} \tag{5.25}$$

$$\frac{1}{q^2+m^2} \rightarrow \frac{1}{q^2+m^2} - \frac{1}{q^2+\Lambda^2} = \frac{\Lambda^2-m^2}{(q^2+m^2)(q^2+\Lambda^2)} \tag{5.26}$$

The first modified propagator corresponds to the introduction of higher derivative terms like $(\partial^2_\mu\phi)^2/\Lambda^2$ in the Lagrangian, while the second modified propagator corresponds to *subtracting* the corresponding propagation of a heavy particle. Note that both modifications of the action only affect the gaussian part. They both lead to euclidean invariant theories where the propagators are better behaved as $q \to \infty$. Of course the propagators can be made arbitrary convergent as $|q| \to \infty$ if wanted, by for instance:

$$\frac{1}{q^2+m^2} \rightarrow \frac{1}{q^2+m^2} - \frac{\sum_{i=1}^N C_i(\Lambda,m)}{q^2+\Lambda^2} \qquad \text{(Pauli Villars regularization)} \tag{5.27}$$

where the constants C_i are adjusted such that the behaviour of the propagator for large $|q|$ will be as Λ^{2N}/q^{2N+2}. We should stress that the modification of the propagators serve the same purpose as the explicit cut-off $|q| < \Lambda$, in the sense that it limits the "volume" of momentum space from where we get the important contributions to our integrals. It is just a somewhat "smoother" cut-off than $|q| < \Lambda$.

However, we might want the regularization to respect more than just the euclidean invariance. In some situations it will be highly desirable if internal symmetries of the lagrangian are compatible with the regularization procedure since that might limit the possible divergences which appear when performing the momentum integration. Internal symmetries will be discussed in detail later, but let us mention that the symmetry we have in mind is local (non-abelian) gauge symmetry. If we want to preserve both euclidean invariance and non-abelian internal symmetries, not many regularizations are available. In fact essentially only one: *dimensional regularization.* As this method is actually quite convenient for higher order calculations too, we will describe it here although it is not so intuitive as the momentum or lattice cut off.

The important observations are simply:

(A) : $\omega(D)$ decreases with dimension d (convergence improves when going to lower dimensions.

(B) : The integrals $\Gamma_D(p_1,\ldots,p_n)$ can be expressed as analytic functions of d away from the integer dimensions $d \geq 1$.

In a certain way one can say that the idea of dimensional regularization is the same as for the other cut-off's: to reduce the volume of momentum space. Intuitively the decrease of volume with dimension is clear and it increases the convergence of certain type of integrals if we decrease dimensions. The one-dimensional integral

$$\int \frac{dq_1}{q_1^2+m^2}$$

is convergent while the corresponding two dimensional integral

$$\int \frac{dq_1 dq_2}{q_1^2+q_2^2+m^2}$$

is divergent, as are all higher dimensional integrals of the same kind. Of course it depends on the kind of functions $f(q)$ we want to integrate, whether the convergence will improve with decreasing dimension d, since $f(q)$ in general will have some dependence of d. However, it is typically rational functions in scalar product of integration momenta and external momenta we want to integrate. They have a natural extension to any positive integer dimension without changing their fall off at infinity as the above example illustrates, while on the other hand the measure of integration will change (and "improve") for lower dimensions since

$$d^d q \sim |q|^{d-1} d|q|$$

We introduce the method by a few examples.

Example 3: The Γ-function

The Γ function can be defined by the following integral representation

$$\Gamma(z) = \int_0^\infty d\alpha\, \alpha^{z-1}\, e^{-\alpha} \tag{5.28}$$

and we have when z is a positive integer: $\Gamma(z) = (z-1)!$. The integral representation (5.28) obviously defines an analytic function $\Gamma(z)$ in the half plane Re $z > 0$. By partial integration it is possible to set

$$\Gamma(z) = \frac{1}{z}\int_0^\infty d\alpha\, \alpha^{z}\, e^{-\alpha} \tag{5.29}$$

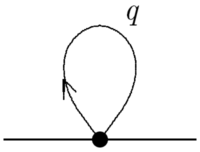
thereby extending the definition to Re $z > -1$ and proving the functional relation $z\Gamma(z) = \Gamma(z+1)$ also in this region. (5.29) also shows that $\Gamma(z)$ has a simple pole at $z = 0$. This procedure can now be repeated, and the result is that $\Gamma(z)$ can be defined in the whole complex plane except at $z = -n, n = 0, 1, \ldots$ where it has simple poles. The pole structure can be analyzed further by introducing the *digamma function* $\psi(z)$

$$\begin{aligned}
\psi(z) &\equiv \frac{1}{\Gamma(z)} \frac{d\,\Gamma(z)}{dz} \\
\psi(n+1) &= 1 + \frac{1}{2} + \ldots + \frac{1}{n} - \gamma \\
\psi'(n+1) &= \frac{\pi^2}{6} - \sum_{k=1}^{n} \frac{1}{k^2}
\end{aligned}$$

where $\gamma = 0.57721\cdots$ denotes Eulers constant. Near $z = -n$ we have the expansion:

$$\Gamma(-n+\varepsilon) = \frac{(-1)^n}{n!} \left[\frac{1}{\varepsilon} + \psi(n+1) + \frac{1}{2}\varepsilon[\frac{\pi^2}{3} + \psi^2(n+1) - \psi'(n+1)] + \mathcal{O}(\varepsilon^2) \right] \tag{5.30}$$

Example 4: The tadpole diagram



$$I = \int \frac{d^d q}{(2\pi)^d} \frac{1}{q^2 + m^2} \tag{5.31}$$

The integral is convergent for $d = 1$. Formally we can calculate it in any d as follows:

$$\frac{1}{q^2 + m^2} = \int_0^\infty d\alpha\; e^{-\alpha(q^2+m^2)} \tag{5.32}$$

We note that for positive integer d we have

$$\int \frac{d^d q}{(2\pi)^d}\, e^{-\alpha q^2} = (4\pi\alpha)^{-d/2}\,. \tag{5.33}$$

and we promote this relation to be valid for all complex d. In fact it follows from general theorems that an analytic function coinciding with (5.33) for positive integer d is unique if we require it to grow slower that a certain exponential function in the complex plan.

By using the definition of the Γ function given in the previous example and making formal interchange of integrations (valid in $d = 1$ where I is well defined) we get

$$\begin{aligned}
\int \frac{d^d q}{(2\pi)^d} \frac{1}{q^2 + m^2} &= \int \frac{d^d q}{(2\pi)^d} \int d\alpha\; e^{-\alpha(q^2+m^2)} = \int d\alpha\; e^{-\alpha m^2} \int \frac{d^d q}{(2\pi)^d}\, e^{-\alpha q^2} \\
&= \int d\alpha\; e^{-\alpha m^2}\, \alpha^{-d/2}/(4\pi)^{d/2} = \frac{\Gamma(1-d/2)(m^2)^{d/2-1}}{(4\pi)^{d/2}}
\end{aligned}$$

Differentiation after m^2 we get

$$\int \frac{d^d q}{(2\pi)^d} \frac{1}{(q^2+m^2)^n} = \frac{\Gamma(n-d/2)}{(4\pi)^n \Gamma(n)} \left(\frac{m^2}{4\pi}\right)^{d/2-n} \tag{5.34}$$

For a given (integer) d the integral in (5.34) will be divergent for $n < d/2$. *We see that the onset of divergence according to naive power counting is associated with a pole in the* Γ*-function* $\Gamma(n - d/2)$ *in the dimensional regularized integral.*

Example 5: The self-energy diagram

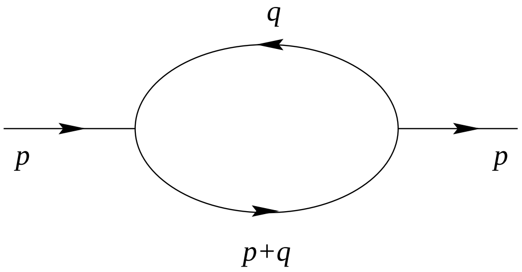


Figure 5.8: Momentum assignment for the self-energy diagram.

The diagram is shown in fig.8. According to the momentum assignment shown in the figure we have the following integral

$$I(p) = \int \frac{d^d q}{(2\pi)^d} \frac{1}{q^2+m^2} \frac{1}{(q+p)^2+m^2} \tag{5.35}$$

We can now proceed as for the tadpole. If we denote the two quadratic forms in p and q by Δ_1 and Δ_2 we have

$$\frac{1}{\Delta_1}\frac{1}{\Delta_2} = \int_0^\infty d\alpha_1 d\alpha_2 \; e^{-(\alpha_1\Delta_1+\alpha_2\Delta_2)}$$

By a change in variables $\alpha_1 = t(1-\alpha), \alpha_2 = t\alpha$ where $\alpha \in [0,1]$ the double integral can be written as

$$\int_0^1 d\alpha \int_0^\infty dt \; t \, e^{-((1-\alpha)\Delta_1+\alpha\Delta_2)t} = \int_0^1 d\alpha \frac{1}{[(1-\alpha)\Delta_1+\alpha\Delta_2]^2}$$

If we shift q by $q = \tilde{q} - \alpha p$ we have

$$(1-\alpha)\Delta_1 + \alpha\Delta_2 = q^2 + 2\alpha qp + \alpha p^2 + m^2 = \tilde{q}^2 + [\alpha(1-\alpha)p^2 + m^2]$$

and our integral (5.35) can finally be written as

$$I(p) = \int_0^1 d\alpha \int \frac{d^d\tilde{q}}{(2\pi)^d} \frac{1}{(\tilde{q}^2 + [\alpha(1-\alpha)p^2+m^2])^2} \tag{5.36}$$

The $\tilde{q}$ integral in (5.36) is precisely of the form (5.34) and we can use this result to write

$$I(p) = \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}} \int_0^1 d\alpha \; [\alpha(1-\alpha)p^2 + m^2]^{d/2-2} \tag{5.37}$$

We see again that *the dimensional regularized integral will have a pole of a* Γ*-function when* $d \to 4$*, the lower critical dimension beyond which the integral (5.35) will be divergent according to naive power counting.*

Example 6: General one-loop integrals and Feynman parameters

A possible momentum assignment to the general one-loop diagram is shown in fig.9. The integral corresponding to the Feynman diagram of fig.10 is:

$$I(p_1,\ldots,p_n) = \int \frac{d^d q_1}{(2\pi)^d} \prod_{i=1}^n \frac{1}{q_i^2+m^2} = \int \frac{d^d q}{(2\pi)^d} \prod_{i=1}^n \frac{1}{(q+\tilde{p}_i)^2+m^2} \tag{5.38}$$

We can now transform the q integration to a α-parametric integral the same way as was done in the last two examples. Again the purpose is to exhibit the d dependence more

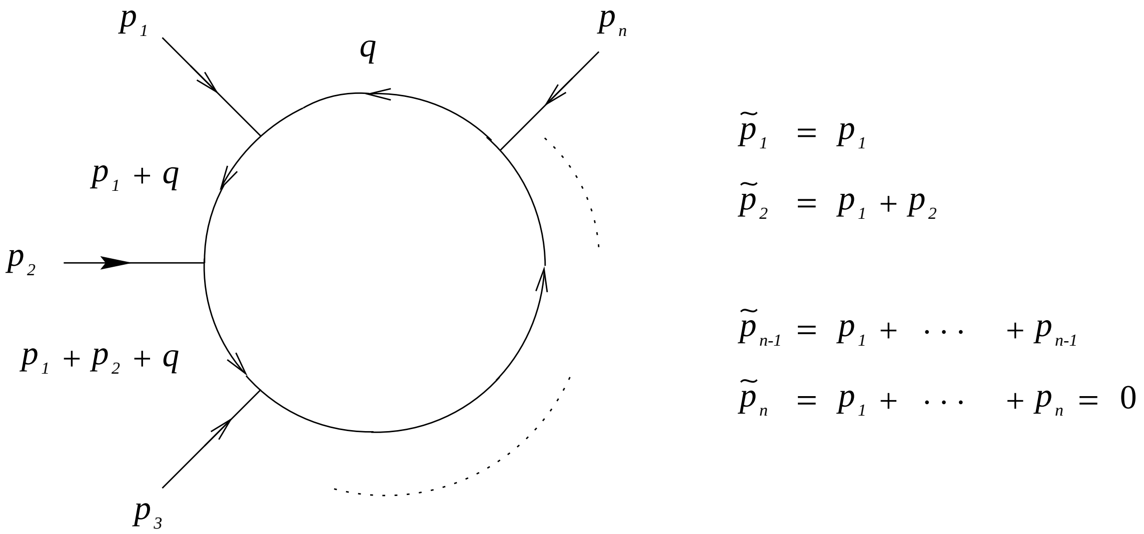


Figure 5.9: Momentum assignment for a general one-loop diagram.

explicitly. The steps are the same: First we represent the product of propagators by the exponential integral. In this way we get a *sum* of terms in the exponent. Next we perform one of the parametric integrals by a change of variables $\tilde{\alpha}_i = t\alpha_i$ where $\sum \alpha_i = 1$:

$$\begin{aligned}
\frac{1}{\Delta_1 \cdots \Delta_n} &= \int_0^\infty \prod_{i=1}^n d\tilde{\alpha}_i e^{-(\tilde{\alpha}_1 \Delta_1 + \cdots + \tilde{\alpha}_n \Delta_n)} \\
&= \int \prod_{i=1}^n d\alpha_i \; \delta(1 - \sum \alpha_i) \int_0^\infty dt \; t^{n-1} e^{-t(\alpha_1 \Delta_1 + \cdots + \alpha_n \Delta_n)} \\
&= \int \prod_{i=1}^n d\alpha_i \; \delta(1 - \sum \alpha_i) \frac{\Gamma(n)}{(\alpha_1 \Delta_1 + \cdots + \alpha_n \Delta_n)^n}
\end{aligned}$$

We use this in (5.38) to get

$$I(p_1, \ldots, p_n) = \int \prod_{i=1}^n d\alpha_i \; \delta(1 - \sum \alpha_i) \int \frac{d^d q}{(2\pi)^d} \frac{\Gamma(n)}{[\sum_{i=1}^n \alpha_i (q + \tilde{p}_i)^2 + m^2]^n}$$

The q integral is again of the form (5.34) since a simple shift $q = \tilde{q} - \sum \alpha_i \tilde{p}_i$ will make the denominator a function of $\tilde{q}^2$:

$$\begin{aligned}
\sum_{i=1}^n \alpha_i (q + \tilde{p}_i)^2 + m^2 &= \tilde{q}^2 + m^2 + F(\alpha_i, \tilde{p}_i) \\
F(\alpha_i, \tilde{p}_i) &= \sum \alpha_i \tilde{p}_i^2 - (\sum \alpha_i \tilde{p}_i)^2
\end{aligned}$$

The $\tilde{q}$-integral can now be performed and (5.38) can be written as the following parametric integral:

$$I(p_1, \ldots, p_n; m) = \frac{\Gamma(n - d/2)}{(4\pi)^{d/2}} \int \prod^n d\alpha_i \; \delta(1 - \sum^n \alpha_i)[m^2 + F(\alpha_i, \tilde{p}_i)]^{d/2-n} \tag{5.39}$$

The α_i parameters are called Feynman parameters.

For $n > d/2$ (5.38) is actually convergent and no analytic continuation is needed to get (5.39) provided d is an positive integer. The special case of $n = 2$ of course agrees with the self-energy calculation in the former example.

5.4 One-loop renormalization

In the last section we saw that the divergences in the perturbation expansion in four dimensions showed up as poles in the dimensional regularized integrals. Had we chosen another regularization, like for instance the Pauli-Villars regularization, power-like divergences would have appeared as powers of the cut-off Λ, while logarithmic divergences would manifest themselves as terms of the form $\log(\Lambda/m)$. It is the purpose of this section to show that the infinities one encounters to lowest order perturbation theory can be absorbed in a redefinition of the coupling constants. We will do this by considering explicitly the lowest order perturbation for two renormalizable theories: ϕ^4 in four dimensions and ϕ^3 in six dimensions. The reason for also considering the somewhat artificial ϕ^3-theory in six dimension is that the socalled wave function renormalization is absent in the lowest order ϕ^4-theory in four dimensions.

Since we have already dealt with the general l-loop diagram in the last section we have all the machinery needed. In both cases the superficial divergence $\omega(D)$ of a 1PI diagram D depends only on the number of external lines. We have

$$\omega(D) = 4 - E \quad (d = 4), \qquad \omega(D) = 6 - 2E \quad (d = 6)$$

For ϕ^4 the 2- and 4-point function are superficial divergent, while for ϕ^3 and $d = 6$ the 2- and the 3-point functions are superficially divergent.

In both cases the coupling constants have dimensions when we leave the dimension where the theory is renormalizable: $d = 4$ *for* ϕ^4 *and* $d = 6$ *for* ϕ^3. A coupling $\sim \lambda\phi^n$ in d dimensions has the following mass dimension (recall (5.20))

$$[\lambda] = d - (d-2)n/2 \tag{5.40}$$

It is convenient to write λ away from the *critical* dimension (where$[\lambda] = 0$) as a *dimensionless* coupling times a parameter μ with the dimension of mass. We therefore introduce the notation $\tilde{\lambda}$ for the coupling constant λ away from the critical dimension and write

$$\tilde{\lambda} = \mu^{d-(d-2)n/2}\,\lambda$$

In the following we will be somewhat sloppy when we write expressions like $\lambda\phi^4$. According to the notation introduced we should only write this in $d = 4$, but sometimes we might use it instead of the correct expression $\tilde{\lambda}\phi^4$, in order to avoid too cumbersome a notation. We hope it will lead to no confusion.

We have now introduced a new mass parameter μ in the theory. At the moment the appearance of this mass parameter is mysterious, especially since one would not expect any physical mass or scattering cross section to depend on it. The existence of such a parameter might be less surprising by recalling that we could have used one of the other regularization procedures where a cut-off Λ with dimension of mass is explicitly introduced. Also in that case any physical observable better not depend on Λ even if we have used Λ in some intermediate steps of the calculation. We will later describe in detail what happens to the parameter μ, let us only mention here that it will enter in the so called renormalization group equations which describe scaling properties of the Green functions. In this way the true reason for the appearance of this mass parameter can be traced to necessity of breaking scale invariance when renormalizing relativistic field theories.

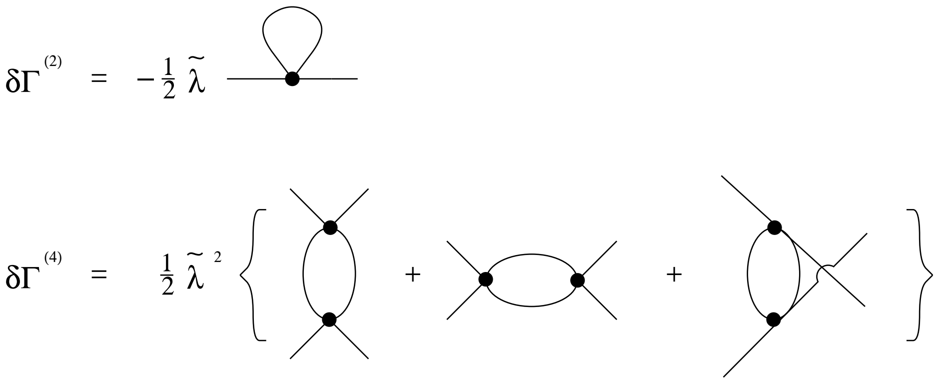


Figure 5.10: The lowest order 1PI diagrams for a ϕ^4 theory.

In the next two subsections we calculate the divergent 1-loop 1PI diagrams for ϕ^4- and ϕ^3-theories.

5.4.1 ϕ^4 in d=4

We introduce the notation

$$\varepsilon = 4 - d \qquad \tilde{\lambda} = \mu^{\varepsilon}\lambda$$

and consider the lowest order perturbative corrections to the 1PI Green functions. As explained above only the 2-point functions and the 4-point functions are superficially divergent. By using either Wick's theorem and the expansion of chapter 2 or the Dyson-Schwinger equations of chapter 4 we get the one loop corrections illustrated in fig.10. If we introduce the following notation for the momenta in $\Gamma^{(4)}(p_1,p_2,p_3,p_4)$:

$$s = (p_1+p_2)^2, \;\; t = (p_1+p_4)^2, \;\; u = (p_1+p_3)^2$$

we get for the one-loop contribution:

$$\begin{aligned} \delta\Gamma^{(2)}(p) &= -\frac{\lambda\Gamma(\frac{\varepsilon}{2}-1)}{2(4\pi)^2}\left(\frac{4\pi\mu^2}{m^2}\right)^{\frac{\varepsilon}{2}} m^2 \\ \delta\Gamma^{(4)}(s,t,u) &= \mu^{\varepsilon}\,\frac{\lambda^2\Gamma(\frac{\varepsilon}{2})}{2(4\pi)^2}\left(\frac{4\pi\mu^2}{m^2}\right)^{\frac{\varepsilon}{2}}\int_0^1 d\alpha \sum_{z=s,t,u}(1+\alpha(1-\alpha)\frac{z}{m^2})^{-\frac{\varepsilon}{2}} \end{aligned}$$

We can now expand to order $\varepsilon = 4-d$ as $d \to 4$ using the pole structure of $\Gamma(z)$ described in example 3:

$$\begin{aligned} \delta\Gamma^{(2)}(p) &= \frac{\lambda}{(4\pi)^2}\, m^2\left(\frac{1}{\varepsilon} + F^{(2)}(m^2,\mu^2,\varepsilon)\right) \\ \delta\Gamma^{(4)}(s,t,u) &= \mu^{\varepsilon}\frac{\lambda^2}{(4\pi)^2}\left(\frac{3}{\varepsilon} + F^{(4)}(s,t,u,m,\mu,\varepsilon)\right) \end{aligned}$$

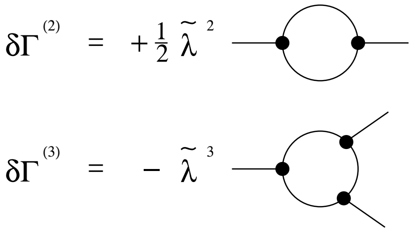


Figure 5.11: The lowest order 1PI Green functions for a ϕ^3 theory.

where the F functions are given by:

$$\begin{aligned} F^{(2)}(\,\cdot\,) &= -\frac{1}{2}\ln\frac{m^2}{4\pi\mu^2}+\frac{1}{2}\psi(2)+\mathcal{O}(\varepsilon) \\ F^{(4)}(\,\cdot\,) &= \frac{3}{2}\psi(1)-\frac{3}{2}\ln\frac{m^2}{4\pi\mu^2}-\frac{1}{2}\int_0^1 d\alpha\sum_{z=u,t,s}\ln(1+\alpha(1-\alpha)\frac{z}{m^2})+\mathcal{O}(\varepsilon) \end{aligned}$$

This separates the divergent pole structure to one-loop in ϕ^4 for $d=4$. We have kept a factor μ^ε in front of $\Gamma^{(4)}$ for dimensional reasons.

For completeness we note:

$$\int_0^1 d\alpha\ln(1+\alpha(1-\alpha)\frac{z}{m^2}) = -2+\sqrt{1+4m^2/z}\ln\left(\frac{\sqrt{1+4m^2/z}+1}{\sqrt{1+4m^2/z}-1}\right) \tag{5.41}$$

The main part of the physics is actually to be found in this non-trivial momentum dependence of the two-point function. We are however concentrating on the pole part, since our goal is to understand renormalization.

5.4.2 ϕ^3 in d=6

Like for the ϕ^4 in $d=4$ we introduce

$$\varepsilon=(6-d)/2, \qquad \tilde{\lambda}=\mu^\varepsilon\,\lambda$$

$\Gamma^{(2)}$ diagrams have $\omega(D)=2$, $\Gamma^{(3)}$ diagrams $\omega(D)=0$, all other diagrams D have $\omega(D)<0$ and they are convergent at l-loop level for $d=6$.

By use of the perturbative machinery of chapter 2 or 4 we get the 1-loop corrections to the 2- and 3-point functions shown in fig.11[1] and we get from (5.34),(5.37) and (5.39)

$$\begin{aligned} \delta\Gamma^{(2)}(p) &= \frac{\lambda^2\Gamma(\varepsilon-1)}{2(4\pi)^3}\left(\frac{4\pi\mu^2}{m^2}\right)^\varepsilon m^2\int_0^1 d\alpha\left(1+\alpha(1-\alpha)\frac{p^2}{m^2}\right)^{1-\varepsilon} \\ \delta\Gamma^{(3)}(p_i) &= -\mu^\varepsilon\,\frac{\lambda^3\Gamma(\varepsilon)}{(4\pi)^3}\left(\frac{4\pi\mu^2}{m^2}\right)^\varepsilon\int\prod_{i=1}^{3}d\alpha_i\delta(1-\sum\alpha_i)\left(1+\frac{F(\alpha_i,p_i)}{m^2}\right)^{-\varepsilon} \end{aligned}$$

where F is defined in example 6.

---

[1] We have ignored tadpole diagrams for simplicity. They will contribute to the mass renormalization in a trivial manner, which we leave to the reader to work out.

We can expand in ε for $d \to 6$:

$$\begin{aligned}
\delta\Gamma^{(2)}(p) &= -\frac{\lambda^2}{2(4\pi)^3}\Big(\frac{1}{\varepsilon}(m^2+\frac{1}{6}p^2)+F^{(2)}(p^2,m^2,\mu^2,\varepsilon)\Big) \\
\delta\Gamma^{(3)}(p_i) &= -\mu^{\varepsilon}\,\frac{\lambda^3}{2(4\pi)^3}\Big(\frac{1}{\varepsilon}+F^{(3)}(p_i^2,m^2,\mu^2,\varepsilon)\Big)
\end{aligned}$$

where $F^{(2)}$ and $F^{(3)}$ are finite for $\varepsilon \to 0$.

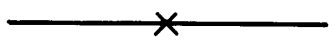
5.5 Counter terms

The lesson learned from these two calculations is that the infinities can be hidden away by adding to $\mathcal{L}(\phi)$ additional terms, *so-called counter terms*, since the pole terms have exactly the same structure as the terms in the original lagrangian. The coefficients multiplying the $1/\varepsilon$ terms are powers of the coupling constants times constants or m^2 and p^2 ($p_\mu^2 \sim -\partial_\mu^2$ by Fourier transformation) . All non-trivial momentum behaviour, which does not correspond to any momentum dependence in the Lagrangian is not associated with the pole terms. The same would have been true if we had used another regularization like Pauli-Villars, only would the poles be replaced by powers and logarithms of the cut-off Λ.

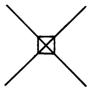
If we add (for ϕ^4):

$$\mathcal{L}_{c.t.}(\phi) = \left[\frac{\lambda}{16\pi^2}\,\frac{1}{\varepsilon}\right]\,\frac{1}{2}m^2\phi^2 + \mu^{\varepsilon}\left[\frac{\lambda^2}{16\pi^2}\,\frac{3}{\varepsilon}\right]\,\frac{\phi^4}{4!} \tag{5.42}$$

and thereby introduce two new vertices:



$$-\frac{\lambda}{16\pi^2}\,\frac{1}{\varepsilon}\,m^2$$



$$-\mu^{\varepsilon}\frac{\lambda^2}{16\pi^2}\,\frac{3}{\varepsilon}$$

by definition we get a finite result to one-loop since these vertices will cancel the poles. The diagrammatic expansion is shown in fig.12

In the same way we get for ϕ^3 in $d = 6$: (note the different sign compared to ϕ^4. We will return to the significance of this later)

$$\mathcal{L}_{c.t.}(\phi) = -\left[\frac{\lambda^2}{12(4\pi)^3}\,\frac{1}{\varepsilon}\right]\,\frac{1}{2}(\partial_\mu\phi)^2 - \left[\frac{\lambda^2}{2(4\pi)^3}\,\frac{1}{\varepsilon}\right]\,\frac{1}{2}m^2\phi^2 - \mu^{\varepsilon}\,\left[\frac{\lambda^3}{2(4\pi)^3}\,\frac{1}{2\varepsilon}\right]\,\frac{\phi^3}{3!}$$

If we write (lets take the ϕ^3-theory as an ex. and recall that $\tilde{\lambda} = \mu^{\varepsilon}\lambda$):

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m^2\,\phi^2 + \frac{\mu^{\varepsilon}\lambda}{3!}\phi^3 \tag{5.43}$$

$$\mathcal{L}_{ct} = C\,\frac{1}{2}(\partial_\mu\phi)^2 + B\,\frac{1}{2}m^2\phi^2 + A\,\frac{\mu^{\varepsilon}\lambda}{3!}\phi^3 \tag{5.44}$$

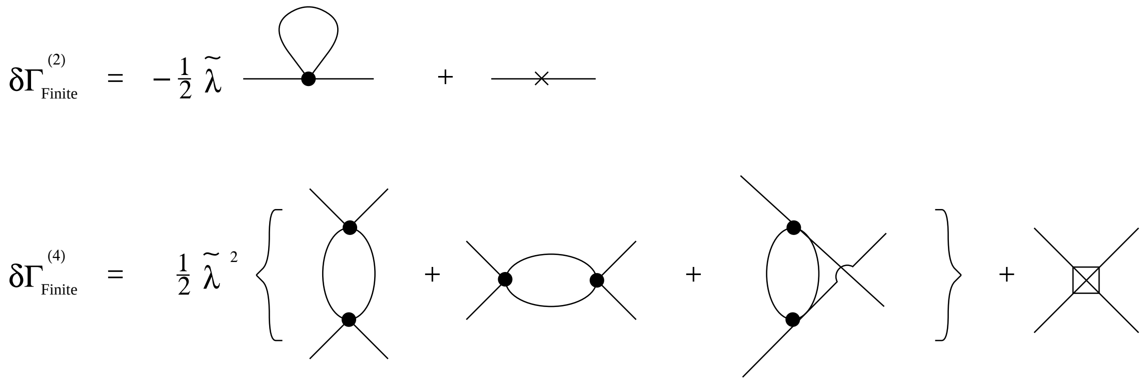


Figure 5.12: The one loop expansion for ϕ^4. The rhs will be finite since the counter terms cancel the poles in the loop diagrams.

we can write:

$$\begin{aligned}
\mathcal{L}^{ren}(\phi) &\equiv \mathcal{L}(\phi) + \mathcal{L}_{c.t.}(\phi) && (5.45)\\
&= \frac{1}{2}(\partial_\mu \phi_0)^2 + \frac{1}{2} m_0^2\ \phi_0^2 + \frac{\lambda_0}{3!}\ \phi_0^3 && (5.46)
\end{aligned}$$

where:

$$\begin{aligned}
\phi_0 &= \sqrt{1+C}\ \phi \equiv Z_\phi^{1/2}\ \phi && (5.47)\\
m_0^2 &= m^2(1+B)/Z_\phi && (5.48)\\
\lambda_0 &= \mu^\varepsilon \lambda\ (1+A)/Z_\phi^{n/2} && (5.49)
\end{aligned}$$

where the last relation is valid for a ϕ^n-theory. ϕ_0, m_0 and λ_0 are called *bare* fields, *bare* masses and *bare* coupling constants. $\mathcal{L}^{ren}$ is called the *renormalized* Lagrangian, an unfortunate notation since its coefficients are infinite as $\varepsilon \to 0$. The bare quantities all diverge as $\varepsilon \to 0$ while the *renormalized* quantities λ, m are finite in this limit: They can be identified with some physical parameters of the theory.

It is convenient for later use to write explicitly the relations between the bare quantities and the renormalized ones (λ and m). To one-loop we found:

$$\begin{aligned}
\lambda_0 &= \mu^\varepsilon \left(\lambda + \frac{a_1(\lambda)}{\varepsilon}\right) && (5.50)\\
m_0^2 &= m^2 \left(1 + \frac{b_1(\lambda)}{\varepsilon}\right) && (5.51)\\
Z_\phi &= 1 + \frac{c_1(\lambda)}{\varepsilon} && (5.52)
\end{aligned}$$

where the lowest order expressions for the coefficient functions a_1, b_1 and c_1 are found in example 7 for the ϕ^4- and the ϕ^3-theory.

Example 7: Relation between bare and renormalized quantities.

For ϕ^4 we have the following lowest order expressions for the coefficient functions:

$$a_1(\lambda) = \frac{3\lambda^2}{(4\pi)^2}, \quad b_1(\lambda) = \frac{\lambda}{(4\pi)^2}, \quad c_1(\lambda) = 0,$$

while we for ϕ^3 get

$$a_1(\lambda) = -\frac{3\lambda^3}{8(4\pi)^3}, \quad b_1(\lambda) = -\frac{5\lambda^2}{12(4\pi)^3}, \quad c_1(\lambda) = -\frac{\lambda^2}{12(4\pi)^3}$$

For higher order perturbative calculations the above formulas generalize in two ways: The coefficient functions $a_1(\lambda), b_1(\lambda)$ and $c_1(\lambda)$ become higher order polynomials of λ, the degree depending on the order to which we expand. But in addition we get higher order singularities of the type $1/\varepsilon^k$, $k = 2, 3, \ldots$, the highest order again depending on the order to which we expand. These higher order terms are not really independent and the associated coefficient functions can in fact all be found from the coefficient functions $a_1(\lambda), b_1(\lambda), c_1(\lambda)$ of the simple pole terms. An example of such a recursion relation is given in example 13, but we will not discuss it further, although it is quite important from a practical point of view. We can summarize this description by stating (without proof) that the one-loop formulas (5.50)-(5.52) generalize to arbitrary order in perturbation theory in the following way (where we note again that the higher pole coefficient functions $a_2(\lambda), \ldots$ are not really independent functions):

$$\lambda_0 = \mu^\varepsilon \left(\lambda + \sum_{k=1}^{\infty} \frac{a_k(\lambda)}{\varepsilon^k}) \right) \tag{5.53}$$

$$m_0^2 = m^2 \left(1 + \sum_{k=1}^{\infty} \frac{b_k(\lambda)}{\varepsilon^k} \right) \tag{5.54}$$

$$Z_\phi = 1 + \sum_{k=1}^{\infty} \frac{c_k(\lambda)}{\varepsilon^k} \tag{5.55}$$

For fixed m_0, λ_0 and ε eq.(5.53) determines λ as a function of μ, and eqs. (5.54)-(5.55) determine m and Z_ϕ as functions of λ, and in this way as functions of μ. The dependence of μ will play an important role in the following. A dual interpretation is also useful: Assume m and λ (and in addition μ) are held fixed. Then the bare mass and coupling constant will be a function of the cut off ε.

The fact that renormalization, acording to (5.47)-(5.49) and (5.53)-(5.55), is nothing but a redefinition of the parameters of the theory implies that we have relation:

$$G^{(n)}(p_i, \ldots, p_n, m, \lambda, \mu, \epsilon) = Z_\phi^{-n/2} \; G_0^{(n)}(p_1, \ldots, p_n, m_0, \lambda_0, \epsilon) \tag{5.56}$$

for the connected n-point function. The lhs of (5.56) is finite when $\varepsilon \to 0$ by construction, while the finiteness of the rhs is formal and depends on cancellations between Z_ϕ and G_0 of quantities going to infinity when $\varepsilon \to 0$.

Eq. (5.56) follows from the definition of the connected Green functions as derivatives of the free energy. We have defined the connected Green functions $G^{(n)}$ from the functional integral, which uses as lagrangian $\mathcal{L}^{ren}(\phi) - \phi \cdot J$, by functional differentiation of the free energy $F[J] = -\ln Z[J]$ after J. If we want to express the renormalized lagrangian in terms of *bare* quantities, as in (5.46), but also in the presence of external sources, we clearly have to scale the external *bare* sources inversely to ϕ_0 such that $\phi J = \phi_0 J_0$:

$$\phi_0 = Z_\phi^{1/2} \phi, \qquad J_0 = Z_\phi^{-1/2} J \tag{5.57}$$

We can now express $Z[J]$ and $F[J]$ in terms of bare quantities. Since the lagrangians have the same functional form and takes the same value: $\mathcal{L}^{ren}(\phi, J) = \mathcal{L}_0(\phi_0, J_0)$, a change in variables $\phi \to \phi_0$ inside the functional integration for $Z[J]$ gives, except for an infinite measure factor $C_Z = \prod_{x \in R^d} Z_\phi^{1/2}$, precisely the functional integral $Z_0[J_0]$ for the *bare* theory:

$$C_Z \, Z[J] = Z_0[J_0], \qquad F[J] = F_0[J_0] + \ln C_Z. \tag{5.58}$$

Since the bare connected Green functions are defined by functional differentiation of $F_0[J_0]$ with respect to J_0 (5.56) follows from (5.58). We can continue this analogue and define the generating functional for the *bare* 1PI Green functions from $F_0[J_0]$ as usual:

$$\Gamma_0[\phi_0^{cl}] = F_0[J_0] + J_0 \cdot \phi_0^{cl}, \qquad \phi_0^{cl} \equiv -\frac{\delta F_0}{\delta J_0}, \tag{5.59}$$

and it follows that

$$\phi_0^{cl} = Z_\phi^{1/2} \phi^{cl}, \qquad \Gamma_0[\phi_0^{cl}] = \Gamma[\phi^{cl}] - \ln C_Z. \tag{5.60}$$

Since the *bare* 1PI Green functions are derived by functional differentiation after ϕ_0^{cl} we deduce from (5.60) that the 1PI Green functions will satisfy an equation similar to (5.56):

$$\Gamma^{(n)}(p_i, \ldots, p_n, m, \lambda, \mu, \epsilon) = Z_\phi^{n/2} \; \Gamma_0^{(n)}(p_1, \ldots, p_n, m_0, \lambda_0, \epsilon) \tag{5.61}$$

The difference between (5.56) and (5.61) is the power of Z_ϕ. Heuristically this follows from the observation that the 1PI Green function $\Gamma^{(n)}$ is the 1PI irreducible component of the connected Green function $G^{(n)}$ where all external propagators have been removed. Since the bare and the renormalized propagators by (5.56) are related by

$$G^{(2)}(p) = Z_\phi^{-1} G_0^{(2)}(p)$$

the removal of an n external propagators implies an addition multiplication with Z_ϕ^n on the rhs of (5.56) when changing from $G^{(n)}$ to $\Gamma^{(n)}$.

Eqs. (5.61) and (5.46)-(5.56) show *the multiplicative nature of renormalization.* We will discuss that aspect later.

Let us end this section by discussing once more the renormalization procedure in light of the results derived. Let us assume that $\varepsilon > 0$. For a given μ we have relations between the "bare" quantities λ_0, m_0, ϕ_0 and the renormalized ones λ, μ, ϕ and between the related Green functions. The relation is operational: At the tree level (no loops) this operation is the identity, but already at the one-loop level the transformation between bare and renormalized quantities is nontrivial. Higher order corrections will move the bare and renormalized quantities away from each other. If we stop the perturbation expansion at finite order and *then* turn the cut off ε to zero this difference is singular due to the poles in ε. It is less clear whether the relation between bare and renormalized quantities remains singular if we could perform the summation to all orders of perturbation theory and then take ε to zero. But in what sense do the bare and renormalized quantities represent the same theory ? One way to answer this question, is as follows: The bare parameters can be viewed as a short distance approximation to the renormalized parameters, in the sense that no interactions have yet had the chance to take place. As we measure the correlation between ϕ's separated over larger distances in space and time interactions will modify the values of masses and couplings we attempt to measure. We can always imagine the renormalized masses and coupling constants as being related to measurements at a chosen length scale, and consequently they will differ from the bare masses and couplings. However, in order to *define* the theory at all, we have to introduce a regularization and thereby a mass parameter μ, and the interacting, renormalized theory is defined as a *limit* where the cut-off is removed. This limit is is not unique in the sense that it is possible to find a whole family of renormalized couplings, masses and wave function renormalizations $\lambda(\mu, \varepsilon), m(\mu, \varepsilon), Z_\phi(\mu, \varepsilon)$ which, for a given choice of of bare parameters, by the process of renormalization, leads to finite theories when the cut-off ε is removed. This point of view has a nice realization in the statistical theory of critical phenomena, as will be discussed in detail in the next chapter. It is also the origin of the renormalization group equations which formulate in a precise manner that the family of theories we get by starting with a set of bare parameters and implementing the process of renormalizations are related. We will discuss in detail the renormalization group equations in section 7.

A dual point of view on the process of renormalization is also very useful. The starting point is the well defined renormalized expressions when a cut off ε is present. The bare and renormalized couplings, masses etc are related by expressions involving the cut off, which is needed in order to define our theory. We now want to remove the cut off, but it should be done in such a way that renormalized couplings, masses etc are kept fixed (and preferable related to some physical observables). This can only be done by adjusting the bare parameters as functions of the cut off in a specific way, namely the one given above which relates bare and renormalized parameters. This *fine tuning* of the bare parameters in order to get to a renormalized theory with no reference to the short distance cut off also has a simple interpretation in the statistical theory of critical phenomena.

5.6 Renormalization conditions and finite renormalization

The procedure of subtracting only the pole terms is called *minimal subtraction* (MS). We got finite results but depending on the mass parameter μ. This specific subtraction is convenient for calculations, but otherwise arbitrary. We could have subtracted any *finite* constants too. It might be more transparent if we fixes the parameters in a different way. Let us use the ϕ^4 theory as an example. We could demand that the superficially divergent, 1PI Green functions are equal to their *tree (≡ no loop) values* at $p_i = 0$. In the case of the ϕ^4 interaction the superficially divergent diagrams are associated with two- and four-point functions and the tree values are just the terms we read off from the Lagrangian:

$$\Gamma^{(2)}_{tree}(p) = p^2 + m^2, \qquad \Gamma^{(4)}_{tree}(p_i) = \lambda.$$

The above mentioned conditions therefore translate to

$$\Gamma^{(2)}(0) = m^2, \quad \frac{d}{dp^2}\,\Gamma^{(2)}(p^2)|_{p^2=0} = 1, \quad \Gamma^{(4)}(p_i = 0) = \lambda. \tag{5.62}$$

Clearly these requirements would fix the counter terms uniquely (also the finite parts) and they have the virtue of showing that *the arbitrariness of the theory is exactly equal to the number parameters in the lagrangian*, here m^2,λ and the factor in front $(\partial_\mu\phi)^2$. This means that the mass and the coupling constant cannot be calculated within the theory. However, the non-trivial p_i-dependence present in the two- and four-point functions after renormalization allows in principle a test against experiments.

There is nothing magic about the choice of *normalization conditions* (5.62). For instance one could use other points in momentum space for the normalization:

$$\Gamma^{(2)}(\bar{\mu}^2) = (\bar{\mu}^2 + m^2), \quad \frac{d\Gamma^{(2)}(p^2)}{dp^2}|_{p^2=\bar{\mu}^2} = 1, \quad \Gamma^{(4)}(p_i)|_{p_i p_j = \bar{\mu}^2(\delta_{ij}-1/4)} = \lambda \,, \tag{5.63}$$

the latter prescription chosen such that $s = t = u = \bar{\mu}^2$. But one can, of course, choose any value of s, t, u and p^2 as normalization point.

Two renormalized theories which only differs by the choice of normalization conditions can be related through the introduction of *finite* counter terms (no poles!) and exactly as for the infinite renormalization this is equivalent to a redefinition of m and λ and a finite wave function renormalization Z_ϕ. One will therefore get relations like (5.56) and (5.61) between the Green functions, but this time with a finite Z_ϕ.

A final point should be mentioned here. The relations (5.53)-(5.55) are particular simple in the sense that the coefficient functions a_k etc. are functions only of λ. This is specific for MS. Other normalization condition would in general lead to coefficient functions $a_k(\lambda, m/\mu)$ depending also on the dimensionless ratio m/μ. Nothing is wrong with this, but from the point of view of practical calculations the MS- and closely related prescriptions are very convenient. We will use MS in the next section.

5.7 The renormalization group

We will now return to the magic mass parameter μ introduced by dimensional regularization. We have seen that we can trade it away for an arbitrary subtraction point $\bar{\mu}$. Eventually this arbitrariness will be fixed by the requirement that the mass m(the pole of the propagator) agrees with some physical observed mass and the "charge" λ agrees with some physical observed charge. In this way the μ dependence will eventually disappear from the theory[2]. However, the freedom to choose μ (or a subtraction point) at the intermediate steps of the calculation is convenient, since it allows us to explore *scaling properties* of the Green functions. Recall that a convergent 1PI ϕ^4 diagram D with n external lines contributes to the full 1PI Green function with an amplitude Γ_D having the following scaling behaviour:

$$\Gamma_D(tp_i; tm) = t^{4-n}\Gamma_D(p_i, m) \tag{5.64}$$

Had it not been for the divergences this relation would also be true for the full 1PI Green function since it would be a sum of terms satisfying (5.64). However, the necessity of renormalization introduces a new mass scale (μ or the subtraction point $\bar{\mu}$). By including μ in the scaling we can still write:

$$\Gamma^{(n)}(tp_i, tm, t\mu) = t^{(4-n)}\Gamma^{(n)}(p_i, m, \mu) \tag{5.65}$$

---

[2] For massless theories the situation is somewhat more complicated since there is no physical mass m to which the μ dependence can be transferred

This means that $\Gamma^{(n)}$ satisfies the following equation:

$$t\frac{d}{dt}\Gamma^{(n)}(tp_i,tm,t\mu)=(4-n)\Gamma^{(n)}(tp_i,tm,t\mu) \tag{5.66}$$

or, absorbing the t dependence in m and μ:

$$\left(t\frac{\partial}{\partial t}+m\frac{\partial}{\partial m}+\mu\frac{\partial}{\partial\mu}+n-4\right)\Gamma^{(n)}(tp_i,m,\mu)=0 \tag{5.67}$$

It is now possible to get rid of the μ derivative by use of the relation (5.56) or (5.61) between the bare Green functions and the renormalized ones. Since the bare Green functions have no dependence of μ we get:

$$\frac{d}{d\mu}\Gamma_0^{(n)}(p_i,m_0,\lambda_0,\varepsilon)=0 \quad\Rightarrow \frac{d}{d\mu}\left\{Z_\phi^{-n/2}(\lambda,\varepsilon)\Gamma^{(n)}(p_i,\lambda,m,\mu,\varepsilon)\right\}=0. \tag{5.68}$$

This formula should be understood in the following way: For fixed bare couplings m_0 and λ_0 and an $\varepsilon>0$ we have a well defined (regularized) theory where the bare Green functions have no reference to μ. However, the renormalized coupling and mass and Z_ϕ become functions of μ as described in connection with equations (5.53)-(5.55): λ becomes a function $\lambda(\mu,\varepsilon;\lambda_0,m_0)$, and $m(\lambda,\varepsilon)$, $Z_\phi(\lambda,\varepsilon)$ functions of μ through λ.[3]
When acting on $\Gamma^{(n)}(p_i,\lambda(\mu),m(\mu),\mu))$ we can write

$$\mu\frac{d}{d\mu}=\mu\frac{\partial}{\partial\mu}+\mu\frac{d\lambda}{d\mu}\frac{\partial}{\partial\lambda}+\mu\frac{dm}{d\mu}\frac{\partial}{\partial m} \tag{5.69}$$

Using this, and the following definitions

$$\begin{aligned}\beta(\lambda,\varepsilon) &\equiv \mu\frac{d\lambda}{d\mu}|_{m_0,\lambda_0,\varepsilon} && (5.70)\\ \gamma_d(\lambda,\varepsilon) &\equiv \frac{\mu}{2Z_\phi}\frac{dZ_\phi}{d\mu}|_{m_0,\lambda_0,\varepsilon} && (5.71)\\ \gamma_m(\lambda,\varepsilon) &\equiv \frac{\mu}{m}\frac{dm}{d\mu}|_{m_0,\lambda_0,\varepsilon} && (5.72)\end{aligned}$$

some algebra allow us to write (5.68) as

$$\left(\mu\frac{\partial}{\partial\mu}+\beta(\lambda)\frac{\partial}{\partial\lambda}+\gamma_m(\lambda)m\frac{\partial}{\partial m}-n\gamma_d(\lambda)\right)\Gamma^{(n)}(p_i,\lambda,m,\mu)=0 \tag{5.73}$$

In this equation we have taken the limit $\varepsilon\to 0$ since there is only reference to renormalized quantities. The coefficient functions $\beta(\lambda),\gamma_d(\lambda),\gamma_m(\lambda)$ are called the *β-function and the anomalous field and mass dimensions.* The reason for the latter names will become clear in a moment. (5.67) and (5.73) allow us to eliminate $\mu\partial/\partial\mu$ and we get the *renormalization group equation*

$$\left(-t\frac{\partial}{\partial t}+\beta(\lambda)\frac{\partial}{\partial\lambda}+[\gamma_m(\lambda)-1]m\frac{\partial}{\partial m}-n\gamma_d(\lambda)+4-n\right)\Gamma^{(n)}(tp_i,\lambda,m,\mu)=0 \tag{5.74}$$

---

[3] We note again that it is due to the special simple properties of MS that Z_ϕ is not a function of m/μ. For more general normalization conditions, like the ones described in the last section, $Z_\phi=Z_\phi(\lambda,m/\mu,\varepsilon)$.

This equation describes the behaviour of the $\Gamma^{(n)}(p_i)$ when one scales the momenta from p_i to tp_i.

It is a first order partial differential equation with characteristic equations:

$$t\frac{\partial\bar{\lambda}(t)}{\partial t} = \beta(\bar{\lambda}(t)) \qquad \bar{\lambda}(t=1)=\lambda \tag{5.75}$$

$$t\frac{\partial\bar{m}(t)}{\partial t} = \bar{m}(t)(\gamma_m(\bar{\lambda}(t))-1) \qquad \bar{m}(t=1)=m \tag{5.76}$$

From the theory of partial differential equations we know that the the solution to (5.74) can be expressed by the solutions to the characteristic equations as follows:

$$\boxed{\Gamma^{(n)}(tp_i,\lambda,m;\mu) = t^{4-n}\, e^{-n\int_1^t \frac{ds}{s}\gamma_d(\bar{\lambda}(s))}\,\Gamma^{(n)}(p_i,\bar{\lambda}(t),\bar{m}(t);\mu)} \tag{5.77}$$

Proof:

One way to prove (5.77) without appealing to the general theory of partial differential equations is simply to insert the rhs in (5.74) and perform the differentiations. In order to do that it is important to note that $\bar{\lambda}(t)$ as defined by (5.75) is actual a function of λ because of the boundary condition $\bar{\lambda}(t=1)=\lambda$. Similar remarks apply for $\bar{m}$ which depends both on m and λ. To be entirely correct we should write:

$$\bar{\lambda}=\bar{\lambda}(t,\lambda), \qquad \bar{m}=\bar{m}(t,m,\lambda).$$

From (5.75) we get:

$$\ln t = \int_\lambda^{\bar{\lambda}} \frac{d\lambda'}{\beta(\lambda')}$$

and differentiating after λ gives

$$\frac{\partial\bar{\lambda}}{\partial\lambda}|_t = \frac{\beta(\bar{\lambda})}{\beta(\lambda)}. \tag{5.78}$$

In the same way (5.76) can be integrated to

$$\ln\frac{\bar{m}}{m} = \int_1^t \frac{dt}{t}(\gamma_m(\bar{\lambda}(t,\lambda))-1) = \int_\lambda^{\bar{\lambda}} \frac{d\lambda'}{\beta(\lambda')}(\gamma_m(\lambda')-1). \tag{5.79}$$

Differentiating the first two expressions in (5.79) with respect to m gives:

$$\frac{\partial\bar{m}}{\partial m}|_{t,\lambda} = \frac{\bar{m}}{m} \tag{5.80}$$

while differentiation of the first and the last expression in (5.79) by means of (5.78) leads to

$$\frac{\beta(\lambda)}{\bar{m}}\frac{\partial\bar{m}}{\partial\lambda} = \gamma_m(\bar{\lambda})-\gamma_m(\lambda). \tag{5.81}$$

Using (5.78),(5.80) and (5.81) it is now straight forward to verify that the rhs of (5.77) satisfies (5.74). This completes the proof.

Equation (5.77) is one of the most important in this chapter. In order to understand better the structure recall the "naive" scaling behaviour (5.64) which would be valid if we had no divergences. Let us rewrite it as in (5.77)

$$\Gamma_D^{(n)}(tp_i,\lambda,m) = t^{4-n}\Gamma_D^{(n)}(p_i,\lambda,\frac{m}{t}) \tag{5.82}$$

Assume for a moment that anomalous scaling dimensions $\gamma_m(\bar{\lambda}(t))$ and $\gamma_d(\bar{\lambda}(t))$ entering in (5.77) are (approximately) constants as functions of t. Then the characteristic equations (5.75)-(5.76) can be solved and we get

$$\bar{m}(t) = \frac{m}{t^{1-\gamma_m}}, \qquad t^{4-n}\; e^{-n\int_1^t ds/s\, \gamma_d(\bar{\lambda}(s))} = t^{4-n(1+\gamma_d)}$$

By inserting these solutions in (5.77) and comparing with (5.82) it is clear why γ_d and γ_m are called the anomalous scaling dimensions. In general $\gamma_{d,m}(\bar{\lambda}(t))$ will not be constants as functions of t, since the so called *running coupling constant* $\bar{\lambda}(t)$ will not be constant. Only in the neighbourhood of a zero of the β-function will this be the case, as is clear from the definition (5.75) of the running coupling constant. *The zeroes of the β-function play therefore a special role and are called fixed points of the renormalization group.* The reason for this will be clear in a moment, when we discuss the possible behaviour of the running coupling constant. The appearance of a running coupling constant in (5.77) is maybe the most surprising result of the structure of renormalization group. Since $\bar{\lambda}(t)$ may either decrease or increase with t (5.77) shows that *the validity of perturbation theory may depend on the momentum range we are probing*, a feature which is absent in the "naive" scaling relation (5.82). As an example let us assume that the running coupling constant goes to zero for $t \to \infty$. In this way perturbation theory will be reliable for large momenta (i.e. small distances), while it might break down for soft processes. This is exactly the situation which occur for the field theories which seem to be most relevant for the description of Nature: the non-abelian gauge theories. We will return to these in chapter 7 and 8.

It is important to emphasize that (5.77) does not provide us with any means of actually calculating the Green functions, since we do not know the β-function. We will end this chapter by showing how to calculate the lowest order term in the β-function within perturbation theory. In perturbation theory the β-function will appear as a power series in the coupling constant and $\beta(\lambda = 0) = 0$. So $\lambda = 0$ *is a fixed point*, called *the gaussian fixed point* since the theory has a trivial gaussian action (i.e. free action) if $\lambda = 0$. As λ increases too far from zero a perturbative calculation of the first few terms of the β-function will not give a reliable expression for the β-function. It is, however, very instructive to classify the possible behaviour of $\beta(\lambda)$ outside perturbation theory, and study the associated behaviour of the running coupling constant $\bar{\lambda}(t)$.

In fig. 13 we have showed four possible beta functions, all starting from the gaussian fixed point.

(I) : $\beta(\lambda)$ is positive and increases faster than λ. In this case $\bar{\lambda}(t)$ will start to increase from λ and eventually it will increase to ∞ *for a finite value of t.* This follows from the integrated version of (5.75):

$$t = \exp \int_{\lambda}^{\bar{\lambda}} \frac{d\lambda'}{\beta(\lambda')} \tag{5.83}$$

Since the integral is convergent we see that $\bar{\lambda} \to \infty$ for $t_0 = \exp(\int_{\lambda}^{\infty} d\lambda'/\beta(\lambda'))$. This singularity is called the *Landau pole.* It signals a definite break down of perturbation theory and maybe of the theory itself.

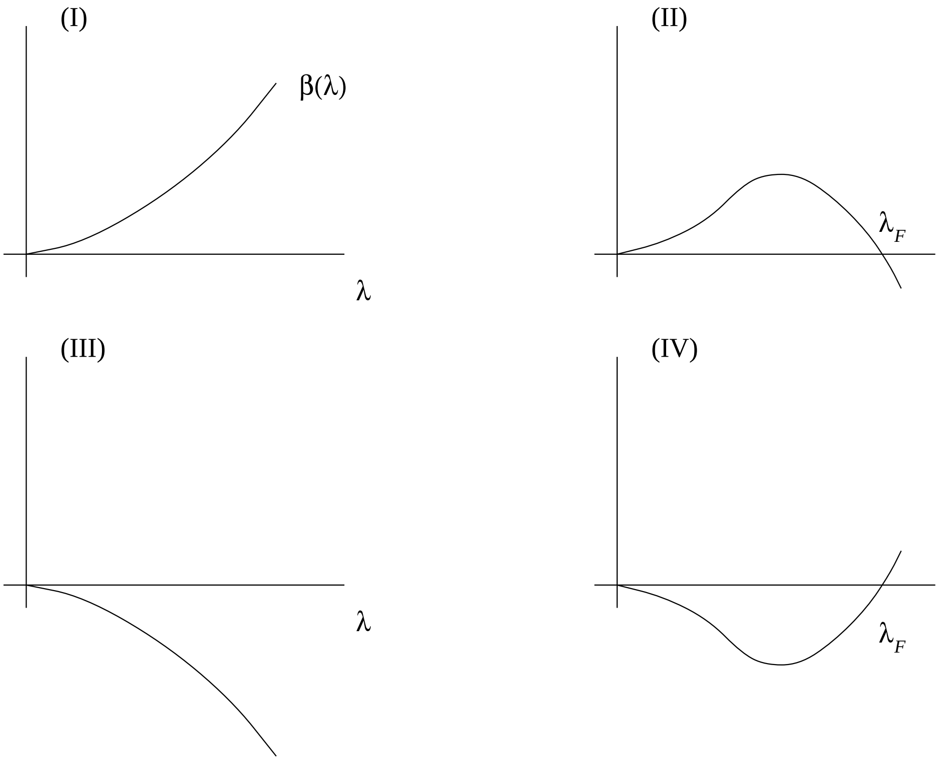


Figure 5.13: Four different β-functions

(II) : $\beta(\lambda)$ is positive, but turns around and has a zero, *i.e. a fixed point* $\lambda_F \neq 0$. We now see the reason for calling λ_F a fixed point. Close to the fixed point we have

$$\beta(\lambda) = \beta'(\lambda_F)(\lambda - \lambda_F) + \cdots \tag{5.84}$$

and the integral (5.83) will diverge as $\lambda \to \lambda_F$. This means that for $t \to \infty$ we have $\lambda \to \lambda_F$. *By scaling the momenta to infinity the running coupling constant will be driven to the fixed point* λ_F. Such a fixed point is called an *ultraviolet stable fixed point* while the fixed point at zero is in this case is called *infrared stable* since $\bar{\lambda}(t)$ is driven to this fixed point by letting the momentum scale $t \to 0$. By changing the momentum scale t we see that $\bar{\lambda}(t)$ will always move between an infrared and an ultraviolet stable fixed point, provided such a point exists (in (I) we had no ultraviolet stable fixed point).

(III) : $\beta(\lambda)$ decreases faster than λ. We have the same situation as in (I), except that everything is turned around: As $t \to \infty$ $\bar{\lambda}(t) \to 0$: *in this case the gaussian fixed point is ultraviolet stable*, and we see that perturbation theory becomes more reliable at short distances (high momenta). This phenomenon is called *asymptotic freedom.* As already mentioned the non-abelian gauge theories seem to fall in this class of field theories. The existence of a Landau pole and the break down of perturbation theory, when we try to probe large distances is even welcome in these theories since it is a common belief that a phenomenon like quark confinement is of non-perturbative nature. It should also be mentioned that it is a common belief (but no proof exists), that only asymptotic free theories makes sense as rigorously defined theories outside perturbation theory. This might be the reason Nature seems so fond of non-abelian gauge theories, since these are the only renormalizable, asymptotic free theories in $d = 4$.

(IV) : In this case the situation is reverse from the one of (II). The gaussian fixed point is ultraviolet stable, while the fixed point at λ_F is infrared stable. Note that a fixed point is infrared (ultraviolet) stable if the first non-zero derivative of the β-function at the fixed point is positive (negative).

We end this section with two examples. The first illustrates the calculation of the β-function to lowest order in perturbation theory for the ϕ^4 theory in $d = 4$ and the ϕ^3 theory in $d = 3$. The first turns out to have an infrared stable gaussian fixed point, while the other has an ultraviolet stable gaussian fixed point. In addition the above mentioned relations between the lowest order pole terms and the higher order pole terms are proven in the examples.

Example 13: Perturbative calculation of the β-function.

ϕ^4: The β-function was defined as

$$\beta(\lambda, \varepsilon) = \mu \frac{\partial \lambda}{\partial \mu}$$

where we have put in the explicit *analytic* dependence on ε away from $d = 4$. Recall the relation between the renormalized coupling constant λ and the bare one λ_0:

$$\lambda_0 = \mu^{\varepsilon} \left(\lambda + \sum_{k=1}^{\infty} \frac{a_k(\lambda)}{\varepsilon^k} \right)$$

Differentiation with respect to μ for fixed λ_0 and use of the definition of the β-function above leads to

$$0 = \varepsilon \left(\lambda + \sum_{k=1}^{\infty} \frac{a_k(\lambda)}{\varepsilon^k} \right) + \beta(\lambda, \varepsilon) \left(1 + \sum_{k=1}^{\infty} \frac{a'_k(\lambda)}{\varepsilon^k} \right) \tag{5.85}$$

and the solution for $\beta(\lambda, \varepsilon)$ compatible with the limit $\varepsilon \to 0$ is

$$\beta(\lambda, \varepsilon) = -\varepsilon\, \lambda - a_1(\lambda) + \lambda\, a'_1(\lambda)$$

or (taking $\varepsilon \to 0$)

$$\beta(\lambda) \equiv \lim_{\varepsilon \to 0} \mu \frac{\partial \lambda}{\partial \mu} = (\lambda \frac{d}{d\lambda} - 1) a_1(\lambda) \tag{5.86}$$

The β-function only depends on the simple pole function $a_1(\lambda)$. This is true to all orders in perturbation theory. In fact all the other coefficient functions $a_k(\lambda)$ can be expressed entirely in terms of this first functions since it is not hard to derive from the consistency of (5.85) the following recursion relation:

$$(\lambda \frac{d}{d\lambda} - 1) a_{k+1}(\lambda) = \beta(\lambda)\, \frac{d a_k(\lambda)}{d\lambda}$$

Since we have already determined the first term in a_1 we now get the following lowest order β-function:

$$\beta(\lambda) = \frac{3\lambda^2}{(4\pi)^2} + \mathcal{O}(\lambda^4) \tag{5.87}$$

and we can use this to expression to find the running coupling constant:

$$\bar{\lambda}(t) = \frac{\lambda}{1 - \frac{3\lambda}{(4\pi)^2} \ln(t)} \tag{5.88}$$

To this lowest order we are in case (I) above, and the running coupling constant increases with the scale until we hit the Landau pole. Of course lowest order pertubation theory becomes invalid long before. We can only expect the above formula to be valid as long as $\frac{3\lambda}{16\pi^2}\ln(t) \ll 1$.

ϕ^3: We use (5.86) and the one loop result for $a_1(\lambda)$ to get

$$\beta(\lambda) = -\frac{3}{4}\frac{\lambda^3}{(4\pi)^3} \tag{5.89}$$

and the running coupling constant will be given by

$$\bar{\lambda}^2(t) = \frac{\lambda^2}{1+\frac{3}{4}\frac{\lambda^2}{(4\pi)^2}\ \ln t^2} \tag{5.90}$$

We see that the theory has asymptotic freedom ! To the one loop approximation it belongs to case (III) discussed above. Unfortunately it is not clear how the ϕ^3-theory can be defined outside perturbation theory, since the functional integral is ill defined (the action is unbounded from below).

Example 14: Perturbative calculation of $\gamma_d(\lambda)$ and $\gamma_m(\lambda)$.

γ_m : From the definition we have

$$m_0^2 = m^2(\lambda)\left(1+\sum_{k=1}^{\infty}\frac{b_k(\lambda)}{\varepsilon^k}\right). \tag{5.91}$$

If we differentiate after μ the lhs is zero while the rhs depend on μ only though λ (in MS):

$$0 = \mu\frac{dm^2(\lambda)}{d\mu}\left(1+\sum_k\frac{b_k(\lambda)}{\varepsilon^k}\right)+m^2(\lambda)\left(\sum_k\frac{b'_k(\lambda)}{\varepsilon^k}\right)\mu\frac{d\lambda}{d\mu}, \tag{5.92}$$

or, using the definition of $\gamma_m(\lambda,\varepsilon)$ and $\beta(\lambda,\varepsilon)$ and the result from ex.13 for $\beta(\lambda,\varepsilon)$:

$$0 = 2\gamma_m(\lambda,\varepsilon)\left(1+\sum_k\frac{b_k(\lambda)}{\varepsilon^k}\right)+\left(\sum_k\frac{b'_k(\lambda)}{\varepsilon^k}\right)\ [-\varepsilon\lambda+\beta(\lambda)]. \tag{5.93}$$

Assuming a power expansion of $\gamma_m(\lambda,\varepsilon)$ in ε leads to the absence of any such powers and instead we get a relation between $\gamma_m(\lambda)$ and $b'_1(\lambda)$ and recursion relations which determines the higher b_k's in terms of b_1 (analogous to the situation in ex.13):

$$\gamma_m(\lambda) = \frac{\lambda}{2}b'_1(\lambda) \tag{5.94}$$

$$\lambda b'_{k+1} = \beta(\lambda)b'_k + \lambda b'_1\ b_k \tag{5.95}$$

γ_d : From the definition we have

$$Z_\phi = 1+\sum_{k=1}^{\infty}\frac{c_k(\lambda)}{\varepsilon^k} \tag{5.96}$$

and differentiating as above after μ and using the definitions of β and γ_d leads to

$$2Z_\phi\gamma_d(\lambda,\varepsilon) = \beta(\lambda,\varepsilon)\sum_k\frac{c'_k(\lambda)}{\varepsilon^k}$$

or

$$2\gamma_d(\lambda,\varepsilon)\left(1+\sum_k \frac{c_k(\lambda)}{\varepsilon^k}\right) = (-\varepsilon\lambda+\beta(\lambda))\sum_k \frac{c'_k(\lambda)}{\varepsilon^k} \tag{5.97}$$

and as above consistency as a Laurent series in ε implies that $\gamma_d(\lambda,\varepsilon)$ can have no dependence on ε. It further leads to a relation between $\gamma_d(\lambda)$ and $c_1(\lambda)$ and recursion relations between the higher c_k coefficients:

$$\gamma_d(\lambda) = -\frac{\lambda}{2}c'_1(\lambda) \tag{5.98}$$

$$\lambda c'_{k+1} = \beta(\lambda)c'_k + \lambda c'_1\, c_k. \tag{5.99}$$

From ex.7 we have the expressions for $b_1(\lambda)$ and $c_1(\lambda)$ for a ϕ^3 in $d=6$ and a ϕ^4 theory in $d=4$ and we can calculate γ_m and γ_d for these theories. We leave that as a trivial exercise.

Chapter 6

Quantum Electrodynamics

6.1 Preliminary remarks

The lagrangian for the abelian gauge theories (in minkowskian space-time) consists of two parts. The first part is the lagrangian for the electromagnetic field and the second part is the matter part, minimally coupled to the electromagnetic field. "Matter" can be either fermionic or bosonic fields:

$$\mathcal{L}(A,\psi,\varphi) = \mathcal{L}_{em}(A) + \mathcal{L}_{fer}(\psi, A) + \mathcal{L}_{bos}(\varphi, A) \tag{6.1}$$

where

$$\begin{aligned}
\mathcal{L}_{em}(A_\mu) &= -\frac{1}{4}F^2_{\mu\nu}(A) && (6.2)\\
\mathcal{L}_{fer}(\psi, A) &= -\bar{\psi}(x)(\not{D} + m)\psi(x) && (6.3)\\
\mathcal{L}_{bos}(\varphi, A) &= -(D_\mu\varphi)^*(D^\mu\varphi) - V(\varphi^*\varphi) && (6.4)
\end{aligned}$$

In these formulae

$$\begin{aligned}
F_{\mu\nu}(A) &= \partial_\mu A_\nu - \partial_\nu A_\mu && (6.5)\\
D_\mu &= \partial_\mu - ieA_\mu && (6.6)\\
\not{D} &= \gamma_\mu D^\mu && (6.7)
\end{aligned}$$

and the γ-matrices satisfy

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu} \quad , \quad \bar{\psi} = \psi^\dagger\gamma_0 \tag{6.8}$$

Local gauge invariance of the lagrangian is the following transformation, which in fact leaves each of the three parts of the lagrangian (6.1) invariant:

$$\begin{aligned}
A_\mu(x) \to A'_\mu(x) &= A_\mu(x) + \partial_\mu\chi(x) && (6.9)\\
\psi_\mu(x) \to \psi'(x) &= e^{ie\chi(x)}\psi(x) && (6.10)\\
\varphi(x) \to \varphi'(x) &= e^{ie\chi(x)}\varphi(x) && (6.11)
\end{aligned}$$

We can view this invariance as an invariance under the action of the group $U(1)$ since it corresponds to multiplying the matter fields by a phase factor. Since the phase factor varies with space-time points x_μ we have an independent $U(1)$ group associated with each space-time point and formally the total invariance group is

$$G_{inv} = \prod_{x \in R^d} U_x(1) \tag{6.12}$$

We shall see the significance of G_{inv} shortly, but for a geometrical interpretation of local gauge invariance, both abelian and non-abelian, we have to refer to the chapter on classical gauge theories.

When we rotate to euclidean space the Lorentz group goes into $SO(d)$ where d is the dimension of the euclidean space. In minkowskian space-time the ψ fields transform as spinors. After the rotation they will transform as spinors of $SO(d)$. The spinor representations of $SO(d)$ can be characterized by hermitean γ-matrices, now satisfying

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} \qquad \mu = 1, \dots d \tag{6.13}$$

The algebra defined by the γ_μ's is called the Clifford algebra and the dimension of its fundamental representation is $2^{[d/2]}$ ($[d/2] \equiv$ integer part of $d/2$). From the γ_μ's we get the generators $s_{\mu\nu}$ of the rotation group $SO(d)$:

$$s_{\mu\nu} = \frac{1}{2}\sigma_{\mu\nu} = \frac{i}{4}[\gamma_\mu, \gamma_\nu] \tag{6.14}$$

There is an important difference between odd and even dimensions, however. In even dimensions $d = 2n$ we can form

$$\gamma_{d+1} = -(-i)^n \ \gamma_1 \dots \gamma_{2n} \tag{6.15}$$

γ_{d+1} is a hermitean matrix $\neq 1$, which anticommutes with all γ_μ, $\mu = 1, \dots d$:

$$\{\gamma_{d+1}, \gamma_\mu\} = 0 \quad , \quad \gamma_{d+1}^2 = 1 \tag{6.16}$$

This means that the representations of $SO(d)$ corresponding to the generators $s_{\mu\nu}$ are *not* irreducible since γ_{d+1} commutes with all $s_{\mu\nu}$. In fact the irreducible representations can be labelled by the eigenvalues of γ_{d+1} : *the chirality.* In odd dimensions we do not have such a γ_{d+1} at our disposal ($\gamma_{d+1} \propto I$) and the concept of chirality does not exist. For instance we have for $d = 3$:

$$\gamma_1 = \sigma_1 \quad , \quad \gamma_2 = \sigma_2 \quad , \quad \gamma_3 = \sigma_3 \quad , \gamma_{d+1} = -i \ \sigma_1\sigma_2\sigma_3 = 1 \tag{6.17}$$

We have already defined the Feynman path integral in terms of anticommuting "classical" Fermi fields, and noticed that $\bar\psi$ and ψ have to be treated as independent variables in the functional integral. We cannot have a relation like (6.8) in euclidean space since it clearly singles out a direction ($\sim \gamma_0$). We will define $\bar\psi$ to transform like the adjoint of ψ with respect to $SO(d)$ rotations. Then $\bar\psi \ \psi$ is a scalar, $\bar\psi\gamma_\mu\psi$ a vector, $\bar\psi\gamma_{d+1}\psi$ a pseudo-scalar (in even dimensions) etc.

With these definitions we have (with $M \equiv$ *minkowskian* and $E \equiv$ *euclidean*)

$$(\gamma^\mu\partial_\mu)_M \to (\gamma_\mu\partial_\mu)_E \tag{6.18}$$

and the action for a free Dirac fermion transform as follows under rotation to euclidean space:

$$\begin{aligned} iS_M[\psi] &= -i\int dt d^{d-1}x\ \bar{\psi}(\gamma^\mu\partial_\mu + m)\psi] \longrightarrow \\ -S_E[\psi\,\bar{\psi}] &= -\int d^dx\ \bar{\psi}(\gamma_\mu\partial_\mu + m)\psi \end{aligned} \tag{6.19}$$

so the euclidean lagrangian can be written as

$$\mathcal{L}_E(\bar{\psi},\psi) = \bar{\psi}(\gamma_\mu\partial_\mu + m)\psi \tag{6.20}$$

Also the rotation of the electromagnetic field to euclidean space deserves some discussions. Since A_μ has a Lorentz index μ, the rotation $t \to -ix_4$ should be supplemented by a rotation of A_0 if we want to keep the terms F_{i0}^2 in the lagrangian real. Since A_μ transforms as x_μ it is natural to transform A_0 as x_0 when we rotate to euclidean space. The "electric" field will then be given by:

$$E_i^{(E)} = \frac{\partial}{\partial x_4}A_i - \frac{\partial}{\partial x_i}A_4 = -iE_i^{(M)} \tag{6.21}$$

and the euclidean lagrangian and action will be:

$$\mathcal{L}_{(em)}^{(E)}(A) = \frac{1}{4}F_{\mu\nu}^{(E)}F_{\mu\nu}^{(E)} = \frac{1}{2}((E_i^{(E)})^2 + B_i^2) \tag{6.22}$$

$$S_{(em)}^{(E)}[A] = \int d^4x\ \mathcal{L}_{(em)}^{(E)}(A) \tag{6.23}$$

With these conventions for rotation to euclidean space the electromagnetic lagrangian satisfies the same rule as a scalar field: the functional form of the euclidean lagrangian is the same as the minkowskian hamiltonian.

In the following we will always work with the euclidean versions of $\mathcal{L}_{(fer)}(\psi,\bar{\psi},A)$, $\mathcal{L}_{(bos)}(\varphi,\varphi^*,A)$ and $\mathcal{L}_{(em)}(A)$.

6.2 Definition of the functional integral

One main problem in quantization of gauge theories is exposed by considering the free lagrangian $\mathcal{L}_{(em)}(A)$. The corresponding action (in euclidean space) is

$$S_{(em)}[A] = \frac{1}{2}\int d^dx\ A_\mu(x)(-\partial^2\delta_{\mu\nu} + \partial_\mu\partial_\nu)A_\nu(x) \tag{6.24}$$

It is a gaussian functional in A_μ, and from (6.22)-(6.23) it is clear that it is positive semi-definite functional. However, due to the gauge invariance (6.9) it has zero modes:

$$(-\partial^2\delta_{\mu\nu} + \partial_\mu\partial_\nu)\partial_\nu\chi = 0 \tag{6.25}$$

for any function χ and (6.25) means that we cannot define the Green function $G_{\mu\nu}(x,y)$ satisfying

$$[-\partial^2\delta_{\mu\nu} + \partial_\mu\partial_\nu]G_{\nu\lambda}(x,y) = \delta_{\mu\lambda}\delta^{(d)}(x-y) \tag{6.26}$$

The existence of such a Green function is necessary if we want to perform gaussian integrals of the kind

$$\begin{aligned} Z[J] &= \int \mathcal{D}A_\mu(x) \exp -\frac{1}{2}\int d^dx\ (A_\mu(-\partial^2\delta_{\mu\nu}+\partial_\mu\partial_\nu)A_\nu - J_\mu A_\mu) \\ \text{"} = \text{"} &\ Z[0]\ \exp\frac{1}{2}\int d^dx\ d^dy J_\mu(x)G_{\mu\nu}(x,y)J_\nu(y) \end{aligned} \tag{6.27}$$

We will now formulate the quantization procedure in a way which is slightly more complicated than necessary, but which allows us to generalize the procedure to non-abelian gauge groups. Let us *define* expectation values of *gauge invariant* observables $\mathcal{O}(A,\bar{\psi},\psi,\varphi^*,\varphi)$ by the *formal* expression:

$$\langle\mathcal{O}(A,\bar{\psi},\psi,\varphi^*,\varphi)\rangle = \frac{\int\mathcal{D}(A_\mu\psi\bar{\psi}\varphi\varphi^*)\ \mathcal{O}(A,\bar{\psi},\psi,\varphi^*,\varphi)e^{-S[A,\bar{\psi},\psi,\varphi^*,\varphi]}}{\int\mathcal{D}(A_\mu\psi\bar{\psi}\varphi\varphi^*)\ e^{-S[A,\bar{\psi},\psi,\varphi^*,\varphi]}} \tag{6.28}$$

For the reason mentioned above, both numerator and denominator are ill defined. In the following we will show that it is possible, by a number of formal manipulations, to write both numerator and denominator as well defined functional integrals times a formal product

$$\mathrm{Vol}(G_{inv}) = \prod_{x\in R^d}\int d\chi(x) \tag{6.29}$$

G_{inv} is the invariance group (6.12) and $\int d\chi(x)$ can be viewed as "volume" of the group of gauge transformations in x. This interpretation will be even more transparent in the non-abelian case.

In this way we will manage to cancel an infinite factor, independent of any dynamics, between numerator and denominator and we will be left with an effective action

$$S_{eff}[A,\bar{\psi},\psi,\varphi^*,\varphi]$$

which is no longer invariant under local gauge transformations, but to which we can add source terms and define generating functionals in standard fashion.

The virtue of this approach is that since the starting point, (6.28), is gauge invariant we know that we are bound to get gauge invariant results when we use $S_{eff}[A]$ in the calculation of expectation values of gauge invariant observables, even if $S_{eff}[A]$ is not gauge invariant. The disadvantage of the approach is its formal nature. We could have avoided that by first eliminating the superficial gauge degrees of freedom ("fixing the gauge"), but we would then have lost manifest gauge invariance. We will consider here only the first approach.

The first step is to impose a gauge condition. We will be interested in *covariant* conditions like

$$\partial_\mu A_\mu(x) = c(x) \tag{6.30}$$

in order not to break euclidean invariance. The main requirement for a gauge condition like (6.30) is that there for every field $A_\mu(x)$ exists one and only one field $A'_\mu(x) = A_\mu(x) + \partial_\mu\chi(x)$ among the class of gauge transformed fields of $A_\mu(x)$, such that (6.30) is satisfied [1]. This is illustrated in fig. 1. The gauge condition defines a subspace $\mathcal{M}[A]$

---

[1] This statement should be interpreted with some care: assume (6.30) for $A_\mu(x)$. If a gauge transformed $A'_\mu(x) \equiv A_\mu(x)+\partial_\mu\chi$ also satisfies (6.30) we have: $\partial^2\chi = 0$. Clearly this equation has solutions. However, it has *none* in euclidean space which fall off at infinity.

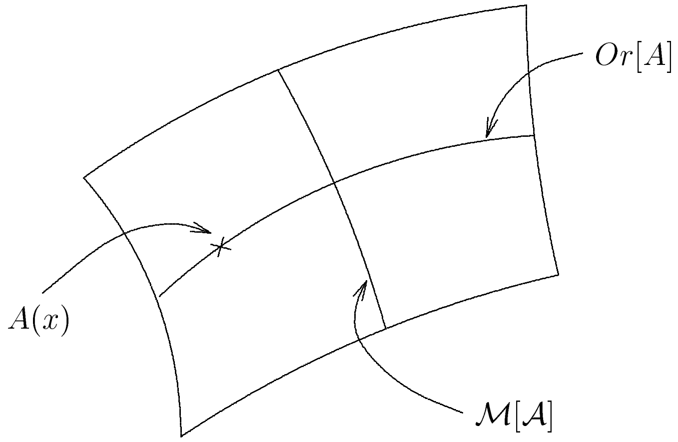


Figure 6.1: For a given configuration $A(x)$ the orbit $Or[A]$ should intersect $\mathcal{M}[A]$ only once.

in the space of field configurations. For a given field configuration $A_\mu(x)$ the *orbit* of A_μ ($Or[A]$) denotes the subspace of field configurations which can be obtained by a gauge transformation of A_μ. The requirement is that $Or[A]$ and $\mathcal{M}[A]$ intersect in one and only one point which we will denote $\chi^{(A)}$. This determines $\chi^{(A)}(x)$ as a functional of A. For all $A_\mu(x)$ we have

$$\partial^2\chi^{(A)}(x) = -\partial_\mu A_\mu(x) + c(x) \tag{6.31}$$

which has a unique solution under the usual assumption in euclidean space that $A_\mu(x)$ and $c(x)$ falls off sufficiently fast of infinity. Let ${}^\chi A_\mu$ denote the gauge transformed of $A_\mu(x)$ by $\chi(x)$:

$${}^\chi A_\mu(x) \equiv A_\mu(x) + \partial_\mu\chi(x). \tag{6.32}$$

Lemma: $$\int \prod_{x\in R^d} d\chi(x) \cdot \prod_{y\in R^d} \delta\left(-\partial_\mu({}^\chi A_\mu(y)) + c(y)\right) = \frac{1}{\det(-\partial^2)}$$

Proof:

By definition

$$-\partial_\mu({}^\chi A_\mu) + c = -\partial^2\chi - \partial_\mu A_\mu + c$$

and, by (6.31), a shift of integration variables $\chi \to \chi + \chi^A$ changes the l.h.s. of the equation in the lemma to

$$\int \prod_x d\chi(x) \prod_y \delta(-\partial^2\chi(y)) \tag{6.33}$$

For a finite, positive definite, symmetric matrix Δ_{ij} we have:

$$\int \prod_{i=1}^N d\chi_i \prod_i^N \delta(\Delta_{ij}\chi_j) = \frac{1}{\det\Delta} \tag{6.34}$$

as is seen by choosing an orthonormal basis. The proof formally generalizes to the infinite dimensional matrix $-\partial^2(x,y) \equiv -\partial_x^2\delta^{(d)}(x-y)$. This completes the proof.

As a result of the lemma we can write

$$\text{"1"} = \int \prod_x \; d\chi(x) \det(-\partial^2) \prod_y \delta\left(-\partial_\mu({}^\chi A_\mu(y)) + c(y)\right) \tag{6.35}$$

and we can get rid of the δ-function by integrating over the function $c(y)$. *Note that we in this way take a kind of average over many different gauge conditions of the form (6.30)*:

$$\begin{aligned} \text{"1}(\alpha)\text{"} &= \int \mathcal{D}\; c(x)\; e^{-\frac{1}{2\alpha}\int d^dx c^2(x)} \;.\; \text{"1"} \\ &= \int \prod_x d\chi(x) \det(-\partial^2)\; e^{-\frac{1}{2\alpha}\int d^dx(\partial_\mu({}^\chi A_\mu(x)))^2} \end{aligned} \tag{6.36}$$

The factor "1" from (6.35) have been inserted in the second line of (6.36) and the $c(x)$ integration performed.

By a number of formal manipulations we have produced an awfully complicated factor "$1(\alpha)''$, which depends on nothing but a constant α. Clearly we can multiply path integrals like $\int \mathcal{D}A\; e^{-S_{em}[A]}$ by this factor without changing expectation values like (6.28) and we can write (suppressing any dependence on ψ, φ for notational simplicity):

$$\int \mathcal{D}Ae^{-S[A]}\mathcal{O}(A)\;\text{"}1(\alpha)'' = \det(-\partial^2)\int\prod_x d\chi(x)\cdot\int \mathcal{D}A\; e^{-S[A]-\frac{1}{2\alpha}\int d^dx(\partial^\chi A)^2}\mathcal{O}(A) \tag{6.37}$$

If we consider only gauge invariant observables as in (6.28), a change of variables: $A_\mu \to A_\mu - \partial\chi$ (i.e. a gauge transformation without $-\chi$) will not affect $S[A]$ and $\mathcal{O}(A)$ since they are gauge invariant. But $\partial_\mu({}^\chi A_\mu) \to \partial_\mu A_\mu$. All reference to χ has disappeared inside the functional and we have managed to extract an infinite factor

$$\det(-\partial^2)\int\prod_x d\chi(x) \tag{6.38}$$

which is independent of any dynamics.

For the purpose of calculating gauge invariant quantities we can write:

$$\boxed{\langle\mathcal{O}(A,\bar\psi,\psi,\varphi^*,\varphi)\rangle = \frac{\int\mathcal{D}(A_\psi\bar\psi\varphi\varphi^*)\mathcal{O}(A,\bar\psi,\psi,\varphi^*,\varphi)e^{-S_{eff}(A,\bar\psi,\psi,\varphi^*\varphi)}}{\int\mathcal{D}(A\psi\bar\psi\varphi\varphi^*)e^{-S_{eff}(A,\bar\psi,\psi,\varphi^*,\varphi)}}} \tag{6.39}$$

where

$$\boxed{S_{eff}(A,\bar\psi,\psi,\varphi^*,\varphi) = \int d^dx\mathcal{L}_{em}(A) + \frac{1}{2\alpha}(\partial_\mu A_\mu)^2 + \mathcal{L}_{fer}(A,\bar\psi,\psi) + \mathcal{L}_{bos}(A,\varphi^*,\varphi)} \tag{6.40}$$

These formulae represent the final result of our formal manipulations[2]. The "effective" lagrangian

$$\mathcal{L}_{em}^{(eff)}(A) = \frac{1}{4}F_{\mu\nu}^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu)^2 \tag{6.41}$$

---

[2] At this point the difference between abelian and non-abelian gauge theories should be mentioned. In the case of non-abelian gauge theories one can repeat many of the above steps. The difference is that $\int\prod_x d\chi(x)$ is replaced by $\int\prod_x dU(x)$, where the group element U refers to the non-abelian gauge group and dU is the Haar measure for the group. Further the factor $\det(-\partial^2)$ will be replaced by $\det(-\partial_\mu D_\mu)$ where $D_\mu = \partial_\mu - iA_\mu$ is the covariant derivative. The determinant now has an explicit dependence on A_μ and cannot be taken outside the functional integral. This will lead to additional interactions in the effective action, which are conveniently handled by the introduction of so-called "ghosts".

and the corresponding "effective" action

$$S_{em}^{(eff)}[A] = \frac{1}{2}\int d^dx A_\mu \left(-\partial^2\delta_{\mu\nu} + (1-\frac{1}{\alpha})\partial_\mu\partial_\nu\right) A_\nu \tag{6.42}$$

is no longer gauge invariant, but as already discussed it will not affect the calculation of gauge invariant observables. *The important point is that* $S_{em}^{(eff)}[A]$ *is a gaussian functional which is invertible.* In fact we can write:

$$S_{em}^{(eff)}(A) = \frac{1}{2}\int \frac{d^dk}{(2\pi)^d} A_\mu(-k)(G^{(0)})^{-1}_{\mu\nu}(k)A_\nu(k) \tag{6.43}$$

$$(G^{(0)})^{-1}_{\mu\nu}(k) = k^2\delta_{\mu\nu} - (1-\frac{1}{\alpha})k_\mu k_\nu \tag{6.44}$$

and $G^{(0)}_{\mu\nu}(k)$ has to satisfy

$$\left(k^2\delta_{\mu\nu} - (1-\frac{1}{\alpha})k_\mu k_\nu\right) G^{(0)}_{\nu\lambda} = \delta_{\mu\lambda} \tag{6.45}$$

The solution is

$$G^{(0)}_{\mu\nu}(k) = \frac{1}{k^2}\left(\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right) \tag{6.46}$$

It is seen that the propagator simplifies if we choose $\alpha = 1$ (a choice called Feynman gauge, although it does strictly speaking not correspond to a gauge choice in the sense (6.30)). In practical calculations we will use $\alpha = 1$.

6.3 The Ward-Takahashi identities

The obstacle for defining perturbation theory was removed in the last section and using the effective action (6.40) we can proceed in the standard fashion and define generating functionals for the full -, the connected - and the $1PI$ Green functions. For notational simplicity we will confine ourselves to consider the coupling between fermions and gauge fields, i.e. quantum electrodynamics (QED). The additional coupling of the gauge field to a scalar field can be worked out in a straight forward manner. The partition function, or the generating functional for the full Green functions is

$$Z[J,\xi,\bar{\xi}] = \int \mathcal{D}A\mathcal{D}\psi\mathcal{D}\bar{\psi}e^{-S_{eff}[A,\psi,\bar{\psi};J,\xi,\bar{\xi}]} \tag{6.47}$$

$$S_{eff} = \int d^dx \, \left[\mathcal{L}_{eff}(A,\bar{\psi},\psi) - J_\mu A_\mu - \bar{\xi}\psi - \bar{\psi}\xi\right] \tag{6.48}$$

$$\mathcal{L}_{eff}(A,\bar{\psi},\psi) = \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu)^2 + \bar{\psi}(\not{D} + m)\psi \tag{6.49}$$

As usual $\psi, \bar{\psi}, \xi, \bar{\xi}$ are anticommuting Grassmann variables which carry spinor indices which we have not written explicitly,

$$Z[J,\xi,\bar{\xi}] = e^{-F[J,\xi,\bar{\xi}]} \tag{6.50}$$

$$\Gamma[A^{(cl)},\bar{\psi}^{(cl)},\psi^{(cl)}] = F[J,\xi,\bar{\xi}] + \int d^dx \, \left[J_\mu(x)A^{(cl)}_\mu(x) + \bar{\xi}(x)\psi^{(cl)}(x) + \bar{\psi}^{(cl)}(x)\xi(x)\right] \tag{6.51}$$

for the generating functional for connected Green functions(F) and $1PI$ Green functions (Γ). We have (left derivatives only)

$$\frac{\delta F}{\delta J_\mu} = -A_\mu^{(cl)} \quad , \quad \frac{\delta F}{\delta \bar{\xi}} = -\psi^{(cl)} \quad , \quad \frac{\delta F}{\delta \xi} = \bar{\psi}^{(cl)} \tag{6.52}$$

$$\frac{\delta \Gamma}{\delta A_\mu^{(cl)}} = J_\mu \quad , \quad \frac{\delta \Gamma}{\delta \psi^{(cl)}} = -\bar{\xi} \quad , \quad \frac{\delta \Gamma}{\delta \bar{\psi}^{(cl)}} = \xi \tag{6.53}$$

It is important to realize that the Green functions which we construct by functional derivatives are in general not gauge invariant, since Z, F and Γ are not gauge invariant due to the presence of source terms and the gauge fixing term. They are nevertheless important quantities, both as tools for calculating gauge invariant observables and for proving renormalizability of the theory in $d \leq 4$. The explicit breaking of gauge invariance leads to relations between various Green functions, known as Ward-Takahashi identities. They are easily derived from (6.47) by noting that although (6.49) is no longer invariant under local gauge transformations, the measure $\mathcal{D}A\mathcal{D}\psi\mathcal{D}\bar{\psi}$ *is* invariant under such transformations [3]. By performing a change in integration variables corresponding to an infinitesimal gauge transformation $e^{ie\epsilon(x)} \approx 1 + ie\epsilon(x) + 0(\epsilon(x)^2)$:

$$\begin{aligned} \psi'(x) &\approx (1 + ie\ \epsilon(x))\,\psi(x) && (6.54) \\ A'_\mu(x) &= A_\mu(x) + \partial_\mu\ \epsilon(x) && (6.55) \end{aligned}$$

$Z[J, \xi, \bar{\xi}]$ itself will not change since we only change integration variables. On the other hand the parts of S_{eff} which involve the source terms and $(\partial_\mu A_\mu)^2$ *will* change, and to first order in $\epsilon(x)$ we have

$$0 = \delta Z = \int \mathcal{D}(A\psi\bar{\psi})\ e^{-S_{eff}} \int d^dx\ \left[J_\mu \partial_\mu \epsilon + ie\ \epsilon\ (\bar{\xi}\psi - \bar{\psi}\xi) - \frac{1}{\alpha}(\partial_\mu A_\mu)\partial^2\epsilon\right] \tag{6.56}$$

or (performing the functional differentiation with respect to $\epsilon(x)$)

$$0 = \frac{\delta Z}{\delta \epsilon(x)} = -\frac{1}{\alpha}\ \partial^2 \partial_\mu\ \frac{\delta Z}{\delta J_\mu} - \partial_\mu J_\mu(x) \cdot Z + ie\left(\bar{\xi}\ \frac{\delta Z}{\delta \bar{\xi}} + \frac{\delta Z}{\delta \xi}\xi\right). \tag{6.57}$$

This equation can be expressed in terms of F:

$$\boxed{\frac{1}{\alpha}\ \partial^2 \partial_\mu \frac{\delta F}{\delta J_\mu} - \partial_\mu J_\mu - ie\left(\bar{\xi}\ \frac{\delta F}{\delta \bar{\xi}} + \frac{\delta F}{\delta \xi}\xi\right) = 0} \tag{6.58}$$

or (using (6.52) and (6.53))

$$\boxed{-\frac{1}{\alpha}\ \partial^2 \partial_\mu A_\mu^{(cl)} - \partial_\mu\ \frac{\delta \Gamma}{\delta A_\mu^{(cl)}} - ie\left(\frac{\delta \Gamma}{\delta \psi^{(cl)}}\ \psi^{(cl)} + \bar{\psi}^{(cl)} \frac{\delta \Gamma}{\delta \bar{\psi}^{(cl)}}\right) = 0} \tag{6.59}$$

Eq. (6.58) is the generalized form of the so-called Ward-Takahashi identity in QED. The corresponding identities for $1PI$ functions are sometimes called the Lee–Zinn-Justin identities.

We can use (for instance) (6.59) to derive relations between different $1PI$ Green functions by successive functional differentiation after $A_\mu, \psi, \bar{\psi}$.

---

[3] $\psi'(x) = e^{ie\chi(x)}\psi(x)$ is a unitary transformation, while $A'_\mu(x) = A_\mu(x) + \partial_\mu \chi(x)$ is a translation.

6.3.1 The photon propagator

Differentiating (6.59) once after $A_\nu^{(cl)}(y)$ leave us with

$$-\frac{1}{\alpha}\frac{\partial^2}{\partial x_\mu^2}\frac{\partial}{\partial x_\nu}\delta^{(d)}(x-y)-\frac{\partial}{\partial x_\mu}\left.\frac{\delta^2\Gamma}{\delta A_\mu^{(cl)}(x)\delta A_\nu^{(cl)}(y)}\right|_{A^{(cl)},\psi^{(cl)},\bar{\psi}^{(cl)}}=0. \tag{6.60}$$

Recall from chapter 4 that the inverse of the photon propagator $G_{\mu\nu}(x,y)$ is

$$G_{\mu\nu}^{-1}(x,y)=\frac{\delta^2\Gamma}{\delta A_\mu^{(cl)}(x)A_\nu^{(cl)}(y)} \tag{6.61}$$

and let us write

$$G_{\mu\nu}^{-1}=\left(G^{(0)}\right)_{\mu\nu}^{-1}+\Pi_{\mu\nu} \tag{6.62}$$

where $\Pi_{\mu\nu}$ is the *real* $1PI$ part of the two-point function[4] and $G_{\mu\nu}^{(0)}$ is the free propagator, given in momentum space by (6.46). By Fourier transformation (6.60) reads

$$k_\mu\left(\frac{1}{\alpha}\,k^2\delta_{\mu\nu}-[(G^{(0)})_{\mu\nu}^{-1}(k)+\Pi_{\mu\nu}(k)]\right)=0 \tag{6.63}$$

If we recall (6.44): $(G^{(0)})_{\mu\nu}^{-1}=k^2\delta_{\mu\nu}-k_\mu k_\nu+\frac{1}{\alpha}k_\mu k_\nu$ we see that the gauge dependent part which refers to α cancels and we get:

$$\boxed{k_\mu\Pi_{\mu\nu}(k)=0} \tag{6.64}$$

This equation tells us that the radiative corrections to the photon propagator are *purely transverse* and we can write

$$\boxed{\Pi_{\mu\nu}(k)=(k^2\delta_{\mu\nu}-k_\mu k_\nu)\Pi(k^2)} \tag{6.65}$$

The transversality of $\Pi_{\mu\nu}$ further shows that it is *gauge invariant* since a gauge transformation in momentum space has the form $A'_\mu(k)=A_\mu(k)-ik_\mu\chi(k)$. We will return to the importance of (6.65) when counting divergencies in QED.

6.3.2 The Ward-Takahashi identity

By differentiating (6.59) with respect to $\bar{\psi}_\alpha(y)$ and $\psi_\beta(z)$ we get

$$\begin{aligned}0 \;=\; & \left[-\frac{\partial}{\partial x_\mu}\frac{\delta^3\Gamma}{\delta\psi_\beta^{(cl)}(z)\delta\bar{\psi}_\alpha^{(cl)}(y)\delta A_\mu(x)}-\right.\\ & \left.\left.-\,ie\delta^{(d)}(x-y)\frac{\delta\Gamma}{\delta\psi_\beta^{(cl)}(z)\delta\bar{\psi}_\alpha^{(cl)}(x)}+ie\delta^{(d)}(x-z)\frac{\delta\Gamma}{\delta\psi_\beta^{(cl)}(x)\delta\bar{\psi}_\alpha^{(cl)}(y)}\right]\right|_{\psi=\bar{\psi}=A=0}\end{aligned} \tag{6.66}$$

It is a useful exercise to check that (6.66) is actually satisfied if we use the lowest order approximation $\Gamma^{(0)}$ for the generating functional. The lowest order approximation is of course nothing but the action itself:

$$\Gamma^{(0)}[A^{(cl)},\bar{\psi}^{(cl)},\psi^{(cl)}]=S[A^{(cl)},\bar{\psi}^{(cl)},\psi^{(cl)}] \tag{6.67}$$

---

[4] Recall this peculiarity of Γ. Precisely for the two-point function $\delta^2\Gamma/\delta A^2$ it does *not* generate the $1PI$ part but the inverse propagator which has the decomposition (6.62).

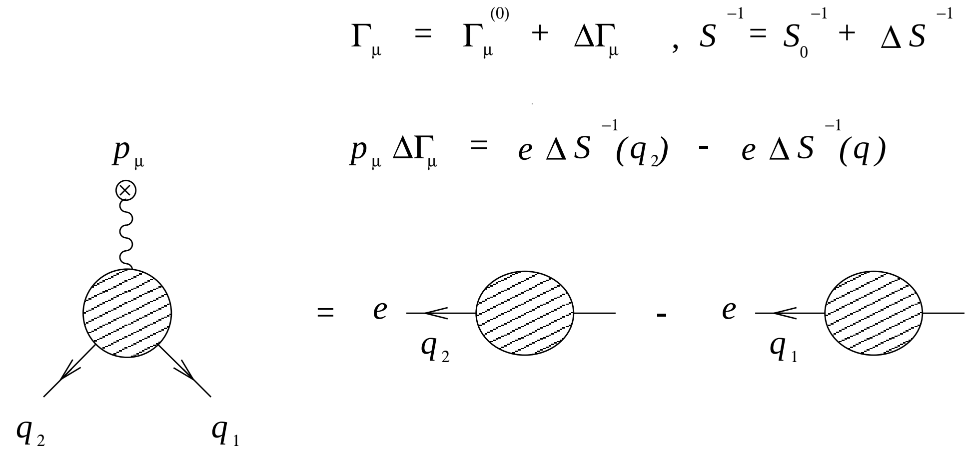


Figure 6.2: The Ward-Takahashi identity for 1PI Green functions

However, (6.66) is valid to all orders and if we use the definition of three-point vertex function and two-point function (the inverse propagator)

$$S_F^{-1}(x,y)_{\alpha\beta} \equiv \left.\frac{\delta^2\Gamma}{\delta\psi_\beta^{(cl)}(y)\delta\bar{\psi}_\alpha^{(cl)}(x)}\right|_{\psi^{(cl)},\bar{\psi}^{(cl)},A^{(cl)}=0} \tag{6.68}$$

$$\Gamma_{\mu;\alpha\beta}(x;y,z) \equiv \left.\frac{\delta^3\Gamma}{\delta\psi_\beta^{(cl)}(z)\delta\bar{\psi}_\alpha^{(cl)}(y)\delta A_\mu^{(cl)}(x)}\right|_{\psi^{(cl)},\psi^{(cl)},A^{(cl)}=0} \tag{6.69}$$

we can write:

$$-\frac{\partial}{\partial x_\mu}\Gamma_{\mu;\alpha\beta}(x;y,z) = -ie\delta^{(d)}(x-z)S_F^{-1}(y-x) + ie\delta^{(d)}(x-y)S_F^{-1}(x-z) \tag{6.70}$$

or by Fourier transformation:

$$\boxed{p_\mu\Gamma_{\mu\alpha\beta}(p;q_2,q_1) = e\left(S_F^{-1}(q_2)_{\alpha\beta} - S_F^{-1}(q_1)_{\alpha\beta}\right)} \tag{6.71}$$

where the photon momentum $p = q_2 - q_1$ by momentum conservation. This relation is illustrated in fig. 2 and it allows us to relate the divergent part of the fermion propagator to the divergent part of the photon-fermion-fermion vertex. We will return to this shortly.

6.3.3 The n-photon vertex function ($n \geq 3$).

Differentiating (6.59) after $A_{\mu_2}^{(cl)}(x_1), \cdots, A_{\mu_n}^{(cl)}(x_n)$ leads to

$$\frac{\partial}{\partial x_{\mu_1}} \left.\frac{\delta^{(n)}\Gamma}{\delta A_{\mu_1}^{(cl)}(x_1)\cdots\delta A_{\mu_n}^{(cl)}(x_n)}\right|_{A^{(cl)}=\bar{\psi}^{(cl)}=\psi^{(cl)}=0} = 0 \tag{6.72}$$

or by Fourier transformation

$$k_{\mu_1}^{(1)}\Gamma_{\mu_1,\cdots,\mu_n}^{(n)}(k^{(1)},\cdots,k^{(n)}) = 0 \tag{6.73}$$

Since $\Gamma^{(n)}_{\mu_1,\cdots,\mu_n}$ is symmetric such a relation is valid for any $k^{(i)}$ and it has importance when counting the superficial divergence of n-photon vertex diagrams. The superficial degree of divergence of the n-photon vertex function is in $d = 4$ (see later)

$$\omega(D) = 4 - n \tag{6.74}$$

which indicates that not only the two-point function but also for instance the four point function should be superficial divergent. This fact would be worrisome for the program of renormalization since we have no candidate for counterterms of order $A^4_\mu(x)$ in the lagrangian. However, precisely as the transversality constraint (6.64) for the two-photon function allowed us to extract two "kinematical" powers of k_μ, as shown explicitly in eq. (6.65), and in this way reduce the divergence of $\Pi_{\mu\nu}$ by two, it can be shown that (6.73) allows us to extract a "kinematical" power of n and in this way reduce the effective superficial divergence from $4-n$ to $4-2n$. We conclude that only the two-point function is superficially divergent. Note that individual four-photon diagrams could be superficially divergent, but then we add all diagrams to a given order these divergencies will cancel provided that every step in the about formal manipulations has been well defined. *This obviously requires that the regularization we use is gauge invariant*, since we have used the "naive" transformation properties of the effective action under gauge transformations. Dimensional regularization satisfies this requirement.

Finally we note that in fact the $(2n+1)$*-photon vertex function vanishes identically* (Furry's theorem). This is a consequence of the invariance of the QED lagrangian under *charge conjugation*. In order to define charge conjugation we note the following theorem: For any two $[d/2]$-dimensional (hermitean) representations γ_μ and γ'_μ of the Clifford algebra there exists a unitary transformation (i.e. a $[d/2] \times [d/2]$ matrix) U such that

$$\gamma_\mu = U\gamma'_\mu U^{-1} \tag{6.75}$$

If γ_μ's satisfy the Clifford algebra the same will be true for minus the transposed matrices $-\gamma_\mu^T$'s. Let us denote the corresponding matrix U from (6.75) by C:

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu. \tag{6.76}$$

We now define the charge conjugation transformation as follows:

$$\begin{aligned} \psi &\to \psi' &&= C\bar{\psi}^T \\ \bar{\psi} &\to \bar{\psi}' &&= -\psi^T C^{-1} &&(6.77)\\ A_\mu &\to A'_\mu &&= -A_\mu. &&(6.78)\end{aligned}$$

From this we get, using the anticommuting nature of the fermionic variables, the following transformations:

$$\begin{aligned} \bar{\psi}_2\psi_1 &\to \bar{\psi}'_2\psi'_1 = -\psi_2^T\bar{\psi}_1^T = \bar{\psi}_1\psi_2 &&(6.79)\\ \bar{\psi}_2\gamma_\mu\psi_1 &\to \bar{\psi}'_2\gamma_\mu\psi'_1 = \psi_2^T\gamma_\mu^T\bar{\psi}_1^T = -\bar{\psi}_1\gamma_\mu\psi_2. &&(6.80)\end{aligned}$$

This implies the following transformations of the non-trivial terms in the lagrangian: $\bar{\psi}\psi$ and $\bar{\psi}\not{A}\psi$ are invariant, while

$$\bar{\psi}\gamma_\mu(\partial_\mu\psi) \to -(\partial_\mu\bar{\psi})\gamma_\mu\psi. \tag{6.81}$$

But a partial integration leads to invariance of the *action* term associated with $\bar{\psi}\gamma_\mu\partial_\mu\psi$ [5]. It is now clear that the action is invariant under charge conjugation. The measure in the functional integral is also invariant up to an (infinite) product of -1's which however is independent of any dynamics and cancels between numerator and denominator in the functional integrals of $\langle A_{\mu_1}A_{\mu_2}\cdots A_{\mu_n}\rangle$. The change of variables (6.77) now gives:

$$\langle A_{\mu_1}(x_1)\cdots A_{\mu_n}(x_n)\rangle = (-1)^n\langle A_{\mu_1}(x_1)\cdots A_{\mu_n}(x_n)\rangle \tag{6.82}$$

which tells us that Green functions with an odd number of external photon lines and no external fermion lines must vanish.

Let us remark that in the case of non-abelian gauge theories charge conjugation does not act as simple on the gauge fields as $A_\mu \to -A_\mu$. Consequently, the $(2n+1)$-gauge field vertex function does not vanish. In addition relations like (6.73) are not satisfied and both the 3-gauge field vertex function and the 4-gauge field vertex functions are superficially divergent. But in this case it is not a disaster since the lagrangian contains both A^3 and A^4 interactions.

6.4 Feynman rules and one-loop renormalizability

6.4.1 General remarks

After the genenal words of wisdom in the last section we will turn to the actual calculations at one-loop level.

The free propagators for QED are read off from the gaussian part of $S_{eff}[A,\bar{\psi},\psi]$:

$$\begin{aligned} S^{(2)}_{em}[A] &= \frac{1}{2}\int \frac{d^dk}{(2\pi)^d}\, A_\mu(-k)(G^{(0)})^{-1}_{\mu\nu}(k)A_\mu(k) && (6.83)\\ S^{(2)}_{fer}[\bar{\psi},\psi] &= \int \frac{d^dk}{(2\pi)^d}\, \bar{\psi}(-k)_\alpha(S^{(0)})^{-1}_{\alpha\beta}(k)\psi_\beta(k) && (6.84)\end{aligned}$$

where

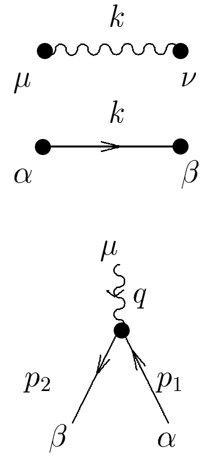
$$\begin{aligned} (S^{(0)})^{-1}(k) &= -i\, \not{k} + m, && (6.85)\\ (G^{(0)})^{-1}_{\mu\nu}(k) &= k^2\delta_{\mu\nu} - (1-\frac{1}{\alpha})k_\mu k_\nu && (6.86)\end{aligned}$$

or (as already discussed)

$$\begin{aligned} S^{(0)}(k) &= \frac{i}{\not{k}+im} = \frac{i\,\not{k}+m}{k^2+m^2} && (6.87)\\ G^{(0)}_{\mu\nu}(k) &= \frac{1}{k^2}\left(\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right) && (6.88)\end{aligned}$$

The interaction is

$$S^{(int)}[A,\bar{\psi},\psi] = -ie\int d^dx\; \bar{\psi}(x)\gamma_\mu A_\mu(x)\psi(x) \tag{6.89}$$



$$G^{(0)}_{\mu\nu}(k) = \frac{1}{k^2}\left(\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right)$$

$$S^{(0)}_F(k)_{\beta\alpha} = \left(\frac{i}{\not{k} + im}\right)_{\beta\alpha}$$

$$ie(\gamma_\mu)_{\beta\alpha}$$

Figure 6.3: Feynman rules for QED in euclidean space.

and the corresponding interaction vertex of the Feynman graphs will get a factor $(ie\gamma_\mu)_{\alpha\beta}$. The rules for Feynman graphs are shown in fig. 3.

We can analyze the divergencies in the same way as for scalar theories. Let us here concentrate on the theory when the dimension of space-time is four. The coupling constant e is dimensionless in four dimensions:

$$[e] = 0 \tag{6.90}$$

According to our analysis for scalar theories this is a *necessary* condition for renormalizability and this conclusion is not changed by the presence of fermions. The counting of superficial divergencies is slightly different, though. The fermion propagator falls off like $1/k$, not $1/k^2$ as the bosonic propagator. Repeating the arguments in chapter 5 we say that the superficial divergence $\omega(D)$ of a $1PI$ diagram D with I_{ph} internal photon lines and I_f internal fermions lines is

$$\omega(D) = dL - 2I_{ph} - I_f \tag{6.91}$$

L denotes the number of loops and d as usual the dimension of space-time. Eq. (6.91) is nothing but powercounting in the corresponding Feynman integral. If E_{ph} and E_f denote the number of external photon and fermion lines and V the number vertices in the diagram D we have

$$2I_{ph} + E_{ph} = V \qquad , \qquad 2I_f + E_f = 2V \tag{6.92}$$

since each vertex has one photon line and two fermion lines, and since internal lines will be counted twice when counting lines connect to vertices. In addition we have

$$L - (I_{ph} + I_f) + V = 1 \tag{6.93}$$

For $d = 4$ eqs. (6.91), (6.92), (6.93) imply:

$$\boxed{\omega(D) = 4 - E_{ph} - \frac{3}{2}E_f} \tag{6.94}$$

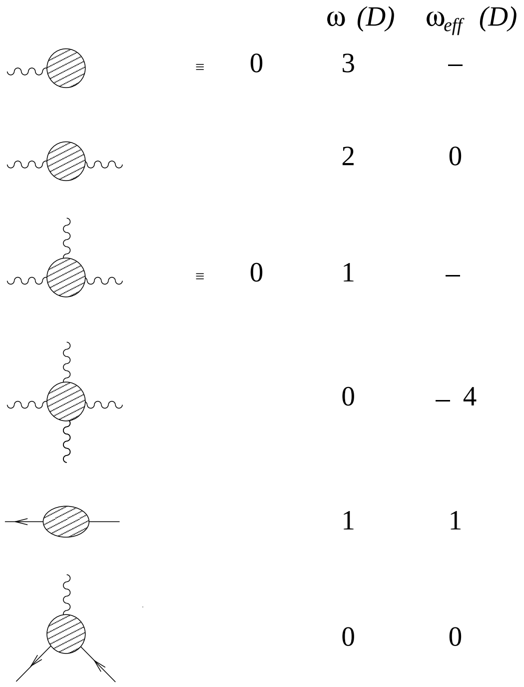


Figure 6.4: Superficial divergent 1PI graphs in QED

This leaves us the superficially divergent $1PI$ diagrams shown in fig. 4. From charge conjugation invariance it follows that the one- and three-photon vertex functions are identical zero, as already mentioned. Further we know from the Ward-Takahashi identities discussed in the last section that the real superficial divergence of the photon self-energy is not two, but zero and the real superficial divergence of the four-photon vertex functions is not zero but minus four. This reduction of divergence was proven only for the sum of all diagrams to a given order, and in addition we have implicitly assumed that the regularization used, in order to make sense of the Feynman integral, respects the gauge invariance of the original lagrangian. This will be fulfilled by dimensional regularization. Finally the Ward-Takahashi showed that the radiative corrections to the photon energy self-energy are purely transverse.

In the next subsection we will calculate the divergent one-loop diagrams and show that the pole terms can be cancelled by adding counterterms which are local polynomials in fields $A_\mu, \bar{\psi}, \psi$ of a structure already present in the original lagrangian. The transversality of the photon self-energy shows that there is no counterterms associated with the gauge fixing part $(\partial_\mu A_\mu)^2$ of $S^{(eff)}[A, \bar{\psi}, \psi]$. Anticipating the results of the next subsection we can therefore write for the counterterms: ($\varepsilon \equiv 4 - d$)

$$\delta\mathcal{L}_{A^2} = (Z_3 - 1)\tfrac{1}{4}F_{\mu\nu}^2, \quad Z_3 - 1 = -\frac{e^2}{12\pi^2}\frac{2}{\varepsilon} \tag{6.95}$$

$$\delta\mathcal{L}_{\bar{\psi}\partial\!\!\!/\psi} = (Z_2 - 1)\bar{\psi}\,\partial\!\!\!/\psi, \quad Z_2 - 1 = -\frac{e^2}{16\pi^2}\frac{2}{\varepsilon} \tag{6.96}$$

$$\delta\mathcal{L}_{\bar{\psi}A\!\!\!/\psi} = (Z_1 - 1)\bar{\psi}\,A\!\!\!/\psi, \quad Z_1 - 1 = -\frac{e^2}{16\pi^2}\frac{2}{\varepsilon} \tag{6.97}$$

---

[5] Alterntively one can write the kinetic term as $\frac{1}{2}(\bar{\psi}\gamma_\mu\partial_\mu\psi - (\partial_\mu\bar{\psi})\gamma_\mu\psi)$. It is then explicitly invariant under charge conjugation, and it differs only from $\bar{\psi}\,\partial\!\!\!/\psi$ by a total derivative term $\frac{1}{2}\partial_\mu(\bar{\psi}\gamma_\mu\psi)$, which does not contribute to the action due to the boundary conditions

$$\delta\mathcal{L}_{m\bar{\psi}\psi} = (Z_m - 1)m\bar{\psi}\psi, \quad Z_m - 1 = -\frac{e^2}{4\pi^2}\frac{2}{\varepsilon} \tag{6.98}$$

and the renormalized 1-loop QED lagrangian will be given by

$$\begin{aligned} \mathcal{L}^{(ren)} &= \mathcal{L}^{(eff)} + \delta\mathcal{L} \\ &= \frac{Z_3}{4}F_{\mu\nu}^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu)^2 + Z_2\bar{\psi}\,\partial\!\!\!/\psi + Z_m m\bar{\psi}\psi + Z_1(-ie\bar{\psi}\,A\!\!\!/\psi) \end{aligned} \tag{6.99}$$

It follows that we can regard $\mathcal{L}^{(ren)}$ as a *bare* lagrangian $\mathcal{L}_0$:

$$\mathcal{L}^{(ren)}(A, \bar{\psi}, \psi, e, \alpha, m) = \mathcal{L}_0^{(eff)}(A_0, \bar{\psi}_0, \psi_0, e_0, \alpha_0, m_0) \tag{6.100}$$

by the following multiplicative renormalization:

$$\begin{aligned} A_0 &= \sqrt{Z_3}A \\ \psi_0 &= \sqrt{Z_2}\psi \\ \bar{\psi}_0 &= \sqrt{Z_2}\bar{\psi} \\ e_0 &= \frac{Z_1}{Z_2\sqrt{Z_3}}\cdot e \\ \alpha_0 &= Z_3\alpha \\ m_0 &= \frac{Z_m}{Z_2}\cdot m \end{aligned} \tag{6.101}$$

Note that the explicit calculation (6.96) and (6.97) shows that

$$Z_1 = Z_2 \tag{6.102}$$

This result is valid beyond the one-loop approximation. The Ward-Takahashi identity (6.71) shows the *divergent* parts of Z_1 and Z_2 by consistency have to satisfy

$$Z_1^{[div]} = Z_2^{[div]} \tag{6.103}$$

and since MS only use the divergent part in its definition of $Z_{1,2}$ we could have anticipated (6.102) and we get

$$e_0 = \frac{1}{Z_3^{\frac{1}{2}}}e \tag{6.104}$$

The correct way to write this relation in $d = 4 - \varepsilon$ dimensions, where e strictly speaking should be replaced by $\tilde{e} = \mu^{\varepsilon/2}e$, $[e] = 0$, is (as discussed in detail in chapter 5)

$$e_0 = \mu^{\varepsilon/2}e\left(1 + \frac{e^2}{24\pi^2}\,\frac{2}{\varepsilon} + \mathcal{O}(e^3)\right) \tag{6.105}$$

from which we get, using $\mu\frac{\partial}{\partial\mu}e_0 = 0$,

$$\mu\frac{\partial e}{\partial\mu} \equiv \beta(e) = \frac{e^3}{12\pi^2} + \mathcal{O}(e^5) \tag{6.106}$$

As discussed in chapter 5 this shows that $e = 0$ is an infrared fixpoint and we can solve (6.106) to get the *effective* coupling constant $\bar{e}^2(t)$ at scale $t\mu_0$ expressed in terms of the renormalized coupling constant $e(\mu_0)$ at scale μ_0

$$\bar{e}^2(t) = \frac{e_0^2}{1 - \frac{e_0^2}{12\pi^2}\ln t^2} \quad , \qquad e_0 = e(\mu_0) \tag{6.107}$$

We see that $\bar{e}(t)$ grows with t. We can of course only trust (6.107) as long as $\frac{e_0^2}{12\pi}\ln t^2 \ll 1$, but if the result is nevertheless extrapolated to large t it hints the breakdown of perturbation theory when $e(t) \to \infty$ at the *Landau pole* $t_0 = e^{6\pi^2/e_0^2}$. We leave it as an exercise to assume the validity of perturbation theory at (say) ev scale and with $e_0^2/4\pi = 1/137$, and calculate the energy scale at which the Landau pole is located.

6.4.2 γ matrices in d dimensions

We will use dimensional regularization in the calculation and one problem which arises is the analytic continuation of the γ_μ-matrices and the associated Clifford algebra to d dimensions. As mentioned above the lowest dimensional representation of d γ-matrices satisfying the Clifford algebra is $2^{[d/2]}$ when d is an integer ≥ 2, and this dimension is not analytic. Two approaches have been used: to keep the dimension of the γ_μ matrices fixed, i.e. to four if we work in four dimensions or to take it to $2^{d/2}$. We will choose the first convention and keep the dimension of the γ_μ matrices fixed to some appropriate value, depending on the dimension of space-time in which we want to consider the theory, i.e. if we want to do calculations for two-dimensional QED we would take the dimension to be two etc. We will now derive some useful formulae.

The basic relation is

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} \cdot I \tag{6.108}$$

where I is the unit-matrix of appropriate dimension as discussed above. If we consider four dimensions we have Tr $I = 4$ even when continued outside $d = 4$. Two relations follow immediately from (6.108):

$$\gamma_\nu\gamma_\nu = d \cdot I$$

$$\gamma_\nu(\gamma_{\mu_1}\cdots\gamma_{\mu_n})\gamma_\nu = 2\gamma_{\mu_n}\gamma_{\mu_1}\cdots\gamma_{\mu_{n-1}} - \gamma_\nu(\gamma_{\mu_1}\cdots\gamma_{\mu_{n-1}})\gamma_\nu\gamma_n$$

Combining these relations for $n = 1, 2$ and 3 leads to

$$\begin{aligned}
\gamma_\nu\gamma_\mu\gamma_\nu &= (2-d)\gamma_\mu \\
\gamma_\nu\gamma_{\mu_1}\gamma_{\mu_2}\gamma_\nu &= 4\delta_{\mu_1\mu_2} \cdot I + (d-4)\gamma_{\mu_1}\gamma_{\mu_2} \\
\gamma_\nu\gamma_{\mu_1}\gamma_{\mu_2}\gamma_{\mu_3}\gamma_\nu &= -2\gamma_{\mu_3}\gamma_{\mu_2}\gamma_{\mu_1} - (d-4)\gamma_{\mu_1}\gamma_{\mu_2}\gamma_3
\end{aligned}$$

Another relation with can be derived from (6.108) is

$$\begin{aligned}
\mathrm{Tr}\,\gamma_{\mu_1}\cdots\gamma_{\mu_{2n}} &= \delta_{\mu_1\mu_2}\mathrm{Tr}\,\gamma_{\mu_3}\cdots\gamma_{\mu_{2n}} - \delta_{\mu_1\mu_3}\mathrm{Tr}\,\gamma_{\mu_2}\gamma_{\mu_4}\cdots\gamma_{\mu_{2n}} + \cdots \\
&\quad + \delta_{\mu_1\mu_{2n}}\mathrm{Tr}\,\gamma_{\mu_2}\cdots\gamma_{\mu_{2n-1}}
\end{aligned}$$

and iterating this equation leads to

$$\begin{aligned}
\mathrm{Tr}\,\gamma_{\mu_1}\gamma_{\mu_2} &= \delta_{\mu_1\mu_2}\mathrm{Tr}\,I \\
\mathrm{Tr}\,\gamma_{\mu_1}\gamma_{\mu_2}\gamma_{\mu_3}\gamma_{\mu_4} &= (\delta_{\mu_1\mu_2}\delta_{\mu_3\mu_4} - \delta_{\mu_1\mu_3}\delta_{\mu_2\mu_4} + \delta_{\mu_1\mu_4}\delta_{\mu_2\mu_3})\ \mathrm{Tr}\,I
\end{aligned}$$

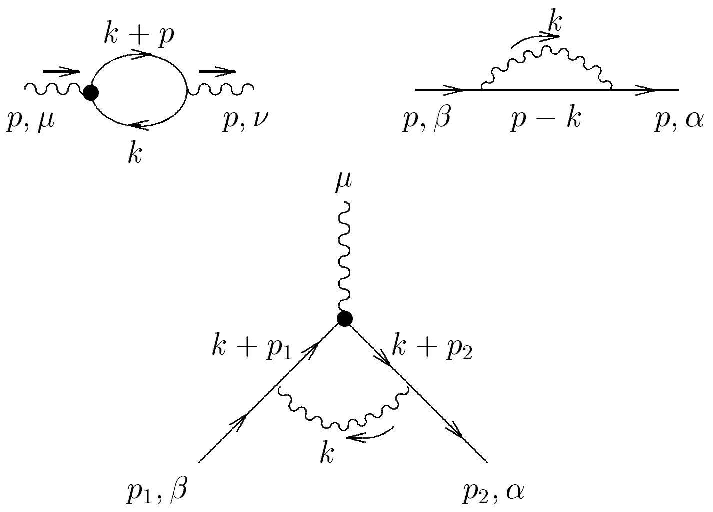


Figure 6.5: The divergent one-loop diagrams in *QED*

The final useful observation is that

$$\mathrm{Tr}\ \gamma_{\mu_1} \cdots \gamma_{\mu_{2n+1}} = 0$$

since repeated application of the basic relation (6.108) allows us to reduce the trace of an odd number of γ matrices to that of single matrices. But they have trace zero since $\gamma_\mu \gamma_\nu = -\gamma_\nu \gamma_\mu$ for $\nu \neq \mu$ implies (no summation over repeated indices):

$$\mathrm{Tr}\ \gamma_\nu \gamma_\mu \gamma_\nu = -\mathrm{Tr}\ \gamma_\mu \gamma_\nu^2 = -\mathrm{Tr}\ \gamma_\mu$$

while the cyclic property of traces gives

$$\mathrm{Tr}\ \gamma_\nu \gamma_\mu \gamma_\nu = \mathrm{Tr}\ \gamma_\mu \gamma_\nu^2 = \mathrm{Tr}\ \gamma_\mu .$$

6.4.3 The photon self-energy $\Pi_{\mu\nu}(p)$ and Z_3

The three divergent one-loop diagrams are shown in fig. 5. We will now discuss how to extract the divergent part of the diagrams, starting with the photon self-energy.

From the momentum assignment of fig. 5 and the Feynman rules we have

$$\begin{aligned} \Pi_{\mu\nu}(p) &= (-1) \cdot \int \frac{d^d k}{(2\pi)^d} (i\tilde{e}\gamma_\mu)_{\alpha\beta} \left(\frac{i}{\not{k} + \not{p} + im} \right)_{\beta\gamma} (i\tilde{e}\gamma_\nu)_{\gamma\delta} \left(\frac{i}{\not{k} + im} \right)_{\delta\alpha} \\ &= \tilde{e}^2 \int \frac{d^d k}{(2\pi)^d} \frac{Tr\ \gamma_\mu [i(\not{k} + \not{p}) + m] \gamma_\nu [i\not{k} + m]}{[(k+p)^2 + m^2][k^2 + m^2]} \end{aligned} \qquad (6.109)$$

In this formula we have as usual that $\tilde{e} = \mu^{\varepsilon/2} e$ where e is dimensionless and $\varepsilon = 4 - d$. We now follow the same path as for scalar theories and introduce a Feynman parameter α and a shift $k = q - \alpha p$ in loop momentum:

$$\Pi_{\mu\nu}(p) = e^2 \mu^\varepsilon \int_0^1 d\alpha \int \frac{d^d q}{(2\pi)^d} \frac{\mathrm{Tr}\ \gamma_\mu [i(\not{q} + (1-\alpha)\not{p}) + m] \gamma_\nu [i(\not{q} - \alpha \not{p}) + m]}{[q^2 + \alpha(1-\alpha)p^2 + m^2]^2} \qquad (6.110)$$

By symmetry the odd powers of q in the numerator vanish, and traces of odd power of γ-matrices vanish according the discussion above. In this way we are left with a numerator

$$-\left(q_\alpha q_\beta - \alpha(1-\alpha)p_\alpha p_\beta\right) \text{ Tr } \gamma_\mu\gamma_\alpha\gamma_\nu\gamma_\beta + m^2 \text{ Tr } \gamma_\mu\gamma_\nu. \tag{6.111}$$

We need to calculate two q integrals

$$\begin{aligned}
\int \frac{d^dq}{(2\pi)^d} \frac{1}{(q^2+\alpha(1-\alpha)p^2+m^2)^2} &= \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}[\alpha(1-\alpha)p^2+m^2]^{2-d/2}} \\
\int \frac{d^dq}{(2\pi)^d} \frac{q_\alpha q_\beta}{(q^2+\alpha(1-\alpha)p^2+m^2)^2} &= \frac{\delta_{\alpha\beta}}{d}\int \frac{d^dq}{(2\pi)^d} \frac{q^2}{(q^2+\alpha(1-\alpha)p^2+m^2)^2} \\
&= \frac{\delta_{\alpha\beta}\ [\Gamma(1-d/2)-\Gamma(2-d/2)]}{(4\pi)^{d/2}\ d\ [\alpha(1-\alpha)p^2+m^2]^{1-d/2}} \\
&= \frac{\delta_{\alpha\beta}\ \Gamma(2-d/2)}{(4\pi)^{d/2}\ (2-d)\ [\alpha(1-\alpha)p^2+m^2]^{1-d/2}}
\end{aligned}$$

In the second integral we have used the replacement $q_\alpha q_\beta \to \delta_{\alpha\beta}q^2/d$ which follows from symmetry.

If we combine the result of integration with the trace identities above we finally get (after some algebra)

$$\Pi_{\mu\nu}(p) = -(p^2\delta_{\mu\nu} - p_\mu p_\nu)\ 2\text{Tr } I\ \Gamma(\frac{\varepsilon}{2})\ \frac{e^2}{(4\pi)^2}\int d\alpha\ \frac{\alpha(1-\alpha)}{[(\alpha(1-\alpha)p^2+m^2)/(4\pi\mu^2)]^{\varepsilon/2}} \tag{6.112}$$

We see that the photon self-energy is transverse in accordance with the general Ward-Takahashi identities. We can now expand in ε and get (to order $\mathcal{O}(\varepsilon)$)

$$\Pi_{\mu\nu}(p) = -(p^2\delta_{\mu\nu} - p_\mu p_\nu)\frac{e^2}{2\pi^2}\left[\frac{1}{6}\frac{2}{\epsilon} - \frac{\gamma}{6} - \int_0^1 d\alpha\ \alpha(1-\alpha)\ln\frac{m^2+p^2(1-\alpha)\alpha}{4\pi\mu^2}\right] \tag{6.113}$$

The finite part of (6.113) contains a lot of physics and we will discuss this further in the next section. Here we concentrate on the determination of the counterterm in the renormalized lagrangian which will cancel the pole term in (6.113) if we calculate the effective action Γ to one loop.

$$\delta\mathcal{L}_A^{(1)} = -\frac{e^2}{12\pi^2}\ \frac{2}{\varepsilon}\ (\partial_\mu A_\nu - \partial_\nu A_\mu)^2 \equiv \frac{1}{4}(Z_3-1)(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 \tag{6.114}$$

will precisely do this job and we get finally

$$Z_3 = 1 - \frac{e^2}{12\pi^2}\frac{2}{\varepsilon}$$

in the minimal subtraction scheme.

6.4.4 The electron self-energy $\Sigma(p)$, Z_2 and Z_m

The free inverse fermion propagator is

$$(S_F^{(0)})^{-1} = -i(\not{p} + im)$$

Let us write the inverse propagator, including radiative corrections, as

$$S_F^{-1} = -i(\not p + im + \Sigma(p)) \tag{6.115}$$

$\Sigma(p)$ is the fermion self-energy part of the propagator, and from the Feynman rules and fig. 5 gives

$$\begin{aligned} -i\Sigma(p)_{\alpha\beta} &= \int \frac{d^dk}{(2\pi)^d}\,\frac{\delta_{\mu\nu}}{k^2}\,(i\tilde{e}\gamma_\mu)_{\alpha\gamma}\left(\frac{i}{\not p - \not k + im}\right)_{\gamma\delta}(i\tilde{e}\gamma_\mu)_{\delta\beta} \\ &= -e^2\mu^\varepsilon \int \frac{d^dk}{(2\pi)^d}\,\frac{[\gamma_\mu(i(\not p - \not k) + m)\gamma_\mu]_{\alpha\beta}}{k^2((p-k)^2 + m^2)} \end{aligned} \tag{6.116}$$

We should note here that the propagator S_F is gauge dependent, contrary to $\Pi_{\mu\nu}(p)$. The calculation will depend on our gauge choice, reflecting that the propagator itself is not a physical observable. There we will only be interested in the aspect which concerns renormalization, and to simplify the calculations as much as possible we will always use the Feynman gauge, as is also done in (6.116).

Again we introduce a Feynman parameter α and make a shift $k = q + \alpha p$ of momentum in (6.116):

$$-\,i\Sigma(p) = -e^2\mu^\varepsilon \int_0^1 d\alpha \int \frac{d^dq}{(2\pi)^d}\,\frac{i\gamma_\mu\left[(1-\alpha)\,\not p - \not q\right]\gamma_\mu + m\gamma_\mu^2}{(q^2 + \alpha(1-\alpha)p^2 + \alpha m^2)^2} \tag{6.117}$$

The linear term in q drops out and from the trace identities we get: $\gamma_\mu\ \not p\gamma_\mu = (-2+\varepsilon)\ \not p$. Integration over q leave us with

$$-\,i\Sigma(p) = \frac{e^2}{(4\pi)^2}\,\Gamma(\varepsilon/2)\int_0^1 d\alpha\ \frac{(2-\epsilon)(1-\alpha)i\ \not p - (4-\epsilon)m}{\left[(\alpha(1-\alpha)p^2 + \alpha m^2)/(4\pi\mu^2)\right]^{\varepsilon/2}} \tag{6.118}$$

The pole part can readily be extracted:

$$-i\Sigma(p)|_{div} = i\frac{e^2}{(4\pi)^2}\ \frac{2}{\varepsilon}\cdot(\not p + i4m) \tag{6.119}$$

This part can be cancelled by counterterms:

$$\delta\mathcal{L}_\psi^{(1)} = -\frac{e^2}{(4\pi)^2}\ \frac{2}{\varepsilon}\bar\psi\ \not\partial\psi - \frac{e^2}{(4\pi)^2}\cdot 4\cdot\frac{2}{\varepsilon}\ m\bar\psi\psi \tag{6.120}$$

in the renormalized lagrangian and we conclude that

$$Z_2 = 1 - \frac{e^2}{(4\pi)^2}\ \frac{2}{\varepsilon} \quad , \quad Z_m = 1 - 4\ \frac{e^2}{(4\pi)^2}\ \frac{2}{\varepsilon} \tag{6.121}$$

6.4.5 The vertex correction and Z_1

We denote the radiative correction to the "free" vertex $\Gamma^{(0)}_{\mu,\alpha\beta} = -i\tilde{e}(\gamma_\mu)_{\alpha\beta}$ by $\Gamma^{(1)}_{\mu;\alpha\beta}(q;p_1,p_2)$ (where $q = p_2 - p_1$ by momentum conservation). From the Feynman rules and fig. 5 we have:

$$\Gamma^{(1)}_{\mu;\alpha\beta}(q;p_1,p_2) = \int \frac{d^dk}{(2\pi)^d}\ \frac{\delta_{\nu\nu'}}{k^2}\left[i\tilde{e}\gamma_\nu\frac{i}{\not p_2 - \not k + im}i\tilde{e}\gamma_\mu\frac{i}{\not p_1 + \not k + im}i\tilde{e}\gamma_{\nu'}\right]_{\alpha\beta} \tag{6.122}$$

or

$$\Gamma_\mu^{(1)}(q;p_1,p_2) = -i\tilde{e}^3 \int \frac{d^dk}{(2\pi)^d} \frac{\gamma_\nu[i(\not{p}_2+\not{k})+m]\gamma_\mu[i(\not{p}_1+\not{k})+m]\gamma_\nu}{k^2[(p_2-k)^2+m^2][(p_1+k)^2+m^2]} \tag{6.123}$$

The same remarks apply here as did for the fermionic self-energy: $\Gamma_\mu^{(1)}(p_1,p_2)$ is gauge dependent. We here use the Feynman gauge.

As in the scalar case we have to introduce $n-1$ Feynman parameters if the 1-loop diagram has n propagators and after a shift $k = q - p_1\alpha_1 - p_2\alpha_2$ we get

$$\begin{aligned}\Gamma_\mu^{(1)}(p_1,p_2) \;=\; & -i\tilde{e}^3 \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \int \frac{d^dq}{(2\pi)^d} \quad \times \\ & \frac{\gamma_\nu[i(\not{q}+(1-\alpha_2)\not{p}_2-\alpha_1\not{p}_1)+m]\gamma_\mu[i(\not{q}+(1-\alpha_1)\not{p}_1-\alpha_2\not{p}_2)+m]\gamma_\nu}{[q^2+\alpha_1(1-\alpha_1)p_1^2+\alpha_2(1-\alpha_2)p_2^2-2p_1p_2\alpha_1\alpha_2+(\alpha_1+\alpha_2)m^2]^3}\end{aligned} \tag{6.124}$$

We see that only the term quadratic in q in the numerator leads to a divergence in $d=4$. We split $\Gamma_\mu^{(1)}$ in two parts $\Gamma_\mu^{(1a)}$ and $\Gamma_\mu^{(1b)}$, where $\Gamma_\mu^{(1a)}$ contains the terms quadratic in q in the numerator and $\Gamma_\mu^{(1b)}$ the rest.

$$\Gamma_\mu^{(1a)}(p_1p_2) = i\tilde{e}^3 \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \int \frac{d^dq}{(2\pi)^d} \frac{(\gamma_\nu\gamma_\alpha\gamma_\mu\gamma_\beta\gamma_\nu)q_\alpha q_\beta}{[q^2+F(\alpha_1,\alpha_2,p_1,p_2)]^3} \tag{6.125}$$

where

$$F(\alpha_1,\alpha_2,p_1,p_2) = \alpha_1(1-\alpha_1)p_1^2+\alpha_2(1-\alpha_2)p_2^2-2p_1p_2\alpha_1\alpha_2+m^2(\alpha_1+\alpha_2) \tag{6.126}$$

As in the case of the photon self-energy the integration over $q_\alpha q_\beta$ produces a factor $\delta_{\alpha\beta}q^2/d$ by symmetry. From the trace identities we have

$$\gamma_\nu\gamma_\alpha\gamma_\mu\gamma_\alpha\gamma_\nu = (2-d)^2\gamma_\mu$$

and we get

$$\Gamma_\mu^{(1a)}(p_1,p_2) = i\tilde{e}^3\gamma_\mu \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \; \frac{(2-d)^2}{d} \int \frac{d^dq}{(2\pi)^d} \frac{q^2}{[q^2+F(\alpha_1,\alpha_2,p_1,p_2)]^3} \tag{6.127}$$

It is seen that $\Gamma_\mu^{(1a)}(q;p_1,p_2) \sim \gamma_\mu$ and that the divergent part of $\Gamma_\mu^{(1a)}(q;p_1p_2)$ is

$$\Gamma_\mu^{(1a)}(q;p_1,p_2)\Big|_{div} = i\tilde{e}\gamma_\mu \frac{e^2}{(4\pi)^2}\frac{2}{\varepsilon} \tag{6.128}$$

since it comes entirely from the following part of $\Gamma_\mu^{(1a)}$:

$$\begin{aligned} i\tilde{e}^3\gamma_\mu \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \; \frac{(2-d)^2}{d} \cdot \int \frac{d^dq}{(2\pi)^2} \frac{1}{[q^2+F(\alpha_1,\alpha_2,p_1,p_2)]^2} = \\ + \frac{i\tilde{e}^3}{(4\pi)^{d/2}} \; \gamma_\mu \frac{(2-d)^2}{d} \cdot \Gamma(\varepsilon/2) \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \; F(\alpha_1,\alpha_2,p_1,p_2)^{-\varepsilon/2}\end{aligned}$$

The pole part of this expression is given by (6.128) and we can cancel it by adding a counterterm

$$\delta\mathcal{L}^{(1)}_{\psi\gamma_\mu A_\mu\psi} = \left[\frac{e^2}{(4\pi)^2}\;\frac{2}{\varepsilon}\right]\; i\tilde{e}\bar{\psi}\not{A}\psi \tag{6.129}$$

to the renormalized lagrangian. This means that

$$Z_1 = 1 - \frac{e^2}{(4\pi)^2}\,\frac{2}{\varepsilon} \tag{6.130}$$

For further reference we note that $\Gamma_\mu^{(1b)}$ is given by

$$\begin{aligned}\Gamma_\mu^{(1b)}(p_1,p_2) &= -i\frac{e^3}{(4\pi)^2}\int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2 \quad \times \\ &\times \quad \frac{\gamma_\nu\left[i\left((1-\alpha_2)\ \not{p}_2 - \alpha_1\ \not{p}_1\right)+m\right]\gamma_\mu\left[i\left((1-\alpha_1)\ \not{p}_1 - \alpha_2\ \not{p}_2\right)+m\right]\gamma_\nu}{F(\alpha_1,\alpha_2,p_1,p_2,m)}\end{aligned} \tag{6.131}$$

where we have taken the limit $\varepsilon \to 0$ since $\Gamma_\mu^{(1b)}(p_1 p_2)$ by definition is finite in this limit.

6.5 Physical applications

We want to discuss some of the physics contained in the one loop results derived so far.

6.5.1 The vacuum polarization

Recall from classical electromagnetism that the *effective* action in a material with a dielectric constant $\epsilon \neq 1$ and a magnetic permeability $\mu \neq 1$ is given by

$$S_{eff}[E,B] = \frac{1}{2}\int dt \int d^3x(\epsilon E^2 - \frac{1}{\mu}B^2) \tag{6.132}$$

The velocity of light in such a material is $c_{material} = 1/\sqrt{\epsilon\mu}$ and the response to external currents J_μ is changed because of the polarizability of the material. For instance Coulombs law in the presence of such a material is changed to:

$$eA_\mu^{Coulomb}(\vec{x}) = \frac{e^2}{4\pi\epsilon|\vec{x}|}\delta_{0\mu} \tag{6.133}$$

A more general expression than (6.132) involves a frequency dependence of the dielectric "constant": $\epsilon \to \epsilon(\omega)$.

Let us assume that the material is such that we can *define* a dielectric function $\epsilon(\omega,\vec{k})$ as the ratio between Fourier components of the displacement field D_i and the electric field E_i (independent of direction i):

$$\epsilon(\omega,\vec{k}) = D(\omega,\vec{k})/E(\omega,\vec{k}) \tag{6.134}$$

and similar for the permeability function $\mu(\omega,\vec{k})$. In that case we would get

$$S_{eff}[E,B] = \frac{1}{2}\int \frac{d\omega}{2\pi}\int \frac{d^3k}{(2\pi)^3}\left[\epsilon(\omega,\vec{k})|E(\omega,k)|^2 - \frac{1}{\mu(\omega,k)}\,|B(\vec{\omega},k)|^2\right] \tag{6.135}$$

and the modified Coulomb law would in momentum space look like:

$$eA_\mu^{Coulomb}(\vec{k}) = \frac{e^2}{\vec{k}^2\epsilon(0,\vec{k})}\delta_{0\mu} \tag{6.136}$$

We remind the reader that expressions like (6.135) are only effective expressions valid in a certain frequency range, but we will show that the vacuum itself behaves very much like a material with $\epsilon(k)$ and $\mu(k) \neq 0$, only is $\epsilon(k) \cdot \mu(k) = 1$ since the theory is *Lorentz invariant*.

The effective action

$$\Gamma(A^{cl}, \bar{\psi}^{cl}, \psi^{cl})\Big|_{\bar{\psi}^{cl}=\psi^{cl}=0}$$

has a general expansion

$$\Gamma(A^{cl}) = \frac{1}{2!}\Gamma_{ij}A_i^{cl}A_j^{cl} + \frac{1}{4!}\Gamma_{ijkl}A_i^{cl}A_j^{cl}A_k^{cl}A_l^{cl} + \ldots \tag{6.137}$$

In this expansion $\Gamma_{i_i \ldots i_{2n}} = \mathcal{O}(e^{2n})$ $(n > 2)$ and to order $\mathcal{O}(e^2)$ we only have to consider the term quadratic in A. Recall that

$$\Gamma_{ij} = G_{ij}^{-1} = (G^{(0)})_{ij}^{-1} + \Pi_{ij} \tag{6.138}$$

which allows us to write

$$\Gamma^{(2)}[A^{cl}] = \frac{1}{2}\int \frac{d^4k}{(2\pi)^4}\; A_\mu^{(cl)}(-k)G_{\mu\nu}^{-1}(k)A_\nu^{(cl)}(k) \tag{6.139}$$

$$G_{\mu\nu}^{-1}(k) = G_{\mu\nu}^{(0)^{-1}}(k) + \Pi_{\mu\nu}(k) = (k^2\delta_{\mu\nu} - k_\mu k_\nu)(1 + \Pi(k^2)) + \frac{1}{\alpha}k_\mu k_\nu \tag{6.140}$$

This expression looks gauge dependent, in the sense that it seems to change under a gauge transformation $A_\mu^{(cl)}(k) \to A_\mu^{(cl)}(k) + k_\mu \Lambda(k)$, but it is not the case if $A_\mu^{(cl)}$ is generated by an external current J_μ which satisfies current conservation:

$$\partial_\mu J_\mu = 0 \Rightarrow \Gamma^{(2)}[A^{(cl)}] = \frac{1}{2}\int \frac{d^4k}{(2\pi)^4}\; (1 + \Pi(k^2))A_\mu^{(cl)}(-k)(k^2\delta_{\mu\nu} - k_\mu k_\nu)A_\nu^{(cl)} \tag{6.141}$$

The proof follows from the fact that all coefficients in the expansion (6.137) are transverse, i.e. vanish when contracted with k_μ, except the gauge fixing term $1/\alpha k_\mu k_\nu$ which multiplies $A_\mu(-k)A_\nu(k)$, as shown in the section which discussed the Ward-Takahashi identities. This leads to

$$k_\mu J_\mu = k_\mu \frac{\delta\Gamma}{\delta A_\mu^{cl}} = \frac{1}{\alpha}k^2 k_\mu A_\mu^{(cl)} \tag{6.142}$$

and we see that $k_\mu J_\mu = 0$ implies $k_\mu A_\mu^{(cl)} = 0$. For further reference we note that in the truncation (6.141) we have

$$J_\mu = G_{\mu\nu}^{-1}A_\nu^{(cl)} \quad \text{or} \quad A_\mu^{(cl)}(k) = G_{\mu\nu}(k)J_\nu(k) = \frac{J_\mu(k)}{k^2(1 + \Pi(k^2))} \tag{6.143}$$

where the last equation is correct if $k_\mu J_\mu = 0$.

It is now an appropriate time to rotate back to minkowskian spacetime since we want to compare with real physics. This is done simply by replacing $\delta_{\mu\nu} \to g_{\mu\nu}$ in (6.141) and we see that $A^\mu(-k)(k^2 g_{\mu\nu} - k_\mu k_\nu)A^\nu(k)$ can be written as $E_i(-k)E_i(k) - B_i(-k)B_i(k)$ in (6.141). We can now write

$$\Gamma^{(2)}[A_\mu^{(cl)}] = \frac{1}{2}\int \frac{dk_0}{2\pi}\;\frac{d^3k}{(2\pi)^3}(1 + \Pi(k^2))\left(|E(k_0, \vec{k})|^2 - |B(k_0, \vec{k})|^2\right) \tag{6.144}$$

A comparison with (6.135) leads to the conclusion

$$\epsilon(k^2) = \frac{1}{\mu(k^2)} = 1 + \Pi(k^2) \qquad \text{(i.e.}\quad \epsilon(k^2)\mu(k^2) = 1) \tag{6.145}$$

and further eq. (6.143) in the case of a static external charge e leads to

$$eA_0^{(cl)}(\vec{k}, k_0) = \frac{e^2}{\vec{k}^2(1+\Pi(\vec{k}^2))} \cdot 2\pi\delta(k_0) \tag{6.146}$$

since $J_\mu(x,t) = e\ \delta^{(3)}(x)\delta_{\mu 0}$ and $J_\mu(k) = 2\pi e\delta(k_0)\delta_{\mu 0}$. This should be compared to (6.136).

Let us recall that $\Pi(k^2)$ in MS is given by:

$$\Pi(k^2) = -\frac{e^2}{2\pi^2}\left[\frac{\gamma}{6} + \int_0^1 d\alpha\ \alpha(1-\alpha)\ln\left[\frac{k^2\alpha(1-\alpha)+m^2-i\epsilon}{4\pi\mu^2}\right]\right] \tag{6.147}$$

where the $i\epsilon$ indicates that we have rotated back to minkowskian space. For $k^2 = \vec{k}^2$ we can drop the $-i\epsilon$ and we see that $\Pi(\vec{k}^2)$ *is a decreasing function* of $\vec{k}^2$. Loosely speaking we have the relation

$$|\vec{x}| \sim \frac{1}{|\vec{k}|}$$

between distance from the charge e and momenta $\vec{k}$ in (6.146) or (6.136), and we see that Coulombs law is modified in such a way that

$$e_{eff}^2(r) \sim \frac{e^2}{1+\Pi(\vec{k}^2)} \quad , \quad |\vec{k}| \sim \frac{1}{r} \tag{6.148}$$

This is *screening.* The polarization of the vacuum makes $e_{eff}^2(r)$ smaller as $r \to \infty$. Especially we get

$$\boxed{e_{eff}^2(r\to\infty) = \frac{e^2}{1+\Pi(0)}} \tag{6.149}$$

As discussed in chapter 5 we have the freedom of performing *finite renormalizations.* Since $e_{eff}(r\to\infty)$ is effectively the charge we observe it is very convenient to choose the renormalization parameter μ such that

$$\Pi(0) = 0 \tag{6.150}$$

With this choice we can identify the ordinary electrostatic charge e in low energy experiments with the renormalized charge e in our lagrangian $\mathcal{L}^{(ren)}(A,\bar{\psi},\psi,m,e)$ and we have

$$\Pi(k^2) = -\frac{e^2}{2\pi^2}\int_0^1 d\alpha(1-\alpha)\alpha\ln\left(\frac{k^2\alpha(1-\alpha)+m^2-i\epsilon}{m^2}\right) \tag{6.151}$$

Note that when $-k^2 \geq 4m^2$, $\Pi(k^2)$ has an imaginary part:

$$\begin{aligned} \text{Im}\,\Pi(k^2) &= \frac{e^2}{2\pi}\int_0^1 d\alpha(1-\alpha)\alpha\ \theta\left(\alpha(1-\alpha)+\frac{m^2}{k^2}\right) \\ &= \frac{e^2}{12\pi}\left(1-\frac{2m^2}{k^2}\right)\sqrt{1+\frac{4m^2}{k^2}}\ \theta\left(1+\frac{4m^2}{k^2}\right) \end{aligned} \tag{6.152}$$

The interpretation of this is precisely like the interpretation of the imaginary part of the dielectric function $\epsilon(\omega)$. In that case the material is able to absorb electromagnetic energy (and eventually convert it to heat). In this case electromagnetic energy disappears too: it is converted into *real* electron and positrons since the condition $-k^2 \geq 4m^2$ means that the virtual photon with k^2 carries sufficiently energy to put the electron-positron pair in the virtual fermion loop responsible for $\Pi(k)$ on their mass shell, i.e. make them *real* propagating particles.

For a given electromagnetic field we can use (6.144) to calculate the total pair-creation rate since

$$\text{Im}\,\Gamma^{(2)}[A^{(cl)}] = \frac{e^2}{24\pi}\int \frac{d^4k}{(2\pi)^4}(1-\frac{2m}{k^2})\sqrt{1+\frac{4m^2}{k^2}}\;\theta(1+\frac{4m^2}{k^2})\left(|E(k)|^2-|B(k)|^2\right) \quad (6.153)$$

Recall that the minkowskian path integral has the interpretation as vacuum to vacuum amplitude in the presence of external sources J_μ (which generates the $A_\mu^{(cl)}$):

$$\langle 0|0\rangle_J = Z[J] = e^{i\Gamma[A^{(cl)}]+iJ\cdot A^{(cl)}} \quad (6.154)$$

The decay rate of the vacuum in the presence of the external source J is given by the square of the vacuum to vacuum amplitude and the rate of creation of electron-positron pair is therefore given by:

$$R \equiv 1-|\langle 0|0\rangle_J|^2 = 1-e^{-2\text{Im}\,\Gamma[A^{(cl)}]} \approx 2\text{Im}\,\Gamma^{(2)}[A_{(cl)}]+\mathcal{O}(e^4) \quad (6.155)$$

We note that pair creation is entirely an electric effect. A given four momentum k_μ contributes only if $k^2 \leq -4m^2 \leq 0$. This means that we can find a Lorentz frame where $\vec{k}=0$ and since the magnetic Fourier component is given by

$$\vec{B}(k) = -i\vec{k}\times\vec{A}(k)$$

the magnetic component vanish in that frame. A pure magnetic field will not be able to create pairs. This is in accordance with classical intuition since a magnetic field cannot perform any work on charge particles as the Lorentz force is perpendicular to the velocity. Energy conservation would be violated if such a field could create particle pairs from the vacuum.

Let us end this section by actually converting (6.146) to a modified Coulomb potential in $\vec{x}$-space. For this purpose it is convenient to rewrite eq. (6.151) for $\Pi(\vec{k}^{\,2})$ as[6]

$$\Pi(\vec{k}^{\,2}) = -\frac{e^2}{12\pi^2}\vec{k}^{\,2}\int_{4m^2}^{\infty}\frac{dq^2}{q^2}\,\frac{1}{q^2+\vec{k}^{\,2}}\left(1+\frac{2m^2}{q^2}\right)\sqrt{1-\frac{4m^2}{q^2}} \quad (6.156)$$

To order e^2 we have from (6.146):

$$A_0^{(cl)}(\vec{k},k_0) = \frac{e}{\vec{k}^{\,2}}\,\left(1-\Pi(\vec{k}^{\,2})\right)2\pi\delta(k_0) \quad (6.157)$$

---

[6] We leave it as an exercise to show the equivalence between (6.156) and (6.151) for $-k^2 \leq 4m^2$.

After Fourier transformation we get:

$$A_0^{(cl)}(\vec{x},t) = e\int \frac{d^3k}{(2\pi)^3}\,\frac{e^{-i\vec{k}\cdot\vec{x}}}{\vec{k}^{\,2}}\,\left(1+\frac{e^2}{12\pi^2}\,\vec{k}^{\,2}\int_{4m^2}^{\infty}\frac{dq^2}{q^2}\,\frac{1}{q^2+\vec{k}^{\,2}}\,\left(1+\frac{2m^2}{q^2}\right)\sqrt{1-\frac{4m^2}{q^2}}\right) \tag{6.158}$$

We now use ($r \equiv |\vec{x}|$):

$$\int \frac{d^3k}{(2\pi)^3}\,\frac{e^{-i\vec{k}\cdot\vec{x}}}{\vec{k}^{\,2}} = \frac{1}{4\pi r} \quad , \quad \int \frac{d^3k}{(2\pi)^3}\,\frac{e^{-i\vec{k}\cdot\vec{x}}}{q^2+\vec{k}^{\,2}} = \frac{e^{-qr}}{4\pi r} \tag{6.159}$$

and we can write

$$\begin{aligned} A_0^{(cl)}(\vec{x},t) &= \frac{e}{4\pi r}\cdot\left(1+\frac{e^2}{6\pi^2}\int_1^{\infty}\frac{du}{u^2}\,e^{-2mru}\left(1+\frac{1}{2u^2}\right)\sqrt{u^2-1}\right) \\ &= \frac{e}{4\pi r}\left(1+\frac{e^2}{16}\,\frac{e^{-2mr}}{(\pi mr)^{3/2}}+\cdots\right) \qquad \text{for } mr \gg 1 \\ &= \frac{e}{4\pi r}\left(1+\frac{e^2}{12\pi^2}\ln\frac{1}{(mr)^2}+\text{const}+\cdots\right) \qquad \text{for } mr \ll 1 \end{aligned} \tag{6.160}$$

According to our discussion above we have chosen the finite renormalization of $\Pi(k^2)$ such that $A_0^{(cl)}(\vec{x},t) \to \frac{e}{4\pi r}$ for $r \to \infty$ and we see that the potential increases relative to the Coulomb potential as r decreases. The interpretation is that the "bare" charge e_0 is larger than the measured charge e, due to the polarization of the vacuum by virtual e^+e^- pairs. The factor multiplying $e/4\pi r$ in (6.160) actually diverges as $r \to 0$ indicating an infinite e_0, but we can not trust this lowest order calculation unless $e^2/12\pi^2\ \ln\ 1/(mr)^2 \ll 1$.

6.5.2 The anomalous magnetic moment of the electron

The classical relation between (orbital) the angular momentum of a point particle with charge q and the magnetic moment is

$$\vec{\mu} = \frac{q}{2m}\vec{L} \quad \text{(orbit)} \tag{6.161}$$

while the contribution from the electron spin $\vec{S}$ is given by

$$\vec{\mu} = \frac{e}{m}\vec{S} = g(\frac{e}{2m})\vec{S} \tag{6.162}$$

The factor g, called the Landé g-factor, describes the deviation from classical physics. For the electron $g = 2$ (to very high precision). In ordinary quantum mechanics this was just an experimental fact. It was one of the non-trivial predictions of the Dirac equation that $g = 2$ and it is a *highly non-trivial* prediction of QED that there are radiative corrections which change this value from $g = 2$ to ($\alpha \equiv \frac{e^2}{4\pi}$)

$$\begin{aligned} g &= 2\left(1+\frac{\alpha}{2\pi}-0.328478445(\frac{\alpha}{\pi})^2+1.18311(\frac{\alpha}{\pi})^3+\cdots\right) \\ &= 2\,(1+0,0011596\ 52359(282)) \end{aligned} \tag{6.163}$$

while the experimental value is

$$g = 2\,(1+0.0011596\ 52359(200)) \tag{6.164}$$

It is a remarkable fact that g can be measured with this precision and the agreement between experiment and theory is most impressive.

Let us recall how the Dirac equation predicts $g = 2$.

$$(\not{D} + m)\psi = 0 \quad \Rightarrow \quad (-\not{D} + m)(\not{D} + m)\psi = 0 \tag{6.165}$$

A little algebra using $i[\gamma_\mu, \gamma_\nu]/2 = \sigma_{\mu\nu}$ and $[D_\mu, D_\nu] = -ieF_{\mu\nu}$ leads to

$$[-D^2 + m^2 - \frac{e}{2}\,\sigma^{\mu\nu}F_{\mu\nu}]\psi = 0 \tag{6.166}$$

We see that the difference between a scalar particle minimally coupled to A_μ and a spin-$\frac{1}{2}$ particle, also minimally coupled, is manifest in the term $\frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}$. In the case of a constant magnetic field B this term is

$$-\frac{e}{2}\,\sigma^{\mu\nu}F_{\mu\nu} = -e\vec{\sigma}\cdot\vec{B} = -2e\,\vec{S}\cdot\vec{B} \tag{6.167}$$

This term precisely explains the non-relativistic result (6.162). If we solve (6.166) for weak fields where $eB \ll m^2$ we get for the energy eigenvalues:

$$\begin{aligned} E^2_{spinor} &= E^2_{scalar} - 2e\,\vec{S}\cdot\vec{B} & (6.168)\\ E_{spinor} &\approx E_{scalar} - \frac{e}{m}\,\vec{S}\cdot\vec{B} = E_{scalar} - \vec{\mu}\cdot\vec{B} & (6.169)\end{aligned}$$

In (6.168) and (6.169) $\vec{S}$ is a vector rather than a spin matrix as in (6.167).

It is possible to add a gauge invariant term to the Dirac equation which changes the magnetic moment of the spinor $\frac{1}{2}$ particle to an arbitrary value:

$$\begin{aligned} \mathcal{L}_{fer}(\bar{\psi}, \psi, A) &= -\bar{\psi}(\not{D} + m)\psi \\ &\to -\bar{\psi}(\not{D} + m)\psi + \frac{\Delta g}{4}\,\frac{e}{2m}\bar{\psi}(\sigma^{\mu\nu}F_{\mu\nu})\psi & (6.170)\end{aligned}$$

This so-called Pauli term would change the calculation (6.169) into:

$$E_{spinor} \approx E_{scalar} - (2 + \Delta g)\,\frac{e}{2m}\,\vec{S}\cdot\vec{B} \tag{6.171}$$

It would not be pleasant to have such a Pauli term in our fundamental lagrangian in order to explain the observed value (6.164) even if Δg would be as small as $0.0011\cdots$. The reason is that the term does not correspond to a renormalizable interaction. The dimension of the coupling constant e/m is the inverse of mass and QED would not be a renormalizable theory. If we discard the presence of such a term in the bare lagrangian our only chance is that it is generated (as a *finite* term) in the effective action $\Gamma(A, \bar{\psi}, \psi)$. In that case the value Δg is uniquely fixed and the theory should be able to *predict* the observed value. We will now show that it is indeed the case.

In sec.4 we calculated the first order radiative correction to the "free" vertex $\Gamma^{(0)}_{\mu,\alpha\beta} = -ie(\gamma_\mu)_{\alpha\beta}$:

$$\Gamma_{\mu,\alpha\beta}(p_1, p_2) = \Gamma^{(0)}_{\mu,\alpha\beta} + \Gamma^{(1)}_{\mu,\alpha\beta}(p_1, p_2) \tag{6.172}$$

where

$$\Gamma^{(1)}_{\mu,\alpha\beta} = \Gamma^{(1a)}_{\mu,\alpha\beta} + \Gamma^{(1b)}_{\mu,\alpha\beta} \tag{6.173}$$

and $\Gamma^{(1a)}_{\mu;\alpha\beta} \propto (\gamma_\mu)_{\alpha\beta}$ while $\Gamma^{(1b)}_{\mu;\alpha\beta}$ was ultraviolet finite and given by (6.131). *As for the vacuum polarization it is now time to rotate to minkowskian spacetime since we want to extract real physics.* In order to simplify the calculations let us further assume that the external momenta q, p_1 and p_2 are on the “mass-shell”, i.e. that they satisfy

$$q^2 = 0, \qquad p_1^2 = m^2, \qquad p_2^2 = m^2 \tag{6.174}$$

and let us calculate $\Gamma^{(2b)}_{\mu,\alpha\beta}(p_1, p_2)$ when dressed between two spinors $\psi(p_1)$ and $\psi(p_2)$ which satisfy the Dirac equation:

$$(-i \not{p}_2 + m)\psi(p_1) = 0 \quad , \quad (-i \not{p}_2 + m)\psi(p_2) = 0 \tag{6.175}$$

In order to calculate $\bar{\psi}(p_2)\Gamma^{(2b)}_\mu \psi(p_1)$ we have to calculate

$$\bar{\psi}(p_2)\gamma_\nu \left[i\left((1-\alpha_2)\not{p}_2 - \alpha_1 \not{p}\right) + m\right]\gamma_\mu \left[i\left((1-\alpha_1)\not{p}_1 - \alpha_2 \not{p}_2\right) + m\right]\gamma_\nu \psi(p_1) \tag{6.176}$$

under the constraints (6.174) and (6.175). In order to use these constraints we have to commutep γ_ν and $\not{p}_2$, and in the same way $\not{p}_1$ and γ_n. In this process we will create commutators $[\gamma_\mu, \gamma_\nu] = -2i\sigma_{\mu\nu}$. After a nasty bit of algebra, which we leave as an exercise to the reader, (6.177) can be written as ($q = p_2 - p_1$):

$$\bar{\psi}(p_2)\left[m^2\gamma_\mu\Big((\alpha_1+\alpha_2)^2 - 2(1-\alpha_1-\alpha_2)\Big) + i\, 8m\, q^\nu \sigma_{\mu\nu}\left(\alpha_1 - \alpha_2(\alpha_1+\alpha_2)\right)\right]\psi(p_1) \tag{6.177}$$

Further the denominator (6.126) simplifies drastically due to the constraint (6.174) :

$$F(\alpha_1, \alpha_2, p_1, p_2, m) = m^2(\alpha_1+\alpha_2)^2 \tag{6.178}$$

This leaves us with the final expression for the part of $\bar{\psi}(p_2)\Gamma^{(1)}_\mu(p_1, p_2)\psi(p_1)$ which involves $\sigma_{\mu\nu}$:

$$\begin{aligned} & \bar{\psi}(p_2)\, \Gamma^{(1)}_\mu(p_1, p_2)\psi(p_1)\Big|_{\sim\sigma_{\mu\nu}} \\ = & \frac{ie^3}{2m\pi^2}\bar{\psi}(p_2)\sigma_{\mu\nu}q^\nu\psi(p_1) \int_0^1 d\alpha_1 \int_0^{1-\alpha_1} d\alpha_2\; \frac{\alpha_1 - \alpha_2(a_1+\alpha_2)}{(\alpha_1+\alpha_2)^2} \\ = & \frac{ie^3}{16m\pi^2}\bar{\psi}(p_2)\sigma_{\mu\nu}q^\nu\psi(p_1) \end{aligned} \tag{6.179}$$

(6.179) will contribute to the effective action $\Gamma_\mu(A_\mu, \bar{\psi}, \psi)$ with a term:

$$\frac{e}{2m}\,\frac{ie^2}{8\pi^2}\;\bar{\psi}(p_2)\sigma^{\mu\nu}q_\nu A_\mu(q)\psi(p_1) \tag{6.180}$$

or in position space

$$\frac{e}{2m}\,\frac{e^2}{16\pi^2}\int d^4x\;\bar{\psi}(x)\sigma_{\mu\nu}F^{\mu\nu}(x)\psi(x) \tag{6.181}$$

If we compare with (6.170) we see that it corresponds to an anomalous magnetic moment

$$\Delta g = \frac{e^2}{4\pi^2} = \frac{\alpha}{\pi} \quad ; \quad \alpha \equiv \frac{e^2}{4\pi} \tag{6.182}$$

This is the first term in the expansion (6.163) mentioned above.

6.5.3 The Lamb shift

The Lamb shift (the lifting of the degeneracy between $2S_{1/2}$ and $2P_{1/2}$ orbits in the hydrogen atom) convinced in the begining of the 50'ties many physicists about the reality of radiative corrections in quantum field theory, in spite of the conceptional difficulties of the calculations, leading, as we have seen, to infinities. In this section we will only sketch the ingrediences in the calculation.

Recall that the non-relativistic energy levels of the hydrogen atom are obtained from the Schrödinger equation

$$\left[-\frac{1}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}-\frac{l(l+1)}{r^2}\right)-\frac{\alpha}{r}\right]\psi_{n,l}(r)=\varepsilon_{n,l}\psi_{n,l}(r) \tag{6.183}$$

where

$$\alpha\equiv\frac{e^2}{4\pi}. \tag{6.184}$$

The mass m is the reduced mass of the electron-nucleus system, which is approximately equal to the electon mass:

$$\frac{1}{m}=\frac{1}{m_e}+\frac{1}{m_N}\approx\frac{1}{m_e}. \tag{6.185}$$

The famous solution, a triumph for early equantum mechanics, was

$$\varepsilon_{n,l}=-\frac{m\alpha}{2n^2} \tag{6.186}$$

and numerically the Rydberg constant $m\alpha^2/2 = 13.6eV$. The important point for us to note is that for a given n the angular momentum variable l can take values $0, 1, \ldots, n-1$, and each level is therefore $\sum_0^{n-1}(2l+1)=n^2$ times degenerate.

The next triumph came from the predictions of the Dirac equation. Recall that the spinor ψ will satisfy the equation:

$$-(\not{D}-m)(\not{D}+m)\psi=0 \tag{6.187}$$

or

$$(-D^2+m^2-\frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu})\psi=0. \tag{6.188}$$

In the case of the coulomb potential of the hydrogen atom we have

$$eA_0=-\frac{\alpha}{r},\qquad eE_i=-\frac{\hat{r}_i}{r^2},\qquad \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu}=\mp i\alpha\frac{\sigma_i\hat{r}_i}{r^2} \tag{6.189}$$

where $\hat{r}=\vec{r}/r$, σ_i denote the Pauli matrices and where we in the last equation have used a representation of the γ-matrices where

$$\sigma^{0i}=\frac{1}{2i}[\gamma^0,\gamma^i]=i\begin{pmatrix}\sigma_i & 0\\ 0 & -\sigma_i\end{pmatrix}. \tag{6.190}$$

If we use a stationary ansatz

$$\psi(x_i,t)=e^{-iEt}\psi_\pm(r,\theta,\varphi) \tag{6.191}$$

we get

$$\left[-\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+\frac{L^2-\alpha^2\pm i\alpha\sigma_i\hat{r}_i}{r^2}-\frac{2\alpha E}{r}-(E^2+m^2)\right]\psi_\pm=0. \tag{6.192}$$

The total angular momentum $\vec{J}=\vec{L}+\vec{S}=\vec{L}+\vec{\sigma}/2$ is a constant of motion and commutes with L^2, the square of the orbital angular momentum, which is also a constant of motion. We can thus label energy eigenvalues with l and j. In the subspace where $J^2=j(j+1)$ and $J_z=m$ the integer l in $L^2=l(l+1)$ takes the two values $l=j\pm\frac{1}{2}$ and $L^2-\alpha^2\pm i\alpha\sigma_i\hat{r}_i$ becomes a 2×2 matrix with eigenvalues $\lambda(\lambda+1)$, where

$$\lambda_\pm=j\pm\frac{1}{2}-\delta_j,\qquad \delta_j=j+\frac{1}{2}-\sqrt{(j+\frac{1}{2})^2-\alpha^2}. \tag{6.193}$$

With this result the equation is formally the same as the Schrödinger equation (6.183) after a few substitutions:

$$\begin{aligned} L^2 &\to L^2-\alpha^2\pm i\alpha\sigma_i\hat{r}_i\\ l(l+1) &\to \lambda(\lambda+1)\\ \alpha &\to \alpha E/m\\ \varepsilon &\to (E^2-m^2)/2m \end{aligned}$$

Therefore the energy levels are given by

$$\frac{E_{nj}-m^2}{2m}=-\frac{m\alpha^2}{2}\frac{E_{nj}^2}{m^2}\frac{1}{(n-\delta_j)^2} \tag{6.194}$$

or

$$E_{nj}=\frac{m}{\sqrt{1+\frac{\alpha^2}{(n-\delta_j)^2}}}=m-\frac{m\alpha^2}{2n^2}-\frac{m\alpha^4}{n^3(2j+1)}+\frac{3}{8}\frac{m\alpha^4}{n^4}+\mathcal{O}(\alpha^6) \tag{6.195}$$

with $n=1,2,\ldots$ and $j=1/2,3/2,\ldots,n-1/2$. What should be noted here is that the n^2-fold degeneracy present for the non-relativistic Scrödinger has been lifted due to the term $m\alpha^4/(n^3(2j+1))$. This splitting is called the *fine structure.* If we use the standard, non-relativistic spectroscopic notation nl_j (e.g. $2S_{1/2}$ or $5P_{3/2}$), which is possible since l and j are constants of motion, we have for instance:

$$E(2P_{3/2})-E(2P_{1/2})\approx\frac{m\alpha^4}{32}=4.5\cdot 10^{-5}eV=10.9GHz. \tag{6.196}$$

Note that we still have a degeneracy for a fixed value of j corresponding to the two values of l $(l=j\pm 1/2)$ which could lead to the same value of j. For instance:

$$E(2S_{1/2})=E(2P_{1/2}). \tag{6.197}$$

Before one can compare with experiments several effects should be taken into account. One is the finite size of the nucleus (the proton in our case). This will shift slightly the energy levels of s-waves where $\psi(0)\neq 0$. Next a correct treatment should also include the recoil of proton. Further we have neglected the magnetic field induced by the magnetic

Figure 6.6: Low-lying energy levels in the hydrogen atom

moment of the proton. All of these effects can be treated semiclassically to the desired approximation and we will not discuss them further, except for noting that the interaction induced by the magnetic moment of the proton:

$$V_{hf} = -\frac{e}{2m}\sigma_i^{(e)} B_i, \tag{6.198}$$

where B_i is the magnetic dipole field from the proton and $\sigma_i^{(e)}/2$ the spin of the electron, gives rise to the so-called *hyperfine splitting* of each electron energy level of a typical size:

$$\Delta E_{hf}(S) = 5.9 \cdot 10^{-6} eV = 1.4 GHz \tag{6.199}$$

The lowlying energy levels now have the approximate form shown in fig.6.

One puzzle remained, the shift of 1057 MHz between the energy levels $2S_{1/2}$ and $2P_{1/2}$, which modifies (6.197) to

$$E(2S_{1/2}) = E(2P_{1/2}) + 1057 MHz. \tag{6.200}$$

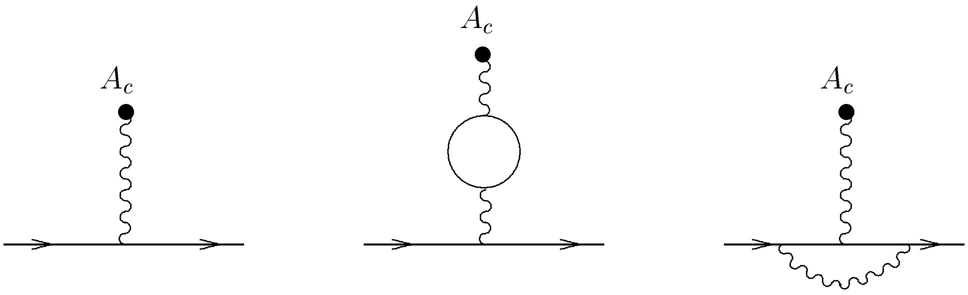


Figure 6.7: The lowest order radiative corrections to the Coulomb potential

This energy difference, called the *Lamb shift*, was explained by the radiative corrections coming from the electron interacting with the quantum fluctuations of the electromagnetic (coulomb) field surrounding the proton. In fig.7 we have shown the lowest order quantum field corrections to the Coulomb potential. This means that the elementary interaction $e\gamma_\mu A_\mu$ which we used in the Dirac equation has to be replaced by:

$$e\gamma_\mu A_\mu \to (e\gamma_\mu + \Gamma_\mu + \Pi_{\mu\nu} G^{\nu\lambda}\gamma_\lambda)A_\mu \tag{6.201}$$

where Γ_μ and $\Pi_{\mu\nu}$ in principle have been computed in the proceeding sections. The modified Dirac equation reads:

$$\left[\not{\partial} + m - i(e\gamma_\mu + \Gamma_\mu + \Pi_{\mu\nu} G^{\nu\lambda}\gamma_\lambda)A_\mu)\right]\psi = 0 \tag{6.202}$$

In order to be consistent with a lowest order calculation the added term should only be taken into account to first order and ordinary lowest order perturbation theory is sufficient for calculating the corrections to the Dirac levels. We can identify three contributions: (1): the one coming from the vacuum polarization, (2): the part of Γ_μ proportional to $\sigma_{\mu\nu}q^\nu$, which was also responsible for the anomalous magnetic moment, and (3): the part of Γ_μ proportional to γ_μ. The last part is the nasty one. It contains infrared divergencies when the external lines are on the mass shell. The reason for this is that the photon is massless. It is impossible to distinguish between an electron and an electron with a very low momentum photon, and one first get a finite result when an integration over photon energies in a finite energy range is performed. In the case of the hydrogen atom the infrared cut off which dictates the energy resolution is determined by the Bohr radius of the atom. For longer wavelengths the radiative corrections (6.201), which were derived under the assumption that we have a free electron in an external field, must be modified to take into account the bound state nature of the electron.. We will not discuss this calculation in detail, but have to refer to other, more extensive textbooks. However, the main effect comes from this term, namely 1010 MHz of the observed 1057 MHz. When we include the contribution from the anomalous magnetic term we add further 68 MHz and the $2P_{1/2}$ has now been lowered too much compared to the observed value. But the vacuum polarization will clearly lower the S state relative to the P state due to screening and the amount is 27 MHz, which leads to a total of 1051 MHz, in very good agreement with experiment. By taking into account higher order radiative corrections this number is increased to 1057.864 ± 0.014 MHz (Mohr, 1975), while the present experimental value is 1057.86 ± 0.02 MHz (Andrews and Newton, 1976).

It should be clear from the above discussion that QED is an unusually succesful theory.

Chapter 7

Quantization of Gauge Theories

7.1 Definition of the functional integral

The problems with the definition of the functional integral for non-abelian gauge theories is the same as for abelian theories: Due to gauge invariance the action is stationary under the space-time dependent gauge transformations $U(x)$

$$A_\mu(x) \to {}^U A_\mu(x) = U(x) A_\mu(x) U^{-1}(x) - i\partial_\mu U(x)\, U^{-1}(x). \tag{7.1}$$

(For a detailed discussion of theses non-abelian gauge transformations we have to refer the the chapter on classical gauge theory). This large invariance means that the functional integral is ill defined. An analogue for an ordinary one-dimensional integral would be to integrate a periodic function from $-\infty$ to $+\infty$. The way we get around this difficulty will be very similar to the approach in the abelian case. We fix the gauge, which in the analogue with the periodic function would correspond to integration over only one period. It is, however, done in such a way that gauge invariance of physical observables is manifest at each step.

Let us fix the notation (for a more careful discussion of non-abelian gauge groups we again have to refer to the chapter which discusses classical gauge theories). We denote the gauge group G. The elements of the group will be denoted U. The *local* gauge invariance we have in mind means that there is a group of gauge transformations associated with each space-time point. In this way we can formally say that the total invariance group is

$$G_{inv} = \prod_{x \in R^d} G_x \tag{7.2}$$

The gauge fields A_μ are elements of the Lie algebra associated with the group G:

$$A_\mu = A_\mu^a T^a \tag{7.3}$$

where T^a are the generators of the Lie algebra. We will denote the Lie algebra associated with the (Lie) group G by $\mathcal{G}$. For a given group G and a given representation R of the group the elements U will be $N_R \times N_R$ dimensional matrices and the same will be true for the generators T^a of the Lie algebra $\mathcal{G}$, since the connection between elements in the group and in the algebra is given by the exponential map (which is 1-1 in a neighbourhood of the identity element of the group):

$$U = e^{i\alpha^a T^a}, \quad \alpha^a \in R \tag{7.4}$$

We can choose generators T^a such that

$$[T_R^a, T_R^b] = ic^{abc}\, T_R^c, \tag{7.5}$$

$$\text{Tr}\; T_R^a T_R^b = T_R \delta^{ab}, \qquad \sum_a (T_R^a)^2 = C_R I. \tag{7.6}$$

C_R is called the value of the Casimir operator in the given representation of $\mathcal{G}$. If N_G denotes the order of the group, i.e. the number of generators in the Lie algebra, we obviously have, by taking the trace in the last equation of (7.6)

$$T_R N_G = C_R N_R \tag{7.7}$$

The constants c^{abc} are called the structure constants of the Lie algebra $\mathcal{G}$. We can define $N_G \times N_G$ matrices

$$(T^b)_{ac} = ic^{abc} \tag{7.8}$$

and these matrices satisfy themselves (7.5), which in this case is becomes the Jacobi identities for the Lie algebra. The Jacobi identity for arbitrary generators in an arbitrary representation can be written:

$$[T^a, [T^b, T^c]] + [T^b, [T^c, T^a]] + [T^c, [T^a, T^b]] = 0 \tag{7.9}$$

and expresses that the associativity of the mapping $\mathcal{G} \times \mathcal{G} \to \mathcal{G}$ defined by the $(A, B) \to [A, B]$ for $A, B \in \mathcal{G}$. The matrix representation of the Lie algebra (and the induced representation of the group G) defined by the structure constants c^{abc} is called the adjoint representation of the algebra $\mathcal{G}$ (and of the group G). For the adjoint representation we have $N_R = N_G$ and $T_R = C_R \equiv C_2(G)$. For $SU(N)$ we have $N_G = N^2 - 1$ while $N_R = N$ for quarks in the fundamental representation. The usual normalization of T_R for quarks is $T_R = 1/2$ and we get $C_R = (N^2 - 1)/2N$, while the value of $C_2(G) = N$. Usually we think of $G = SU(3)$ as the color symmetry group. However, in grand unified theories (GUT) one considers more complicated groups like $SU(5), SO(10), E_6$ etc, in an attempt to unify all known interactions except gravity.

If we want to couple the non-abelian gauge field the matter fields we assume that these matter fields transform in a definite way under the action of the gauge group:

$$\psi_i(x) \to (U_f)_{ij}(x)\psi_j(x) \tag{7.10}$$

where U_f will be $N_f \times N_f$ dimensional matrices which form a certain representation of the gauge group G, and we further assume that the original lagrangian of the ψ-field is invariant under the action of *global* gauge transformations of the kind (7.10), i.e. transformations where the U_f's are independent of x.The so-called minimal coupling, which makes the lagrangian with matter fields invariant under the *local* gauge transformation (7.10), is obtained by simply replacing the ordinary derivatives with covariant derivatives:

$$\partial_\mu \psi_i(x) \to (D_\mu)_{ij}\psi_j(x) \tag{7.11}$$

where

$$(D_\mu)_{ij} \equiv \partial_\mu \delta_{ij} - ig(T_f^a)_{ij} A_\mu^a(x) \tag{7.12}$$

In this formula g denotes a gauge coupling constant which we will discuss further below. In the case of the adjoint representation we have

$$D_\mu^{ab} = \partial_\mu \delta^{ab} + gc^{acb} A_\mu^c \tag{7.13}$$

and we note for future reference that for an infinitesimal gauge transformation

$$U(x) = e^{ig\alpha^a(x)T^a} \approx 1 + ig\alpha(x)T^a \tag{7.14}$$

we have

$$A_\mu^a(x) \to A_\mu^a(x) + D_\mu^{ab}\,\alpha^a(x). \tag{7.15}$$

If we specialize this to the case of space-time independent transformations it shows that the *components* A^a transforms as a vector in the adjoint representation under space-time independent gauge transformations, a fact which also follows directly from the transformation (7.1), if one is familiar with the more abstract group-theoretical definition of the adjoint representation.

The lagrangian non-abelian (or abelian) gauge theories is given by

$$\mathcal{L}(A_\mu) = -\frac{1}{4}\ \mathrm{Tr}\ F_{\mu\nu}^2 \tag{7.16}$$

where we have introduced a coupling constant g such that

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - g\ [A_\mu\ ,\ A_\nu] \tag{7.17}$$

and where we have used the notation (7.3) for the gauge field and further assumed that the normalisation of the generators has been chosen such that $T_R = 1$.

The gauge invariance is now:

$$A'_\mu = U(x)A_\mu U^{-1}(x) - \frac{i}{g}(\partial_\mu U)\ U^{-1}(x) \tag{7.18}$$

If we compare with (7.1) it differs by the coupling constant g. Sometimes it is convenient to absorb the coupling constant in the definition of the gauge field A_μ. It that case g would be absent in (7.12),(7.13),(7.17) and (7.18), but would appear in front of (7.16) as a factor $1/g^2$.

As always we imagine that a rotation to euclidean space has been performed when we want to define our path integrals. The same remarks apply here as in the abelian case: The rotation $x_4 = \tau = it$ should now be supplemented with a rotation of A_0 , the time component of the vector potential, if we want to keep $\mathcal{L}(A)$ real. Since A_μ transforms in the same way as x_μ it is natural to rotate A_0 as x_0. The "electric" field will then be given by

$$E_i^{(Euclid)} = \frac{\partial}{\partial x_4}A_i - \frac{\partial}{\partial x_i}A_4 = -iE_i^{(Minkowski)} \tag{7.19}$$

and the action will be

$$\mathcal{L}^{Euclid}(A) \equiv \frac{1}{4}\mathrm{Tr}\ F_{\mu\nu}^2 = \frac{1}{2}\mathrm{Tr}\ (E^2 + B^2) \tag{7.20}$$

$$S^{(Euclid)}[A] = \int d^dx\ \mathcal{L}^{Euclid}(A) \tag{7.21}$$

We can formulate the quantization problem precisely as we did in the abelian case: We define expectation values by the *formal* expression:

$$\boxed{< \mathcal{O}_1(A)\cdots\mathcal{O}_n(A) > = \frac{\int \mathcal{D}A\ e^{-S[A]}\ \mathcal{O}_1(A)\cdots\mathcal{O}_n(A)}{\int \mathcal{D}A\ e^{-S[A]}}}. \tag{7.22}$$

where $\mathcal{O}_1(A)\cdots\mathcal{O}_n(A)$ are *gauge invariant* observables. As discussed in the chapter on classical gauge theories the path-ordered integrals of various kinds constitute a complete set of observables. However, they are difficult to use in practise and we will in the end consider Green functions which, (as already seen in QED), are strictly speaking not gauge invariant, but which can be used in intermediate steps toward constructing gauge invariant quantities. In the following formal arguments we assume nevertheless that the $\mathcal{O}_k$'s are gauge invariant.The action $S[A]$ is defined by (7.21). We now want to show that it is possible by, a number of formal manipulations to write both numerator and denominator as a well defined functional integral times a formal product

$$\prod_{x\in R^d} V(G), \qquad V(G)=\int_G dU \tag{7.23}$$

where $V(G)$ denotes the "volume" of the compact gauge group G, and the integration in (7.23) is over all group element. The measure dU is the unique measure[1] on the group manifold which is invariant under left- and right translations, i.e. $U\to U_0U$ and $U\to UU_0$. It is called the Haar measure. The infinite product $\prod_{x\in R^d} V(G)$ is independent of any dynamics and cancel between numerator and denominator. The remaining integrals are well defined and can be treated as we did for a scalar field (Dyson-Schwinger equations, perturbative expansions etc.). Heuristically we can say, that we manage to "fix the gauge", i.e. to restrict the space of all gauge configurations, $\mathcal{C}[A]$ to a submanifold $\mathcal{M}[A]$ where gauge equivalent configurations are only counted once. The integration over $\mathcal{C}$ now factorizes in an integration over $\mathcal{M}[A]$, which contains the physics, and an integration over the gauge equivalent configurations, which just produces the infinite product (7.116)[2]. The virtue of this approach is that gauge-invariance formally is manifest (since we have only factored out a product $\prod_{x\in R^d} V(G)$ independent of any dynamics) when we consider expectation values of gauge invariant observables. Further we will show later, that if we define the gauge theory on a lattice, we will be led to precisely this kind of expressions. Since the lattice approach for a finite volume will provide us with well defined finite dimensional integrals, the close relation of (7.22) and (7.23) to the lattice approach is reassuring. The disadvantage of an expression like (7.22) is that it is purely formal: Neither numerator nor denominator in (7.22) exist in a strict sense. One could have avoided this by *first* defining $\mathcal{M}(A)$ by fixing the gauge, but we would then have lost manifest gauge invariance. We will consider here only the first approach.

The first step is to impose a gauge condition which defines $\mathcal{M}(A)$. In general we will be interested in a *covariant* condition like

$$\partial_\mu A_\mu(x)=0\ , \qquad A_\mu=A_\mu^a T^a \tag{7.24}$$

in order not to break euclidean invariance. We will denote the general gauge condition by:

$$\mathcal{F}^a(A)=c^a(x)\ , \quad a=1,\cdots,N \tag{7.25}$$

---

[1] For *compact* groups the Haar measure is unique up to an over all factor

[2] It should be stressed that it is not precisely want we will due, since we will take an average over many such gauge fixing conditions. But that is only done for convenience. In principle one could do with a "genuine" gauge fixing, only will the propagators we get in the end be more singular

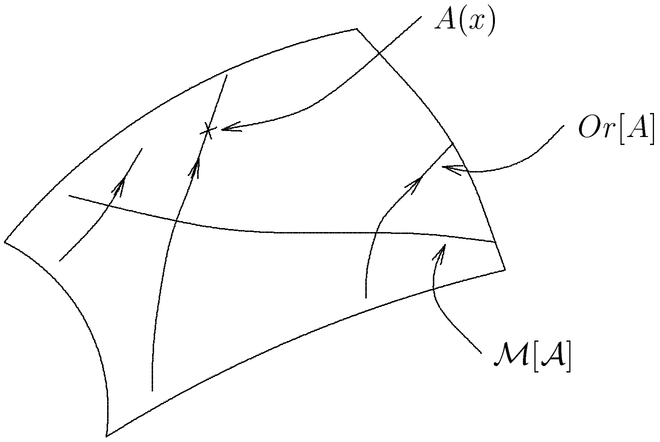


Figure 7.1: For a given configuration $A(x)$ the orbit Or$[A]$ should intersect $\mathcal{M}[A]$ only once

where N is the number of generators in G. The main requirement for $\mathcal{F}$ is the following: Given an $A_\mu(x)$ we define the *orbit* of $A_\mu(x)$ by:

$$\mathrm{Or}[A_\mu(x)] = \{\tilde{A}_\mu(x) \mid \exists\, U(x) \in G \;:\; \tilde{A}_\mu(x) = {}^{U(x)}A_\mu(x)\} \tag{7.26}$$

where ${}^UA$ denotes the gauge transformed of A defined by (7.18). *We assume $\mathcal{F}^a$ is chosen such that for any orbit there is one and only one $\tilde{A}_\mu$ which satisfy (7.25).* The "submanifold" of the configuration space $\{A_\mu(x)\}$ which satisfies (7.25) is denoted $\mathcal{M}$, and its relation to the orbits is shown in fig.7.1.

Let us define the following functional of $A_\mu(x)$ (depending on $\mathcal{F}^a$ and c^a):

$$\frac{1}{\Delta_{\mathcal{F},c}[A]} \equiv \int \prod_x dU_x \prod_{y,a} \delta\left(\mathcal{F}^a({}^UA) - c^a\right) \tag{7.27}$$

where we integrate over the gauge group G at each space-time point x. dU_x denotes the Haar measure on G and ${}^UA(x)$ the gauge transform of $A(x)$:

$${}^UA_\mu(x) = U(x)A_\mu(x)U^{-1}(x) - \frac{i}{g}\partial_\mu U(x)\, U^{-1}(x) \tag{7.28}$$

By the assumption made above there exists for a given configuration $A(x)$ a unique gauge transformation $U(x) = U^{(A)}(x)$ depending on the given $A(x)$ and such that

$$\mathcal{F}^a({}^{U^{(A)}}A) = c^a. \tag{7.29}$$

Let us introduce the following notation:

$$M^{ab}(x,y,A) \equiv \frac{\partial \mathcal{F}^a}{\partial A^c_\mu(x)} D^{cb}_\mu(x;A)\delta^{(d)}(x-y) \tag{7.30}$$

$$D^{ab}_\mu(x;A) \equiv \frac{\partial}{\partial x_\mu}\delta^{ab} + gc^{acb}A^c_\mu(x) \tag{7.31}$$

where D^{ab}_μ is nothing but the matrix components of the covariant derivative D_μ in the adjoint representation. By the determinant $\det M$ of M, we mean the determinant in all indices: a, b and x, y. The last two indices are continuous, and some regularization will usually be needed. We will assume that this can be done and refer to example 2 for a more detailed discussion.

Lemma:

$$\begin{aligned} (1):&\quad \Delta_{\mathcal{F},c}[A] \quad \text{is gauge invariant} \\ (2):&\quad \Delta_{\mathcal{F},c}[A] = \det M\left({}^{U^{(A)}}A\right) \end{aligned}$$

Proof

(1): We use translational invariance of the Haar measure:

$$\begin{aligned} dU &= d(U\,U_0). \\ \Delta^{-1}[{}^{U_0}A] &\equiv \int \prod_x dU_x \prod_{x,a} \delta(\mathcal{F}^a({}^{U}({}^{U_0}A) - c^a) = \\ & \int \prod_x d(U_x\,U_0) \prod_{x,a} \delta(\mathcal{F}^a({}^{UU_0}A) - c^a) = \\ & \int \prod_x d\tilde{U} \prod_{x,a} \delta(\mathcal{F}^a({}^{\tilde{U}}A) - c^a) = \Delta^{-1}[A]. \end{aligned}$$

(2): By (1): $\Delta(A) = \Delta({}^{U^{(A)}}A)$. Call ${}^{U^{(A)}}A = \tilde{A}$. The geometrical interpretation of $\tilde{A}$ is the following: For a given configuration we follow the orbit ${}^{U}A$ until we intersect the submanifold $\mathcal{M}[A]$, see fig.7.1. The point of intersection defines both the configuration $\tilde{A}(x)$ and the gauge transformation $U^{(A)}(x)$. By the definition of the gauge transformation $U^{(A)}(x)$ we have $\mathcal{F}^a(\tilde{A}) = c^a$. We want to calculate $\int \prod_x dU_x \delta(\mathcal{F}^a({}^{U}\tilde{A}) - c^a)$, but need only to expand around $U_x = 1$ (the identity) because of the δ-function and $\mathcal{F}(\tilde{A}) = c^a$. For such expansion:

$$\begin{aligned} U(x) &= e^{i\alpha^a(x)T^a} \simeq 1 + i\alpha^a(x)T^a + \mathcal{O}(\alpha^2) && (7.32) \\ dU_x &= \prod_{a=1}^{N} d\alpha^a(x)(1 + \mathcal{O}(\alpha^2)) && (7.33) \end{aligned}$$

and for such infinitesimal gauge transformations:

$$({}^{U}\tilde{A})^a \simeq \tilde{A}^a + D^{ab}_{\mu}(\tilde{A})\alpha^b + \mathcal{O}(\alpha^2). \tag{7.34}$$

From (7.32)-(7.34) we have:

$$\mathcal{F}^a({}^{U}\tilde{A})(x) = \mathcal{F}^a(\tilde{A})(x) + \frac{\partial \mathcal{F}^a(\tilde{A}(x))}{\partial \tilde{A}^c_{\mu}(x))} D^{cb}_{\mu}(x, \tilde{A})\alpha^b(x) + \mathcal{O}(\alpha^2). \tag{7.35}$$

$$\begin{aligned} \Delta^{-1}[\tilde{A}] &\simeq \int \prod_{x,a'} d\alpha^{a'}_x \prod_{y,a} \delta\left(\frac{\partial \mathcal{F}^a(\tilde{A}(y))}{\partial \tilde{A}^c_{\mu}(y)} D^{cb}_{\mu}(y, \tilde{A})\alpha^b(y)\right) \\ &= \int \prod_{x,a'} d\alpha^{a'}_x \prod_{y,a} \delta\left(\int d^d z M^{ab}(y, z, \tilde{A})\alpha^b(z)\right) \\ &= (\det M(\tilde{A}))^{-1}. \end{aligned}$$

> The last relation is formal and follows from a similar one for finite dimensional real, symmetric matrices M_{ij}:
>
> $$\int \prod_i^n d\alpha_i \prod_j^n \delta(M_{ij}\alpha_j) = \frac{1}{\det M}$$
>
> which is readily proven by an orthonormal transformation on the α's which diagonalizes M.

By means of the lemma we can write (7.27) as

$$1 = \int \prod_x dU_x \left[\det M \left({}^{U(A)}A\right)\right] \prod_{y,a} \delta\left(\mathcal{F}^a({}^U A) - c^a(y)\right) \tag{7.36}$$

The presence of the δ-functions in (7.36) means that we can replace ${}^{U(A)}A$ with ${}^U A$ where U_x is the integration variable:

$$1 = \int \prod_x dU_x \det M\,({}^g A) \prod_{y,a} \delta\left(\mathcal{F}^a({}^U A) - c^a(y)\right) \tag{7.37}$$

The lhs of eq. (7.37) is clearly independent of $c^a(x)$. We can then get rid of the δ-function by multiplying both sides of (7.37) with

$$"1[\alpha]" = \int \prod_{x,a} dc^a(x)\; e^{-\frac{1}{2\alpha}\int d^dx c^a(x)^2} \tag{7.38}$$

where "$1[\alpha]$" means that the normalization of the Gaussian integral will depend on the parameter α. However, again it is a constant independent of any dynamics which will cancel between numerator and denominator in (7.22). From (7.37) and (7.38):

$$"1[\alpha]" = \int \prod dU_x \det M({}^U A) e^{-\frac{1}{2\alpha}\mathcal{F}^2[{}^U A]} \tag{7.39}$$

$$\mathcal{F}^2[A] \equiv \int d^dx \mathcal{F}^a(A)\mathcal{F}^a(A) \tag{7.40}$$

We can finally insert the constant "$1[\alpha]$" in our formal expression (7.22) for expectation values of gauge invariant observables. We have

$$\begin{aligned}\int \mathcal{D}A\, e^{-S[A]} &= \int \mathcal{D}A \int \prod_x dU_x \det M({}^U A) e^{-S[A]-\frac{1}{2\alpha}\mathcal{F}^2[{}^U A]} \\ &= \int \prod_x dU_x \int \mathcal{D}A\;\; \det M({}^U A)\; e^{-S[A]-\frac{1}{2\alpha}\mathcal{F}^2[{}^U A]}\end{aligned}$$

By the gauge invariance[3] of $\mathcal{D}A$ and the gauge invariance of the Yang-Mills action $S[A]$ we can change integration variables to ${}^U A$ and we get the wanted factorization of the gauge group:

$$\boxed{\int \mathcal{D}A e^{-S[A]} = \left[\int \prod_x dU_x\right] \cdot \int \mathcal{D}A e^{-S[A]-\frac{1}{2\alpha}\mathcal{F}^2[A]} \det M(A)} \tag{7.41}$$

---

[3] A gauge transformation (7.28) is a rotation + translation of A_μ, therefore $\mathcal{D}A$ is invariant

The same argument is clearly true for the numerator in (7.22) *because of the assumed gauge invariance of the observables $\mathcal{O}_n(A)$ and the gauge group factor $[\int \prod_x dU_x]$ cancels in (7.22)*:

$$\boxed{< \prod_{i=1}^{n} \mathcal{O}_i(A) >= \frac{\int \mathcal{D}A \ \prod_{i=1}^{n} \mathcal{O}_i(A) \ e^{-S_{eff}[A]}}{\int \mathcal{D}A \ e^{-S_{eff}[A]}}} \tag{7.42}$$

$$\boxed{S_{eff}[A] = S[A] + \frac{1}{2\alpha}\mathcal{F}^2[A] - \log\det M(A)} \tag{7.43}$$

These formulae represent the final result of our formal manipulations. We have managed to get a cancellation of an infinite factor between the numerator and denominator and $S_{eff}[A]$ will have a gaussian part which is invertible. $S_{eff}[A]$ is of course no longer invariant under local gauge transformations, but since our starting point (7.22) *was* gauge invariant we expect that expectation values of gauge invariant operators should still respect gauge invariance even if they are calculated by means of an action which is not gauge invariant. In section 7.7 we shall see that (7.43) has a very interesting symmetry which resembles local gauge invariance sufficiently to insure the gauge invariance of the perturbative expansion generated by S_{eff}.

Before we state the Feynman rules it is convenient to rewrite the term $\log\det M(A)$.

7.2 Gaussian propagators

The determinant term in (7.43) pose a slight problem for perturbation theory: It is non-local. This means it is an infinite series in the coupling constant. To be more explicit let us discuss the case where $\mathcal{F}^a(A)$ is given by:

$$\mathcal{F}^a(A) = \partial_\mu A^a_\mu \tag{7.44}$$

From this we get that the "matrix" $M(A)$ is given by

$$\mathcal{M}^{ab}(x,y,A) = \partial_\mu D^{ab}_\mu \delta(x-y) = \left(\partial^2_\mu \delta^{ab} + g c^{acb}\partial_\mu A^c_\mu(x)\right)\delta(x-y) \tag{7.45}$$

It is convenient to divide by a trivial factor det ∂^2 which has no reference to the dynamics: If we as usual denote ∂^{-2} as $\Delta(x-y)$ we get:

$$\frac{\det \ M(A)}{\det \ \partial^2} = \det \ \left[\frac{M(A)}{\partial^2}\right] = \det \ (1+L) \tag{7.46}$$

where the "matrix" $1+L$, which is strictly speaking the kernel of an operator, is given by

$$(1+L)^{ab}(x,y) = \delta^{ab}\delta(x-y) + g\int d^dz \Delta(x-z)\frac{\partial}{\partial z_\mu} c^{acb} A^c_\mu(z)\delta^{(d)}(z-y) \tag{7.47}$$

We now use the following rewriting of a determinant, which can easily be checked if we can diagonalize the matrix, but which is valid under more general circumstances:

$$\det(1+L) = e^{\mathrm{Tr}\ \log(1+L)} \tag{7.48}$$

Finally the power expansion of the logarithm gives

$$\mathrm{Tr}\ \log(1+L) = \mathrm{Tr}\ \left(\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\ L^n\right) \tag{7.49}$$

and we can write

$$\log \frac{\det\ M(A)}{\det\ \partial^2} = \mathrm{Tr}\ \left(\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\ L^n\right) \tag{7.50}$$

This is the effective term which enters in (7.43). From the expression (7.47) we see that it is an infinite power series in the coupling constant g, and for each new order in the perturbation expansion we will have to introduce new vertices. The trace Tr in (7.48)-(7.50) means trace over everything: sum over the group indices, and integration over spacetime points, as already discussed. It is instructive to write out in detail the first few terms in the expansion (7.50).

Example 2: The power expansion of $\det M(A)/\partial^2$.

$$\begin{aligned}
&\mathrm{Tr}\ \log(1+L) = \\
&\quad +g\int dz_1 c^{aba}\left[\Delta(x-z_1)\frac{\partial}{\partial z_1}A^c(z_1)\delta(z_1-x)\right] \\
&\quad -\frac{g^2}{2}\int dz_1dz_2 c^{abc}\left[\Delta(x-z_1)\frac{\partial}{\partial z_1}A^b(z_1)\delta(z_1-z_2)\right]c^{cda}\left[\Delta(z_2-z_1)\frac{\partial}{\partial z_2}A^d(z_2)\delta(z_2-x)\right] \\
&\quad +\frac{g^3}{3}\int dz_1dz_2dz_3[\cdots]-\cdots
\end{aligned}$$

The first term in (7.50) vanish since one can show that the structure constants can be chosen antisymmetric. (recall their connection with commutators).

To circumvent the problem of an infinite set of vertices in the effective action (7.43) we will use some auxiliary fermionic variables, called *ghosts*, to represent the determinant as a gaussian fermionic "path" integral. Recall that we have:

$$\int \mathcal{D}\phi\mathcal{D}\bar{\phi}\ e^{-\bar{\phi}M\phi}\ \propto (\det M)^{\pm 1}$$

according to whether $\phi, \bar{\phi}$ are anticommuting or commuting variables. To represent $\det M$ we need anticommuting variables which we denote $\eta, \bar{\eta}$. There is a certain ambiguity in this fermionic representation. For a finite dimensional determinant we have:

$$\int \prod_{i=1}^{N} d\eta_i d\bar{\eta}_i\ e^{\pm\bar{\eta}M\eta} = (\pm 1)^N \det M$$

$$\int \prod_{i=1}^{N} d\eta_i d\bar{\eta}_i\ e^{\pm i\bar{\eta}M\eta} = (\pm i)^N \det M$$

Since $\det M$ appears both in numerator and denominator when we calculate expectation values, we can take any of these choices. If we (arbitrarily) choose $e^{\bar{\eta}M\eta}$, we get a total effective action:

$$\mathcal{F}^a(A) = \partial_\mu A^a_\mu \tag{7.51}$$

$$\boxed{S_{eff}[A,\eta,\bar\eta] = \int d^dx \left(\frac{1}{4}(F^a_{\mu\nu}(A))^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu)^2 - \bar\eta^a\partial_\mu D^{ab}_\mu\eta^b\right)} \tag{7.52}$$

We get a similar expression for any function $\mathcal{F}^a(A)$ only:

$$\frac{1}{2\alpha}(\partial_\mu A^a_\mu)^2 \to \frac{1}{2\alpha}\mathcal{F}^a(A)^2 \tag{7.53}$$

$$\bar\eta^a\partial_\mu D^{ab}_\mu\eta^b \to \bar\eta^a\frac{\partial\mathcal{F}^a(A)}{\partial A^c_\mu}D^{cb}_\mu(A)\eta^b \tag{7.54}$$

In the following we will restrict ourself to the choice $\mathcal{F}^a(A) = \partial_\mu A^a_\mu$ as given by (7.51), but all the general results we derive will be valid for a general linear gauge condition $\mathcal{F}^a(A) = \varphi^{ab}_\mu A^b_\mu$ (our choice is $\varphi^{ab}_\mu = \partial_\mu\delta^{ab}$, but other choices like $\varphi^{ab}_\mu = n_\mu\delta^{ab}$, where n_μ is a four-vector, have some interest and are denoted *axial gauges*). Many of the results, but not all, will be valid for a general $\mathcal{F}^a(A)$.

Since

$$D^{ab}_\mu = \partial_\mu\delta^{ab} + gc^{acb}A^c_\mu \tag{7.55}$$

$$F^a_{\mu\nu}(A) = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gc^{abc}A^b_\mu A^c_\nu \tag{7.56}$$

the quadratic part of the effective action (7.43) is given by:

$$S^{(2)}[A,\eta,\bar\eta] = \int d^dx\frac{1}{2}A^a_v\left\{\delta^{ab}\left(-\partial^2\delta_{\mu\nu} + \partial_\mu\partial_\nu(1-\frac{1}{\alpha})\right)\right\}A^b_\mu + \bar\eta^a(-\delta^{ab}\partial^2)\eta^b. \tag{7.57}$$

It follows that the quadratic forms are non-singular. In fact, by Fourier expanding we find that we have to invert $\delta^{ab}(k^2\delta_{\mu\nu} - (1-\frac{1}{\alpha})k_\mu k_\nu)$ and this is readily done. The propagator is:

$$\Delta^{ab}_{\mu\nu}(k) = \delta^{ab}\frac{1}{k^2}\left(\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right) \tag{7.58}$$

Similarly the ghost propagator becomes:

$$\Delta^{ab}(k) = \delta^{ab}\frac{1}{k^2} \tag{7.59}$$

7.3 Feynman rules

Since the gaussian part $S^{(2)}[A,\eta,\bar\eta]$ of $S_{eff}[A,\eta,\bar\eta]$ is invertible, it is possible to define a perturbation expansion in a standard fashion. $S_{eff}[A,\eta,\bar\eta] - S^{(2)}[A,\eta,\bar\eta]$ is called the interaction:

$$\begin{aligned} S_{int}[A,\eta,\bar\eta] &= g\,c^{abc}\int d^dx\ [(\partial_\mu A^a_\nu)A^b_\mu A^c_\nu + \partial_\mu\bar\eta^a A^b_\mu\eta^c] \\ &\quad +\frac{1}{4}g^2c^{abc}c^{ade}\int d^dx\ A^b_\mu A^c_\nu A^d_\mu A^e_\nu \end{aligned} \tag{7.60}$$

We can now expand $e^{-S_{int}[A,\eta,\bar\eta]}$ in power series in g and make Wick contractions. The only additional rule, as always when we have anticommuting variables, is that *each closed ghost loop should be assigned an additional* (-1) *factor.*

We can proceed in standard fashion and define the generating functional for connected Green functions by:

$$e^{-F[J,\bar{\xi},\xi]} = \int \mathcal{D}A\mathcal{D}\eta\mathcal{D}\bar{\eta}\ e^{-S_{eff}[A,\eta,\bar{\eta}]+\int d^dx(JA+\bar{\xi}\eta+\bar{\eta}\xi)} \tag{7.61}$$

and the generating functional for 1PI-Green functions:

$$\Gamma(A^{cl},\eta^{cl},\bar{\eta}^{cl}) = F(J,\xi,\bar{\xi}) + \int d^dx(JA^{cl} + \bar{\xi}\eta^{cl} + \bar{\eta}^{cl}\xi) \tag{7.62}$$

We have the same relations between sources $J,\bar{\xi},\xi$ and the "classical" fields $A^{cl},\eta^{cl},\bar{\eta}^{cl}$ as for the scalar field theories and the abelian gauge theories, except that we have to take into account the anticommuting nature of η^{cl}, $\bar{\eta}^{cl}$ and $\bar{\xi},\xi$:

$$A^{cl}_\mu(x) = -\frac{\delta F}{\delta J_\mu(x)}, \quad \eta^{cl}(x) = -\frac{\delta F}{\delta\bar{\xi}(x)}, \quad \bar{\eta}^{cl}(x) = +\frac{\delta F}{\delta\xi(x)} \tag{7.63}$$

$$J_\mu(x) = +\frac{\delta\Gamma}{\delta A^{cl}_\mu(x)}, \quad \bar{\xi}(x) = -\frac{\delta\Gamma}{\delta\eta^{cl}(x)}, \quad \xi(x) = +\frac{\delta\Gamma}{\delta\bar{\eta}^{cl}(x)} \tag{7.64}$$

In order to get the complete set of rules, which allow us to write down the Feynman integral for any graph which appears in the diagrammatic expansion, we have to know what weight to assign to the vertices too. The vertex functions are slightly more complicated than for the scalar fields since A^a_μ have a,μ indices and derivatives appear, but in principle the rules are given in chapter 3, and the recipe is the same: We rewrite the interaction terms in momentum space and rearrange them as polynomials in A^a_μ and $\eta,\bar{\eta}$. In the following example we illustrate the procedure for the three-point vertex.

Example 3: The three-point vertex

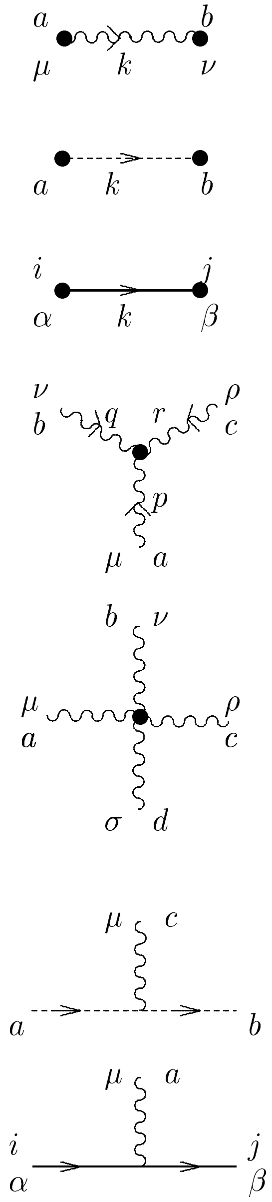
$$g\ c^{abc}\int d^dx\,(\partial_\mu A^a_\nu)\,A^b_\mu A^c_\nu = \int \frac{d^dkd^dpd^dq}{(2\pi)^{3d}}\ \frac{1}{3!}A^a_\mu(k)A^b_\nu(p)A^c_\lambda(q)\tilde{\Gamma}^{abc}_{\mu\nu\lambda}(k,p,q)$$

where a Fourier transformation results in the following expression for V:

$$\begin{aligned}\tilde{\Gamma}^{abc}_{\mu\nu\lambda}(k,p,q) &= (2\pi)^d\delta^{(d)}(k+p+q)\ V^{abc}_{\mu\nu\lambda}(k,p,q)\\ V^{abc}_{\mu\nu\lambda}(k,p,q) &= -igc^{abc}[(q-p)_\mu\delta_{\nu\lambda} + (p-k)_\lambda\delta_{\mu\nu} + (k-q)_\nu\delta_{\mu\lambda}]\end{aligned}$$

In this way we can derive all the rules and a list of these is given in fig.7.2. Only two comments are necessary: The propagators and interactions for fermions (quarks) are included. They will be introduced later. Next, we assume that dimensional regularization is used. The coupling constant g will get a dimension as we leave $d = 4$. However, we prefer to keep g dimensionless and introduce the scale parameter μ, discussed in detail in connection with renormalization of the scalar theories. The coupling constant $\tilde{g}$ appearing in fig.7.2 is therefore:

$$\tilde{g} = \mu^{\varepsilon/2}g, \qquad \varepsilon = 4-d \tag{7.65}$$



$$\frac{\delta^{ab}}{k^2}\left(\delta_{\mu\nu}-(1-\alpha)\frac{k_\mu k_\nu}{k^2}\right)$$

$$\frac{\delta^{ab}}{k^2}$$

$$\left[\frac{i\delta_{ij}}{\not{k}+im}\right]_{\beta\alpha}$$

$$-i\tilde{g}C^{abc}[(r-q)_\mu\delta_{\nu\rho}+(q-p)_\rho\delta_{\mu\nu}+(p-r)_\nu\delta_{\rho\mu}]$$

$$(p+q+r=0)$$

$$\begin{aligned}-\tilde{g}^2\,[&C^{abe}C^{cde}(\delta_{\mu\rho}\delta_{\nu\sigma}-\delta_{\nu\rho}\delta_{\mu\sigma})+\\ &C^{cbe}C^{ade}(\delta_{\mu\rho}\delta_{\nu\sigma}-\delta_{\nu\mu}\delta_{\rho\sigma})+\\ &C^{dbe}C^{cae}(\delta_{\sigma\rho}\delta_{\nu\mu}-\delta_{\nu\rho}\delta_{\mu\sigma})]\end{aligned}$$

$$i\tilde{g}C^{abc}q_\mu$$

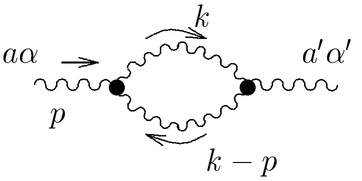
$$i\tilde{g}(T_f^a)_{ji}(\gamma_\mu)_{\beta\alpha}$$

Figure 7.2: Euclidean space Feynman rules in covariant gauge

Before starting any calculation it should be made clear that the Green functions which one calculate by summing Feynman diagrams *are not gauge invariant.* They do not belong to the class of observables which we discuss in (7.42). This is clear from the defining equation for the generating functionals, (7.61) or (7.62). The source term $\int d^dx J\cdot A$ is not invariant under local *non-abelian* gauge transformations. The Green functions themselves are not directly related to physical observables. They are useful quantities when we discuss renormalization of the gauge theories and they can be used as an important tool (essentially the only one we have) when we try to extract gauge invariant information from the theories.

To get an idea of the structure of one loop calculations we will provide some details for one of the diagrams.

Example 4: Calculation of



The Feynman integral is

$$\int \frac{d^dk}{(2\pi)^d}\, V^{abc}_{\alpha\beta\gamma}(p,k-p,-k)\Delta^{cc'}_{\gamma\gamma'}(k)V^{a'b'c'}_{\alpha'\beta'\gamma'}(-p,p-k,k)\Delta^{bb'}_{\beta'\beta}\ (k-p)$$

where the two vertex functions which enters are given by:

$$V^{abc}_{\alpha\beta\gamma}(p,k-p,-k) = -i\tilde{g}c^{abc}\left[(-2k+p)_\alpha\delta_{\beta\gamma} + (p+k)_\beta\delta_{\alpha\gamma} + (k-2p)_\gamma\ \delta_{\alpha\beta}\right]$$

$$V^{a'b'c'}_{\alpha'\beta'\gamma'}(-p,p-k,k) = -i\tilde{g}c^{a'b'c'}\ \left[(2k-p)_{\alpha'}\delta_{\beta'\gamma'} + (-p-k)_{\beta'}\delta_{\alpha'\gamma'} + (-k+2p)_{\gamma'}\delta_{\alpha'\beta'}\right]$$

The propagators are given by:

$$\frac{1}{k^2}\left(\delta_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right)\delta^{ab}$$

and in order to simplify the calculation we choose $\alpha = 1$ (Feynman gauge), but recall that the result will be gauge dependent: it will depend on the chosen α. After some algebra we get:

$$-\tilde{g}^2 c^{abc}c^{a'bc}\ \int \frac{d^dk}{(2\pi)^d}\,\frac{F(k,p)}{k^2(k-p)^2}$$

where

$$F(k,p) = (-4d+6)k_\alpha k_{\alpha'} + (-d+6)p_\alpha p_{\alpha'} + (2d-3)(k_\alpha p_{\alpha'} + k_{\alpha'}p_\alpha) - (5p^2 + 2k^2 - 2pk)\delta_{\alpha\alpha'}$$

In deriving this result we have used: $\delta_{\alpha\alpha} = d$. Further we note that $c^{abc}c^{a'bc} = \delta^{aa'}C_2(G)$ where $C_2(G)$ is the Casimir of the group G in the adjoint representation.

In order to calculate the integral we need to know:

$$\int \frac{d^dk}{(2\pi)^d}\,\frac{1,\ k_\alpha,\ k_\alpha k_{\alpha'}}{k^2(k-p)^2}$$

and from the formula of 1-loop integrals already given in chapter 5 we have:

$$\begin{aligned}
\int \frac{d^dk}{(2\pi)^d}\,\frac{1}{k^2(k-p)^2} &= \frac{\Gamma(2-\frac{d}{2})B(\frac{d}{2}-1,\frac{d}{2}-1)}{(4\pi)^{d/2}}(p^2)^{d/2-2} \equiv I(d,p)\\
\int \frac{d^dk}{(2\pi)^d}\,\frac{k_\mu}{k^2(k-p)^2} &= \frac{1}{2}p_\mu I(d,p)\\
\int \frac{d^dk}{(2\pi)^d}\,\frac{k_\mu k_\nu}{k^2(k-p)^2} &= \left(\frac{d}{4(d-1)}p_\mu p_\nu - \frac{1}{4(d-1)}p^2\delta_{\mu\nu}\right)I(d,p)
\end{aligned}$$

In the above equations the B-function is defined by

$$B(x,y) = \int_0^1 d\alpha \alpha^{x-1}(1-\alpha)^{y-1}$$

and we have

$$B(x,y) = \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)}$$

which at the same time provides us with an analytic continuation of the B-function.

Finally, after some algebra, we get:

$$\left[\left\{\frac{(-4d+6)}{2(d-1)} + (-d+6) + (2d-3)\right\} p_\alpha p_{\alpha'} + \left\{\frac{4d-6}{4(d-1)} - 4\right\} p^2 \delta_{\alpha\alpha'}\right] \cdot$$
$$\left(-\delta^{aa'} C_2(G) \cdot \Gamma(2-\frac{d}{2}) B(\frac{d}{2}-1, \frac{d}{2}-1)\right) \frac{g^2}{16\pi^2} \cdot \left(\frac{p^2}{4\pi\mu^2}\right)^{d/2-2}$$

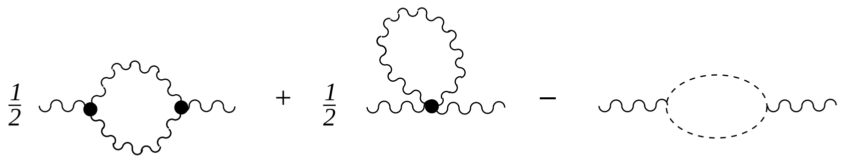
For $\varepsilon \equiv 4 - d \to 0$ the pole term (the divergent part!) is:

$$-\delta^{ab} C_2(G) \frac{g^2}{16\pi^2} \left(\frac{11}{3} p_\alpha p_{\alpha'} - \frac{19}{6} \delta_{\alpha\alpha'} p^2\right) \frac{2}{\varepsilon}$$

In a similar way all one loop correction to the propagators and vertex functions can be calculated. We list for completeness the correct combinational weights of the diagrams, which can either be derived either by making Wick contractions directly or by working out the Dyson-Schwinger equations as for the scalar case (see chapter 4). Further, the divergent contributions to the propagators and vertex functions are given. We encourage the reader to check the results.

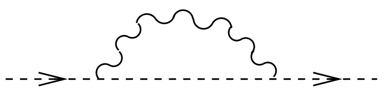
Example 5: Divergent parts of one-loop diagrams.

(1): Gauge propagators self-energy



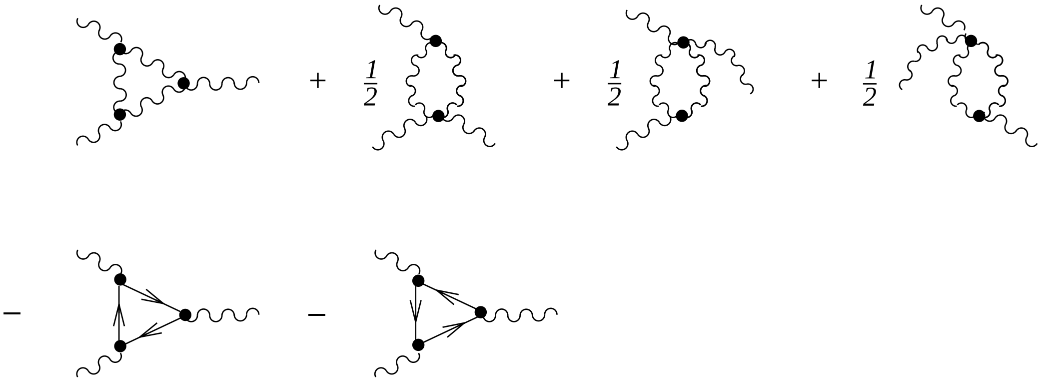
Divergent part: $-\frac{g^2 C_2(G)}{16\pi^2} \left(\frac{5}{3} + \frac{1}{2}(1-\alpha)\right) \frac{2}{\varepsilon} \cdot \{p^2 \delta_{\mu\nu} - p_\mu p_\nu\} \delta^{ab}$

(2): The ghost propagators self-energy



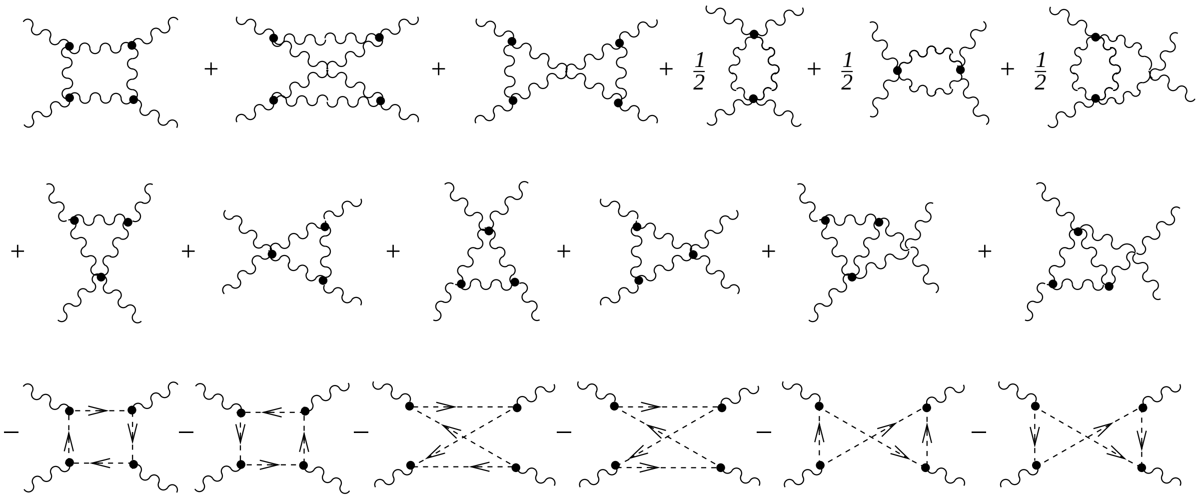
Divergent part: $-\frac{g^2 C_2(G)}{16\pi^2} \left(\frac{1}{2} + \frac{1}{4}(1-\alpha)\right) \frac{2}{\varepsilon} \cdot p^2 \delta^{ab}$

(3): Three-point vertex



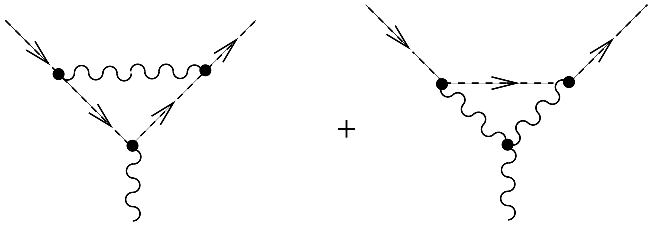
Divergent part: $-\frac{g^2C_2(G)}{16\pi^2}[\frac{2}{3}+\frac{3}{4}(1-\alpha)]\frac{2}{\varepsilon}\cdot V^{abc}_{\mu\nu\lambda}(k,p,q)$

(4): Four-point vertex



Divergent part: $-\frac{g^2C_2(G)}{16\pi^2}[-\frac{1}{3}+(1-\alpha)]\frac{2}{\varepsilon}\cdot V^{abcd}_{\mu\nu\lambda\eta}$

(5): Ghost-vector vertex



Divergent part: $\frac{g^2C_2(G)}{16\pi^2}\frac{\alpha}{2}\cdot\frac{2}{\varepsilon}\cdot V^{abc}_{\mu}(p)$.

7.4 One-loop renormalization

We can analyse the divergences in the same way as for the scalar theory. We will be interested in the theory when the dimension of spacetime is four. The coupling constant g is dimensionless in four dimensions: $[g] = 0$ and according to our analysis this means that the theory is renormalizable (the fact that we have ghosts does not change this conclusion), but the question is more complicated than for the scalar theories and the abelian gauge theories we have considered earlier. The reason is that we want to maintain the invariance under local gauge transformations. If local gauge invariance is not maintained we have obviously modified the theory in a drastic way. At the formal level there is no reason why we should break gauge invariance if we use dimensional regularization. However, the perturbation expansion itself cannot use the full gauge invariant action but use the factorization (7.41) and the corresponding effective action given by (7.43) or (7.52) does not have the local non-abelian symmetry. Further we split the action in a gaussian part and the rest, which we call the interaction. This split does not respect non-abelian gauge invariance either. These two bad features of the perturbative expansion complicate the proof that the theory can be renormalized in a sensible way. In a later section we will prove that $S_{eff}[A, \eta, \bar{\eta}]$ has a hidden invariance, the so-called BRS invariance, which ensures that we can perform a multiplicative renormalization which preserves the gauge structure of the theory. In this section we will limit ourselves to show that the explicit calculations of the last section allow a multiplicative renormalization at the one-loop level.

We can proceed as for the scalar theory and classify the diagrams according to their superficial divergence $\omega(D)$ of diagram D:

$$\omega(D) = 4 - E_A - \frac{3}{2} E_{\eta\bar{\eta}} \tag{7.66}$$

where E_A and $E_{\eta\bar{\eta}}$ denote the external lines in the diagram. Due to the nature of the ghost interaction we have an identical number of η and $\bar{\eta}$ lines. This expression (valid in four dimensions) is the same for the scalar theory except for term $\frac{3}{2}E_{\eta\bar{\eta}}$. The reason for the factor $\frac{3}{2}$ is that ghost-gauge field vertex $V^{A\bar{\eta}\eta}$ contains a factor p_μ proportional to the momentum associated with the ghost line (see the table of Feynman rules). This factor is associated with the derivative ∂_μ acting on $\bar{\eta}$ in the term $g(\partial_\mu \bar{\eta}^a) c^{abc} A^b_\mu \eta^c$. Note the asymmetry: the derivative only acts on $\bar{\eta}$. As half of the external $E_{\eta\bar{\eta}}$ lines in a 1PI diagram are $\bar{\eta}$ "lines", effectively $\frac{1}{2}E_{\eta\bar{\eta}}$ momenta p_i are external momenta and therefore not effective in the in the power counting of divergences of the loop integral. Consequently $\omega(D)$ is reduced with $\frac{1}{2}E_{\eta\bar{\eta}}$. Effectively this means: $E_{\eta\bar{\eta}} \to E_{\eta\bar{\eta}} + \frac{1}{2}E_{\eta\bar{\eta}}$ in (7.66).

From (7.66) it is seen that the superficially divergent diagrams, i.e. the diagrams D where $\omega(D) \geq 0$, precisely correspond to the vertex functions in the lagrangian and to the propagators, as illustrated in fig.7.3.

The 1-loop diagrams mentioned in the last section are the only divergent 1-loop diagrams according to this analysis and the counter terms which we have to add to $\mathcal{L}_{eff}(A, \eta, \bar{\eta})$ in order to get finite 1-loop results is fixed by the above 1-loop calculation. If we use MS (minimal subtraction) we only have to add pole terms which, as is seen explicitly from the divergent part, can be chosen as local polynomials in the fields $A^a_\mu, \eta^a, \bar{\eta}^a$:

$$\delta\mathcal{L}_{A^2} \;=\; (Z_3 - 1)\, \frac{1}{4}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)^2 \tag{7.67}$$

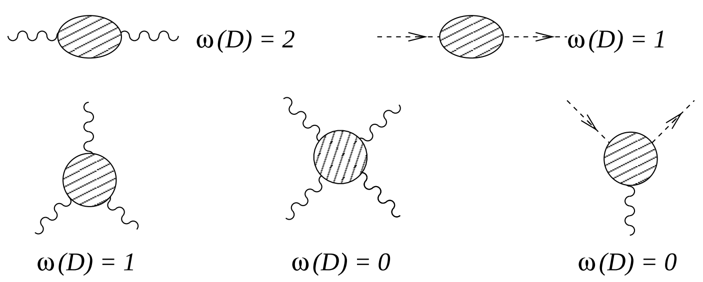


Figure 7.3: The superficially divergent 1PI diagrams for non-abelian gauge theories

$$Z_3 = 1 - \frac{g^2 C_2(G)}{16\pi^2}\left(\frac{5}{3} + \frac{1}{2}(1-\alpha)\right)\frac{2}{\varepsilon}$$

$$\begin{aligned} \delta\mathcal{L}_{\bar\eta,\eta} &= (\tilde Z_3 - 1)(-\bar\eta^a \partial^2 \eta^a) \qquad (7.68)\\ \tilde Z_3 &= 1 - \frac{g^2 C_2(G)}{16\pi^2}\left(\frac{1}{2} + \frac{1}{4}(1-\alpha)\right)\frac{2}{\varepsilon} \end{aligned}$$

$$\begin{aligned} \delta\mathcal{L}_{A^3} &= (Z_1 - 1)\; g\partial_\mu A^a_\nu c^{abc} A^b_\mu A^c_\nu \qquad (7.69)\\ Z_1 &= 1 - \frac{g^2 C_2(G)}{16\pi^2}\left(\frac{2}{3} + \frac{3}{4}(1-\alpha)\right)\frac{2}{\varepsilon} \end{aligned}$$

$$\begin{aligned} \delta\mathcal{L}_{A^4} &= (Z_4 - 1)\; \frac{g^2 1}{4} c^{abc} c^{ade} A^b_\mu A^c_\nu A^d_\mu A^e_\nu \qquad (7.70)\\ Z_4 &= 1 - \frac{g^2 C_2(G)}{16\pi^2}\left(\frac{1}{3} + (1-\alpha)\right)\frac{2}{\varepsilon} \end{aligned}$$

$$\begin{aligned} \delta\mathcal{L}_{\bar\eta A\eta} &= (\tilde Z_1 - 1)\; g c^{abc} \partial_\mu \bar\eta^a A^b_\mu \eta^c \qquad (7.71)\\ \tilde Z_1 &= 1 + \frac{g^2 C_2(G)}{16\pi^2}\frac{\alpha}{2}\cdot\frac{2}{\varepsilon} \end{aligned}$$

From (7.67)-(7.71) it follows that

$$\begin{aligned} \mathcal{L} + \delta\mathcal{L} \;=\; & \frac{1}{4}\, Z_3(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)^2 + \frac{1}{2\alpha}(\partial_\mu A^a_\mu)^2 \\ & + g c^{abc} Z_1\; \partial_\mu A^a_\nu A^b_\mu A^c_\nu + \frac{g^2}{4} c^{abc} c^{ade}\; Z_4\; A^b_\mu A^c_\nu A^d_\mu A^e_\nu \qquad (7.72)\\ & + \tilde Z_3 \partial_\mu \bar\eta^a \partial_\mu \eta^a + g c^{abc}\; \tilde Z_1\; \partial_\mu \bar\eta^a A^b_\mu \eta^c \end{aligned}$$

A priori this lagrange function has nothing to do with the one we started with. Originally we had one coupling constant g, now we have three: $Z_1 g, Z_4 g^2$ and $\tilde Z_1 g$. We can ask: Is it possible by a (multiplicative) renormalization to regard $\mathcal{L} + \delta\mathcal{L}$ as a (bare) lagrangian $\mathcal{L}_0(A_0, \eta_0\bar\eta_0, g_0, \alpha_0)$ (recall the discussion for scalar theories):

$$\mathcal{L}(A, \bar\eta, \eta, g, \alpha) + \delta\mathcal{L}(A, \bar\eta, \eta, g, \alpha) = \mathcal{L}_0(A_0, \eta_0, \bar\eta_0, g_0, \alpha_0) \qquad (7.73)$$

From (7.67)-(7.71) and the equations (7.72) and (7.73) we get the following consistency relations:

$$Z_3\;(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)^2 = \;(\partial_\mu A^a_{0\nu} - \partial_\nu A^a_{0\mu})^2 \;\;\Rightarrow A_0 = Z_3^{\frac{1}{2}} A$$

$$\begin{array}{rcl} \tilde{Z}_3\ (\partial_\mu \bar{\eta} \partial_\mu \eta^a) = & \partial_\mu \bar{\eta}_0^a \partial \eta_0^a & \Rightarrow \eta_0,\ \bar{\eta}_0 = \tilde{Z}_3^{\frac{1}{2}}\ \eta,\ \bar{\eta} \\ \dfrac{1}{2\alpha}(\partial_\mu A_\mu^a)^2 = & \dfrac{1}{2\alpha}(\partial_\mu A_{0\mu}^a)^2 & \Rightarrow \alpha_0 = Z_3\ \alpha \\ Z_1\ g \partial_\mu A_\nu^a A_\mu^b A_\nu^c = & g_0 \partial_\mu A_{0\nu}^a A_{0\mu}^b A_{0\nu}^c & \Rightarrow g_0 = \dfrac{Z_1}{Z_3^{\frac{3}{2}}}\ g \end{array} \tag{7.74}$$

But in addition to the relations (7.74) we still have the following two relations:

$$\begin{array}{rcl} g^2\ Z_4 A^4 = & g_0^2\ A_0^4 & \Rightarrow \dfrac{Z_4}{Z_1} = \dfrac{Z_1}{Z_3} \\ g \tilde{Z}_1\ \bar{\eta} A \eta = & g_0 \bar{\eta}_0 A_0 \eta_0 & \Rightarrow \dfrac{\tilde{Z}_1}{\tilde{Z}_3} = \dfrac{Z_1}{Z_3} \end{array} \tag{7.75}$$

These relations are called the *Slavnov-Taylor identities*, and at the present stage we have no real understanding of the identities. We can only check from (7.67)-(7.71) whether (7.75) is satisfied of not. Since the Z_i's to 1-loop order have the form $1 + c_i g^2/\varepsilon + (g^4)$ we can to a 1-loop approximation rewrite (7.75) as

$$Z_4 = 2Z_1 - Z_3 + \mathcal{O}(g^4), \qquad \tilde{Z}_1 = Z_1 + \tilde{Z}_3 - Z_3 + \mathcal{O}(g^4). \tag{7.76}$$

From (7.67)-(7.71) we see that (7.76) indeed is satisfied to one loop accuracy. A similar result is true to two loop. This indicates that non-abelian gauge theories are multiplicative renormalizable. In the section on BRS-invariance we will prove (7.76) without relying on direct calculations. However, as already remarked, the results are not that *surprising*: As long as we use a regularization which respects the symmetries of the lagrangian, there is no need to introduce counter terms which break the symmetry. Before turning to the rather technical aspects of BRS-symmetry we will discuss how to include fermions (quarks) in the above calculations, and also discuss some implications of the one-loop results.

7.5 Fermions

The formalism can readily be extended to include matter fields. We know that gluons couple to quarks. The coupling in *minkowskian* spacetime can be written as

$$\mathcal{L}_{qg}(A, \psi) = -\bar{\psi}(x)(\gamma^\mu D_\mu + m)\psi \tag{7.77}$$

The metric $g_{\mu\nu}$ in d-dimensional minkowskian space-time is the usual one: $g_{00} = -1$ and $g_{ij} = \delta_{ij},\ i, j = 1, ..., d-1$. The γ-matrices satisfy

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu} \tag{7.78}$$

and $\bar{\psi}$ is defined by

$$\bar{\psi} = \psi^\dagger \gamma_0 \tag{7.79}$$

Finally, the covariant derivative is defined by (as already mentioned)

$$D_\mu \equiv \partial_\mu - ig(T_f^a)_{ij} A_\mu^a \tag{7.80}$$

In this formula for the covariant derivative the matrices T_f^a constitute a N_f-dimensional representation of the non-abelian Lie group G and the fermions form a column vector

$$\psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_{Nf} \end{pmatrix}, \tag{7.81}$$

but at the same time they also constitute a representation of the Lorentz group and have a spinor index α ranging from 1 to N_{spin}, the dimension of the spinor representation of the Lorentz group:

$$\psi \sim \psi_{i\alpha}, \quad i = 1, \cdots, \quad N_f, \quad \alpha = 1, \cdots, N_{spin} \tag{7.82}$$

When we rotate to euclidean space the Lorentz group goes into $SO(d)$ where d is the dimension of the euclidean space. The spinor representations of $SO(d)$ can still be characterized by γ-matrices now satisfying (as already mentioned when we discussed QED)

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}, \qquad \mu = 1, \ldots, d. \tag{7.83}$$

Further we recall from the discussion of euclidean fermions in QED that $\bar{\psi}$ will have to be treated as an independent variable, not related to ψ by $\bar{\psi} = \psi^\dagger \gamma_0$. We will define $\bar{\psi}$ to transform like the adjoint of ψ with respect to all transformations. Then $\bar{\psi}\psi$ is a scalar, $\bar{\psi}\gamma_\mu\psi$ a vector etc.

The euclidean action for the Dirac fields will now be:

$$S = \int d^d x \bar{\psi}(\gamma_\mu D_\mu + m)\psi \tag{7.84}$$

and this action is invariant under local gauge transformations, $SO(d)$ rotations and translations. The Feynman rules can be found for this new term and have already been given in the general table of Feynman rules. They are separately stated in fig.7.4, where we have also shown the new class of superficially divergent 1PI diagrams. The diagrams are, as usually for renormalizable theories, associated with the propagators and the vertex functions already present in the lagrangian. Explicitly we have here fermionic self-energy diagrams and diagrams associated with the $\bar{\psi}\gamma_\mu A_\mu \psi$ vertex. The diagrams have superficial divergence $\omega(D) = 1$ and $\omega(D) = 0$, respectively. The divergences can be cancelled by counter terms compatible with a multiplicative renormalization. We have

$$\delta\mathcal{L}(A, \bar{\psi}, \psi) = (Z_{f2} - 1)\, \bar{\psi}\gamma_\mu\partial_\mu\psi - i(Z_{f1} - 1)\, g\bar{\psi}\gamma_\mu A_\mu\psi + (Z_m - 1)m\bar{\psi}\psi \tag{7.85}$$

where

$$\begin{aligned} Z_{f2} &= 1 - \frac{g^2 C_f}{16\pi^2}\alpha \cdot \frac{2}{\varepsilon} \\ Z_{f1} &= 1 - \frac{g^2 C_f}{16\pi^2}\left(\alpha C_f + C_2(G)\left[1 - \frac{1-\alpha}{4}\right]\right)\frac{2}{\varepsilon} \\ Z_m &= 1 - \frac{g^2 C_f}{16\pi^2}(4 - (1 - \alpha))\,\frac{2}{\varepsilon} \end{aligned} \tag{7.86}$$

We note that these expressions reduce to the ones of abelian electrodynamics if one substitutes $C_f = 1, C_2(G) = 0$. Of course the new interaction also adds new divergent

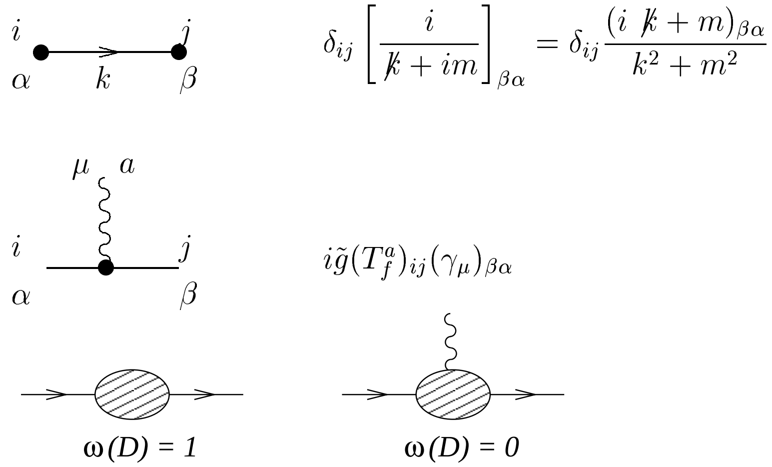


Figure 7.4: New propagators, new vertices, and the new superficially divergent 1PI Green functions introduced by the fermionic interaction

diagrams to the 1PI Green functions we have already considered in the pure gauge theory. The one loop diagrams along with their divergent parts are shown in fig.7.5. These diagrams will change the values of Z_3 and Z_1 but with these new values we can still attempt to write:

$$\mathcal{L}(A,\bar\psi,\psi,g,m) + \delta\mathcal{L}(A,\bar\psi,\psi,g,m) = \mathcal{L}_0(A_0,\bar\psi_0,\psi_0,g_0,m_0) \tag{7.87}$$

or

$$Z_{2f}\bar\psi\gamma_\mu\partial_\mu\psi - iZ_{f1}g\bar\psi\gamma_\mu A_\mu\psi + m\bar\psi\psi = \bar\psi_0\gamma_\mu\partial_\mu\psi_0 - ig_0\bar\psi_0\gamma_\mu A_{0\mu}\psi_0 + m_0\bar\psi_0\psi_0 \tag{7.88}$$

We have now an extended problem of renormalization: We still have only one coupling constant g and in order to maintain the multiplicative nature of the renormalization we have to have further relations between renormalization constants: *extended Slavnov-Taylor identities.* It is easy to derive them. We still have the old relations coming from the vertex functions in the pure gauge sector. Especially $A_0 = Z_3^{1/2}A$ and $g_0 = gZ_1/Z_3^{3/2}$.

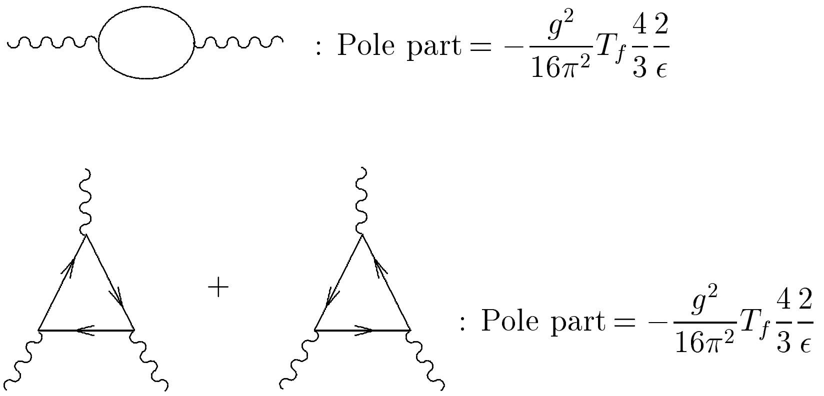


Figure 7.5: The fermionic contributions to 1PI two- and three-point functions of the gauge fields

Using these relations and the wave-function renormalization

$$\psi_0 = Z_{f2}^{1/2}\psi, \qquad \bar{\psi}_0 = Z_{f2}^{1/2}\bar{\psi} \tag{7.89}$$

which follows from (7.88), multiplicative renormalization demands

$$Z_{f1} g \bar{\psi} A \psi = g_0 \bar{\psi}_0 A_0 \psi_0 \Rightarrow Z_{1f} = Z_{2f}\frac{Z_1}{Z_3}$$

We conclude that the Slavnov-Taylor identities should be extended to

$$\boxed{\frac{Z_{f1}}{Z_{f2}} = \frac{Z_1}{Z_3} = \frac{\tilde{Z}_1}{\tilde{Z}_3} = \frac{Z_4}{Z_1}} \tag{7.90}$$

Again we can check that these extended relations are satisfied to one loop, but a proof which does not appeal to explicit calculation must wait until we have introduced the notation of BRS-invariance. At this point we will content ourself to just state that non-abelian theory of gauge field coupled in addition to Dirac fermions seems to provide us with a renormalizable theory, which preserves the gauge symmetries of the original theory.

Finally we can couple the system to bosonic fields too by adding

$$\mathcal{L}(\phi, A) = (D_\mu\phi)^\dagger(D_\mu\phi) + V(\phi^\dagger\phi) \tag{7.91}$$

Needless to say multiplicative renormalization can be extended to this case. Further Slavnov-Taylor identities result, and we will still have only one gauge coupling g for the whole theory thanks to the Slavnov-Taylor identities.

Before we turn to a systematic study of the Slavnov-Taylor identities and the BRS-invariance we will discuss some of the physical implication of our one-loop results.

7.6 Asymptotic freedom

We consider now the effects of a one-loop calculation. We found

$$g_0 = g\mu^{\varepsilon/2}\frac{Z_1}{Z_3^{3/2}}\ , \ \varepsilon = 4 - d \tag{7.92}$$

Other combinations could be used by means of the Slavnov-Taylor identities. μ is the mass scale which is introduced by dimensional regularization. As we have seen it can be related to other mass scales like a subtraction point $p^2 = \mu^2$ in momentum space by a finite renormalization. Heuristically we can view μ as a typical mass scale where we perform our calculations. Using the values for Z_1 and Z_3 to one loop, calculated in the minimal subtraction scheme (MS) where only the pole term is included we find

$$g_0 = \mu^{\varepsilon/2} g\left(1 - \frac{g^2}{16\pi^2}\left(\frac{11}{6}C_2(G) - \frac{2}{3}T_f\right)\frac{2}{\epsilon}\right) \tag{7.93}$$

In (7.93) we have included the effect of fermion loops in Z_1 and Z_3. $C_2(G)$ denotes the Casimir for the adjoint representation and T_f for the fermionic representation is defined by (7.6).

Since the bare quantities: g_0, A_0, ψ_0 etc. are independent of μ (but not of ϵ, of course) we get from (7.93):

$$0 = \mu \frac{dg_0}{d\mu} = \frac{\varepsilon}{2}\mu^{\varepsilon/2}\left(g - \frac{g^3}{16\pi^2}(\cdot)\frac{2}{\varepsilon}\right) + \mu^{\varepsilon/2}\left(1 - \frac{3g^2}{16\pi^2}(\cdot)\frac{2}{\varepsilon}\right)\mu\frac{\partial g}{\partial\mu} + \mathcal{O}(g^5) \tag{7.94}$$

or

$$\mu\frac{\partial g}{\partial \mu} = \frac{-\frac{\varepsilon}{2}g + \frac{g^3}{16\pi^2}(\cdot) + \mathcal{O}(g^5)}{1 - \frac{3g^2}{16\pi^2}(\cdot)\frac{2}{\varepsilon} + \mathcal{O}(g^4)} = -\frac{\varepsilon}{2}g - \frac{2g^3}{16\pi^2}(\cdot) + \mathcal{O}(g^5) \tag{7.95}$$

We conclude that the limit $\epsilon \to 0$ is finite to one loop and *independent of the gauge parameter* α. In general we can write: (taking $\varepsilon \to o$)

$$\mu\frac{\partial g}{\partial\mu}|_{\varepsilon\to 0} = \beta(g) \tag{7.96}$$

where *the β-function* has a perturbative expansion

$$\beta(g) = \beta_1 g^3 + \beta_2 g^5 + \beta_3 g^7 + \cdots \tag{7.97}$$

It can be shown that the coefficients β_i are non-singular for $\varepsilon \to 0$, and that the two first coefficients are universal: they are gauge independent and they are independent of the cut off used (here MS-scheme). As already discussed in chapter 5 for scalar theories we can solve (7.96) (truncating (7.97) to the first term $\beta_1 g^3$):

$$\boxed{g^2(\mu) = \frac{g^2(\mu_0)}{1 - \beta_1 g^2(\mu_0)\ln(\mu^2/\mu_0^2)}} \tag{7.98}$$

$g^2(\mu)$ can be understood as an effective coupling constant relevant for the scale μ if $g^2(\mu_0)$ is relevant for the energy scale μ_0. This aspect is discussed more carefully in chapter 5. Of course we can strictly speaking only trust (7.98) as long as $\beta_1 g^2(\mu_0)\ln^2\frac{\mu}{\mu_0} \ll 1$, but in the case where $\beta_1 < 0$ the situation is actually better in the following way: If we are in a perturbative regime for one choice $g^2(\mu_0)$ the same will be true for all higher energy scales. In fact the effective coupling constant will decrease as $1/\ln(\frac{\mu}{\mu_0})$. This is illustrated in fig.7.6. Now the effective coupling constant will grow at low energy scales $\mu < \mu_0$, i.e. at long distances, a phenomenon called *infrared slavery*. This is a pleasant surprise: not only is perturbation theory reliable at short distances (high energy), but the perturbation expansion hint itself that something drastic could take place at long distances. And we *want* something drastic to happen, since our theory is formulated in terms of quarks and gluons, but all we see at long distances are hadrons which are bound states of quarks, the interaction being mediated by the gluons. We have *confinement* of quarks and gluons. Heuristically we can imagine that the growth of the effective coupling constant prevents the existence of free quarks. Clearly an answer to such questions cannot be given within perturbation theory. Unfortunately we have to say that not much progress has been made in analyzing the confinement problem by means of non-perturbative methods. The use of lattice gauge theories, which we will discuss later, seems at the moment the most promising way to address the non-perturbative question of confinement.

It is seen from (7.95) that the non abelian theories lead to a $\beta_1 < 0$:

$$\boxed{\beta_1 = -\frac{1}{16\pi^2}\left(\frac{11}{3}C_2(G) - \frac{4}{3}T_f\right)} \tag{7.99}$$

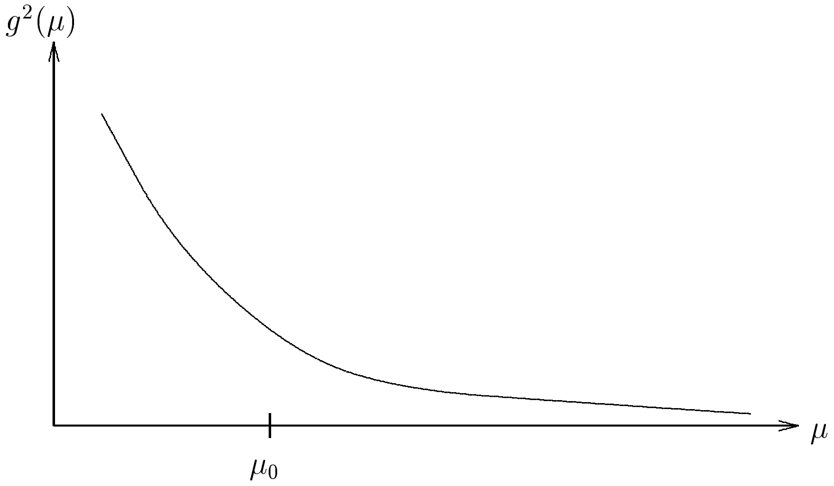


Figure 7.6: The behaviour of the running coupling constant for an asymptotic free theory

The non-abelian gauge theories are the only known theories which in a natural way provides a negative β-function in four dimensions. Any content of fermions or scalar particles counteracts this. In particular will abelian gauge theories with scalar or fermion particles have $\beta_1 > 0$ ((7.99) with $C_2(G) = 0, T_f = 1$).

In QCD we imagine that the gauge group is $SU(3)$ and we have

$$C_2(SU(3)) = 3 \tag{7.100}$$

and

$$T_f = \frac{1}{2} \tag{7.101}$$

for each Dirac fermion in the fundamental representation of $SU(3)$. Thus the total β_1 is

$$\beta_1 = -\frac{1}{16\pi^2}\left(11 - \frac{2}{3}n_f\right) \tag{7.102}$$

where n_f denote the number of flavours (i.e. the number of different quark families). At the moment we have observed $n_f = 5$ (up, down, strange, charm, bottom), but expect to see at least one more (top) because of the so-called anomalies (see next chapter). We conclude: *QCD is asymptotically free.* Note also that the demand of asymptotic freedom put a (rather weak) constraint on possible grand unified models.

7.7 BRS-invariance

7.7.1 The Lee–Zinn-Justin identities

The full understanding of the relations like the Slavnov-Taylor identities was obtained by Becchi, Rouet and Stora. They discovered an additional invariance, called BRS-invariance, of the effective Lagrange function given by (7.52). It is a *global* invariance, but as we shall see it resembles the local gauge invariance sufficiently to ensure relations like the Slavnov-Taylor identities, and consequently multiplicative renormalization of non-abelian gauge theories. The only additional requirement needed is a regularization which respects the invariances of the effective action (7.52). Dimensioned regularization preserve gauge invariance and we will always imagine that we use dimensional regularization to make sense of the formal expressions we write down.

Recall that under an infinitesimal local gauge transformation $U(x) \sim e^{ig\alpha^a(x)T^a} \simeq 1 + ig\alpha^a(x)T^a$ we have

$$\begin{aligned} A'(x) = A(x) + \delta A(x) &= U(x)A(x)U^{-1}(x) - \frac{i}{g}\partial U(x)U^{-1}(x) \\ &= A(x) + (\partial\alpha^a)T^a - ig[A, \alpha^a T^a] + O(\alpha^2) \end{aligned}$$

or

$$\delta_\alpha A_\mu(x) \simeq [D_\mu, \alpha(x)]; \qquad (\delta_\alpha A^a_\mu \simeq D^{ab}_\mu \alpha^b(x)) \tag{7.103}$$

In the following we will use the short-hand notation $D_\mu\alpha$ for the vector $V^a = D^{ab}\alpha^b$ and we will change somewhat inconsistently between the interpretation of a variable like A as being an element $A^a T^a$ in the Lie algebra and being the vector with components A^a with transform as a vector in the adjoint representation. This is done in order to avoid too cumbersome a notation and will hopefully not lead to any misunderstanding.

The effective lagrangian (7.52):

$$\mathcal{L}_{eff}(A, \eta, \bar{\eta}) = \frac{1}{4}F^{a^2}_{\mu\nu} + \frac{1}{2\alpha}(\partial_\mu A^a_\mu)^2 - \bar{\eta}^a \partial_\mu D^{ab}_\mu \eta^b \tag{7.104}$$

is *not* invariant under the local gauge transformation (7.103). The whole point of constructing the effective action was to break this invariance in an "organized" way and make a perturbative expansion about a gaussian extremum possible. However, $\mathcal{L}_{eff}$ *is* invariant under the following infinitesimal transformation (BRS-transformation) which mixes gauge fields and ghost fields, and which resembles the local gauge transformation (7.103) quite a lot:

$$\begin{aligned} \delta_\varepsilon A^a &= D^{ab}_\mu \eta^b \cdot \delta\varepsilon \\ \delta_\varepsilon \bar{\eta}^a &= \frac{1}{\alpha}\partial_\mu A^a_\mu \cdot \delta\varepsilon \\ \delta_\varepsilon \eta^a &= -\frac{g}{2}c^{abc}\eta^b\eta^c \cdot \delta\varepsilon \end{aligned} \tag{7.105}$$

$\delta\varepsilon$ is an infinitesimal *global* parameter which is a *grassmann variable*: it anticommutes with $\eta, \bar{\eta}$ and commutes with A^a_μ. Clearly the anticommuting nature of $\delta\varepsilon$ is needed in order that (7.105) makes sense (A stays bosonic also after transformation, etc.). Although the transformation for A looks like a local gauge transformation with $\eta^a(x)\cdot\delta\varepsilon$ playing the role of $\alpha^a(x)$ in (7.103), it is $\delta\varepsilon$ which is the infinitesimal parameter, and the transformation is a *global* transformation. If we write the BRS-transformation on the fields as $e^{Q_B\delta\varepsilon}$ the operator Q_B will precisely be the generator of the BRS-transformation and we denote it the BRS-charge, in analogue with other conserved charges associated with internal symmetries, like the electric charge associate with (global) abelian gauge invariance. From (7.105) we get:

$$\begin{aligned} Q_B A &= [D_\mu, \eta] \\ Q_B \bar{\eta} &= \frac{1}{\alpha}\partial_\mu A_\mu \\ Q_B \eta &= ig\eta^2 \end{aligned} \tag{7.106}$$

where η^2 is a short hand notation for $T^a\eta^a T^b\eta^b$ which can be written as $[T^a, T^b]\eta^a\eta^b/2$. It is seen that the charge Q_B carries ghost number 1 and maps commuting variables into anticommuting ones and vice versa.

We will now prove the invariance of (7.104) under the transformation (7.105). First we note the following properties of the BRS transformation:

lemma: $Q_B[D_\mu, \eta] = 0; \quad Q_B\eta^2 = 0; \quad Q_B\partial_\mu A_\mu = \partial_\mu[D_\mu, \eta]$

> **proof:**
>
> We prove only the first of the relations. The technique is the same for the two other relations.
>
> $$\begin{aligned}\delta_\varepsilon[D_\mu, \eta] &= \partial_\mu\delta_\varepsilon\eta - ig[\delta_\varepsilon A, \eta] - ig[A, \delta_\varepsilon\eta] \\ &= ig\partial_\mu\eta^2\delta\varepsilon - ig[(\partial_\mu\eta - ig[A, \eta])\ \delta\varepsilon, \eta] - ig[A, ig\eta^2\ \delta\varepsilon]\end{aligned}$$
>
> The derivative term is zero:
>
> $$-\frac{g}{2}c^{abc}((\partial_\mu\eta^b)\eta^c + \eta^b\partial_\mu\eta^c) + gc^{abc}(\partial_\mu\eta^b)\eta^c = 0$$
>
> The term: $-[[A, \eta]\delta\varepsilon, \eta] + [A, \eta^2\delta\varepsilon]$ can be written in components:
>
> $$-[[T^a, T^b], T^c]A^a\eta^b\delta\varepsilon\eta^c + \frac{1}{2}[T^a, [T^b, T^c]]A^a\eta^b\eta^c\delta\varepsilon.$$
>
> Using the anticommuting nature of η^b, η^c and $\delta\varepsilon$ we can rearrange it:
>
> $$\frac{1}{2}\left\{[T^a[T^b, T^c]] + [T^b, [T^c, T^a]] + [T^c, [T^a, T^b]\right\} A^a\eta^b\eta^c\delta\epsilon$$
>
> The Jacobi identity for the Lie algebra gives $\{\cdot\} = 0$.

Since (7.104) looks very much like a local gauge transformation $(F^a_{\mu\nu})^2$ is invariant under (7.105) for the same reasons that it is invariant under (7.103). The rest of $\mathcal{L}_{eff}$ transforms under (7.105) as follows:

$$\delta_\varepsilon\left[\frac{1}{2\alpha}(\partial_\mu A_\mu)^2 - \bar{\eta}\partial_\mu D_\mu\eta\right] = \left\{\frac{1}{\alpha}(\partial_\mu A_\mu)\partial_\mu\delta_\varepsilon A_\mu - (\delta_\varepsilon\bar{\eta})\partial_\mu D_\mu\eta\right\} - \bar{\eta}\partial_\mu\delta_\varepsilon(D_\mu\eta) = 0$$

The curly bracket is zero by definition (7.105), while the last term vanish by the lemma. We have now proven that $\mathcal{L}_{eff}(A, \eta\bar{\eta})$ is invariant under the BRS transformation. From the lemma it follows further that:

$$Q_B^2 A = 0; \quad Q_B^2\eta = 0; \quad Q_B^2\bar{\eta} = \partial_\mu D_\mu\eta \tag{7.107}$$

The rhs of last equation is zero by the classical equations of motion ($\partial D\eta = 0$ is just the classical equation of motion for η). From the point of view of quantum field theory we cannot put it equal zero, since the path integral includes field configurations which do not satisfy the classical equations of motion. It is, however, possible to reformulate the field content of effective lagrangian in such a way that:

$$\boxed{Q_B^2 = 0 \quad (Q_B \text{ nilpotent})} \tag{7.108}$$

The way to get $Q_B^2 = 0$ is to introduce the so-called *Lautrup-Nakanishi field B*: We use the relation

$$e^{-\frac{1}{2}A^2} = \int \frac{dB}{\sqrt{2\pi}}\, e^{iAB-\frac{1}{2}B^2}$$

which in the functional integral dressing looks like:

$$e^{-\int d^dx \frac{1}{2\alpha}(\partial_\mu A_\mu^1)^2} = \int \mathcal{D}B\; e^{\int d^dx\; [iB^a(\partial_\mu A_\mu)^a - \frac{\alpha}{2}B^a(x)^2]} \tag{7.109}$$

This means that we can replace $(\partial_\mu A_\mu^a)^2/2\alpha$ with $-iB^a\partial_\mu A_\mu^a + \alpha B^2/2$ and we get a new effective lagrangian:

$$\boxed{\mathcal{L}_{eff}(A,\eta,\bar{\eta},B) = \frac{1}{4}(F_{\mu\nu}^a)^2 - iB^a\partial_\mu A_\mu^a + \frac{\alpha}{2}B^a(x)^2 - \bar{\eta}\partial_\mu D_\mu\eta} \tag{7.110}$$

The BRS transformation can now be written:

$$\begin{aligned} Q_B A &= [D_\mu, \eta] \\ Q_B\bar{\eta} &= -iB \qquad (7.111) \\ Q_B\eta &= ig\eta^2 \\ Q_B B &= 0 \qquad (7.112) \end{aligned}$$

The last equation in (7.107) is now replaced by $Q_B^2\bar{\eta} = -Q_B B = 0$ and (7.108) is indeed satisfied. Clearly (7.110) is invariant under (7.111). In many respects (7.110)-(7.111) are more convenient than (7.104)-(7.105). For instance it is readily seen that

$$\mathcal{L}_{eff}(A,\eta,\bar{\eta},B) = \frac{1}{4}(F_{\mu\nu}^a)^2 + Q_B\left(\bar{\eta}\left(\partial_\mu A_\mu + i\alpha B/2\right)\right) \tag{7.113}$$

and the invariance under (7.111) is now a trivial consequence of $Q_B^2 = 0$. However for the purpose of discussing renormalizability (7.104)-(7.105) are as good and we will use them as they involve one field less.

Our effective Lagrangian is invariant under BRS. The functional measure $\mathcal{D}A\mathcal{D}\eta\mathcal{D}\bar{\eta}$ is also invariant under BRS. This is clear from (7.105) when we consider the fields as generalized coordinates: For A we just have a translation plus a rotation, for $\bar{\eta}$ a pure translation while we for η has a rotation. Only the source term in the generating functional (7.61) is not invariant under BRS transformations:

$$\delta_\varepsilon(JA + \bar{\xi}\eta + \bar{\eta}\xi) = (J\cdot Q_B A + \bar{\xi}\cdot Q_B\eta + Q_B\bar{\eta}\cdot\xi)\;\delta\varepsilon \tag{7.114}$$

If we view the BRS transformation as a simple change in integration variables and use the invariance of the measure we get to first order in ε:

$$0 = \int \mathcal{D}A\mathcal{D}\eta\mathcal{D}\bar{\eta}\;\; \left\{\int d^dx(J\cdot Q_B A + \bar{\xi}\cdot Q_B\eta + Q_B\bar{\eta}\cdot\xi)\right\}\; e^{-S_{eff}[A,\eta,\bar{\eta}]+\int d^dx(JA+\bar{\xi}\eta+\bar{\eta}\xi)} \tag{7.115}$$

In this expression $Q_B A$ and $Q_B\eta$ are relatively complicated functions. In order to simplify expressions it is convenient to introduce sources for $Q_B A$ and $Q_B\eta$. Since both these terms are BRS invariant we can assume the sources are BRS invariant and introduce the following BRS invariant coupling:

$$\int d^dx \left\{ K^a_\mu (Q_B A)^a_\mu - L^a (Q_B \eta)^a \right\} \tag{7.116}$$

Nothing is changed in (7.115), except that the source term (82) is added to the exponent, and we can rewrite (7.115) as follows:

$$0 = \int d^dx \left(J\frac{\delta}{\delta K} - \bar{\xi}\frac{\delta}{\delta L} + \frac{1}{\alpha}\partial_\mu \left(\frac{\delta}{\delta J_\mu}\right) \xi \right) \; e^{-F[J,\xi,\bar{\xi},K,L]}$$

or

$$\boxed{0 = \int d^dx \left(J^a_\mu \frac{\delta}{\delta K^a_\mu} - \bar{\xi}^a \frac{\delta}{\delta L^a} + \frac{1}{\alpha}\partial_\mu (\frac{\delta}{\delta J^a_\mu})\xi^a \right) \; F[J,\xi,\bar{\xi},K,L].} \tag{7.117}$$

This relation is somewhat similar to the ordinary Dyson Schwinger (DS) equations: It relates different *connected Green functions* and its origin is the invariance of the measure and $S_{eff}[A,\eta,\bar{\eta}]$ under a field transformation (the BRS transformation).

As for the DS equations it is convenient to transform (7.117) into an equation for the generation functional Γ for 1PI Green functions, related to F by the Legendre transformation (7.62) (we do not perform a Legendre transformation in K,L). From (7.64) we get:

$$\int d^dx \left\{ \frac{\delta\Gamma}{\delta A^a_\mu}\frac{\delta\Gamma}{\delta K^a_\mu} + \frac{\delta\Gamma}{\delta\eta^a}\frac{\delta\Gamma}{\delta L^a} + \frac{\delta\Gamma}{\delta\bar{\eta}^a}\frac{1}{\alpha}(\partial_\mu A_\mu) \right\} = 0. \tag{7.118}$$

In order to simplify the notation somewhat we drop from now the notation A^{cl}_μ *and write simply* A_μ. Hopefully no confusion is possible. Whenever we talk about fields in connection with the generating functionals it is never the genuine quantum fields we have in mind.

Another useful relation is derived by using the invariance of the measure under $\bar{\eta} \to \bar{\eta} + \delta\bar{\eta}$. Such a change is precisely what we used to derive the DS-equation for a scalar field. The only difference here is that we use it for the "unphysical" ghost field. The action and the source term changes:

$$\begin{aligned} \delta \Big(\frac{1}{4}(F^a_{\mu\nu})^2 + \frac{1}{2\alpha}(\partial_\mu A^a_\mu)^2 - \bar{\eta}\partial_\mu D_\mu \eta - JA - \bar{\xi}\eta - \bar{\eta}\xi - K\cdot Q_B A + LQ_B\eta \Big) \\ = -\delta\bar{\eta}\left(\partial_\mu D_\mu \eta) + \xi\right) = -\delta\bar{\eta}\left(\partial_\mu (Q_B A)_\mu + \xi\right) \end{aligned}$$

and we get as before, by considering $\bar{\eta} \to \bar{\eta} + \delta\bar{\eta}$ as a change of integration variables in the functional integral and expanding to first order in $\delta\bar{\eta}$:

$$0 = \int \mathcal{D}A\mathcal{D}\eta\mathcal{D}\bar{\eta} \int d^dx \; \delta\bar{\eta}(x) \left\{ \partial_\mu (Q_B A)_\mu + \xi \right\} e^{-S_{eff}[A,\eta,\bar{\eta}] + \int d^dx (JA + \bar{\xi}\eta + \bar{\eta}\xi + K(Q_B A) - LQ_B\eta)} \tag{7.119}$$

Since this is true for any $\delta\bar{\eta}(x)$ we can drop $\int d^dx$ and have (7.119) fulfilled as a local identity (contrary to (7.118)):

$$0 = \left(\partial_\mu \frac{\delta}{\delta K^a_\mu} + \xi^a \right) e^{-F[J,\xi,\bar{\xi},K,L]} \quad \text{or} \quad \partial_\mu \frac{\delta}{\delta K^a_\mu} F[J,\xi,\bar{\xi},K,L] - \xi^a = 0 \tag{7.120}$$

By the Legendre transformation we change from $J, \xi, \bar{\xi}$ to the classical fields $A, \bar{\eta}, \eta$ and we can write (7.120) as

$$\partial_\mu \frac{\delta \Gamma}{\delta K_\mu^a} - \frac{\delta \Gamma}{\delta \bar{\eta}^a} = 0 \tag{7.121}$$

Eq. (7.118) and (7.121) are the equations needed to derive the Slavnov-Taylor identities. If we eliminate $\frac{\delta \Gamma}{\delta \bar{\eta}}$ from (7.118) by means of (7.121) and redefine Γ in order to get rid of the gauge fixing part $\frac{1}{2\alpha}(\partial A)^2$, we are led to the final equations:

$$\tilde{\Gamma}[A, \eta, \bar{\eta}, k, L] \equiv \Gamma[A, \eta, \bar{\eta}, K, L] - \frac{1}{2\alpha} \int d^d x (\partial_\mu A_\mu^a)^2 \tag{7.122}$$

$$\boxed{\int d^d x \left\{ \frac{\delta \tilde{\Gamma}}{\delta A_\mu^a} \frac{\delta \tilde{\Gamma}}{\delta K_\mu^a} + \frac{\delta \tilde{\Gamma}}{\delta \eta} \frac{\delta \tilde{\Gamma}}{\delta L} \right\} = 0} \tag{7.123}$$

$$\boxed{\partial_\mu \frac{\delta \tilde{\Gamma}}{\delta K_\mu^a} - \frac{\delta \tilde{\Gamma}}{\delta \bar{\eta}^a} = 0} \tag{7.124}$$

These equations, called the *Lee– Zinn-Justin identities*, together with the tree level formula for the generating functional (the gauge fixing term has dropped out by definition of $\tilde{\Gamma}$):

$$\tilde{\Gamma}^{(0)}[A, \bar{\eta}, \eta, K, L] = \int \left(\frac{1}{4} (F_{\mu\nu}^a)^2 - \bar{\eta} \partial_\mu D_\mu \eta - K Q_B A + L Q_B \eta, \right) \tag{7.125}$$

constitute the basic tool for proving the Slavnov-Taylor identities.

7.7.2 The structure of divergences

Relations (7.123) and (7.124) are very important if we want to understand in a systematic way the structure of the divergencies in the non-abelian theories. Already in the abelian case we saw the corresponding relations were very useful for classification of the divergences, but we could have done without them. In the non-abelian case that is very difficult. Recall that renormalization is much more demanding in the non-abelian case: not only must we argue that divergences can be absorbed by only adding counter terms of the kind already present in the lagrangian, but in order that the renormalized theory can be viewed as a gauge theory, invariant under that action of local gauge transformations, the renormalization constants have to satisfy the Slavnov-Taylor identities.

It is not within the scope of this course to provide a general proof of the renormalizability of the non-abelian gauge theories. Instead we will show how we can derive the Slavnov-Taylor identities to lowest order by the use (7.123) and (7.124) and without actually doing the one-loop calculations. The procedure we outline can by induction be extended to arbitrary order without too many difficulties. We assume as usual that our formal expressions have be regularized in a gauge invariant way, i.e. by dimensional regularization.

We now look for divergent graphs at 1-loop. Apart from the old divergences we have 3 new ones due to the source terms KQ_BA and $LQ_B\eta$. The corresponding graphs are shown in fig.7.7. These graphs give poles and local counter terms corresponding to each

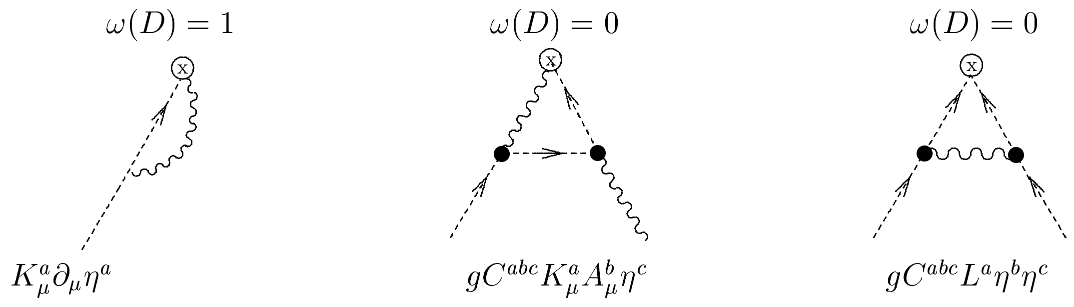


Figure 7.7: The new divergences due to the source terms $K(Q_B A)$ and $L(Q_B \eta)$.

term in $K(Q_B A) - L(Q_B \eta) = K^a_\mu(\partial_\mu \eta^a + g c^{abc} A^b_\mu \eta^c) + L^a \frac{g}{2} c^{abc} \eta^b \eta^c$. We can now write the divergent structure of $\tilde{\Gamma}$ as:

$$\tilde{\Gamma}_{div}(A, \eta, \bar{\eta}, K, L) = \tilde{\Gamma}_{div}(A, \eta, \bar{\eta}) - (\tilde{Z}_3 - 1) \int d^d x K_\mu \partial_\mu \eta$$
$$-(\tilde{Z}_1 - 1) \int d^d x \; g c^{abc} K^a_\mu A^b_\mu \eta^c + (X - 1) \int d^d x \; L^a (-\frac{g}{2} c^{abc} \eta^b \eta^c) \tag{7.126}$$

The idea is to use (7.123) and (7.124) to completely determine the structure of $\tilde{\Gamma}_{div}(A, \eta, \bar{\eta}, K, L)$ starting from (7.126) plus the fact that the loop expansion (the expansion in $\hbar$ or the coupling constant g, since there is only one coupling constant) allow us to write:

$$\tilde{\Gamma} = \tilde{\Gamma}^{(0)} + \hbar \tilde{\Gamma}^{(1)}_{div} + \hbar \tilde{\Gamma}^{(1)}_{finite} + 0(\hbar^2) \tag{7.127}$$

where $\Gamma^{(0)}$ is given by (7.125). The divergent part in (7.127) means the pole part in MS (minimal subtraction). From (7.125):

$$\frac{\delta \tilde{\Gamma}_{div}}{\delta \bar{\eta}^a} = \partial_\mu \frac{\delta \tilde{\Gamma}_{div}}{\delta K^a_\mu} = -(\tilde{Z}_3 - 1) \partial^2 \eta^a - (\tilde{Z}_1 - 1) g c^{abc} \partial_\mu (A^b \eta^c)$$

or:

$$\tilde{\Gamma}_{div}[A, \eta, \bar{\eta}] = \tilde{\Gamma}_{div}[A] - \int d^4 x \;\; (\tilde{Z}_3 - 1) \bar{\eta}^a \partial^2 \eta^a + (\tilde{Z}_1 - 1) \bar{\eta}^a g c^{abc} \partial_\mu (A^b \eta^c) \tag{7.128}$$

This fixes the divergent ghost terms.

In order to fix $\tilde{\Gamma}_{div}[A]$ we isolate the *terms in (7.123) which are simple poles and of order* $\hbar$:

$$\int d^4 x \left\{ \left(\frac{\delta \tilde{\Gamma}^{div}}{\delta A} \frac{\delta \tilde{\Gamma}^{(0)}}{\delta K} + \frac{\delta \tilde{\Gamma}^{(0)}}{\delta A} \frac{\delta \tilde{\Gamma}^{div}}{\delta K} \right) + \left(\frac{\delta \tilde{\Gamma}^{div}}{\delta \eta} \frac{d \tilde{\Gamma}^{(0)}}{\delta L} + \frac{\delta \tilde{\Gamma}^{(0)}}{\delta \eta} \frac{\delta \tilde{\Gamma}^{div}}{\delta L} \right) \right\} = 0 \tag{7.129}$$

where (as already mentioned):

$$\tilde{\Gamma}^{(0)} = - \int d^4 x \left\{ \frac{1}{4} F^2_{\mu\nu} - \bar{\eta}^a (\partial_\mu D_\mu \eta)^a - K^a_\mu (D_\mu \eta)^a + L^a (-\frac{g}{2} c^{abc} \eta^b \eta^c) \right\} \tag{7.130}$$

In principle we can simply insert (7.128), (7.130) and (7.126) in (7.129) and we will get a complete determination of $\tilde{\Gamma}^{div}[A]$.

Let us mention the systematic results:

(1): *Look for terms linear in K.*

Contributions from the middle two terms in (7.129) cancel and from the others:

$$0 = -\int d^4x \left\{(\tilde{Z}_1 - 1)K^a_\mu g c^{abc}\eta^c\right\}\left\{D_\mu\eta^b\right\} + \int d^4x (X-1)\left\{(-D_\mu K^a_\mu)\right\}\left\{-\frac{g}{2}c^{abc}\eta^b\eta^c\right\}$$

We conclude after a few rearrangements:

$$X = \tilde{Z}_1 \tag{7.131}$$

(2): *Look for terms linear in A and not containing K,L*

We expand $\Gamma^{div}[A] = \Gamma^{(2)}_{div}[A] + \Gamma^{(3)}_{div}[A] + \Gamma^{(4)}_{div}[A]$, the superscript referring to the power of the gauge field A. The equation is

$$0 = \int d^4x \frac{\delta\Gamma^{div}[A]}{\delta A^a_\mu}\left\{D_\mu\eta^a\right\} - (D_\mu F^a_{\mu\nu}\left\{(\tilde{Z}_3 - 1)\partial_\nu\eta^a + (\tilde{Z}_1 - 1)g c^{abc}A^b_\nu\eta^c\right\} \tag{7.132}$$

and we can now expand this equation in powers of A. For the terms linear in A we have

$$\partial_\nu D_\mu F_{\mu\nu} \to \partial_\nu\partial_\mu\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) = 0$$

and therefore from (7.132)

$$\int d^4x \frac{\delta\Gamma^{(2)}_{div}[A]}{\delta A^a_\mu}\partial_\mu\eta^a = 0$$

or, by a partial integration

$$\partial_\mu\frac{\delta\Gamma^{(2)}_{div}[A]}{\delta A_\mu} = 0 \tag{7.133}$$

The divergent quadratic part is purely transverse:

$$\Gamma^{(2)}_{div}[A] = (Z_3 - 1)\int d^4x\,\frac{1}{2}\left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu\right)^2 \tag{7.134}$$

(3): *Look for terms quadratic in A and not containing L, K.*

We continue the expansion of (7.132) and get for the quadratic terms:

$$\int d^4x \left\{\frac{\delta\tilde{\Gamma}^{(3)}_{div}}{\delta A^a_\mu}\partial_\mu\eta^a + \frac{\delta\tilde{\Gamma}^{(2)}_{div}}{\delta A^a_\mu}g c^{abc}A^b_\mu\eta^c + \partial_\mu(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)\cdot(\tilde{Z}_1 - 1)g c^{abc}A^b_\nu\eta^c + \right.$$
$$\left. + g c^{abc}A_\mu(\partial_\mu A_\nu - \partial_\nu A_\mu)\cdot(\tilde{Z}_3 - 1)\partial_\nu\eta^c\right\} = 0$$

or

$$\int d^4x \left\{\frac{\delta\tilde{\Gamma}^{(3)}_{div}}{\delta A^a_\mu} + \left(Z_3 + \tilde{Z}_1 - \tilde{Z}_3 - 1\right) g\ c^{abc}(\partial_\mu A^b_\nu - \partial_\nu A^b_\mu)A^c_\nu\right\}\partial_\mu\eta^a = 0$$

or

$$\tilde{\Gamma}^{(3)}_{div}[A] = (Z_3 + \tilde{Z}_1 - \tilde{Z}_3 - 1)\ \int d^4x \frac{1}{2} g c^{abc}A^a_\mu A^b_\nu\left(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu\right) \tag{7.135}$$

(4): *Look for terms cubic in A and not containing L, K*

In the same way as we got (7.135) from (7.132), we get by collecting terms of cubic order from (7.132) the following equation:

$$\tilde{\Gamma}^{(4)}_{div}(A) = (Z_3 + 2\tilde{Z}_1 - \tilde{Z}_3 - 1)\frac{g^2 c^{abc}c^{ade}}{4}\int d^4x\ A^a_\mu A^b_\nu A^d_\mu A^e_\nu \tag{7.136}$$

We conclude from (7.135) and (7.136) that the renormalization constants Z_1 and Z_4 associated with the three- and four-point vertices are given by

$$\begin{aligned} Z_1 &= Z_3 + \tilde{Z}_1 - \tilde{Z}_3 = Z_3 \cdot \frac{\tilde{Z}_1}{\tilde{Z}_3}(1 + \mathcal{O}(\hbar^2)) && (7.137) \\ Z_4 &= Z_3 + 2\tilde{Z}_1 - \tilde{Z}_3 = Z_3 \cdot \frac{\tilde{Z}_1^2}{\tilde{Z}_3}(1 + \mathcal{O}(\hbar^2)) && (7.138) \end{aligned}$$

This is exactly the Slavnov-Taylor identities and they are a consequence of BRS invariance. We finally remark that with these relations and the scaling relations already derived:

$$A_0 = Z_3^{1/2} A, \qquad \eta_0,\ \bar{\eta}_0 = \tilde{Z}_3^{1/2} \eta,\ \bar{\eta}, \qquad \alpha_0 = Z_3 \alpha, \qquad g_0 = \frac{Z_1}{Z_3^{3/2}} g$$

we can deduce the scaling of the source terms K, L at one-loop level:

$$\begin{aligned} K(Q_B A) &\to K_\mu^a \left(\tilde{Z}_3 \partial_\mu \eta^a + \tilde{Z}_1 g c^{abc} A_\mu^b \eta^c \right) \\ &= K_\mu^a\ \tilde{Z}_3^{1/2} \left(\partial_\mu \eta_0^a + g_0 c^{abc} A_{0\mu}^b \eta_0^c \right) = \tilde{Z}_3^{1/2}\ K(Q_B A_0) \end{aligned}$$

$$L(Q_B \eta) \to L^a X \left(-\frac{g}{2} c^{abc} \eta^b \eta^c \right) = Z_3^{1/2} L^a \left(-\frac{g_0}{2} c^{abc} \eta_0^b \eta_0^c \right) = Z_3^{1/2} L(Q_B \eta_0).$$

From these two equations we have

$$K_0 = \tilde{Z}_3^{1/2} K, \qquad L_0 = Z_3^{1/2} L \tag{7.139}$$

and we can finally write

$$\Gamma^{one\ loop}(A, \eta, \bar{\eta}, K, L) = \Gamma_0(A_0, \eta_0, \bar{\eta}, K_0, L_0). \tag{7.140}$$

By induction this proof can be extended to n-loops. But we are not going to discuss the details here. They are tedious, but not extremely complicated.

7.7.3 Inclusion of fermions

Although we will not perform the analysis in detail let us mention the steps involved if we want to include fermions in the above analysis of BRS-invariance:

1. Find the BRS transformation for $\psi, \bar{\psi}$
2. Couple $Q_B \psi$ and $Q_B \bar{\psi}$ to external BRS-invariant sources $\bar{H}$ and H
3. Derive equations for $\Gamma \left(A, \eta, \bar{\eta}, \psi, \bar{\psi}, K, L, H, \bar{H} \right)$ corresponding to (7.123) and(7.124)
4. Isolate the terms which are simple poles and of order $\hbar$ and check the extended Slavnov-Taylor identities.
5. Proceed by induction to order $\hbar^n$

Since an infinitesimal gauge transformation acts on $\bar{\psi}$ and ψ as:

$$\begin{aligned}\psi_i &\to \psi_i + i\alpha^a(x)T^a_{ij}\psi_j\\ \bar{\psi}_i &\to \bar{\psi}_i - i\bar{\psi}_j T^a_{ji}\alpha^a(x)\end{aligned}$$

and we get the BRS transformation by replacing $\alpha^a(x)$ with $\eta^a(x)\delta\varepsilon$ we get the following transformation: (assuming $\{\eta,\psi\} = \{\delta\varepsilon,\psi\} = 0$)

$$\delta\psi_i = -iT^a_{ij}\eta^a\psi_j\delta\varepsilon, \qquad \delta\bar{\psi}_i = -i\bar{\psi}_j\eta^a T_{ji}\delta\varepsilon \tag{7.141}$$

and we conclude:

$$Q_B\psi = -iT^a\eta^a\psi, \qquad Q_B\bar{\psi} = -i\bar{\psi}\eta^a T^a \tag{7.142}$$

It can be checked that

$$Q_B^2\psi = Q_B^2\bar{\psi} = 0 \tag{7.143}$$

and that

$$S_{qg} = \int d^dx \left\{\bar{\psi}(\gamma_\mu D_\mu + m)\psi - \bar{H}_{i\alpha}(Q_B\psi)_{i\alpha} - (Q_B\bar{\psi})_{i\alpha}H_{i\alpha}\right\} \tag{7.144}$$

is BRS invariant if H and $\bar{H}$ are invariant.

As usual we write the partition function Z and the free energy F as a function of the external sources $J, \bar{\xi}, \xi, K, L$ and now $\bar{\zeta}, \zeta, \bar{H}, H$ corresponding to a source term

$$\int d^dx \left(\bar{\zeta}_{i\alpha}\psi_{i\alpha} + \bar{\psi}_{i\alpha}\zeta_{i\alpha} + \bar{H}_{i\alpha}(Q_B\psi)_{i\alpha} + (Q_B\bar{\psi})_{i\alpha}H_{i\alpha}\right) \tag{7.145}$$

A Legendre transformation in $J, \bar{\xi}, \xi, \bar{\zeta}, \zeta$, but *not* in the BRS invariant source terms $K, L, H, \bar{H}$ defines our effective action Γ and $\tilde{\Gamma}$ as functions of $A, \eta, \bar{\eta}, \psi, \bar{\psi}$ and $K, L, H, \bar{H}$. We get finally the generalization of (7.122)-(7.124):

$$\boxed{\int d^dx \left\{\frac{\delta\tilde{\Gamma}}{\delta A^a_\mu}\frac{\delta\tilde{\Gamma}}{\delta K^a_\mu} + \frac{\delta\tilde{\Gamma}}{\delta\eta^a}\frac{\delta\tilde{\Gamma}}{\delta L^a} + \frac{\delta\tilde{\Gamma}}{\delta\psi_{ia}}\frac{\delta\tilde{\Gamma}}{\delta\bar{H}_{ia}} + \frac{\delta\tilde{\Gamma}}{\delta\bar{\psi}_{ia}}\frac{\delta\tilde{\Gamma}}{\delta H_{i\alpha}}\right\} = 0} \tag{7.146}$$

$$\boxed{\partial_\mu\frac{\delta\tilde{\Gamma}}{\delta K^a_\mu} - \frac{\delta\tilde{\Gamma}}{\delta\bar{\eta}^a} = 0} \tag{7.147}$$

where the modified generating functional $\tilde{\Gamma}$ is defined as in (7.122):

$$\tilde{\Gamma}[A,\eta,\bar{\eta},\psi,\bar{\psi},K,L,H,\bar{H}] = \Gamma[A,\eta,\bar{\eta},\psi,\bar{\psi},K,L,H,\bar{H}] - \frac{1}{2\alpha}\int d^dx(\partial_\mu A_\mu)^2 \tag{7.148}$$

The new divergent 1-loop diagrams introduced by the sources are shown in fig.7.8. The divergent part has the form:

$$(Y^{[1]} - 1)\int d^dx\ i\bar{\psi}(\eta T)H + (\tilde{Y}^{[1]} - 1)\int d^dx i\bar{H}(T\eta)\psi \tag{7.149}$$

As before we can expand $\Gamma = \Gamma^{(0)} + \hbar(\Gamma^{div} + \Gamma^{conv})$ where we know $\Gamma^{(0)}$ and we can determine Γ^{div} in a systematic way. Eventually we will find that the generalized Slavnov-Taylor identities are satisfied to one-loop and an induction argument can be performed. Rather than discussing this in detail let us check how this quite involved formalism reduces to the abelian Ward-Takahashi identities which we have already used in QED.

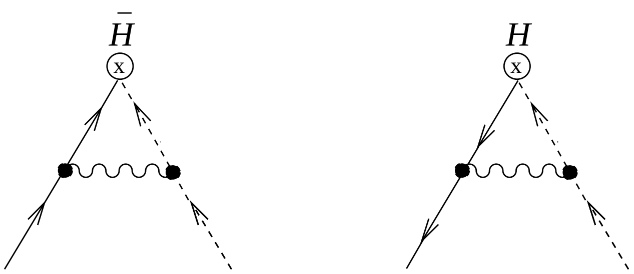


Figure 7.8: The new divergent one-loop diagrams introduced by the source terms $H, \bar{H}$.

7.8 The abelian case

We assume now that the gauge group is abelian. This means that the group structure constants $c^{abc} = 0$. The tree level effective action is:

$$S^{(0)}[A,\eta,\bar{\eta},\psi,\bar{\psi}] = \int d\ dx \left(\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{2\alpha}(\partial_\mu A_\mu)^2 - \bar{\eta}\partial^2\eta + \bar{\psi}(\gamma_\mu D_\mu + m)\psi\right) \tag{7.150}$$

The BRS invariance is rather trivial since $c^{abc} = 0$:

$$\begin{array}{rcl} Q_B A_\mu &=& \partial_\mu \eta \\ Q_B \bar{\eta} &=& \frac{1}{\alpha}\partial_\mu A_\mu \\ Q_B \eta &=& 0 \\ Q_B \psi &=& -i\eta\psi\ , \\ Q_B \bar{\psi} &=& -i\bar{\psi}\eta \end{array} \tag{7.151}$$

Our basic equations (7.146) and (7.147) reduces to

$$\int \left(\frac{\delta\tilde{\Gamma}}{\partial A_\mu}\frac{\delta\tilde{\Gamma}}{\delta K_\mu} + \frac{\delta\tilde{\Gamma}}{\delta\psi_\alpha}\frac{\delta\tilde{\Gamma}}{\delta\bar{H}} + \frac{\delta\tilde{\Gamma}}{\delta\bar{\psi}_\alpha}\frac{\delta\tilde{\Gamma}}{\delta H}\right) = 0 \tag{7.152}$$

$$\partial_\mu \frac{\delta\tilde{\Gamma}}{\delta K_\mu} = \frac{\delta\tilde{\Gamma}}{\delta\bar{\eta}} \tag{7.153}$$

The equations are however greatly simplified by the fact that *the ghosts do not couple to the gauge field* as is clear from (7.150) and (7.151). This implies that *none* of the BRS invariant source term renormalizes and we have

$$\frac{\delta\tilde{\Gamma}}{\delta K_\mu} = \partial_\mu \eta\ , \quad \frac{\delta\tilde{\Gamma}}{\delta H} = i\bar{\psi}\eta\ , \quad \frac{\delta\tilde{\Gamma}}{\delta\bar{H}} = i\eta\psi. \tag{7.154}$$

Eqs. (7.153) and (7.154) imply

$$\boxed{\partial_\mu \frac{\delta\tilde{\Gamma}}{\delta A_\mu} - i\bar{\psi}_\alpha \frac{\delta\tilde{\Gamma}}{\delta\bar{\psi}_\alpha} - i\frac{\delta\tilde{\Gamma}}{\delta\psi_\alpha}\psi_\alpha = 0} \tag{7.155}$$

This is the Ward identity for QED. By functional differentiation we get relations between the different 1PI vertex functions and propagators, as already discussed in detail.

Chapter 8

Chiral Anomalies

8.1 Chiral invariance

We use the word anomaly in a field theoretical context when a symmetry of the classical Lagrangian is not preserved in the quantum theory.

Recall that associated with any continuous *global* classical symmetry

$$\varphi_r(x) \to \{e^{-i\alpha\lambda}\}_{rs}\ \varphi_s(x) \tag{8.1}$$

we have a conserved *Nöether current.* The derivation is as follows: Let us consider an infinitesimal variation of φ in accordance with (8.1):

$$\varphi_r(x) \to\ \varphi_r(x) - i\varepsilon\,\lambda_{rs}\varphi_s(x) \tag{8.2}$$

Since we assume $\mathcal{L}(\varphi, \partial_\mu\varphi)$ is invariant under (8.1) we can write (to first order in the change $\delta\varphi$):

$$0 = \delta\mathcal{L}(\varphi, \partial_\mu\varphi) = \frac{\partial\mathcal{L}}{\partial\varphi_r}\ \delta\varphi_r + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_r)}\ \delta(\partial_\mu\varphi_r) \tag{8.3}$$

From (8.2) the change in φ_r can be written as $\delta\varphi_r = -i\varepsilon\lambda_{rs}\varphi_s$, $\delta(\partial_\mu\varphi_r) = -i\varepsilon\lambda_{rs}\partial_\mu\varphi_s$ and by use of the Euler-Lagrange equations:

$$\frac{\partial\mathcal{L}}{\partial\varphi_r} = \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_r)} \tag{8.4}$$

we get

$$0 = -i\varepsilon\partial_\mu\left\{\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_r)}\ \lambda_{rs}\ \varphi_s\right\} \tag{8.5}$$

We have the conserved Nöether current:

$$\boxed{j_\mu \equiv i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_r)}\ \lambda_{rs}\varphi_s\ , \qquad \partial_\mu j_\mu = 0} \tag{8.6}$$

Ex.1: The electromagnetic current

As the first example we consider the charged scalar field, with Lagrangian

$$\mathcal{L}(\varphi, \varphi^*, A_\mu) = (D_\mu\varphi)^*(D_\mu\varphi) + V\left(|\ \phi\ |^2\right), \tag{8.7}$$

where $D_\mu = \partial_\mu - ieA_\mu$ is the ordinary covariant derivative and we have ignored the kinetic part of the gauge field since it plays no role in the following. We get that $\lambda_{rs} = e$ for φ and $\lambda_{rs} = -e$ for φ^* and this leads to the following conserved current

$$j_\mu = ie\left((D_\mu\varphi)^*\varphi - \varphi^* D_\mu\varphi\right) \tag{8.8}$$

As the next example we consider the charged Dirac fermion, with Lagrangian

$$\mathcal{L}(\psi, \bar{\psi}, A_\mu) = -\bar{\psi}(\gamma_\mu D^\mu + m)\psi \tag{8.9}$$

In this case we have $\lambda_{rs} = e$ for ψ and $\lambda_{rs} = -e$ for $\bar{\psi}$. This leads to the conserved current (where we have dropped an i compared to (8.6))

$$j_\mu = e\bar{\psi}\gamma_\mu\psi \tag{8.10}$$

Ex.2: The chiral Current

The Lagrangian is the same as in the first example for the Dirac fermion, except that the mass is equal to zero:

$$\mathcal{L}(\psi, \bar{\psi}, A_\mu) = i\bar{\psi}\gamma_\mu D^\mu\psi \tag{8.11}$$

Recall from chapter 6 that it is possible in *even* dimensions to define γ_{d+1} by

$$\gamma_{d+1} = (-i)^{d/2+1}\gamma_1\gamma_2\cdots\gamma_d. \tag{8.12}$$

(In euclidean space-time the is one less factor of $-i$). The γ_{d+1}-matrix is hermitian and anticommutes with all γ_μ, $0 \leq \mu \leq d-1$ for d even, but is only proportional to the the identity matrix for odd dimensions d. The Lagrangian is invariant under the so-called *chiral* symmetry (which we now consider in $d = 4$)::

$$\begin{aligned} \psi(x) &\rightarrow e^{i\alpha\gamma_5}\psi \approx (1 + i\alpha\gamma_5)\psi & (8.13)\\ \bar{\psi}(x) &\rightarrow \bar{\psi}e^{i\alpha\gamma_5} \approx \bar{\psi}(1 + i\alpha\gamma_5). & (8.14) \end{aligned}$$

The invariance of $i\psi\gamma_\mu D^\mu\psi$ follows from $\{\gamma_\mu, \gamma_5\} = 0$. Note that a mass term breaks the symmetry (8.13)-(8.14):

$$\delta(\bar{\psi}\psi) \approx 2i\alpha\ \bar{\psi}\ \gamma_5\ \psi\ \neq\ 0. \tag{8.15}$$

Under the chiral symmetry we have $\lambda_{rs} \equiv \lambda_{\alpha\beta} = (\gamma_5)_{\alpha\beta}$ and the conserved Nöether current j^5_μ is

$$\boxed{j^5_\mu = \bar{\psi}\gamma_\mu\ \gamma_5\psi} \tag{8.16}$$

A priori it is not clear to what extend a classical relation like $\partial^\mu j_\mu = 0$ will be preserved in the quantum field theory. We used the classical Euler-Lagrange equations (8.4) to derive the current conservation and in the path integral we are clearly integrating over field configurations where this equation is not satisfied. On the other hand the classical field equations are (modulo problems with gauge fixing) satisfied as expectation values:

$$\left\langle \frac{\delta S[\varphi]}{\delta\varphi(x)} \right\rangle = 0 \tag{8.17}$$

This follows simply by making a change of variables in the functional integral: $\varphi(x) \to \varphi(x)+\varepsilon(x)$ as already discussed in connection with the Dyson–Schwinger equations. Thus it is not unreasonable to expect that we have for a Nöether current

$$\langle \partial^\mu j_\mu(x) \rangle = 0 \tag{8.18}$$

In fact this was one of the Ward identities of QED as already shown in one of the former chapters in the case where j_μ was the electromagnetic current (8.10). It might be worth to repeat the derivation in more general terms: Let us promote the global symmetry (1)–(2) to a *local* transformation by replacing $\alpha \to \alpha(x)$.

$$\varphi_r(x) \to \varphi_r(x) - i\alpha(x)\ \lambda_{rs}\ \varphi_s(x) \tag{8.19}$$

Of course the field transformation (8.19) is no longer a symmetry of the Lagrangian. We have:

$$\begin{aligned} \delta\mathcal{L}(\varphi, \partial_\mu \varphi) = \frac{\partial \mathcal{L}}{\partial \varphi_r}\delta\varphi_r + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_r)}\delta(\partial_\mu \varphi_r) = \\ -i\alpha(x)\left\{ \frac{\partial \mathcal{L}}{\partial(\varphi_r)}\lambda_{rs}\varphi_s + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_r)}\lambda_{rs}\partial_\mu\varphi_s \right\} - i\ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_r)}\lambda_{rs}\varphi_s\partial_\mu\alpha(x) \end{aligned} \tag{8.20}$$

The bracket is zero because (8.2) is a global symmetry and therefore:

$$\boxed{\delta S[\varphi] = \int dx\ \delta\mathcal{L}(\varphi) = \int dx\ \alpha(x)\partial^\mu j_\mu(x)} \tag{8.21}$$

Let us write

$$1 = \frac{\int \mathcal{D}\varphi e^{-S[\varphi]}}{\int \mathcal{D}\varphi e^{-S[\varphi]}}$$

and perform the change of variables (8.19) in the numerator: *Under the assumption that the functional measure $\mathcal{D}\varphi$ is invariant under the (formal) unitary local transformation (8.19) we get* (writing $e^{\delta S} \approx 1 - \delta S$):

$$\langle \delta S \rangle = 0 \ \ \forall \alpha(x) \qquad \text{or:} \qquad \boxed{\langle \partial^\mu j_\mu(x) \rangle = 0} \tag{8.22}$$

Clearly we can get a whole series of Ward identities associated with the global symmetry (8.1) if we change variables in the numerator in

$$\langle \varphi(x) \cdots \varphi(x_n) \rangle = \frac{\int \mathcal{D}\varphi\ \varphi(x_1)\cdots\varphi(x_n)e^{-S[\varphi]}}{\int \mathcal{D}\varphi\ e^{-S[\varphi]}}$$

exactly as we did for above.

If we concentrate on the chiral invariance described in ex. 2, it is an interesting exercise to check $\left\langle \partial^\mu j^5_\mu(x) \right\rangle = 0$ explicitly in perturbation theory. From the arguments given above it follows that:

(1): if $\left\langle \partial^\mu j^5_\mu(x) \right\rangle = 0$ it is a proof that the measure $\mathcal{D}\psi\mathcal{D}\bar\psi\mathcal{D}A$ is invariant under *local* chiral transformations. In fact one would expect this to be the case since formally (8.19) is a unitary transformation.

(2): If $\left\langle \partial^\mu j^5_\mu(x) \right\rangle \neq 0$ there must be a non-trivial Jacobian associated with the *local* chiral transformation.

8.2 Perturbative calculation

We now perform a perturbative calculation of $\langle \partial^\mu j_\mu^5 \rangle$. *As usual we will rotate to euclidean space-time while doing the calculations.* In order to simplify the calculation we will only perform it in 2-dimensions (but we will keep the notation γ_5, rather than change to γ_3).The generalization to four dimensions contains no new surprises. The Lagrangian is (using now euclidean notation)

$$\mathcal{L}(\bar\psi, \psi, A) = -\bar\psi \gamma_\mu D_\mu \psi + \frac{1}{4} F_{\mu\nu}^2 + \text{gaugefix} + \text{ghosts} \tag{8.23}$$

The interaction term is

$$\mathcal{L}_{int}(\bar\psi, \psi, A) = ie\bar\psi \gamma_\mu A_\mu \psi = ie\bar\psi \not{A} \psi \tag{8.24}$$

and since the dynamics of the A_μ field plays no role at l-loop level we will just treat A_μ as an external field and only concentrate on the functional integration over the $\psi, \bar\psi$ fields:

$$\left\langle \partial_\mu j_\mu^5 \right\rangle = \frac{\int \mathcal{D}\psi \mathcal{D}\bar\psi \; \partial_\mu j_\mu^5(x) \; e^{ie\int dx \bar\psi \not{A} \psi} \; e^{-S_0[\psi, \bar\psi]}}{\int \mathcal{D}\psi \mathcal{D}\bar\psi \; e^{-S_0[\psi, \bar\psi]}} \tag{8.25}$$

where

$$S_0[\psi, \bar\psi] = \int dx \; \bar\psi \not\partial \psi. \tag{8.26}$$

By expanding the exponent of the interaction term in a power series in the charge (the coupling constant) e we get to lowest order in e:

$$\left\langle \partial_\mu j_\mu^5(x) \right\rangle = \left\langle \partial_\mu j_\mu^5(x) \right\rangle_0 + ie \int dy A_\nu^{ex}(y) \left\langle \bar\psi(y) \gamma_\nu \psi(y) \partial_\mu j_\mu^5(x) \right\rangle_0 + \mathcal{O}(e^2) \tag{8.27}$$

Here $\langle \;\cdot \rangle_0$ denotes the expectation value with respect to the free action $S_0[\psi, \bar\psi]$. The first term on the rhs is zero, as is easily seen: It involves $\mathrm{tr}\, \gamma_\mu \gamma_5 = 0$. By a Wick contraction we get:

$$\begin{aligned} \left\langle \partial_\mu j_\mu^5(x) \right\rangle &= ie \frac{\partial}{\partial x_\mu} \int dy A_\nu^{ex}(y) \left\langle \psi_\beta(x) \bar\psi_{\alpha'}(y) \right\rangle_0 (\gamma_\nu)_{\alpha'\beta'} \left\langle \psi_{\beta'}(y) \bar\psi_\alpha(x) \right\rangle_0 (\gamma_\mu \gamma_5)_{\alpha\beta} \\ &= ie \frac{\partial}{\partial x_\mu} \int dy A_\nu^{ex}(y) \; \mathrm{tr} \left[S^0(x-y) \gamma_\nu S^0(y-x) \gamma_\mu \gamma_5 \right] \end{aligned} \tag{8.28}$$

where $S^0(x-y)_{\alpha\beta} \equiv \left\langle \psi_\alpha(x) \bar\psi_\beta(y) \right\rangle_0$ denotes the free propagator $\not\partial^{-1}(x-y)$ and tr is a spinor index trace.

Fourier transforming (8.28) leads to:

$$\left\langle p_\mu j_\mu^5(p) \right\rangle = e p_\mu \; A_\nu^{ex}(-p) \int \frac{d^2k}{(2\pi)^2} \mathrm{tr} \left[S^0(k) \gamma_\nu S^0(k+p) \gamma_\mu \gamma_5 \right] \tag{8.29}$$

or:

$$\left\langle \partial_\mu j_\mu^5(p) \right\rangle = e p_\mu A_\nu^{ex}(-p) \; T_{\mu\nu}(p) \tag{8.30}$$

$$T_{\mu\nu}(p) = \mu^{2-d} \int \frac{d^dk}{(2\pi)^d} \frac{\mathrm{tr} \left[(\gamma_\rho k_\rho) \gamma_\nu (k+p)_\lambda \gamma_\lambda \gamma_\mu \gamma_5 \right]}{k^2 (k+p)^2} \tag{8.31}$$

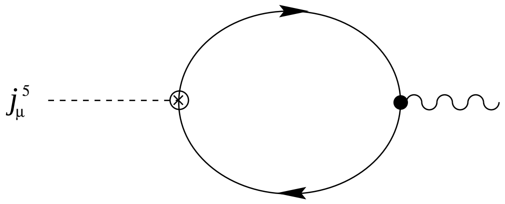


Figure 8.1: The diagram contributing to the chiral anomaly in two dimensions.

where we have converted the (formal) integral in (8.29) into the standard notation of dimensional regularization, i.e. we have written the integral in d dimension instead of two dimensions, and introduced the mass parameter μ which keeps the total mass dimension of $T_{\mu\nu}(p)$ invariant. We have in addition used that the massless fermion propagator is

$$S^0(k) = \frac{i}{\not k} = \frac{i\not k}{k^2} \tag{8.32}$$

Diagrammatically we can represent $T_{\mu\nu}(p)$ as shown in fig.8.1. It is seen that the integral defining $T_{\mu\nu}(p)$ in (8.31) is logarithmically divergent in $d = 2$. It can be evaluated using standard techniques of dimensional regularization:

$$T_{\mu\nu}(p) = \mu^{2-d}\mathrm{tr}\left[\gamma_\mu\gamma_5\gamma_\alpha\ \gamma_\nu\gamma_\beta\right]\int\frac{d^dk}{(2\pi)^d}\,\frac{k_\alpha(k+p)_\beta}{k^2(k+p)^2} \tag{8.33}$$

where

$$\begin{aligned}\int\frac{d^dk}{(2\pi)^d}\frac{k_\alpha(k+p)_\beta}{k^2(k+p)^2} &= \int_0^1 d\alpha\int\frac{d^dk}{(2\pi)^d}\ \frac{k_\alpha(k+p)_\beta}{(k^2+2\alpha k\cdot p+\alpha p^2)^2}\\ &= \int_0^1 d\alpha\left\{\frac{\Gamma(1-d/2)}{(4\pi)^{d/2}}\frac{\delta_{\alpha\beta}}{2(\alpha(1-\alpha)p^2)^{1-d/2}} - \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}}\frac{\alpha(1-\alpha)p_\alpha p_\beta}{(\alpha(1-\alpha)p^2)^{2-d/2}}\right\}\\ &= \left\{\frac{\Gamma(1-d/2)}{(4\pi)^{d/2}}\frac{\delta_{\alpha\beta}}{2} - \frac{\Gamma(2-d/2)}{(4\pi)^{d/2}}\frac{p_\alpha p_\beta}{p^2}\right\}\end{aligned} \tag{8.34}$$

At this point we have the pole term $\Gamma(1-d/2)$ corresponding to the logarithmic singularity. Using a little γ-algebra the second term is zero and we get:

$$\begin{aligned}&\left\langle p_\mu j^5_\mu(p)\right\rangle = \\ &\qquad ep_\mu A^{ex}_\nu(-p)\left(\frac{\mu^2}{p^2}\right)^{1-d/2}\frac{\Gamma(1-d/2)}{2(4\pi)^{d/2}}\mathrm{tr}\left[\gamma_\mu\gamma_5\gamma_\alpha\gamma_\nu\gamma_\alpha\right]\ \cdot\ \int_0^1 d\alpha\frac{1}{\alpha(1-\alpha)^{1-d/2}}\end{aligned} \tag{8.35}$$

and further γ-algebra leads to:

$$\mathrm{tr}\left[\gamma_\mu\gamma_5\gamma_\alpha\gamma_\nu\gamma_\alpha\right] = -\mathrm{tr}\left[\gamma_5\gamma_\mu\left\{2\delta_{\alpha\nu}-\gamma_\nu\gamma_\alpha\right\}\gamma_\alpha\right] = (d-2)\mathrm{tr}\left[\gamma_5\gamma_\mu\gamma_\nu\right] \tag{8.36}$$

where we have used

$$\gamma_\alpha\gamma_\alpha = d \qquad (\text{since } \gamma^2 = 1) \tag{8.37}$$

We can now take the limit $d \to 2$ and get:

$$\left\langle p_\mu j^5_\mu\right\rangle = -\frac{e}{2\pi}\,\mathrm{tr}\left[\gamma_\mu\gamma_\nu\gamma_5\right]p_\mu A^{ex}_\nu(-p) \tag{8.38}$$

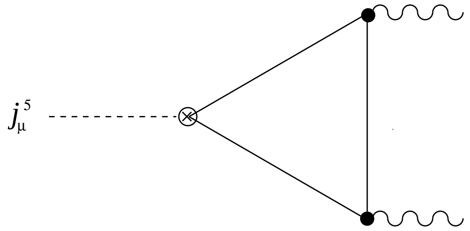


Figure 8.2: The diagram contributing to the chiral anomaly in four dimensions.

or

$$\left\langle \partial_\mu j^5_\mu(x) \right\rangle = -\frac{e}{2\pi}\,\mathrm{tr}\,[\gamma_\mu\gamma_\nu\gamma_5]\,\frac{1}{2}(\partial_\mu A^{ex}_\nu(x) - \partial_\nu A^{ex}_\mu(x)) \tag{8.39}$$

At this point (*but only at this point, in order not to get in conflict with dimensional regularization which has no simple definition of* γ_5 *away from even dimensions*) we can use the special properties of two dimensions:

$$\gamma_\mu\gamma_\nu = i\epsilon_{\mu\nu}\gamma_5 + \delta_{\mu\nu} \tag{8.40}$$

to get

$$\left\langle \partial_\mu j^5_\mu \right\rangle = -i\frac{e}{\pi}\,\tilde{F}^{ex}\ ,\quad \tilde{F} \equiv \frac{1}{2}\epsilon_{\mu\nu}F^{ex}_{\mu\nu}. \tag{8.41}$$

Let us at this point rotate back to Minkowski space-time. Thereby $\tilde{F} \to i\tilde{F}$ and we get

$$\boxed{\left\langle \partial^\mu j^5_\mu \right\rangle = \frac{e}{\pi}\,\tilde{F}^{ex}\ ,\ \tilde{F} \equiv \frac{1}{2}\epsilon^{\mu\nu}F^{ex}_{\mu\nu}} \qquad (d=2) \tag{8.42}$$

We conclude that we have an anomaly for the chiral current.

The same conclusion is true for four dimensions. Because of the properties of γ_μ, γ_5 matrices is 4 dimensions the diagram shown in fig.8.1 is zero. But the triangle diagram shown in fig.8.2 is logarithmically divergent in $d = 4$ and we get expressions quite similar to (8.41) and (8.42):

$$\left\langle \partial_\mu j^5_\mu \right\rangle = -i\frac{e^2}{16\pi^2}\,\tilde{F}^{ex}_{\mu\nu}F^{ex}_{\mu\nu}\ ,\ \tilde{F}_{\mu\nu} \equiv \frac{1}{2}\epsilon_{\mu\nu\lambda\rho}F_{\lambda\rho} \tag{8.43}$$

in euclidean space. By rotating to Minkowski space-time $F\tilde{F} \to iF\tilde{F}$ and we have

$$\boxed{\left\langle \partial^\mu j^5_\mu \right\rangle = \frac{e^2}{16\pi^2}\,\tilde{F}^{ex}_{\mu\nu}(F^{ex})^{\mu\nu}\ ,\ \tilde{F}_{\mu\nu} \equiv \frac{1}{2}\epsilon_{\mu\nu\lambda\rho}F^{\lambda\rho}} \qquad (d=4) \tag{8.44}$$

From the arguments given above a *local* chiral transformation can not leave the measure $\mathcal{D}\psi\mathcal{D}\bar{\psi}$ invariant. That the situation is indeed like this was first realized by Fujikawa (1979). In the next section we will now outline the proof.

8.3 The path integral measure under chiral transformations

Again we will study the effect of the chiral transformations in two dimensions and only mention how it is generalized to four dimensions. (In this case, somewhat contrary to the perturbative calculation above, the generalization is quite trivial).

Again we view the A_μ field as an external field since it plays no role in the anomaly, except as a spectator and again we rotate to euclidean space-time during the calculation.

Let $A^{ex}_\mu(x)$ be given. $i\not{D} = i\gamma_\mu(\partial_\mu - ieA_\mu)$ is formally a hermitian operator and we can expand ψ and $\bar{\psi}$ on its eigenfunctions. In order to simplify the notation we will assume (falsely) that the spectrum of $i\gamma_\mu D_\mu$ is discrete and denote the eigenfunctions $\varphi_n(x)$. We have

$$\psi(x) = \sum_n a_n \varphi_n(x), \qquad \bar{\psi}(x) = \sum_n \varphi_n^+(x)\,\bar{b}_n \tag{8.45}$$

$$i\gamma_\mu D_\mu \varphi_n(x) = \lambda_n \varphi_n(x) \;, \qquad \int dx\; \varphi_n^+ \varphi_m = \delta_{nm} \tag{8.46}$$

Almost by definition we have:

$$\mathcal{D}\bar{\psi}\mathcal{D}\psi = \prod_n d\bar{b}_n \; \prod_m da_m \tag{8.47}$$

We can calculate the effect of a *local* chiral transformation on a_n and $\bar{b}_n$:

$$\begin{array}{l} \psi(x) \to \psi'(x) = e^{i\alpha(x)\gamma_5}\psi(x) = \sum_n a_n e^{i\alpha(x)\gamma_5}\varphi_n(x) \\ \qquad\quad \| \\ \quad \sum_n a'_n \varphi_n(x) \end{array} \tag{8.48}$$

or:

$$a'_n = \sum_m \int dx \varphi_n^+(x) e^{i\alpha(x)\gamma_5} \varphi_m(x) a_m \equiv \sum_m C_{nm} a_m \tag{8.49}$$

$$\bar{b}'_n = \sum_m \int dx \bar{b}_m \varphi_m^+(x) e^{i\alpha(x)\gamma_5} \varphi_n(x) \equiv \sum_m \bar{b}_m C_{mn} \tag{8.50}$$

We conclude that

$$\prod_n d\bar{b}'_n \; \prod_m da'_m = (\det C)^{-2} \; \prod_n d\bar{b}_n \; \prod_m da_m \tag{8.51}$$

The question we have to address is whether $\det C = 1$ for infinitesimal *local* chiral transformations. We can write

$$C = I + \hat{\alpha} + \mathcal{O}(\alpha^2) \tag{8.52}$$

$$\hat{\alpha}_{nm} \equiv \int dx \varphi_n^+ \; \alpha(x)\gamma_5 \varphi_m(x) \;, \qquad (I_{nm} = \delta_{nm}) \tag{8.53}$$

$$(\det C)^{-1} = e^{-\mathrm{tr}\,\log C} = e^{-\mathrm{tr}\,\hat{\alpha} + \mathcal{O}(\alpha^2)} = 1 - \mathrm{tr}\,\hat{\alpha} + \mathcal{O}(\alpha^2) \tag{8.54}$$

and we have

$$\boxed{(\det C)^{-1} = 1 - \int dx\; \alpha(x)\; \sum_n \varphi_n^+(x)\gamma_5\varphi_n(x) + \mathcal{O}(\alpha^2)} \tag{8.55}$$

The expression $\sum_n \varphi_n^+(x)\gamma_5\varphi_n(x)$ is not well defined. We need to regularize it. Since we will insist that our theory is gauge invariant *we demand that the regularization should be* *gauge invariant.* The eigenvalues λ_n are gauge invariant since $i\gamma_\mu D_\mu$ is covariant. Since $|\lambda_n| \to \infty$ for $|n| \to \infty$ two gauge invariant regularizations suggest themselves:

$$(A): \qquad \sum_n \frac{1}{\lambda_n^s}\, \varphi_n^+(x)\gamma_5\, \varphi_n(x)\ , \qquad s \to 0 \tag{8.56}$$

$$(B): \qquad \sum_n \varphi_n^+(x)\ \gamma_5\ e^{-\left(\frac{\lambda_n}{M}\right)^2}\ \varphi_n(x)\ , \quad m \to \infty \tag{8.57}$$

The first method is called ζ-function regularization after Riemann's Zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$. $\zeta(s)$ is analytic for $s > 1$, but can be analytically continued to $s < 1$ and has no poles except at $s = 1$. Here we will follow Fujikawa, however and use (B). It is a *gauge invariant* way to cut away eigenvalues $|\lambda_n| > M$. In order to evaluate (8.57) we reintroduce $i\not{D}$ and write $\varphi_n(x) = \langle x|n\rangle$:

$$\sum_n \varphi_n^+(x)\gamma_5 e^{\frac{\not{D}^2}{M^2}}\varphi_n(x) = \sum_n \langle n \mid x\rangle\gamma_5\ e^{\frac{\not{D}^2}{M^2}}\langle x \mid n\rangle =$$

$$\mathrm{Tr}\left\{\mid x\rangle\gamma_5\ e^{\frac{\not{D}^2}{M^2}}\langle x \mid\right\} = \mathrm{tr}\int \frac{d^2k}{(2\pi)^2}\ e^{-ikx}\ \gamma_5\ e^{\frac{\not{D}^2}{M^2}}\ e^{ikx} \tag{8.58}$$

In (8.58) we have simply performed a change in basis from $\mid n\rangle$ to plane waves $e^{ikx} \times v_\alpha$, where the v_α's are just constant spinors. The first trace "Tr" is over the whole Hilbert space while "tr" just denotes the trace over spinor indices.

lemma: $$e^{-ikx}e^{\frac{\not{D}^2}{M^2}}e^{ikx} = \left(e^{-\frac{k^2}{M^2}}\ e^{\frac{\not{D}^2+2ik_\mu D_\mu}{M^2}}\right)\cdot 1(x)$$

where

$$\not{D}^2 = D^2 - ie\sigma_{\mu\nu}F_\mu\ ,\ \sigma_{\mu\nu} = \frac{i}{4}[\gamma_\mu, \gamma_\nu]$$

and $1(x)$ denotes the constant function.

Proof:

$$\begin{aligned}
\not{D}^2 &= \gamma_\mu D_\mu \gamma_\nu D_\nu = \frac{1}{2}\left\{\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu\right\} D_\mu D_\nu + \frac{1}{2}\left\{\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu\right\} D_\mu D_\nu \\
&= \delta_{\mu\nu} D_\mu D_\nu + \frac{1}{2}\left[\gamma_\mu\gamma_\nu\right] D_\mu D_\nu = D_\mu^2 + \frac{1}{4}\left[\gamma_\mu, \gamma_\nu\right]\left(D_\mu D_\nu - D_\nu D_\mu\right) \\
&= D_\mu^2 + \frac{1}{4}\left[\gamma_\mu, \gamma_\nu\right]\left(-ieF_{\mu\nu}\right) = D^2 - i\sigma_{\mu\nu}F_{\mu\nu}.
\end{aligned}$$

This means that

$$\left(e^{-ikx}\not{D}^2 e^{ikx}\right) f(x) = \left((-k^2 + 2ikD_\mu) + \not{D}^2\right) f(x)$$

From (8.58) and the lemma we now get

$$\sum_n \varphi_n^+(x)\gamma_5\ e^{\frac{\not{D}^2}{M^2}}\varphi_n(x) = \tag{8.59}$$

$$\int \frac{d^2k}{(2\pi)^2}\ e^{-\frac{k^2}{M^2}}\ \mathrm{tr}\,\gamma_5 \left(1 + \frac{\not{D}^2 + 2ik_\mu D_\mu}{1!M^2} + \frac{(\not{D}^2 + 2ik_\mu D_\mu)^2}{2!M^4} + \mathcal{O}\left(\frac{1}{M^6}\right)\right)\cdot 1(x)$$

The gaussian integral leads by dimensional reasons to a term $\sim M^2$ and the first term is potentially divergent. However, $\mathrm{tr}\,\gamma_5 = 0$. The second term is finite and independent of M as $M \to \infty$ while the rest of the terms will go to zero as $\mathcal{O}(1/M^2)$. From the second term only $\mathrm{tr}\,\gamma_5\sigma_{\mu\nu} \neq 0$ ($\mathrm{tr}\,\gamma_5\sigma_{\mu\nu} = \epsilon_{\mu\nu}$ in $d = 2$) and we get:

$$\sum_n \varphi_n^+(x)\gamma_5\, e^{\frac{\not{D}^2}{M^2}}\, \varphi_n(x) \to \int \frac{d^2k}{(2\pi)^2}\, e^{-\frac{k^2}{M^2}}\, \frac{ie\epsilon_{\mu\nu}F_{\mu\nu}(x)}{M^2} = \frac{ie}{2\pi}\, \tilde{F} \tag{8.60}$$

for $M \to \infty$ in $d = 2$. The result may be generalized to $d = 4$. The gaussian integral is then proportional M^4. The two first terms in (8.59) are potentially divergent terms but $\mathrm{tr}\,\gamma_5(\cdots) = 0$ for all terms. The third term in (8.59) leads to a finite contribution for $M \to \infty$ while the terms denoted $\mathcal{O}(1/M^6)$ vanish in this limit. For the third term only $\mathrm{tr}\,\gamma_5\sigma_{\mu\nu}\sigma_{\lambda\xi} \neq 0$ and in fact equal $\frac{i}{2}\epsilon_{\mu\nu\lambda\xi}$. We then get

$$\sum_n \varphi_n^+(x)\gamma_5 e^{\frac{\not{D}^2}{M^2}}\varphi_n(x) \to i\frac{e^2}{32\pi^2}\, F_{\mu\nu}(x)\tilde{F}_{\mu\nu}(x) \tag{8.61}$$

in the limit $M \to \infty$.

We can now derive the chiral ward identities directly from the functional integral. We only have to add that during a *local* chiral transformation $\psi(x) \to e^{\alpha(x)\gamma_5}\psi(x)$, the measure changes as:

$$\boxed{\mathcal{D}\psi\mathcal{D}\bar{\psi} \to \mathcal{D}\psi\mathcal{D}\bar{\psi}e^{\frac{ie}{\pi}\int dx\alpha(x)\tilde{F}(x)}} \qquad (d = 2) \tag{8.62}$$

$$\boxed{\mathcal{D}\psi\mathcal{D}\bar{\psi} \to \mathcal{D}\psi\mathcal{D}\bar{\psi}e^{i\frac{e^2}{16\pi}\int dx\alpha(x)\tilde{F}_{\mu\nu}(x)F_{\mu\nu}(x)}} \qquad (d = 4) \tag{8.63}$$

Therefore the change in action under an infinitesimal local chiral transformation is effectively changed from equation (8.21) to (for $d = 4$)

$$\delta S^{eff} = \int dx\alpha(x)\left(\partial_\mu J^5_\mu(x) + \frac{ie^2}{16\pi^2}\tilde{F}_{\mu\nu}(x)F_{\mu\nu}(x)\right) \tag{8.64}$$

and we get as in (17) that $\langle \delta S^{eff} \rangle = 0$, which just reflects a "trivial" change in variables in the path integral, leads to the chiral anomalies (8.41) ($d = 2$) and (8.43) ($d = 4$). Applying the same transformation to the expectation values

$$\langle \psi(x_1)\cdots\psi(x_n)\bar{\psi}(y_1)\cdots\bar{\psi}(y_m)A_{m_1}(z_1)\cdots A_{m_l}(z_l)\rangle \tag{8.65}$$

and using the transformation (8.63) leads to the so-called chiral ward identities between various Green functions. These were first derived in perturbation theory, and before the derivation of Fujikawa it was a mystery how one could obtain them from formal manipulations of the path integral.

8.4 Extensions of the simple anomaly

In the following we will mention several extensions of the above formulas for the chiral anomaly. No proofs will be given, but in principle the results are straight forward generalizations of the simplest example of an anomaly studied above.

8.4.1 The chiral charge for abelian theories

In the former sections we studied the abelian anomaly. The name anomaly was related to the fact that the Nöether current $j^5_\mu(x)$ of the abelian global chiral transformation

$$\psi \to e^{i\gamma_5\alpha}\psi \tag{8.66}$$

was not conserved at the quantum level. Integrating over space we find (discarding surface terms of the current at spacial infinity)

$$\frac{\partial Q_5}{\partial t} = \frac{e^2}{16\pi^2}\int d^3x \; \mathrm{tr}\, \tilde{F}_{\mu\nu}F^{\mu\nu} \neq 0, \qquad Q_5(t) \equiv \int d^3x \; j^5_0(x_i,t) \tag{8.67}$$

However, in the formulation of Fujikawa it seemed as if the problems were connected to local chiral transformations, and it is not entirely obvious that global chiral symmetry really should be broken. Indeed, it turns out that the term $\tilde{F}F$ which violates the conservation of the current j^5_μ is a total divergence:

$$F_{\mu\nu}\tilde{F}^{\mu\nu} = 4\partial_\mu\left(\varepsilon^{\mu\nu\rho\sigma}A_\nu\partial_\rho A_\sigma\right) \tag{8.68}$$

and we conclude that we can define another current

$$\tilde{j}^5_\mu = j^5_\mu - \frac{e^2}{4\pi^2}\left(\varepsilon_{\mu\nu\rho\sigma}A^\nu\partial^\rho A^\sigma\right) \tag{8.69}$$

which *is* conserved and where

$$\tilde{Q}_5 = \int d^3x \tilde{j}^5_0 \tag{8.70}$$

is a conserved charge. At first sight it looks alarming to attribute any physical significance to the current $\tilde{j}^5_\mu$ since it is not gauge invariant: Under a gauge transformation $A_\mu \to A_\mu - e^{-1}\partial_\mu\alpha$ we have:

$$\tilde{j}^5_\mu \to \tilde{j}^5_\mu + \frac{e}{4\pi^2}\varepsilon_{\mu\nu\rho\sigma}\partial^\rho A^\sigma\partial^\nu\alpha \tag{8.71}$$

However, $\tilde{Q}_5$ *is* gauge invariant:

$$\tilde{Q}_5 = Q_5 - S_{cs}[A_i] \tag{8.72}$$

where $S_{cs}[A_i]$ is the so-called three-dimensional *Chern-Simons action*

$$S_{cs}[A_i] = \frac{e^2}{4\pi^2}\int d^3x \; \varepsilon_{ijk}A_i\partial_j A_k. \tag{8.73}$$

This action plays an important role in solid state physics in the attempts to describe the quantum Hall effect and high T_c superconductivity. Unfortunately we have no space to discuss these interesting aspects here. Let us only show that $S_{cs}[A]$ is gauge invariant: Under a gauge transformation we have:

$$S_{cs}[A] \to S_{cs}[A] - \int d^3x \; \varepsilon_{ijk}\partial_j A_k\partial_i\alpha \tag{8.74}$$

and by a partial integration the last term can by written as a surface integral over a surface S at spatial infinity:

$$\int_S dS_i B_i\alpha \tag{8.75}$$

This integral vanishes in an abelian theory without monopoles where the magnetic field B_i falls off faster than $1/r^2$.

Since $\tilde{Q}_5$ is gauge invariant and conserved it is a potential candidate as the generator of *global* chiral transformations in the quantum theory instead of the non-conserved Q_5 and one can show that it indeed generates the global chiral symmetries in the quantum theory. We conclude that global chiral symmetry is unbroken at the quantum level even if the classical Nöether current is not conserved.

8.4.2 The chiral charge in non-abelian theories

The abelian chiral symmetry can be embedded in a non-abelian theory of massless fermions in a trivial way. Let the action be

$$S(\bar{\psi}, \psi) = -\int d^4x \bar{\psi}\gamma^\mu D_\mu \psi \tag{8.76}$$

$$D_\mu = \partial_\mu - igA_\mu^a T^a. \tag{8.77}$$

(8.76) still has an *abelian* chiral symmetry (8.66). We get as a trivial generalization of our former anomaly equation:

$$\partial^\mu j_\mu^5 = \frac{g^2}{16\pi^2} \mathrm{tr}\, F^{\mu\nu} \tilde{F}_{\mu\nu} \tag{8.78}$$

where the $F_{\mu\nu}$ inside the trace as usual is the Lie algebra element $F_{\mu\nu}^a T^a$. Again it is not clear that the *global* chiral symmetry should be broken. Indeed, even in the non-abelian case one can write $\mathrm{tr}\, F^{\mu\nu}\tilde{F}_{\mu\nu}$ as a total derivative:

$$\mathrm{tr}\, F^{\mu\nu}\tilde{F}_{\mu\nu} = \partial_\mu \left[4\varepsilon^{\mu\nu\rho\sigma} \mathrm{tr}\, (A_\nu \partial_\rho A_\rho + \frac{2}{3} A_\nu A_\rho A_\sigma) \right] \tag{8.79}$$

where $A_\mu = A_\mu^a T^a$. Again we can construct a current:

$$\tilde{j}_\mu^5 = j_\mu^5 - \frac{g^2}{4\pi^2} \varepsilon_{\mu\nu\rho\sigma} (A^\nu \partial^\rho A^\sigma + \frac{2}{3} A^\nu A^\rho A^\sigma) \tag{8.80}$$

which is conserved. It suffers from the same disease as its abelian analogue: It is not gauge invariant. Let us repeat the analysis of the abelian theory and check whether the conserved charge $\tilde{Q}_5$ *is* gauge invariant. Integrating $\tilde{j}_5$ over space we get:

$$\tilde{Q}_5 = Q_5 - S_{cs}[A_i] \tag{8.81}$$

where the $S_{sc}[A]$ denotes the non-abelian Chern-Simons action:

$$S_{cs}[A] = \frac{g^2}{4\pi^2} \int d^3x \; \varepsilon_{ijk} \; \mathrm{tr} \left[A_i \partial_j A_k + \frac{2}{3} A_i A_j A_k \right] \tag{8.82}$$

Let us check how $S_{cs}[A]$ transform under a non-abelian gauge transformation

$$A_i \to U A_i U^{-1} - \frac{i}{g} \partial_i U U^{-1} \tag{8.83}$$

Since the time-dependence is of no importance in the present context we will ignore it. Then we can view the gauge function $U(x) \in G$, G being the gauge group, as a map from R^3 to G. We will further assume that the gauge transformation is trivial at spatial infinity:$U(x) = I$ for $|x| \to \infty$. This means that we can compactify R^3 to S^3 and still consider $U(x)$ as a map from S^3 into G. In order to simplify the discussion we will further assume that $G = SU(2)$, the simplest continuous non-abelian group. However, all conclusions to be reached extend to other semi-simple Lie groups. Any $SU(2)$ matrix can represented in a unique way as

$$U = \sum_{i=1}^{4} a_i \sigma_i, \qquad \sum_{i=1}^{4} a_i^2 = 1 \tag{8.84}$$

where σ_i denote the Pauli matrices for $i = 1, 2, 3$ while $\sigma_4 = I$. We see that $SU(2)$ is topologically equivalent to S^3. With the above mentioned boundary condition on U the topological classification of gauge transformations $U(x)$ from $R^3 \to G$ is identical to the topological classification of maps $U : S^3 \to S^3$. The homotopy classes of such maps are completely characterized by the so-called *winding number* n, which describes how many times the image of S^3 by U winds around S^3. Gauge transformations within one class can be continuously deformed into each other, but it is impossible by a continuous deformation of the map U to move from one homotopy class to another. For a given U there is a closed expression for this winding number:

$$n = \frac{1}{4\pi^2} \int d^3x \; \varepsilon_{ijk} \; \mathrm{tr}\, \partial_i U U^{-1} \partial_j U U^{-1} \partial_k U U^{-1}. \tag{8.85}$$

Let us now return to the Chern-Simons action and perform a gauge transformation (8.83). After some algebra one finds:

$$S_{cs}[A] \to S_{cs}[A] + \frac{1}{4\pi^2} \int d^3x \; \varepsilon_{ijk} \; \mathrm{tr}\, \partial_i U U^{-1} \partial_j U U^{-1} \partial_k U U^{-1}. \tag{8.86}$$

we now see that $S_{cs}[A]$ is invariant under so-called local gauge transformations, i.e. gauge transformations which are topologically trivial. This is analogous to the abelian case. However, a new feature has entered in the non-abelian case. For the so-called *large* gauge transformations $S_{cs}[A]$ *will* change by integer amounts, according to (8.85) and (8.86). *We conclude that* $\tilde{Q}_5$ *is not gauge invariant.*

There is still a possibility for saving $\tilde{Q}_5$ as a gauge invariant quantity. Since the gauge transformations fall in disconnected topological classes it is not a priori obvious that they all belong to the theory. Clearly the ones connected to the identity, the ones with winding number 0, should be included, but the other ones could be irrelevant for dynamical reasons. To make this more precise consider the following situation: Let us work in temporal gauge where $A_0 = 0$. As a residual gauge invariance we still have all gauge transformations which are independent of time, i.e. precisely the ones we considered above. Let us rotate to euclidian space-time and ask whether there is *any* field configuration with finite action which interpolate between the trivial classical vacuum configuration $A_i^{(0)} = 0$ and the vacuum configuration obtained by a large gauge transformation with winding number n: $A_i^{(n)} = -i/g \partial_i U(n) U^{-1}(n)$, where $U(n)$ has winding number n. If A_i^{int} is such an interpolating configuration, i.e. $A_i^{int}(x_4 = -\infty, x_j) = A_i^{(0)}(x_j)$ and $A_i^{int}(x_4 = \infty, x_j) =$

$A_i^{(1)}(x_j)$ we have to lowest semiclassical order that the amplitude for a transition $A^{(0)} \to A^{(n)}$ will be

$$T(A^{(0)} \to A^{(n)}) \sim e^{-\min(\frac{1}{4}\int d^4x \, \mathrm{tr}\, F_{\mu\nu}(A^{int})^2)} \tag{8.87}$$

In case the action for all such interpolation configurations were infinite it would indicate, that it is impossible to move between the different gauge sectors and if we start out with configurations close to the trivial vacuum, the large gauge transformations will be irrelevant and consequently we would be able to consider $\tilde{Q}_5$ as conserved.

The general statement we can make is that there exists such interpolating configurations with finite action. Any such configuration has to satisfy that

$$n = S_{cs}[A^{(n)}] - S_{cs}[A^{(0)}] = \frac{g^2}{16\pi^2}\int d^4x \; \mathrm{tr}\, F_{\mu\nu}\tilde{F}_{\mu\nu}, \tag{8.88}$$

as is clear from the definitions above. In fact the gauge field configurations which falls off sufficiently fast that we can consider them as belonging to the compactification S^4 of euclidean R^4 are classified by their *Pontryagin index n*, defined by

$$n = \frac{g^2}{16\pi^2}\int d^4x \; \mathrm{tr}\, F_{\mu\nu}\tilde{F}_{\mu\nu} \tag{8.89}$$

We see that this n is nothing but the winding number considered above. In addition the configurations which allow a compactification to S^4 fall off sufficiently fast that the action *is* finite and we see that there is no reason not to include the large gauge transformations among our allowed gauge transformations. It is an interesting exercise to calculate the minimum action which can appear in (8.87). The gauge field configurations which saturate the minimum are so-called self-dual field configurations ($F_{\mu\nu} = \tilde{F}_{\mu\nu}$) which means that (8.89) actually gives the lower bound on the action. For $n = 1$ these configurations are called *instantons* and we see that the leading semiclassical contribution to (8.87) will be:

$$T(A^{(0)} \to A^{(1)}) \sim e^{-\frac{16\pi^2}{g^2}}. \tag{8.90}$$

We refer to the chapter on classical gauge fields for details.

We have therefore reached the conclusion that $\tilde{Q}_5$ cannot be gauge invariant, and it has to be discarded as a physical observable. At this point we have no candidate for a conserved chiral charge, and the conclusion which has been drawn is that *the abelian global chiral symmetry has to be considered explicitly broken for non-abelian gauge theories*, in contrast to the situation for abelian gauge theories.

8.4.3 Gauge anomalies

In the above examples the gauge fields have been coupled to the fermions via the vector current $\bar{\psi}\gamma_\mu A_\mu \psi = \mathrm{tr}\, j_\mu A_\mu$ and we could maintain conservation of j_μ at the quantum level. In fact the chiral anomaly arose as a result of insisting on gauge invariance, which implies the conservation of j_μ. It is clear that chiral fermions, i.e. fermions which are either right-handed or left-handed:

$$\psi_{R,L} = \frac{1}{2}(1 \pm \gamma_5)\psi, \tag{8.91}$$

coupled to gauge fields will pose a potential problem since a potential anomaly now might get in conflict with the current conservation required by local gauge invariance. Let us consider a general gauge group G which can have both abelian and non-abelian components and couple the gauge fields to left handed fermions:

$$S(\bar{\psi},\psi,A) = -\int d^4x\bar{\psi}\frac{1}{2}(1+\gamma_5)\not{D}\psi \tag{8.92}$$

This action is invariant under local gauge transformations

$$A_\mu \to U(x)A_\mu U(x)^{-1} - \frac{i}{g}\,\partial_\mu U(x)U(x)^{-1}, \quad \psi(x)\to U(x)\psi(x), \quad \bar{\psi}\to\bar{\psi}U^{-1}(x) \tag{8.93}$$

and of course under the corresponding global transformations, from which we get the *covariantly* conserved Nöether current

$$D^\mu j_\mu = 0,\ (\partial^\mu j_\mu - ig[A^\mu, j_\mu] = 0), \qquad j^a_\mu = -\bar{\psi}\frac{1}{2}(1+\gamma_5)\gamma_\mu T^a\psi, \tag{8.94}$$

We can only construct a consistent gauge invariant theory if the partition function is gauge invariant. This especially means that the part

$$Z[A_\mu] = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}e^{iS(\bar{\psi},\psi,A)} \tag{8.95}$$

must be gauge invariant (since the rest of the complete partition function which involves $\int \mathcal{D}Ae^{i\int F^2}$ is explicitly gauge invariant). Let us now apply an infinitesimal *local* gauge transformation to the A_μ in (8.95): $A_\mu \to A_\mu + D_\mu\alpha$. We get

$$\frac{\delta Z[A]}{\delta\alpha(x)} = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\ \left(D^\mu j_\mu(x)\right)\ e^{iS(\bar{\psi},\psi,A)} \tag{8.96}$$

i.e.

$$\frac{\delta Z[A]}{\delta\alpha(x)} = 0\ \equiv\ \langle D^\mu j_\mu\rangle = 0. \tag{8.97}$$

However, if we instead repeat the Fujikawa analysis of the effect of an infinitesimal change of variables $\psi \to (1+i\alpha)\psi$, $\bar{\psi}\to\bar{\psi}(1-i\alpha)$ we do *not* get (8.97) due to the fact that the measure is not invariant. The same conclusion is reached by a direct calculation of the triangle diagram and we see that the situation is the same as for the abelian chiral (or axial) anomaly: the current j_μ has an anomaly relative to the classical *covariant* equation (8.94), only are the coefficients slightly different from the the ones encountered in the abelian case due to the projection operator $\frac{1}{2}(1-\gamma_5)$ and the (possible) non-abelian nature of the current j_μ:

$$\left\langle D_\mu j^a_\mu\right\rangle = \partial_\mu\left(\frac{1}{24\pi^2}\varepsilon^{\mu\nu\rho\sigma}\mathrm{tr}\,\left[T^a(A_\nu\partial_\rho A_\sigma + \frac{1}{2}A_\nu A_\rho A_\sigma)\right]\right) \tag{8.98}$$

However, contrary to the cases considered above, this anomaly is in direct conflict with the requirement of gauge invariance as expressed by (8.97) and unless the rhs of (8.98) vanishes identically there is an obstruction to the construction of the gauge theory itself! Due to the antisymmetry of $\varepsilon_{\mu\nu\rho\sigma}$ the rhs of (8.98) will vanish if

$$d^{abc} = \mathrm{tr}\,T^a(T^bT^c + T^cT^b) = 0. \tag{8.99}$$

For $SU(2)$ this condition is satisfied since $T^bT^c + T^cT^b = 0$. For $SU(N),\ N \geq 3$ it is not satisfied and we have to combine several species of chiral fermions if we want to achieve $\langle D_\mu j_\mu \rangle = 0$. Such a cancellation is possible if we note that the over-all sign of the term on the rhs of (8.98) change if the left-handed fermions are replaced by right-handed ones, i.e. $(1+\gamma_5)$ with $(1-\gamma_5)$ in the action (8.92). Note also that if we consider Dirac fermions there is never a problem since the massless Dirac fermion decomposes in a left handed and a right handed chiral fermion, and these gives precisely opposite contributions to the rhs of (8.98).

8.5 Physical consequences of the anomaly

We now discuss a number of applications of the anomaly.

8.5.1 The electromagnetic π^0 decay

One of the most famous applications of the abelian anomaly is to the decay $\pi^0 \to \gamma\gamma$. In the quark models the pions π^a, $a = 1,2,3$ are related to the divergence $\partial^\mu j^5_\mu$ of the axial vector currents

$$j^5_\mu(x)^a \sim \bar{\psi}(x)\frac{\sigma^a}{2}\gamma_5\gamma_\mu\psi(x) \tag{8.100}$$

where σ^a denote the Pauli matrices and are associated to the flavor indices. It is seen that the calculation of the decay of π^0 (index 0 is equivalent to $a = 3$) to two γ's will involve a triangle diagram with a photon at two of the vertices and the pion field (represented by the divergence $\partial^\mu j^5_\mu(x)^3$) at the third vertex. This diagram is anomalous and consequently the current j^5_μ has an anomaly. Before the discovery of the anomaly the decay posed a problem since one could argue from chiral symmetry that $\partial^\mu j^5_\mu$ was (partially) conserved[1] and the decay should be suppressed. If one takes into account the anomaly one finds agreement with experiments. In addition it should be noted that one only finds this agreement if we have three independent internal degrees of freedom assigned to each quarks. In this way we also get a beautiful verification of the assignment of three colours to each quark. (Of course the present days measurement of the Z^0 decay verifies this assignment much better).

8.5.2 Non-consevation of baryon number in the electroweak theory

Another example of an abelian anomaly, but now embedded in a non-abelian theory, is the one associated with the baryon-number current in the standard electroweak theory. Let us remind the reader about the fermion content of the standard model. We have three generations of fermions. The fermion content of the first generation consists of two left-handed $SU(2)$ doublets : A leptonic one of the electron neutrino and the left-handed component of the electron and a hadronic one consisting of the left-handed components of the up- and down-quarks:

$$L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \qquad Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \tag{8.101}$$

---

[1] The mass of the pion's lead to an explicit breaking of the chiral symmetry.

where $e_L = \frac{1}{2}(1-\gamma_5)e^-$, while the right-handed part of the electron $R = \frac{1}{2}(1+\gamma_5)e^-$. The chiral nature of the fermions in the standard model forbids explicit mass terms, also for the electron (the mass terms are acquired from the Yukawa coupling to the Higgs field by spontaneous symmetry breaking). A similar decomposition is valid for the quarks. All the right-handed components are singlets under the $SU(2)$-weak gauge transformations. The part of the lagrangian involving fermions is:

$$\begin{aligned} S_{leptons} &= -\int d^4x \left[\bar{R}\gamma_\mu(\partial_\mu - ig'B_\mu)R + \bar{L}\gamma_\mu(\partial_\mu - \frac{1}{2}ig'B_\mu + \frac{1}{2}igA^a_\mu\sigma^a)L\right] && (8.102)\\ S_{quarks} &= -\int d^4x \left[\bar{Q}_L\gamma_m(\partial_\mu + \frac{1}{2}ig'Y_LB_\mu + \frac{1}{2}igA^a_\mu\sigma^a)Q_L\right.\\ &\quad \left. + \sum_{i=1}^{2}\bar{Q}_R(i)\gamma_\mu(\partial_\mu + \frac{1}{2}g'Y_R(i)B_\mu)Q_R(i)\right]. && (8.103)\end{aligned}$$

For the two other generations, $((\nu_\mu,\mu);(s,c))$ and $((\nu_\tau,\tau);(b,t))$, we have similar lagrangians.

The fermion number F is classically a conserved quantum number in the above theory. The same is true for the baryon and the lepton numbers, B and L, and even for the three lepton numbers (L_e, L_μ, L_τ) separately. The symmetries of the lagrangian which leads to these conserved quantum numbers are the global (abelian) transformations:

$$q_j \to e^{i\alpha(B)}q_j, \qquad l_j \to e^{i\alpha(L_j)}l_j \tag{8.104}$$

in an obvious notation. As is well known the quarks have $B = 1/3$ while the leptons have $L = 1$. The Nöether currents of the symmetries (8.104) are classically conserved. Due to chiral nature of the fermions and the asymmetric coupling of the left- and right-handed particles to the weak $SU(2)$ fields, both the B and the L currents have an anomaly with respect to this gauge group. Since $SU(2)$ is non-abelian we have the situation discussed earlier: An abelian anomaly embedded in a non-abelian theory. Due to the existence of field configurations with non-trivial topology (instantons) we have to consider the symmetry (8.104) as explicitly broken and the quantum numbers B and L are not conserved[2]. This amazing situation, which seems in direct contradiction with the experimental reality, was first noticed by ‘t Hooft in 1976. He also resolved the apparent conflict with experiments. If we denote the Nöether current associated with the first symmetry in (8.104) by j^B_μ we have according to (8.80) that

$$\partial^\mu j^B_\mu = \frac{N_{gen}}{16\pi^2}\mathrm{tr}\, F\tilde{F}. \tag{8.105}$$

Integration gives

$$B(t_2) - B(t_1) = \frac{N_{gen}}{16\pi^2}\int_{t_1}^{t_2} dt \int d^3x \;\mathrm{tr}\, F\tilde{F}. \tag{8.106}$$

From this equation we see that the baryon number will change by an integer when the integral on the rhs does. At zero temperature such a change can be related to a tunneling from one classical vacuum configuration to a neighbouring one, the two connected by a large gauge transformation. This tunneling effect is precisely the one calculated in

---

[2] On the other hand it can be shown that the anomalous contributions cancel between B and L which means that $B - L$ *is* conserved also at the quantum level.

(8.90) and the process is therefore exponentially suppressed by $e^{-4\pi/\alpha_w}$ in a semiclassical approximation. Note that the coupling appears in a non-analytic way. (For the calculation of this tunneling amplitude we refer again to the chapter on classical gauge theory). α_w denotes the weak coupling constant $\alpha_w = g_w^2/4\pi \approx 1/30$ and the amplitude for baryon number violation is very small due to the smallness of the electroweak coupling constant. Not many baryons have decayed in the lifetime of the universe according to this estimate, and in this way the baryon anomaly is not in contradiction with experiments.

At this point it seems as if we can happily forget everything about the baryon anomaly. However, it was later recognized that the suppression by tunneling between different gauge vacua is only effective at zero temperature. In the early universe, when the temperature is comparable to the masses in the electroweak theory, thermal fluctuation might mediate a transition between different gauge sectors, which are related by large gauge transformations. At sufficiently high temperature this thermal diffusion will be unsuppressed. This means that the rhs of (8.106) will change rapidly and therefore also the baryon number. *We reach the conclusion that, due to the anomaly, we cannot consider the baryon number a conserved quantum number in the early universe.* This in turn implies that the number of baryons and anti-baryons was essentially equal, no matter what happens at the GUT scale, all the way down to the electroweak phase transition, and we have to explain the observed asymmetry of baryons and anti-baryons *at or after the electroweak phase transition.* This is a major unsolved problem at the moment.

8.5.3 The solution of the U(1) problem

We have no space to discuss this problem and its solution in any detail. We can only try to give the reader some hints. The starting point is the phenomenological observation from the sixties that the chiral symmetry $SU(2)_L \times SU(2)_R$ can be consider as an underlying symmetry for the strong interactions. It has to be considered spontaneously broken to $SU(2)$. The pions are the (almost) massless Goldstone bosons associated with th broken symmetry. However, after the hadron physics has been associated with the specific QCD lagrangian the chiral symmetry group is naturally $U(2) \times U(2)$. The reason is that if we write down the *massless* QCD lagrangian of N_f flavours then *any* unitary transformation between the flavour components of the quarks leave the lagrangian invariant, and if it is massless the chiral variants of the unitary transformation leave it invariant too, i.e. the symmetry group is $U(N_f) \times U(N_f)$, the transformations being:

$$\psi \to e^{i\alpha^a T^a}\psi, \qquad \psi \to e^{i\alpha^a T^a \gamma_5}\psi. \tag{8.107}$$

In (8.107) the matrices T^a are generators of the Lie algebra of $U(N_f)$ and act on the flavor index of ψ, which in addition has a colour index for the colour symmetry group $SU(N_c)$, $N_c = 3$.

In the real world the quarks are not massless and the chiral symmetry is not exact, but the u and the d quarks have relatively small masses and we expect the chiral symmetry of two flavours to be a reasonable approximate underlying symmetry of the strong interactions. In this way we have almost explained the underlying $SU(2) \times SU(2)$ symmetry of the sixties. The only problem is that that we have arrived at $U(2) \times U(2)$ instead of $SU(2) \times SU(2)$. The difference is essentially a group $U(1) \times U(1)$. One of the $U(1)$'s is no problem: it is the hadron number, which is conserved. It was just too trivial to be included in the discussion in the sixties. The other $U(1)$ is associated with the abelian chiral transformations $e^{i\alpha\gamma_5}$ and is *not* observed in nature. One obvious solution is to assume that this symmetry is spontaneously broken like the chiral $SU(2)$ part of $U(2)$. This would however mean that we should observe a fourth low mass Goldstone boson, in addition to the three pions already associated with the spontaneously broken chiral $SU(2)$. The natural candidate, the η, has however a mass m_η much larger than the pions. The $U(1)$ problem was the following: What has happened to the fourth Goldstone boson?

The answer (due to 't Hooft) is that the classically conserved current associated with the abelian chiral transformation has an anomaly. In addition this anomaly is embedded in a non-abelian gauge group ($SU(3)$-color) and due to field configurations with non-trivial topology we have to consider the symmetry as explicitly broken. This means that there is no need for a light fourth particle.

8.5.4 Consistency relations in model building

As discussed above, theories with chiral fermions coupled to gauge fields have potentially a problem due to the possibility of a clash between the requirement of gauge invariance and the anomaly. Let us consider as an example the standard model where all fermions are chiral. The $SU(3)$ part causes no problems since right-handed and left-handed components contribute oppositely. In the electroweak group $SU(2)$ is a safe group according to the remarks following (8.99). Therefore potential problems come from the hypercharge $U(1)$ group. There are two dangerous triangle diagrams, one diagram with one $U(1)$-field and two $SU(2)$-fields and one triangle diagram with three $U(1)$-fields[3] If we first consider the two $SU(2)$ and one $U(1)$ gauge fields only the doublets contribute, since only they couple to $SU(2)$ gauge fields, and from (8.98) we get

$$\sum_{doublets} Y_L \mathrm{tr}\,\sigma^a\sigma^b = 0, \quad \text{i.e.} \quad \sum_{doublets} Y_L = 0. \tag{8.108}$$

The triangle diagram with three $U(1)$ fields yields, again from (8.98), the condition:

$$\sum_{left-handed\ parts} Y_L^3 - \sum_{right-handed\ parts} Y_R^3 = 0, \tag{8.109}$$

where the summation over left-handed parts means that each of the two components of the doublets should be counted and in addition other degrees of freedom like colour should be counted too. The assignment for the standard model is

$$\begin{array}{llll} Y_L = 1/3, & Y_R(1) = 4/3, & Y_R(2) = -2/3 & \text{for quarks} \\ Y_L = -1, & Y_R(1) = 0, & Y_R(2) = -2 & \text{for leptons} \end{array} \tag{8.110}$$

One can check that (8.108) and (8.109) are satisfied provided quarks exists with three degrees of freedom. These are provided by the color quantum numbers. While (8.108) is trivially true, we get for (8.109):

$$\left[3\times 2\times\left(\frac{1}{3}\right)^3 + 2\times(-1)^3\right] - \left[3\times\left(\frac{4}{3}\right)^3 + 3\times\left(-\frac{2}{3}\right)^3 + (-2)^3\right] = 0. \tag{8.111}$$

---

[3] They actually reduce to the same condition due to the relations between the Y_L's and the Y_R's in the standard model.

Chapter 9

Lattice Field Theory

9.1 Field theories as critical classical spin systems

In the following we will always assume that we are working in d-dimensional euclidean space. The lagrangian for a scalar field will be

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial_\mu \phi)^2 + V(\phi) \tag{9.1}$$

and the partition function (the generating functional for Greens functions) can be written:

$$\begin{aligned} Z(J) &= \int \mathcal{D}\phi \;\; e^{S(\phi,J)} \\ &= \int \mathcal{D}\phi \;\; e^{-\int d^dx (\mathcal{L}(\phi)+\phi J)}. \end{aligned} \tag{9.2}$$

The measure "$\mathcal{D}\phi$" is the Feynman path integral measure and is ill defined as it stands in (9.2):

$$\mathcal{D}\phi = \prod_{x \in R^d} d\phi(x) \quad . \tag{9.3}$$

To make sense of the measure one can discretize euclidean spacetime R^d by imposing a hyper-cubic lattice structure:

$$\begin{aligned} x &\rightarrow x_n \equiv a\; n_\mu \hat{e}_\mu \\ \phi(x) &\rightarrow \phi_n \equiv \phi(x_n) \\ \mathcal{D}\phi &\rightarrow \prod_n d\phi_n \\ \partial_\mu \phi &\rightarrow \frac{1}{a}\left(\phi(x_n + \hat{e}_\mu) - \phi(x_n)\right) \equiv \frac{1}{a}(\phi_{n+\mu} - \phi_n) \\ S(\phi, J) &\equiv \int d^dx \; \frac{1}{2}\left((\partial_\mu \phi)^2 + V(\phi) + J\phi\right) \\ &\rightarrow \sum_n a^d \left\{ \frac{1}{a^2} \sum_\mu \frac{1}{2}(\phi_{n+\mu} - \phi_n)^2 + V(\phi_n) + J_n \phi_n \right\} \quad . \end{aligned} \tag{9.4}$$

In these formulae $\hat{e}_\mu$ denotes d orthonormal vectors and "a" the lattice spacing. If one takes a finite volume V of spacetime, the measure $\mathcal{D}\phi$ is converted into a finite dimensional integral $\prod_n^V d\phi_n$ and one can study the limit $V \rightarrow \infty$.

This regularization of the path integral breaks euclidean invariance. But we have obtained a strict control over the short distance singularities of the theory, since we have an ultraviolet cut off $\Lambda = \pi/a$. Furthermore it turns out that internal symmetries, even local ones, can usually be preserved in a natural way. This is especially important if we want to address gauge theories. I the following example it is shown in detail how the lattice regularization modifies the continuum propagator $(p^2+m^2)^{-1}$ and at the same time provides an ultraviolet cut off $\Lambda = \pi/a$.

Example 1

Let us first record the formulas for Fourier transformation on an infinite lattice:

$$\begin{aligned} \phi(x_n) &= \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \frac{d^dp}{(2\pi)^d}\, e^{ipx_n} \tilde{\phi}(p) && (9.5)\\ \tilde{\phi}(p) &= \sum_n a^d\, e^{-ipx_n} \phi(x_n) && (9.6)\end{aligned}$$

where the restriction in the integration range comes from the lattice structure of space-time: Since $x_n = n\ a$ the function $\exp(ipx_n)$ is a periodic function of p_μ under $p_\mu \to p_\mu + 2\pi/a$. The same is therefore true for $\tilde{\phi}(p)$ as define above.

Recall that we got the continuum free propagator from the gaussian part of the action by Fourier transformation:

$$\int d^dx\phi(x)[-\partial_\mu^2 + m^2]\phi(x) = \int \frac{d^dp}{(2\pi)^d}\tilde{\phi}(-p)[p_\mu^2 + m^2]\tilde{\phi}(p) \tag{9.7}$$

From (9.4) it follows that the lattice equivalence to (9.7) is:

$$\sum_{n,m} a^d\ \phi_n(\frac{-\Delta_{n,m}}{a^2} + m^2\delta_{n,m})\phi_m = \int_{\frac{-\pi}{a}}^{\frac{\pi}{a}} \frac{d^dp}{(2\pi)^d}\ \tilde{\phi}(-p)\left[\sum_\mu \frac{2-2\cos ap_\mu}{a^2} + m^2)\right]\tilde{\phi}(p) \tag{9.8}$$

Here the difference operator $\Delta_{n,m}$ on the lhs is the discrete laplacian and is given by

$$\Delta_{n,m} = \sum_\mu (\delta_{n+\hat{\mu},m} + \delta_{n-\hat{\mu},m} - 2\delta_{n,m}) \tag{9.9}$$

and the rhs follows by inserting (9.5)-(9.6) and using

$$\sum_n a^d\ e^{ipx_n} = (2\pi)^d\delta^{(d)}(p) \quad \text{and} \quad \int_{\frac{-\pi}{a}}^{\frac{\pi}{a}} \frac{d^dp}{(2\pi)^d} e^{i(x_n-x_m)p} = \frac{1}{a^d}\delta_{n,m} \tag{9.10}$$

We finally get the following modification of the continuum propagator

$$\frac{1}{p^2+m^2} \to \frac{a^2}{4\sin^2(ap/2) + (am)^2} \tag{9.11}$$

while the momentum p is cut off at π/a.

Let us consider a ϕ^4 theory in d dimensions:

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\phi^4 \quad . \tag{9.12}$$

By scaling the fields and sources and coupling constants:

$$\lambda = g^2 a^{d-4}, \qquad \phi' = g a^{\frac{d}{2}-1}\phi, \qquad J' = g a^{\frac{d}{2}+1} J \tag{9.13}$$

the partition function (9.2) may be written:

$$\begin{aligned} Z(J', g) &= \prod_n^V (g^{-1} a^{-(d/2-1)}) \cdot \int \prod_n^V d\phi'_n \cdot \exp(-\frac{1}{g^2} S(\phi', J')) \\ S(\phi', J') &= \sum_n \left[\frac{1}{2} \sum_\mu (\phi'_{n+\mu} - \phi'_n)^2 + \frac{1}{2} m^2 a^2 \phi'^2_n + \frac{1}{4!} \phi'^4_n + J'_n \phi'_n \right] . \end{aligned} \tag{9.14}$$

The free energy (the generating functional for connected Green functions) of the system, $F(J)$ is defined by:

$$Z(J) = e^{-F(J)} \tag{9.15}$$

and the constant in front of the integral in (9.14) only contributes to $F(J)$ with an additive constant proportional to the volume, but with no reference to the dynamics. It can be dropped.

We can view (9.14) as the partition function of a classical spin system. Indeed, an effective, classical theory of spin-spin coupling in a ferromagnet would have the following hamiltonian:

$$H(s, h) = -\sum_{n,m} v_{n,m} s_n \cdot s_m + h \cdot \sum_n s_n \tag{9.16}$$

and partition function

$$Z(h, \beta) = \int \prod_n ds_n \; \rho(s_n) e^{-\beta H(s,h)} \quad . \tag{9.17}$$

In (9.16) $v_{n,m}$ is the coupling of spins at sites n and m in the lattice. If we assume the lattice is hyper-cubic and we only have the nearest neighbour interactions, we can write

$$-\sum_{n,m} v_{n.m} s_n \cdot s_m = K \cdot \sum_n \left(\sum_\mu (s_{n+\mu} - s_n)^2 - 2d s_n^2 \right) \quad . \tag{9.18}$$

In (9.17) $\beta = 1/kT$ and $\rho(s_n)$ is a weight factor describing the local, microscopic properties of the spin. Since we assume the spins are classical, we have to define the concept of spin. For instance we could assume that the spin is a vector, constraint by $s^2 = l(l+1)$. A model with this constraint is called the *Heisenberg model.* A toy model which has played an important role in theoretical considerations (mainly because it can be solved explicitly in two dimensions) is the *Ising model* where s is a scalar, which can take the values ± 1. A choice like

$$\rho(s_n) \propto \exp(-(\kappa s_n^2 + \lambda s_n^4)) \tag{9.19}$$

gives a convenient effective description. An appropriate choice of of κ and λ (typically $\kappa < 0$) allows us to approximate the model mentioned with any desired precision. The partition function may now be written as:

$$\begin{aligned} Z(K, \mu, \lambda, h) &= \int \prod_n ds_n e^{-H(s; K, \mu, \lambda, h)} \\ H(s; K, \mu, \lambda, h) &= \sum_n \left[K(\beta) \sum_\mu (s_{n+\mu} - s_n)^2 + \mu(\beta) s_n^2 + \lambda(\beta) s_n^4 + h s_n \right] . \end{aligned} \tag{9.20}$$

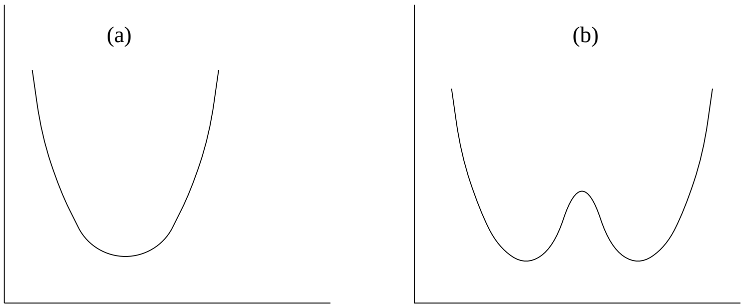


Figure 9.1: Potential $\mu(\beta)s^2 + \lambda(\beta)s^4$ for (a): $\mu(\beta) > 0$ and (b): $\mu(\beta) < 0$

The Ginzburg-Landau theory of ferromagnetic transitions assumes that $K(\beta), \mu(\beta)$ and $\lambda(\beta)$ are smooth functions of the temperature since they depend only on local properties. The ferromagnetic transition occurs at $\mu(\beta_c) = 0$. The value of T, for which $\mu(\beta_c) = 0$, is called the critical temperature T_c. Minimizing the *effective* hamiltonian in (9.20) we get a ground state where all $s_n = 0$ if $\mu > 0$ (see fig.9.1a):

$$< s > \equiv \frac{1}{V} \sum_n s_n = 0 \quad \text{for} \quad \mu(\beta) > 0 \tag{9.21}$$

while the ground state for $\mu(\beta) < 0$ corresponds to all s_n aligned with $s_n = \sqrt{\frac{-\mu}{2\lambda}}$ (see fig.9.1b):

$$< s > \equiv \frac{1}{V} \sum_n s_n = \sqrt{\frac{-\mu}{2\lambda}} \quad \text{for} \quad \mu(\beta) < 0 \quad . \tag{9.22}$$

As we assume $\mu(\beta)$ is a smooth function near β_c we might write:

$$\mu(\beta) \approx c_0(\beta - \beta_c) \quad \text{for} \quad \beta \approx \beta_c \tag{9.23}$$

and we get:

$$< s > \sim \sqrt{\beta - \beta_c} \quad \text{for} \quad \beta > \beta_c \quad . \tag{9.24}$$

This shows the typical non-analytic behaviour at a phase transition.

If we compare our partition function for the scalar field (9.14) with the one for our spin system (9.20) we see that the continuum limit of our regularized field theory ($a \to 0$) corresponds to approaching the critical point of the ferromagnetic transition ($\mu(\beta) \to 0$) because of the identification

$$m^2 a^2 \sim \mu(\beta) \quad . \tag{9.25}$$

It is therefore not surprising that all the machinery and intuition available from the theory of critical phenomena can be taken over to field theory. Let us briefly summarize the notations used: near the critical point the following observables are of interest (among others):

$$
\begin{aligned}
&< s > (h,\beta) \equiv \frac{1}{V}\sum_n < s_n > && \text{(magnetization)}\\
&\chi(h,\beta) \equiv \frac{\partial < s >}{\partial h} = \frac{1}{V}\sum_{n,m} < (s_n - < s >)(s_m - < s >) > && \text{(susceptibility)}\\
&< (s_n - < s >)(s_m - < s >) > \sim \; \exp(-|x_n - x_m|/\xi(h,\beta)), \quad \xi(h,\beta) \ll |x_n - x_m| && \text{(correlation length)}\\
&< (s_n - < s >)(s_m - < s >) > \sim \; |x_n - x_m|^{-(d-2+\eta)}, \quad a \ll |x_n - x_m| \ll \xi(h,\beta) && \text{(anomalous dimension } \eta)
\end{aligned}
\tag{9.26}
$$

The behaviour of these quantities is obviously governed by the spin fluctuations and the correlation length ξ is of crucial importance. The hypothesis that *all singular behaviour* near the phase transition is due to the divergence of the correlation length ξ is called a *scaling hypothesis.* In the gaussian approximation where we only include quadratic fluctuation around the minimum (9.21) or (9.22) in our functional integral it is easily seen that

$$\xi(h,\beta) \sim \frac{1}{\sqrt{|\mu(h,\beta)|}} \tag{9.27}$$

and $\xi(h,\beta)$ diverges near the critical point.

The singular behaviour leads to the definition of critical exponents, characterizing it:

$$
\begin{aligned}
\chi(\beta) &\sim |\beta - \beta_c|^{-\gamma}\\
\xi(\beta) &\sim |\beta - \beta_c|^{-\nu} \quad .
\end{aligned}
\tag{9.28}
$$

and by using the assumption that the behaviour of the correlation function is governed by only one divergent parameter near the critical point it is possible to show that

$$\gamma = \nu(2-\eta) \qquad \text{(Fischer's scaling relation)} \quad . \tag{9.29}$$

We will prove this relation later.

It should be stressed that these exponents are *not* just mathematical definitions. One can measure γ, ν and η in materials like Fe, Ni, $YFeO_3$, Gd, etc. using neutron diffraction and other experimental techniques. The remarkable fact is that they are identical even if the materials mentioned of course have different $\rho(s_n)$ and $\sum_{n,m} V_{n,m} \cdot s_n s_m$ and vastly different T_c. Only few sets of distinctly different values of these critical exponents are observed and each set satisfies scaling relations like (9.29) very well. The *universality* of the critical exponents extends even further. The same critical exponents can be observed in ferromagnetic transitions and certain liquid-vapor transitions. Since these critical exponents are connected with a divergent correlation length ξ, we see that *long range phenomena near the critical points show universality.*

For field theory this translates into the statement that the details of how we regularize the theory at lattice distances are to *a large extent irrelevant* for the continuum limit.

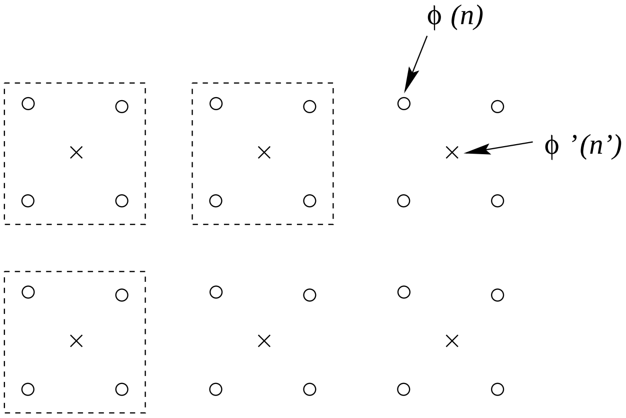


Figure 9.2: A Kadanoff blocking for the scale $s = 2$ on a cubic lattice

Only few parameters are *relevant* and they determine the continuum limit. *Different possibilities of taking the continuum limit of the regularized theory are labelled by different critical exponents and are said to belong to different universality classes.*

For this reason it becomes of major importance in field theory to understand which universality classes can exist, and the basic tool for understanding the whole concept of universality is *the renormalization group equations* (RGE). Today it has been combined with *Monte Carlo techniques* used in large scale computer simulations of these theories. On the configurations generated by the computer it is possible to carry out the renormalization group transformations and one talks about the *Monte Carlo renormalization group approach.* It is fair to say that MC-techniques are the only general tools available, if we want to explore non-perturbative aspects of field theory for dimensions $d > 2$.

9.2 Renormalization group and critical phenomena

9.2.1 Kadanoff blocking

The renormalization group approach to critical phenomena is the simplest way to understand universality. In the following we will drop the distinction between spins and fields and our toy model hamiltonian will be

$$H = \sum_n \left[\sum_\mu \frac{1}{2}(\phi_{n+\mu} - \phi_n)^2 + \mu\phi_n^2 + \lambda\phi_n^4 \right] \quad . \tag{9.30}$$

In a socalled *Kadanoff transformation* we divide our original lattice in blocks of size s^d, where s is an integer, and define an average field in the block $B_s(n')$ labelled n' (see fig.9.2):

$$\phi'_{n'} = s^{-d} \sum_{n \in B_s(n')} \phi_n \quad . \tag{9.31}$$

The distribution of $\phi'_{n'}$ can be determined from the one of ϕ_n :

$$e^{-H'[\phi'_{n'}]} = \int \prod_n d\phi_n e^{-H[\phi]} \prod_{n'} \delta(\phi'_{n'} - s^{-d} \sum_{n \in B_s(n')} \phi_n) \quad . \tag{9.32}$$

We end by scaling the blocks back to the original size:

$$\begin{aligned} x_s &\equiv x/s \\ \phi_s &\equiv s^\alpha \phi' \quad ; \quad \alpha = (d-2+\eta)/2 \\ H_s(\phi_s(x_s)) &\equiv H'(\phi'(x')) \end{aligned} \tag{9.33}$$

If we measure the correlation length ξ in lattice units it has been decreased by s :

$$\xi_s = \xi/s \ . \tag{9.34}$$

The need for the factor $s^{-\alpha}$ on ϕ may not be evident, but recall that near the critical point the correlation length ξ diverges, the theory has "almost" massless excitations, and the correlation function will have a power law fall off (9.26):

$$< \phi_n \phi_0 > \sim \frac{1}{n^{d-2+\eta}} \quad , \quad 1 \ll n \ll \xi \quad . \tag{9.35}$$

Consider now the block transformation (9.31):

$$\begin{aligned} < \phi'_{n'} \phi'_{0'} > &= < s^{-d} \sum_{n \in B_s(n')} \phi_n \cdot s^{-d} \sum_{m \in B_s(0')} \phi_m > \\ &\approx < \phi_{sn'} \phi_0 > \end{aligned} \tag{9.36}$$

since all the s^{2d} correlation functions are at essentially the same distance if n' is very large. From (9.35) we see that the short distance[1] properties of correlation functions are only left unchanged by blocking if we at the same time scale ϕ' by s^α. The need for such a rescaling can be related to the wave function renormalization encountered earlier when we discussed renormalization and we shall later see that the anomalous dimension η in (9.33) or (9.35) is nothing but the anomalous scaling dimension introduced when we discussed the renormalization group earlier.

The form of $H_s(\phi_s)$ is not identical to the one of $H(\phi)$ in (9.30) which was the starting point. Other terms like

$$(\phi_{n+\mu} - \phi_n)^2 \phi_n^2 \ , \quad (\phi_{n+\mu} - 2\phi_n + \phi_{n+\mu})^2 \tag{9.37}$$

will be generated. As we want to repeat the Kadanoff blocking it is therefore natural to start with a completely general action:

$$H[\phi] = \sum_\alpha K_\alpha S_\alpha(\phi) \quad . \tag{9.38}$$

where S_α $(\alpha = 1, 2, \ldots)$ are different actions which should conform with the original symmetries of the action and lattice. The couplings K_α now take values in a multi-dimensional (in principle infinite-dimensional) coupling constant space, and successive blockings can be viewed as a mapping of this space onto itself, called *the renormalization group transformation* $T_{RG}(s)$.

$$T_{RG}(s) \ : \ \{K_\alpha\} \to \{(T_{RG}(s)K)_\alpha\} \quad . \tag{9.39}$$

---

[1] Short distance still means large compared to the lattice spacing a

The argument s refers to the block size and the "group" structure enters since the block size $s_1 \cdot s_2$ can be obtained by successive blockings: $T_{RG}(s_2)T_{RG}(s_1) = T_{RG}(s_1 s_2)$. Strictly speaking we only have a *semi-group*, since there will not necessary be an inverse $T_{RG}^{-1}(s)$. One can understand this from the point of view of physics, since we are reducing the degrees of freedom by blocking.

The operator $T_{RG}(s)$ might have certain fixed points $K^\star_\alpha$. For such a fixed point we have

$$(T_{RG}(s)K^\star)_\alpha = K^\star_\alpha \quad \forall \alpha \tag{9.40}$$

and it follows from (9.34) that the correlation length must be infinite (or zero) for this choice of coupling constants. The fluctuations extend over all scales of the lattice and the system is critical.

To each fixed point $K^\star_\alpha$ we can associate a *critical surface*, namely the points K_α which are attracted by the fixed point $K^\star_\alpha$:

$$(T^n_{RG}(s)K)_\alpha \to K^\star_\alpha \quad \text{for} \quad n \to \infty. \tag{9.41}$$

The important point of the blocking is that we perform a coarse graining of the system. By taking the average over blocks we ignore short distance details, but keep long range phenomena intact.

Every point on the critical surface corresponding to $K^\star_\alpha$ has infinite ξ (since blocking reduces correlation length and the fixed point to which the point converges also has $\xi = \infty$) and the long distance physics for any point on the critical surface is therefore expected to be identical to the long distance physics determined by $K^\star_\alpha$.

*The fundamental hypothesis linking RGE to critical phenomena is that the couplings of the material in question (*Fe, Ni, etc.*):*

$$K_\alpha(\beta) = (K(\beta), \mu(\beta), \lambda(\beta), \ldots) \tag{9.42}$$

belong to a critical surface when $\beta = \beta_c \;\; (T = T_c)$.

If we now assume (as will be justified in the next sections) that (1): Critical surfaces are expected to be large subspaces of the total infinite dimensional coupling constant space $\{K_\alpha\}$ and that (2): The critical exponents are determined *by* T_{RG} *near the fixed point*, we can understand universality: different materials "i" at their critical points $\beta_c^{(i)}$ can be represented by vastly different $K_\alpha^{(i)}$ in the coupling constant space, but *they will belong to the same critical surface* (i.e. have the same fixed point $K^\star_\alpha$) and consequently they have the same long distance physics.

9.2.2 Expansions near a fixed point

Suppose that a point K_α is near a fixed point $K^\star_\alpha$. When T_{RG} acts on coupling constants which are close to the fixed point, we can approximate it by a linear operator, since the changes induced by the operator are only small:

$$\begin{aligned} K_\alpha &= K^\star_\alpha + \delta K_\alpha \\ (T_{RG}(s)K)_\alpha &= K^\star_\alpha + \sum_{\alpha'} T_{\alpha\alpha'} \delta K_{\alpha'} + \mathcal{O}((\delta K)^2) \quad . \end{aligned} \tag{9.43}$$

If we expand δK_α in eigenvectors of the linear operator $T_{\alpha\alpha'}$:

$$\begin{aligned} \delta K_\alpha &= \sum_a h_a v_{a\alpha} \\ \sum_{\alpha'} T_{\alpha\alpha'} v_{a\alpha'} &= \lambda_a v_{a\alpha} \end{aligned} \tag{9.44}$$

the action (9.2.1) can be written:

$$\begin{aligned} H[\phi] &= H^\star[\phi] + \sum_a h_a v^a[\phi] \\ H^\star[\phi] &= \sum_\alpha K^\star_\alpha S_\alpha[\phi] \\ v_a[\phi] &= \sum_\alpha v_{a\alpha} S_\alpha[\phi] \quad . \end{aligned} \tag{9.45}$$

Repeated application of T_{RG} will give:

$$H[\phi] \to H^\star[\phi] + \sum_a \lambda_a^n h_a v^a[\phi] \quad . \tag{9.46}$$

Interactions with $\lambda^a < 1$ are suppressed after a few T_{RG} steps. They are called *irrelevant.* Interactions with $\lambda^a > 1$ are called *relevant* and they will eventually take us away from the critical point provided the decomposition of δK_α contains these components. Finally the interactions with $\lambda^a = 1$ are called *marginal.* Whether they will contribute or not can only be decided by considering higher order corrections to the linearized T_{RG} transformation given by $T_{\alpha\beta}$.

We realize that the critical surface in the neighbourhood of $K^\star_\alpha$ is spanned by the irrelevant operators $v^a(\phi)$.

If we are close to the critical surface, but not exactly on the surface, the coefficients h_a for the relevant operators must be small. If we block we will first move towards the fixed point $K^\star_\alpha$ since the irrelevant operators dominate, but eventually when n, the number of blockings, is large enough we will have $\lambda_a^n h_a > 1$ for the relevant operators and we will be taken away along the direction of the largest relevant operator. This relevant direction is called a *renormalization group trajectory.* The flow near $K^\star_\alpha$ is illustrated in fig.9.3.

If there are n relevant operators at a given fixed point we will denote it by $FP^{(n)}$ and it will require the *tuning* of n parameters to reach the critical surface. If there is only one relevant operator we can reach the critical point by changing any of the coupling constants. In the laboratory the tuning is performed by changing, for instance, the temperature. This will create a flow of $K_\alpha \equiv (K(\beta), \mu(\beta), \lambda(\beta), \ldots)$ which eventually will cross the critical surface if the system has a ferromagnetic transition.

In the context of model calculations on a computer one can check the above mentioned picture. Since we have a detailed knowledge of the configurations which we generate by MC-simulations, we can perform the blocking and actually follow the flow in the coupling constant space $\{K_\alpha\}$.

With respect to actual materials like Fe, Ni, Gd, etc. the coarse graining implemented by T_{RG} is a purely mental process which allows us to *understand* universality. We can, however, give the following qualitative description of the use of T_{RG}. Suppose we look at a sample of a ferromagnet through a microscope and that our eyes can see spin variations down to a certain size. Then T_{RG} represents the operation of decreasing the magnification

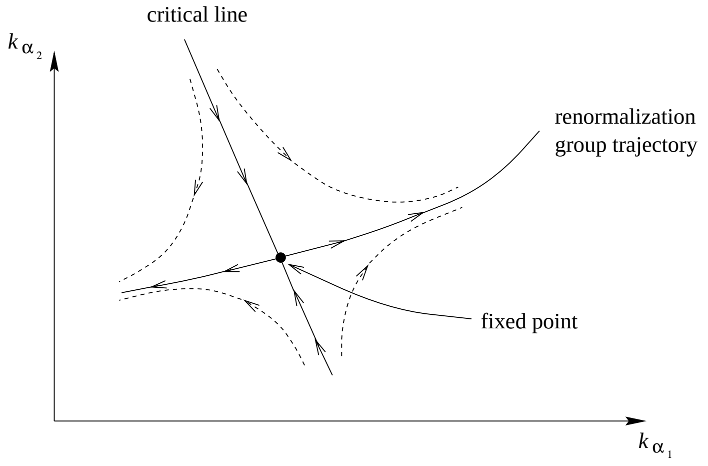


Figure 9.3: RGE-flow in the neighbourhood of a fixed point. The critical surface is usually multi-dimensional

factor of the microscope by the factor s which entered into the definition of block B_s, i.e. the sample observed appears to shrink by a factor s. Assume now that we have brought one of the materials to its critical point. The hypothesis that this critical point lies on a critical surface characterized by a fixed point $\{K^*_\alpha\}$ translates to the statement that if we decrease the magnification by a sufficient amount, we shall not see any change by a further decrease. In addition the *long range spin fluctuations* which we observe at this stage are the same for all materials associated with the same critical surface. The underlying short distance structure depending on a specific lattice and microscopic details of the material has faded away, and only universal long range phenomena survive.

9.2.3 Critical exponents near a fixed point

It was essential for some of the arguments given above in favour of universality that the critical exponents were determined by the fixed points. We will now show that this is the case.

Let us for simplicity assume that there is only one relevant direction. We will denote the correlation function $< \phi(x)\phi(0) >$ by $G(x)$. x always refers to a lattice point. If we are close to the critical surface and to the fixed point itself, we can repeat the discussion in the last subsection and write the linearized renormalization group transformation $T_{RG}(s)$:

$$\begin{aligned} K_\alpha &= K^\star_\alpha + \sum_{a_i} v_{a_i\alpha} h_{a_i}(\beta) \\ (T_{RG}(s)K)_\alpha &= K^\star_\alpha + \lambda_{a_1} h_{a_1}(\beta) v_{a_1\alpha} + \sum_{i=2}^{\infty} \lambda_{a_i} h_{a_i}(\beta) v_{a_i\alpha} + \mathcal{O}(h^2) \qquad (9.47) \\ h_{a_1}(\beta) &= (\beta - \beta_c)\, h^0_{a_1} + \mathcal{O}((\beta-\beta_c)^2). \end{aligned}$$

The notation is as follows: v_{a_i} are eigenvectors for the the linearized operator $T_{RG}(s)$ of blocking with block size s, as defined by (9.43) and (9.44). λ_{a_i} denote the corresponding eigenvectors. The first one is defined as corresponding to the relevant direction and has $\lambda_{a_1} > 1$, while the rest correspond, by assumption, to irrelevant directions and have $\lambda_{a_i} < 1,\ i = 2, 3, ...$ Recall that the critical surface is defined as being spanned by the irrelevant directions, i.e. characterized by $h_{a_1}(\beta) = 0$, a requirement which fixes $\beta = \beta_c$.

Since we, again by assumption, are close to the critical surface the function $h_{a_1}(\beta)$ must be small and we have the expansion of $h_{a_1}(\beta)$ given above.

After n repeated blockings the iteration of (9.47) looks as follows

$$(T^n_{RG}(s)K)_\alpha = K^*_\alpha + \lambda^n_{a_1} h_{a_1}(\beta) v_{a_1\alpha} + \mathcal{O}(\lambda^n_{a_i}) \tag{9.48}$$

Although the coefficient $h_{a_1}(\beta)$ was small compared to the other coefficients $h_{a_i}(\beta)$, repeated application of the renormalization group transformation $T_{RG}(s)$ will ensure that only the relevant operator dominates since only $\lambda_{a_1} > 1$. After sufficiently many RG-steps we will be on the RG-trajectory, and continued RG-steps would move us away from the fixed point. We now counteract this in a well organized way by adjusting $\beta \to \beta_c$. Since the correlation length changes with a factor s for each blocking this will relate the change in β to the (change in) correlation length. First we note that since $\lambda_{a_1} > 1$ we can write

$$\lambda_{a_1} = s^{1/\nu}, \quad \nu > 0. \tag{9.49}$$

The (semi-)group property of $T_{RG}(s)$ ensures that it is a sensible definition, i.e. that ν is independent of s. Next we choose a sequence of β_n converging to β_c such that

$$1 = h_{a_1}(\beta_n)\ s^{n/\nu} \tag{9.50}$$

or (by (9.47) and (9.49)) when we are close to β_c:

$$s^n = \frac{1}{|(\beta_n - \beta_c)\ h^0_{a_1}|^\nu} \tag{9.51}$$

This choice of β_n is made such that that repeated application of the renormalization group transformations keep us at a well defined distance from the fixed point. In fact we have from (9.50) and (9.47)

$$(T^n_{RG}(s)K)_\alpha \to K^*_\alpha + v_{a_1\alpha} \quad \text{for} \quad n \to \infty \tag{9.52}$$

From the definition of blocking we have (when x and $x/s^n \gg a$):

$$G(x;\{K_\alpha\}) = s^{-2\alpha} G(x/s;\{(T_{RG}(s)K)_\alpha)\}) = s^{-2\alpha n} G(x/s^n;\{((T^n_{RG}(s)K)_\alpha\}) \tag{9.53}$$

where the scale factor α is the one associated with the rescaling of the fields after blocking (recall (9.33)-(9.36)). Since $\{((T^n_{RG}(s)K)_\alpha\}$ for sufficiently large n has no n dependence, according to (9.52), the correlation function $G(x;\{K_\alpha\})$ will for such choices of β_n only depend on x/s^n for large x. . If we compare it to the generic form of the two-point function (see (9.26))

$$G(x;\{K_\alpha\}) \sim \exp(-|x|/\xi(\beta))$$

we are led to the conclusion that $\xi(\beta_n) \sim s^n$, or stated differently: *for sufficiently large* n *the condition (9.50) amounts to increasing the correlation length* $\xi(\beta)$ *by a factor* s *when changing* β_n *to* β_{n+1}. This implies

$$\xi(\beta) \propto \frac{1}{|\beta - \beta_c|^\nu} \tag{9.54}$$

and the ν *introduced by (9.49) can be given the interpretation as the critical exponent* ν *defined by (9.28). Further we see that the critical exponent* ν *is related to the (largest) relevant eigenvalue by (9.49):* $\lambda_{a_1} = s^{1/\nu}$

(9.53) can now be written as

$$G(x;\{K_\alpha(\beta)\})|_{\beta\to\beta_c} = \xi^{-2\alpha}(\beta)\; G(x/\xi(\beta);\{K^\star_\alpha + v_{a_1\alpha}\}) \tag{9.55}$$

or by a Fourier transformation, using $\alpha = (d-2+\eta)/2$,

$$G(p;\{K_\alpha(\beta)\})|_{\beta\to\beta_c} = \xi^{2-\eta}(\beta)\; G(\xi(\beta)p;\{K^\star_\alpha + v_{a_1\alpha}\}). \tag{9.56}$$

Also η is determined by the fixed point. In fact, as already mentioned, $s^{-\alpha}$ is the unique scale factor for the field ϕ which leaves the behaviour of the $G(x, K_\alpha)$ invariant *at* the critical surface. Stated differently we can say only one, or very few, choices of η will result in a fixed point for our chosen RG-transformations.

Finally the critical exponent γ is also determined from (9.55). From the definition (9.26) we get the susceptibility by integrating $G(x)$ over x (the magnetization m is assumed zero):

$$\begin{aligned}\chi(\beta) &\equiv \int d^dxG(x;\{K_\alpha(\beta)\}) \\ &= \xi^{-2\alpha}\int d^dxG(x/\xi(\beta);\{K^\star_\alpha + v_{a_1\alpha}\}) \\ &= \xi^{-2\alpha+d}\cdot\int d^dyG(y\;;\;\{K^\star_\alpha + v_{a_1\alpha}\})\end{aligned} \tag{9.57}$$

or, since the last parenthesis has no β dependence:

$$(\beta-\beta_c)^{-\gamma} = \xi(\beta)^{2-\eta} = (\beta-\beta_c)^{-\nu(2-\eta)} \tag{9.58}$$

which is Fischer's scaling relation $\gamma = \nu(2-\eta)$.

A result like this is typical for renormalization group- or scaling arguments. It is possible to derive relations between the critical exponents, but only in a few simple models can the exponents themselves be determined.

9.3 The continuum limit

9.3.1 Definition of the continuum limit

In this section we will discuss in more detail the approach to a critical point belonging to a critical surface, and how this approach relates to the usual renormalization known from field theory.

We use the formalism developed in the last section. As we have argued, the long distance physics of all points on the critical surface is identical and determined by the fixed point. In this connection it should be mentioned that *the position of the critical point is dependent on the renormalization group procedure used.* We have considered here the socalled Kadanoff blocking, since it has a direct physical interpretation as taking the average over fields, and in this way coarse graining the system. However, one is free to choose other procedures which reduce in a systematic way the degrees of freedom associated with short distance fluctuations. The critical surface itself (i.e its position in the infinite dimensional coupling constant space) is independent of the T_{RG} procedure used, as is the long distance physics associated with the surface, but essentially any point on the critical surface can serve as a fixed point if the procedure is chosen appropriately.

The central formulae derived in the last section were (9.55) and (9.56). However, all distances involved were expressed in units of the lattice spacing a, since our starting point, the Hamiltonian (9.30), had no reference to a. Clearly, (9.55) and (9.56) cannot be true if $\eta \neq 0$ and we introduce the length scale a. For this reason it is convenient to distinguish between correlation length measured in lattice units, i.e. number of lattice spacings, and correlation length in physical space. We introduce

$$\xi_{ph} = \xi_l a, \qquad x_{ph} = x_l a, \qquad p_{ph} = p_l / a \tag{9.59}$$

It is now possible to define a continuum limit without any reference to the lattice in the following way: Until now we have assumed a fixed lattice, which we could even associate with some solid state spin system. Let us now return to the starting point and consider the lattice as a cut off device. This means we are free to change the lattice spacing a. *We now require that the physical correlation length ξ_{ph} is kept fixed while we approach the fixed point the way described in the last section.* Since the lattice correlation length ξ_l is increased by a factor s for each step $\beta_n \to \beta_{n+1}$ in the approach to the fixed point, the requirement of a fixed ξ_{ph} imposes a reduction of the lattice spacing $a \to a/s$. In this way the fine tuning $\beta \to \beta_c$ is turned into a scaling $a \to 0$ and we can replace β by a. If we define the mass parameter of the theory as the inverse of the physical correlation length

$$m_{ph} \equiv 1/\xi_{ph} \tag{9.60}$$

we can *define the continuum correlation function* as follows

$$\boxed{G_{cont}(p_{ph}; m_{ph}) \equiv a^{-\eta} G(p_l; \{K(a)\})|_{a \to 0}} \tag{9.61}$$

This relation is nothing but the relation between the renormalized correlation function $G_{cont}(\cdot)$, which has no reference to the cut off $1/a$ and the "bare" correlation function $G(\cdot)$ which is still defined by the lattice parameters. The divergent factor $a^{-\eta}$ is the wave function renormalization, and it has the following origin: If we wanted to introduce a correlation length ξ_{ph} instead of ξ_l in (9.56) we would make a mistake if $\eta \neq 0$, for simple dimensional reasons. In the process of blocking, the parameters with dimensions of length, associated with all the irrelevant coupling constants, *must* sneak in and ensure the correct dimensions on both sides of (9.55) and (9.56). These irrelevant couplings are not important for the long distance physics (as their name tells us) and they are associated with short distance physics, in fact physics at the scale a. The only way they will show up in the long distance physics, is as an overall factor $a^{-\eta}$ in scaling relations like (9.56). The factor $a^{-\eta}$ is precisely what is needed in order to be able to write (9.56) in terms of physical correlation length, since the dimension of $G(p_{ph})$ when the lattice spacing is explicitly introduced *must* be -2:

$$G(p_{ph}; \{K_\alpha(\beta)\})|_{\beta \to \beta_c} = a^{\eta}\, \xi_{ph}^{2-\eta}\, G(\xi_{ph} p_{ph}; \{K_\alpha^\star + v_{a_1 \alpha}\}). \tag{9.62}$$

This equation shows explicitly that the limit (9.61) is well defined and independent of a.

The above construction for the two point correlator can be generalized in a straight forward manner to the n-point correlation function.

Let us emphasize once more the important points in the above construction:

(1): At the fixed point the theory is scale invariant (massless).

(2): We had one *relevant* coupling constant, say (K_0). By choosing some values of $K_1, K_2, \ldots$ we got to the *critical surface* $K_\alpha = \{0, K_1, K_2, \ldots\}$ by *fine tuning* of the relevant coupling. The physical correlation length $\xi_{ph} \equiv 1/m_{ph}$ is related to the correlation length ξ measured in lattice- or cut off units a by

$$\frac{1}{m_{ph}} \equiv \xi_{ph} = \xi_l a \ . \tag{9.63}$$

and by the requirement that m_{ph} was unchanged during a blocking we could relate the renormalization group transformation to a change in cut off a and fix the fine tuning of the relevant parameter: K_0. The important point is that the *massive* continuum theory, i.e. the theory with a finite correlation length, is defined not *at* the critical point but by the fine tuned *approach* of the relevant coupling constant to the critical surface. The same would be true if we had n relevant couplings. By fixing the physical value of these as in (9.63) the requirement that (long distance) physics is invariant under the renormalization group transformation T_{RG}, when we are near a critical surface, would fix the fine tuning of the relevant parameters in terms of the cut off "$1/a$". *Such relations, describing the change in the bare coupling constant under a change of cut off while keeping physics constant, can be viewed as the origin of renormalization in quantum field theory.*

(3): At this point it might be confusing why we in general moved out in the infinite dimensional coupling constant space when we did the blocking in the last section. When we renormalize field theory we usually adjust only a few coupling constants. The renormalization group transformations in the continuum did not lead us to an infinite dimensional coupling constant space. The reason for this difference is that the blocking procedure is much more precise than is needed for describing the long distance behaviour. A blocking as defined here *exactly* reproduces *all* predictions for the variables which are not integrated over by the blocking. The expense is that one has to enter into an infinite dimensional coupling constant space. We could, and that is often done in computer Monte Carlo simulations, approach the continuum limit by just changing the few relevant couplings in the simplest discretized version of the continuum action. We would then have no control over the change in correlation length when changing the couplings, but we could of course either try to measure the correlation length or try to calculate it. It is usually not easy to calculate the change in correlation length, but it might be feasible near a critical point.

9.3.2 The gaussian fixed point

The most important fixed point is the gaussian fixed point. It is the point in coupling constant space where all coupling constants, except the one in front of the pure gaussian term, are tuned to zero. The virtue of this fixed point is that we know it exists, and that one can perform the ordinary perturbation theory around it.

It is instructive to consider the purely gaussian case. If we perform the blocking, the action will after a few steps contain next to nearest neighbour interactions, which will result in higher derivative terms etc.. This illustrates one basic problem with the RGE's on the lattice. It is almost impossible to do any analytic calculations. Their importance is to be found at the conceptual level[2], since they provide us with new insight in the process of renormalization and link the critical phenomena of solid state physics to relativistic field theory. Rather than carrying out the analysis on the lattice (which *can* be done in detail in the gaussian case), it is convenient for the purpose of illustration to return to the continuum formalism, but with the insight provided by the lattice RGE's. By doing this we appeal to the remarks of the last section, where it was stated that the long distance physics associated with a critical surface was independent of the specific blocking procedure used, as long as it provided a systematic reduction of the degrees of freedom associated with short distance physics. We have seen that the lattice introduces a momentum cut off

$$\Lambda = \pi/a \tag{9.64}$$

If we use the same cut off notation in the continuum, a reduction of Λ corresponds to larger lattice spacing and therefore, in a not very precise way, to a coarse graining. In this way we get a kind of blocking by integrating out the momenta between Λ and Λ/s, where $s > 1$ is a scale factor. In the following example we perform this kind of blocking in a generalized gaussian model where higher derivative terms are included. This model is sufficiently general to allow an illustration of many of the concepts introduced in the last section.

Example 2: The generalized gaussian model

We consider the generalized gaussian action:

$$\begin{aligned} H[\tilde{\phi}] &= \frac{1}{2}\int^{\Lambda} d^d p \tilde{\phi}(p) D(p) \tilde{\phi}(-p) \\ D(p) &= \tilde{K}_0 + \tilde{K}_1 \sum_\mu p_\mu^2 + \tilde{K}_2 (\sum_\mu p_\mu^2)^2 + \tilde{K}_3 \sum_\mu p_\mu^4 + \cdots \quad . \end{aligned} \tag{9.65}$$

Note that we have even included terms which break rotational invariance, in order to stay close to the lattice version.

By changing to dimensionless variables

$$q = a \cdot p, \qquad \phi(q) = a^{-(d/2+1)} \tilde{\phi}(p),$$

$$K_0 = a^2 \tilde{K}_0, \qquad K_1 = \tilde{K}_1, \qquad K_2 = a^{-2} \tilde{K}_2, \cdots \tag{9.66}$$

we get

$$\begin{aligned} H[\phi] &= \frac{1}{2}\int^{\pi} d^d q \phi(q) D(q) \phi(-q) \\ D(q) &= K_0 + K_1 \sum_\mu q_\mu^2 + K_2 (\sum_\mu q_\mu^2)^2 + \cdots \end{aligned} \tag{9.67}$$

Let the RG procedure in momentum space be the integral over the high frequency part: $\pi/s < q \leq \pi$. In the gaussian case this is a triviality since different momenta do not

---

[2] It is not true any longer that they only have importance at the conceptual level. As already mentioned the fast computers of today allow us to carry out the blocking in a very concrete way and test with success the ideas involved.

couple. We can write the path integral as follows

$$\begin{aligned} Z &= \int \mathcal{D}\phi(q) e^{-\frac{1}{2}\int^{\pi} dq\ \phi(q)D(q)\phi(-q)} \\ &= \left[\int \mathcal{D}\phi(q) e^{-\frac{1}{2}\int_{\pi/s}^{\pi} dq\ \phi(q)D(q)\phi(-q)}\right] \left[\int \mathcal{D}\phi(q) e^{-\frac{1}{2}\int^{\pi/s} dq\ \phi(q)D(q)\phi(-q)}\right] \\ &= C(s,\{K\}) \int \mathcal{D}\phi(q) e^{-\frac{1}{2}\int^{\pi/s} dq\ \phi(q)D(q)\phi(-q)} \end{aligned}$$

The factor $C(s,\{K\})$ does not have any reference to the field components $\phi(q)$, $|q| < \pi/s$ and will factor out in any correlation function of such field components. We therefore ignore it in the following. If we follow the conventions of the last section and introduce the "blocked" field

$$\phi_s(sq) = s^{-\frac{d+2-\eta}{2}} \phi(q) \tag{9.68}$$

we can write (again ignoring a s dependent normalization factor)

$$Z = \int \mathcal{D}\phi_s(q) e^{-H[\phi_s]} \tag{9.69}$$

where (again in the notation of the last section)

$$\begin{aligned} H_s(\phi_s) &= \frac{1}{2}\int^{\pi} d^d q\ \phi_s(q) D_s(q) \phi_s(-q) \\ D_s(q) &= K_0 s^{2-\eta} + K_1 s^{-\eta} \sum_{\mu} q_\mu^2 + K_2 s^{-2-\eta} (\sum_{\mu} q_\mu^2)^2 + \cdots \end{aligned} \tag{9.70}$$

From the last equation we read off the RG transformation $T_{RG}(s)$:

$$T_{RG}(s) : \{K_\alpha\} \to \{s^{2-\eta} K_0\ , s^{-\eta} K_1\ , s^{-2-\eta} K_2, \cdots\} \tag{9.71}$$

and we have a fixed point at:

$$K^\star_\alpha = \{0, K_1, 0, 0 \cdots\} \tag{9.72}$$

provided $\eta = 0$. This point is called *the gaussian fixed point*. The value of K_1 is arbitrary. If we fix it to one, the action (9.65) -(9.67) defines the massless free field in the continuum in the limit where $a \to 0$.

We observe the general pattern already advocated: The relevant directions are few, in fact there is only one corresponding to K_0, there is one marginal coupling, K_1, and the rest of the couplings $K_2, K_3, \ldots$ are irrelevant. This is a simple illustration of the statement that the critical surface, which in this case is given by $K_\alpha = \{0, 1, K_2, K_3, \ldots\}$ is large, in fact of finite co-dimension. Note also that the *massive* free field theory is obtained by a *fine tuned approach* to the critical surface, by the procedure of the last section. Each "blocking" will increase K_0 by a factor s^2 as is seen from (9.71). We are on a RG trajectory and are taken away from the fixed point. We have to compensate for this, as described in the last section, by adjusting the relevant coupling K_0 closer to the fixed point: $K_0 \to K_0/s^2$. In this way we stay at a fixed distance from the critical surface during the blocking, but since the physical mass $m_{ph} = 1/\xi_{ph}$ is related to K_0 by $K_0 = m_{ph}^2 a^2$, as is seen from (9.65) and (9.66), such a change in K_0 implies a change $a \to a/s$ in the original cut off *a provided the physical mass m_{ph} or the physical correlation length ξ_{ph} is kept fixed.* In this way we get the continuum massive theory when the "lattice spacing" a is taken to zero.

A final point worth noticing is that terms associated with the breaking of rotational invariance ($\sum_\mu p_\mu^4$ etc.) are all irrelevant terms. This is how euclidean invariance is restored when we approach a fixed point.

The message from the above example is that the coupling constant K_α in front of the a generalized gaussian term $K_\alpha S_\alpha$ given by

$$K_\alpha \int d^d x (\partial^k_\mu \phi(x))^2$$

will scale as follows under a blocking of size s:

$$K_\alpha \to s^{-d_{S_\alpha}} K_\alpha, \qquad d_{S_\alpha} = -2 + 2k$$

This d_{S_α} is nothing but the engineering dimension of S_α, and it is clear from the example that it is obtained simply by the replacement $x \to x/s, \phi_s(x/s) = s^{(d-2)/2}\phi(x)$ required by blocking. When we are very close to the gaussian fixed point the same is true for possible interaction terms like $\lambda \int d^d x \phi^4(x)$. To a first approximation, when the coupling constants are very small, one simply replaces $x \to x/s, \phi_s(x/s) = s^{(d-2)/2}\phi(x)$. This is however only true to the very lowest order in λ, as is clear from the functional integral. The action is no longer diagonal in the momenta, and different scales mix. However, to the very lowest order we can ignore this mixing, and we find again that an action like

$$K_n \int d^d x \phi^n(x)$$

will result in the following scaling of K_n under a blocking of size s:

$$K_n \to s^{-d_n} K_n, \qquad d_n = (d-2)n/2 - d$$

Since d_n again is nothing but the engineering dimension of the action and the eigenvalue of the linearized RG-operator $T_{RG}(s)$ in this direction will be

$$\lambda_n = s^{-d_n}$$

we see that *the relevant interactions near a gaussian fixed point are precisely the interactions which make the theory super renormalizable, while the marginal interactions are the ones which make the theory renormalizable. All irrelevant interactions spoil renormalizability.* In the last chapter we classified the renormalizable scalar theories. They were finite in numbers in all dimensions larger than two, and we therefore have a proof that the critical surface of the gaussian fixed point is of finite co-dimension. A similar proof is not known for other critical points, but is believed to be true, and is essential for the whole idea of universality.

9.3.3 Triviality versus asymptotic freedom

In $d = 4$ the operator $\int d^4 x \phi^4$ has engineering dimension zero. An operator of dimension zero will be a marginal operator with respect to the gaussian fixed point, corresponding to eigenvalue $\lambda = s^0 = 1$ under a blocking of size s. One has to go beyond the gaussian approximation in order to discover whether it will become relevant or irrelevant. The corresponding couplings are called *ultraviolet asymptotically free* (or just asymptotically free) and *infrared asymptotically free* (or sometimes non-asymptotically free), respectively. The canonical examples are non-abelian $SU(N)$ gauge theories and scalar ϕ^4 theories.

For infrared free couplings one cannot have a renormalized coupling defined at the gaussian fixed point. (At least not the way discussed here). Let us, for the purpose of illustration, define the renormalized coupling as the value of the bare coupling after applying n blockings such that

$$s^n a = \text{fixed physical distance.} \tag{9.73}$$

When the cut off $1/a \to \infty$ it is clear by definition[3] that the renormalized coupling is smaller than the bare coupling, since the corresponding term in the action was *irrelevant* with respect to the gaussian fixed point. At the gaussian fixed point the bare coupling is taken to be zero and the renormalized coupling will be even closer to zero. Therefore theories with only infrared free couplings cannot define a *non-trivial* continuum field theory at the gaussian fixed point[4]. They could, however, have other fixed points where a non-trivial theory could be defined. A search for such fixed points is therefore of outmost importance in these theories (like ϕ^4, ordinary QED, etc.). At the moment there are no convincing indications that such points can be found in $d = 4$.

For ultraviolet asymptotically free theories the gaussian fixed point is much more interesting since the renormalized coupling is larger than the bare coupling. One therefore has a chance that even if the bare coupling (by definition) is taken to zero when approaching the fixed point, the renormalized coupling might remain finite and in this way define a *non-trivial* interacting theory at the gaussian fixed point. The important function which controls the approach to the continuum limit is the β-function.

9.3.4 The β-function

Let us for simplicity consider a theory with only one coupling constant g^2. For the regularized version on the lattice the change of this coupling constant will move us along a one-parameter curve in the multi-parameter space created by blocking. $g^2 \to 0$ will bring us to the critical surface associated with the gaussian fixed point.

When we are close to the critical surface, the correlation length is large and we can find a change Δg^2 in g^2 such that it increases by a factor s

$$\begin{aligned} g'^2 &= g^2 - \Delta g^2 \\ \xi(g'^2) &= s\xi(g^2) \quad . \end{aligned} \tag{9.74}$$

This means that the *long distance physics* will be the same for g'^2 and g^2 provided we identify

$$a(g'^2) = \frac{1}{s} a(g^2) \quad . \tag{9.75}$$

It is worth emphasizing that repeated RG-transformations (with the block size s) will result in a picture shown in Fig.9.4. When the number of blockings n is sufficiently large, the coupling constant flow starting from g^2 will move along the RG-trajectory and will coincide with the $(n+1)^{th}$ RG-step starting from g'^2: it is *only* the long distance physics which is identical for the choices $g^2, a(g^2)$ and $g'^2, a(g'^2)/s$.

---

[3] Note that n in this way becomes a function of a

[4] There might be other ways the define a non-trivial interacting theory at a gaussian fixed point. Since theories like ϕ^4 (or even *QED*) have a non-trivial perturbative loop expansion it is somewhat strange that this expansion should be irrelevant. However, at present nobody knows how to make sense outside perturbation theory of theories which are not asymptotically free.

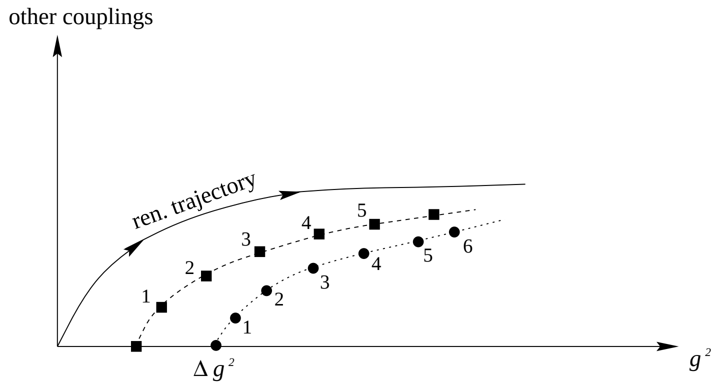


Figure 9.4: Successive blockings starting from g^2 (dots) and $g'^2 = g^2 - \Delta g^2$ (crosses) such that $a(g'^2) = a(g^2)/s$

The equations (9.74)-(9.75) define the relation between g and a which leaves continuum physics invariant when $a \to 0$. *This relation is named the β-function*:

$$\beta(g) = -a \, \frac{d}{da} \, g(a) \tag{9.76}$$

This definition of the β-function is not identical to the one given in the last chapter where we kept the bare coupling constants and the cut-off fixed and varied the renormalized coupling constants. One could as well have chosen the dual point of view and have kept constant the renormalized masses and coupling constants and the subtraction point μ. In this way the independence of the renormalized Green functions of a cut-off Λ (which on the lattice is $1/a$ and in dimensional regularization is associated with the less intuitive parameter ε) would translate into a renormalization group equation for the "bare" Green functions and the β function involved would be given by an expression like (9.76). In a perturbative expansion of this β-function it can be shown that the first two coefficients are the same as for the β-function defined in the last chapter and it can further be shown that the existence and nature of fixed points are independent of the definition.

The nice thing about the gaussian fixed point is that we can calculate $\beta(g)$ for small g^2 by ordinary perturbation theory (we will do that later):

$$\beta(g) = - \, b_0 g^3 - b_1 g^5 + \cdots \tag{9.77}$$

For an asymptotically free theory $b_0 > 0$, since this implies that $g(a)$ is decreasing when a is decreasing.

The *scaling region* is the region where g^2 is so small (a is so small) that within a given, required precision there will be no cut off dependence for physical observables. In the scaling limit any physical quantity with dimension of mass behaves in a definite way as a function of g^2. Since m_{ph} by definitions (9.63)-(9.75) is independent of a we have:

$$a \, \frac{d}{da} \, m_{ph} = 0 \tag{9.78}$$

For dimensional reasons we must have

$$m_{ph} = \frac{1}{a} f(g). \tag{9.79}$$

By inserting in (9.78) we can determine the function f:

$$f(g) + f'(g)\beta(g) = 0 \quad \text{or} \quad \ln(f(g)) = \tilde{c} + \int^{g} \frac{dg'}{\beta(g')} \tag{9.80}$$

and therefore we get

$$m_{ph} = \frac{c}{a} e^{-\int^{g} dg'/\beta(g')} \tag{9.81}$$

The non-perturbative aspect in this formula is the constant c which cannot be calculated within perturbation theory. Similar formulas are valid for any other physical observables in the theory, but the ratios between the different constants c are not accessible in perturbation theory.

9.4 Summary

Renormalization and the renormalization group equations were first discovered in the context of relativistic quantum field theory. For renormalizable theories the infinities in the perturbation theory could be absorbed into a redefinition of the bare coupling constants, while keeping the physical masses and coupling constants fixed. The freedom of choosing the subtraction point where these physical observables were defined, allowed us by some scaling arguments to derive relations between the Green functions at different scales. These relations were called the renormalization group equations.

In this chapter we have seen that the process of renormalization and the concept of the renormalization group has a very concrete interpretation when the lattice is used as a mean of regularization of the functional integral. The use of the lattice allowed us to make contact with statistical mechanics, and the theory of critical phenomena. The field theory could be viewed as a generalized classical spin system. Whenever this spin system became critical by a second order phase transition, i.e. a transition where the correlation length ξ would diverge, it was possible to define a continuum limit of a relativistic field theory. All physical masses in the continuum theory would be expressed in terms of the correlation length (in lattice units) times the lattice spacing. In this way the continuum theory could be defined as the limiting process of approaching the critical point, which means increasing the correlation length, while at the same time diminishing the lattice spacing such that the physical length scale of typical fluctuations stays constant.

The possibility of having a renormalization group equation also becomes more transparent in the statistical interpretation. Since the continuum theory is defined as a limiting process where the correlation length in lattice units diverges it should be allowed to average over "blocks" of spin. No matter how large the block size we choose, we should eventually get the same answer when we approach the critical point. This freedom of choosing the "block"-size of spins allowed us to relate Green functions at different scales and led directly to the renormalization group equations.

9.5 Lattice gauge theories

While the experiments at the large accelerators at CERN, SLAC, Fermilab etc have provided us with impressive experimental verification of the perturbative aspects of the standard model, one of the most interesting sectors of the standard model, the low energy sector of QCD, has remained unaccessible for rigorous theoretical results. Superficially the situation is not so bad: We have been led to a unique theory, the $SU(3)$ non-abelian gauge theory with quarks and gluons. The theory is an asymptotically free theory of fermions (quarks) and massless vector particles (gluons) and, according to our discussion in the chapter on renormalizability of field theories, this is probably a healthy sign, since the general belief at the moment is that only such theories might have a chance to exist as fundamental interacting theories in four dimensional space-time, without a cut-off. If we ignore the current quark masses, which these particles are believed to acquire though spontaneous symmetry breaking in the standard model, the only masses which can appear in this theory are the ones generated dynamically. This means that it should be possible to calculate them as functions of a fixed length scale which we will denote ξ_L, or ,alternatively, a fixed momentum scale $\Lambda_L = \xi_L^{-1}$. This is true for all excitations in the theory. If it is really the correct theory we should be able the predict all mass ratios which appear in the theory, i.e all mass ratios between the hundreds of baryons and mesons which have been observed. In addition we should be able to explain why we have never seen any free quarks and gluons. We have already given a heuristic explanation of confinement in terms of asymptotic freedom in the sense that the asymptotic freedom of the non-abelian theories means that the effective coupling constant goes to zero at short distances, while it grows at large distances, thereby "confining" quarks which carry a color charge. We should however be able to do better than just make these qualitative statements. In fact the real test that QCD is the correct theory of the strong interactions is that it can also predict correctly the low energy excitations which we observe. In this respect we have not been too successful yet.

The lack of success, using standard continuum techniques, in explaining the non-perturbative aspects of QCD has prompted a brute force approach, where one tries directly to calculate the functional integral of this theory by use of the so-called Monte Carlo techniques. The main idea is to formulate a discretized version of the non-abelian gauge theory, suited for computer, and then use the modern fast computers to calculate the non-perturbative aspects of the strong interactions. This approach involves a number of steps

(1): A non-perturbative formulation of the theory. Here a lattice formulation is natural since we have already seen that a lattice formulation of euclidean field theory provides us with a non-perturbative definition of at least scalar field theories. Viewed as statistical systems the possible fixed points in the coupling constant space serve as candidates for interesting continuum limits. The relevant theoretical framework to use in this context is the renormalization group approach.

(2): A natural concept of local gauge invariance which fits the non-perturbative lattice formulation. To find such a formulation appears at first sight to be non-trivial, since the whole concept of "local" gauge invariance seems intimately linked to continuum concepts. Furthermore one would expect that the concept of gauge invariance is of paramount importance. Recall that gauge invariance (in the disguise of BRS-invariance) was crucial for the proof that non-abelian theories could be considered as renormalizable field theories. It is therefore a happy circumstance that there exists a formulation which incorporates in a natural way gauge fields on a space-time lattice. This formulation, which is due to K. Wilson (1974), is also remarkable in that the gauge degree of freedom remains as an exact "local" symmetry at any step. (We will shortly discuss what is meant by "local" on a lattice).

(3): A sensible way of performing numerically the functional integral over gauge field configurations. Since we essentially deal with statistical systems on large lattices this is a problem which is shared with our colleges in statistical mechanics, who want to study critical phenomena of spin systems. There exist a number of ways to deal with integration over such "multi-dimensional" systems. Here we will only describe the most simple minded approach, the so-called Monte Carlo simulations.

We have already described the general philosophy associated with (1). Let us therefore first concentrate on (2).

9.5.1 Gauge invariance on the lattice

Consider a general gauge group G. We will assume we have a unitary, finite dimensional representation of the group. In the real world we have in mind $G = SU(3)$ if we want to model the strong interactions. If we consider the gauge theory of ordinary electromagnetism we will choose $G = U(1)$, while we sometimes for the purpose of simplified illustrations of non-abelian groups will consider $G = SU(2)$. The *total* gauge group of a (continuum) theory with *local* gauge invariance can formally be written as

$$G_{inv} = \prod_{x \in R^d} G_x \tag{9.82}$$

where we have a copy G_x of G associated with each space-time point. (We assume as usual that we have performed a rotation from minkowskian space-time to euclidean space-time). In the chapter on classical gauge theories we saw that the natural gauge invariant observables of the pure gauge theory was the so-called path ordered exponentials:

$$U_C = \mathrm{tr}\, P \exp i \oint_C dx_\mu A_\mu(x) \tag{9.83}$$

where C denotes a closed path, P stand for the path ordering and A_μ is an element in the Lie algebra of G: $A_\mu = A_\mu^a T^a$, T^a being the generators of the Lie algebra. Let us remind the reader that path ordering means the following: For a given curve $C(t) : t \to x_\mu(t)$, $t \in [0,1]$ from $x(0)$ to $x(1)$ $U_C(t)$ is the solution of the differential equation

$$\frac{dU_C(t)}{dt} = i \frac{dx_\mu(t)}{dt} A_\mu(x(t)) U_C(t) \tag{9.84}$$

The solution is written as

$$U_C(t) = P \exp i \left(\int_0^t dt' \frac{dx_\mu(t')}{dt'} A_\mu(x(t')) \right) \tag{9.85}$$

and path ordering refers to the fact that the explicit solution of the matrix equation (9.84) can be written as

$$\begin{aligned} U_C(t) \;\; = \;\; & 1 + i \int_0^t dt_1 \, \dot{x}_{\mu_1}(t_1) A_{\mu_1}(x(t_1)) + \\ & i^2 \int_0^t dt_1 \int_0^{t_1} dt_2 \, \dot{x}_{\mu_1}(t_1) A_{\mu_1}(x(t_1) \, \dot{x}_{\mu_2}(t_2) A_{\mu_2}(x(t_2)) + \cdots \end{aligned} \tag{9.86}$$

where $\dot{x}$ means dx/dt. U_C satisfies the following composition rule:

$$U_C = U_{C_n} U_{C_{n-1}} \cdots U_{C_1} \tag{9.87}$$

if the curve C consists of the parts $C_1, ..., C_n$. In the limit where each part C_i is infinitesimal, (dx_i), we have the following representation

$$U_C(t) = \lim_{n\to\infty} \prod_{k=0}^{n} e^{idx_k A(x_k)} \tag{9.88}$$

where $x_i \in dx_i$, with an abuse of notation. Under a local gauge transformation

$$A(x) \to V(x)A(x)V^{-1}(x) - i\partial_\mu V(x)V^{-1}(x) \tag{9.89}$$

U_C has the following transformation properties

$$U_C(t) \to V(x(t))U_C(t)V^{-1}(x(0)). \tag{9.90}$$

Especially we see that (9.83) is gauge invariant. Let us finally mention a useful formula: Assume we have a closed planar curve C with associated area tensor $a_{\mu\nu}$. In case of an abelian gauge field we have, by Stokes law:

$$\exp i \oint_C dx_\mu A_\mu = \exp \frac{i}{2} F_{\mu\nu} a_{\mu\nu}. \tag{9.91}$$

A similar formula is not valid in the non-abelian case if we use the path ordered exponential, but for curves C of infinitesimal area $da_{\mu\nu}$ we have:

$$P \exp i \oint_C dx_\mu A_\mu = \exp \frac{i}{2} \left(F_{\mu\nu} da_{\mu\nu} + \mathcal{O}(da^2) \right). \tag{9.92}$$

from which we get

$$\text{Re tr} \exp i \oint_C dx_\mu A_\mu = \text{tr}\, I - \frac{1}{8} \text{tr}\, (F_{\mu\nu} da_{\mu\nu})^2 + \mathcal{O}(da^3) \tag{9.93}$$

A path on the lattice is a connected piecewise linear path along links from one lattice site to another. We can therefore take over the continuum formula (9.87) if we associate an element of the gauge group with each (oriented) link (ij) connecting two neighbouring lattice point i and$j$ (the path from j to i):

$$U_{ij} \in G \qquad \text{for link} \quad (ij) \tag{9.94}$$

Again referring to the continuum definition of path ordering we make the assignment:

$$U_{ji} = U_{ij}^{-1}. \tag{9.95}$$

The path ordered continuum integral had the transformation properties (9.90) under a local gauge transformation $V(x)$. We can now introduce the concept of “local” gauge transformations on the lattice by associating the gauge group G to each lattice site i.

If we for simplicity consider a hypercubical lattice $a \cdot Z^d$, where a refers to the lattice spacing, the total gauge group will be the lattice analog of (9.82):

$$G_{total} = \prod_{i \in a \cdot Z^d} G_i. \tag{9.96}$$

A gauge field configuration is an assignment U_{ij} of group elements to each link on the lattice. A gauge transformation is an assignment V_i of group elements to each lattice site i on the lattice and the transformation of the gauge field configuration U_{ij} to another one given by:

$$U_{ij} \to U_{ij}^{(V)} = V_i U_{ij} V_j^{-1} \tag{9.97}$$

By this definition a path ordered integral on the lattice:

$$U_C = U_{(i_n, i_{n-1})} \cdots U_{(i_3, i_2)} U_{(i_2, i_1)} \tag{9.98}$$

where C is a connected path of $n - 1$ links on the lattice from site i_1 to site i_n, has the same gauge transformation properties as in the continuum. This means that the traces of closed loops are gauge invariant observables. (In fact one can prove that they constitute a complete set of gauge invariant observables if there are no additional matter fields).

From the continuum formulas for the path ordered exponentials we have a candidate for the lattice action by (9.93). To see this let us assume that we have a continuum gauge field configuration $A_\mu(x)$ and that we have embedded our hypercubical lattice in R^d such that the lattice sites are denoted $x_\mu(i)$. To each link (ij) connecting $x_\mu(j)$ to $x_\mu(i)$ we can associate

$$U_{ij} = P \exp i \int_0^1 dx_\mu(t) A_\mu(x(t)) = \exp i(a A_\mu(x) + \mathcal{O}(a^2)) \tag{9.99}$$

where $x_\mu(t) = t x_\mu(i) + (1-t) x_\mu(j)$ and $x_\mu = (x_\mu(i) + x_\mu(j))/2$. Let us take the smallest non-trivial loop we can get on the lattice, called a plaquette, which consists of the four links which makes up an elementary square on the hypercubical lattice. For such a plaquette $p = ijkl$ with area a^2, a being the lattice spacing, we have from (9.93)

$$\mathrm{tr}\, I - Re\ \mathrm{tr}\, U_p = \frac{1}{2} a^4 \mathrm{tr}\, F^2_{\mu(p)\nu(p)} + \mathcal{O}(a^6). \tag{9.100}$$

In this formula $\mu(p), \nu(p)$ denote the unit vectors in the hyper-plane containing the plaquette p, x a point in the plaquette (to the given order in a it does not matter which one). This means that

$$S_w[U] = \beta_G \left(1 - \frac{1}{N_G} \sum_p \mathrm{Re}\, \mathrm{tr}\, U_p\right) \to \frac{1}{g_0^2} \int d^d x \ \left(\frac{1}{4} \mathrm{tr}\, F^2_{\mu\nu} + \mathcal{O}(a^2)\right) \tag{9.101}$$

for $a \to 0$, provided

$$\beta_G a^{4-d} = \frac{2N_G}{g_0^2}. \tag{9.102}$$

In these formulas $N_G = \mathrm{tr}\, I$ and g_0 denotes the bare coupling constant of the non-abelian gauge theory.

The action (9.101) is called the Wilson action. It is clear that it in no way is unique. Many other expressions in terms of closed loops of links will have the same limit (9.101) if we start out with the identification (9.99) between lattice link variables and continuum variables. However,(9.101) appears to be the simplest expression and we will use it in the further studies. Note that on the lattice we seem to have no need for the gauge fields A_μ themselves. We can work entirely with group variables. From this point of view it is natural to formulate the integration in the functional integral as an integration over group variables, rather than over the gauge field variables $A_\mu(x)$. The obvious measure dU to use is the (up to a normalization) unique measure which is invariant under left and right translations on the group manifold: $U \to U_0 U$ and $U \to U U_0$. It is called the Haar measure. It further has the property that if U is close to the identity, i.e. $U = \exp iaA$ then

$$dU = \prod_{i=1}^{N_G} dA^a (1 + \mathcal{O}(a)) \tag{9.103}$$

and we would for smooth configurations, where we can make the identification (9.99), not only recover the continuum action by (9.101)-(9.102), but also (at least formally) the usual continuum integration in terms of the gauge fields $A_\mu(x)$. We can therefore define our lattice gauge theory by the following partition function:

$$\boxed{Z(\beta_G) = \int \prod_l dU_l \; e^{-S_w[U]}} \tag{9.104}$$

where the integration is over all links (ij) of the lattice. Since the Haar measure is invariant under left and right group translations it is invariant under the "local" gauge transformations on the lattice, (precisely as the (formal) continuum path integral measure $\mathcal{D}A^a_\mu(x)$ is invariant under local gauge transformation), and we conclude that the partition function $Z(\beta_G)$ is gauge invariant.

We have managed to define in a sensible way the concept of a gauge theory on the lattice by the partition function $Z(\beta_G)$, given by (9.104). It depends on a single coupling constant β_G. We saw above that for a given continuum configuration, "projected" to the lattice, the lattice action converged to the correct continuum expression if the lattice spacing went to zero. We are now interested in the much more subtle question, namely how we recover a continuum limit (and hopefully a non-trivial interacting field theory) by varying β_G. The naive continuum limit is $\beta_G \to \infty$. In this limit each plaquette has only small fluctuations around its maximum value $\mathrm{tr}\, U_p = 1$. This limit can be achieved if each variable U_{ij} is a (lattice version of a) pure gauge configuration:

$$U_{ij} = V_i V_j^{-1}. \tag{9.105}$$

For small fluctuations around such a minimum configuration we can perform a parametrization of the form (9.99):

$$U_{ij} = V_i \exp(iaA_\mu(x)) V_j^{-1} \tag{9.106}$$

were x is identified with the midpoint of the link (ij), and in this way formally repeat the arguments given above for a fixed continuum configuration. Such arguments will however not tell us anything beyond the usual perturbation theory. In order to extract non-perturbative information about the theory we have to follow the general strategy of finding fixed points for critical statistical systems as a function of the coupling constants (here β_G) and taking the continuum limit by approaching such a fixed point. Let us repeat these arguments in a way tuned to our particular problem: We choose a physical length ξ, computed from gauge invariant correlation functions by some method (e.g. by computer simulations, see later) and keep it fixed while the lattice spacing $a \to 0$. This is achieved by tuning the coupling constant (β_G or g_0) to its critical value (β_G^c) in such a way that the renormalization group is satisfied:

$$a\frac{d\xi}{da} \equiv \left(a\frac{\partial}{\partial a} - \beta(g_0)\frac{\partial}{\partial g_0}\right)\xi(a, g_0) = 0 \tag{9.107}$$

and with the β-function (not to be confused with β_G!) defined by

$$\beta(g_0) = -a\frac{\partial g_0}{\partial a}. \tag{9.108}$$

The solution to (9.107) is

$$\xi(a, g_0) = a \ \exp \int_0^{g_0} \frac{dg}{\beta(g)}. \tag{9.109}$$

We see that a zero in the β-function corresponds to a divergent correlation length ξ in terms of the lattice spacing. The fixed point which has our interest is $\beta_G \to \infty$, i..e $g_0 \to 0$. One could in principle perform a calculation of the $\beta(g_0)$-function directly on the lattice in the limit $g_0 \to 0$, since the lattice provide us with an ultraviolet cut-off. However, as mentioned in an earlier chapter, the first two coefficients of the $\beta(g_0)$-function are independent of the particular regularization used, and it is easier to use standard continuum perturbation theory and dimensional regularization (which, like the lattice regularization, respects gauge invariance). The result is (as mentioned earlier) in four dimensional space-time that $g_0 = 0$ is a critical (gaussian) fixed point with a triple zero in g_0 :

$$\beta(g) = -\beta_0 g^3 - \beta_1 g^5 + \mathcal{O}(g^7) \tag{9.110}$$

where the two universal coefficients for the gauge group $G = SU(n)$ are given by

$$\beta_0 = \frac{11}{3}\frac{n}{16\pi^2}, \qquad \beta_1 = \frac{34}{3}\frac{n^2}{16\pi^2}. \tag{9.111}$$

If we introduce the fixed physical length scale ξ_L as mentioned above, we have by integration of (9.109)

$$\xi_L = a \cdot \left(\beta_0 g_0^2\right)^{\beta_1/2\beta_0} \exp\left(\frac{1}{2\beta_0 g_0^2}\right) \tag{9.112}$$

For a fixed ξ_L this equation tells us how the lattice spacing $a(\beta_G)$ scales to zero when we approach the fixed point $\beta_G \to \infty$ ($g_0 \to 0$) . Any mass scale $m(\beta_G)$ in the lattice theory, which we eventually want to associate with a continuum mass should scale such that

$$m(\beta_G)\xi_L(\beta_G) \qquad \text{is finite for} \qquad \beta_G \to \infty \tag{9.113}$$

when both quantities are measured in lattice units, i.e. $a = 1$ in (9.112) which defines $\xi_L(\beta_G)$ since $\beta_G = 2N_G/g_0^2$ (in $d = 4$).

9.5.2 The Wilson loop and the string tension

As an example of a physical quantity which we can measure on the lattice, let us consider the potential between a heavy quark q and a heavy anti-quark $\bar{q}$. Due to the asymptotic freedom of the non-abelian gauge theory, the gaussian fixed point is an ultraviolet stable fixed point, and we can do reliable perturbative calculations at short distances where one finds (almost by definition) that the potential is an ordinary Coulomb potential plus radiative corrections. For larger distances perturbation theory becomes unreliable and the non-perturbative lattice formulation might help us to determine it at these distances. Let us here give a definition of the effective potential between a heavy q-$\bar{q}$ pair which can be used on the lattice without the need to introduce dynamical fermions. To motive the lattice definition let us return to the continuum theory and Minkowkian space-time and consider the simplest case of ordinary QED. If we have added two static charges they have the current

$$j_\mu(x_i,t) = \Big[e\delta^3(x_i - x_i(q)) - e\delta^3(x_i - x_i(\bar{q}))\Big]\,\delta_{\mu 0} \tag{9.114}$$

and the action will change:

$$S_0[A] = \frac{1}{4}\int F^2_{\mu\nu} \to S[A,j] = \frac{1}{4}\int F^2_{\mu\nu} + \int j^\mu A_\mu \tag{9.115}$$

In the path integral the current term will reduce to a line integral due to the $\delta^3(x)$-functions in (9.114)

$$e^{iS[A,j]} = e^{S_0[A]} e^{i\int dt[A_0(x_i(q)) - A_0(x_i(\bar{q})]} \tag{9.116}$$

Since the action is gaussian in the gauge field A_μ one can perform the functional integral (after appropriate gauge fixing) and the result is:

$$\Big\langle e^{i\int dt[A_0(x_i(q)) - A_0(x_i(\bar{q})]}\Big\rangle_0 = e^{iV(R)T} \tag{9.117}$$

where $V(R)$ is the electrostatic potential between the two charged particles separated a distance R, $V(R) = e/R$, T the total time, while

$$\langle(\cdot)\rangle_0 \equiv \frac{\int \mathcal{D}A_\mu\,(\cdot)\,e^{iS_0[A]}}{\int \mathcal{D}A_\mu\,e^{iS_0[A]}}. \tag{9.118}$$

It is an interesting exercise to check this result and we provide some details in the following example:

Example 3: Calculation of Wilson loop

> Recall that we have the general formula for gaussian integration in the case of a free field theory:
>
> $$\begin{aligned} Z[J] &\equiv \int \mathcal{D}A_\mu \exp\Big(i\int d^dx(\frac{1}{2}A_\mu\partial^2A_\mu + J_\mu A_\mu)\Big) \\ &= Z[0]\exp\Big(-i\frac{1}{2}\int\int d^dxd^dyJ_\mu(x)\Delta(x-y)J_\mu(y)\Big) \end{aligned}$$

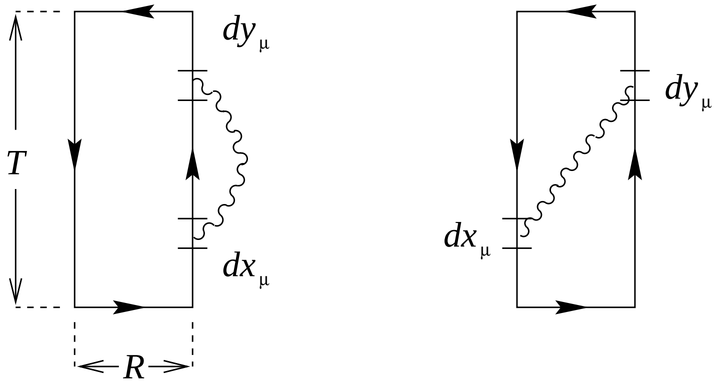


Figure 9.5: The square Wilson loop

The partion function is written in Feynman gauge, where the propagator just is $\delta_{\mu\nu}\partial^{-2}$. It is clear that the expectation value of a Wilson loop can be written as

$$\begin{aligned}\left\langle \exp\left(ie\oint_C dx_\mu A_\mu\right)\right\rangle &= Z[J]/Z[0]\\ &= \exp\left(-i\frac{1}{2}e^2\oint_C\oint_C \Delta(x-y)dx_\mu dy_\mu\right)\end{aligned}$$

provided the current $J_\mu(x) = e\dot{x}_\mu(s)\delta^3(x-x(s))$, where $s \to x_\mu(s)$ denotes the closed curve C. The three-dimensional δ-function should be viewed as a δ-function in the directions orthogonal to the curve C.

Let us perform the calculation in 4d, where the propagator is given by

$$\Delta(x) = \frac{1}{2\pi^2}\,\frac{1}{|x|^2}.$$

We see that the double line integral is singular when $x = y$ and that this singularity is proportional to the lengt of the curve C, which we will denote P. Let us therefore introduced a regularized propagator which is cut off to its value at $|x| = a$ (which we can view as a lattice spacing):

$$\begin{aligned}\Delta(x) &= \Delta(0)\cdot 1(|x|<a) + \Delta'(x)\\ \Delta'(x) &= \Delta(x) \quad \text{for} \quad |x|>a, \quad 0 \quad \text{for} \quad |x|\le a\\ \Delta(0) &= \frac{1}{2\pi^2}\,\frac{1}{a^2}\end{aligned}$$

To evaluate the double line integral explicitly we take a square loop as in fig. 9.5. It is seen that we have the nice interpretion of the double line integral as one where photons propagate from the infinitesimal line-element dx_μ to the infinitesimal line-element dy_μ. Only parallel lines contribute due to the scalar product $dx_\mu dy_\mu$ and the contributions split in two, as illustrated in fig. 9.5. The contributions from a single line in fig. 9.5a is:

$$\int\!\!\int \Delta'(x-y) = \frac{2}{2\pi^2}\int_0^T dy\int_0^{y-a} dx\frac{1}{(y-x)^2} = \frac{2}{2\pi^2}\left[T/a - \ln(T/a)\right].$$

The results to the other edges are similar and in addition we have a contribution from the singular part of the propagator : $\Delta(0)aP$:

$$\int\!\!\int_{\text{part a}} \Delta(x-y)dx_\mu dy_\mu = \left(\Delta(0) + \frac{1}{\pi a^2}\right)aP - \frac{2}{\pi^2}\left(\ln(T/a) + \ln(R/a)\right).$$

> The contribution from fig. 9.5b is
>
> $$\begin{aligned} \int\int_{\text{part b}} \Delta(x-y)dx_\mu dy_\mu &= -\frac{2}{2\pi^2}\int_0^T\int_0^T dy \frac{1}{R^2+(x-y)^2} \\ &= -\frac{2}{\pi^2}\left(\frac{T}{R}\tan^{-1}(T/R) - \frac{1}{2}\ln\left[1+T^2/R^2\right]\right). \end{aligned}$$
>
> A similar contribution is obtained by interchanging R and T.
>
> Let us now consider the limit $T >> R$. In this limit we get:
>
> $$\left\langle \exp\left(ie\oint_C dx_\mu A_\mu\right)\right\rangle \approx \exp i\left(c(a)P - \frac{e^2}{4\pi}\frac{T}{R} - \frac{2e^2}{\pi^2}\ln(R\alpha)\right)$$
>
> where $c(a)$ is a cut off dependent constant. In the where $T \to \infty$ we have, if we define the potential $V(R)$ as the coefficient proportional to iT:
>
> $$V(R) = 2c(a) - \frac{e^2}{4\pi R} \qquad (9.119)$$
>
> and we recognize Coulombs law, except for the constant $c(a)$ which is cut off dependent and represents the “mass” of the infinitely heavy electrons moving along the Wilson loop. In principle we can absorb the term in a mass renormalization if we consider dynamical electrons.

We emphasize again that the result should not be a surprise (except maybe for the cut-off dependent perimeter term which however has a very simple interpretation), since it is just an expression of the difference in vacuum energy with and without static charges. Not also that in case we had a theory where the all particles are massive the contribution from the exchange graphs of fig. 9.5b would fall of exponentially with R since massive propagators fall off exponentially with the distance. This is in agreement with the fact that a Yukawa potential is decreasing exponentially with distance.

The formulation is easy to modify for our use: First we can rotate to Euclidean space-time. The i disappears from the action *but not from the exponential of the line integral* since both dt and A_0 get an i. The rhs of (9.117) is replaced by $\exp(-V(R)T)$. Finally we can approximate the line integral along the two straight lines by a closed line integral around a rectangular loop of size $R \times T$, $T >> R$. It can be viewed as the creation of a heavy $q - \bar{q}$ pair at some early time, which are then separated a distance R where they are kept for a long time T (such that it makes sense to talk about the potential energy), after which they are brought together again and annihilated. Once we have such a closed exponential line integral we also know how to generalize it to a non-abelian case: We use the path ordered exponential and take the trace. We therefore arrive at the following mathematical definition of the effective potential between infinitely heavy $q - \bar{q}$ quarks:

$$\boxed{\left\langle \operatorname{tr} P e^{i\oint_{R\times T} dx_\mu A_\mu}\right\rangle = e^{-V(R)T} \quad \text{for} \quad T \to \infty} \qquad (9.120)$$

The expectation value in (9.120) is with respect to the pure gauge field action and when the path ordered exponentials are used in this context we call them *Wilson loops.* Note that time plays no special role in (9.120) in accordance with the fact that we have rotated to Euclidean space-time, but the shape of the loop is important since only in the limit $T \to \infty$ is the interpretation on the rhs valid.

The most popular model for confinement is one where there effectively is an electric flux string between the quark and the anti-quark. The inspiration for this model is the abelian superconductor where the magnetic flux does not spread. This means that (hypothetical) magnetic monopoles inside such a superconductor would be connected by a flux string, and there would be a linear potential between them, since the flux does not spread, contrary to the situation in empty space. A number of (not so convincing) arguments have been given why the color-electric flux string should behave in the same way in non-abelian gauge theories. The are other reasons which make the linear potential special: It can be shown that it is impossible in a relativistic field theory to have a potential which grows faster with distance and finally such a potential is precisely the one which a relativistic string has to have. In case the potential *is* linear:

$$V(R) = \sigma R \tag{9.121}$$

this implies that the expectation value of the Wilson loop will go like

$$e^{-V(R)T} = e^{-\sigma A} \tag{9.122}$$

where A is the area of the Wilson loop (which we assume is planar). In this case it is a hypothesis that the expectation value of a planar Wilson loop will fall of like the exponential of the area, (essentially) independent of the shape of the loop. The constant σ in front of the area term is call the *string tension.* The notation comes from the fact that a relativistic string has precisely the potential (9.121): the force needed to stretch the string one unit of length is σ (and unlike an ordinary rubber string it is independent of the length we have already stretched the string from equilibrium). If we denote the Wilson loop corresponding to a planar curve C spanning an area $A(C)$ by $W(C)$ the conjecture is that

$$\langle W(C) \rangle = e^{-\sigma A(C)} \quad \text{for} \;\; A(C) \to \infty \tag{9.123}$$

and this may be viewed as a criterion for confinement.

Let us now return to the lattice formulation. The Wilson loops belong to the generic class of observables we have already considered. Explicitly we can write:

$$W_{R,T} = \text{tr} \left[\prod_{l \in \text{boundary} \;\; R \times T} U_l \right] \tag{9.124}$$

where the product is over the $2(R+T)$ links around a rectangular loop enclosing $R \times T$ plaquettes. In principle we can now try to measure the exponential fall off of the expectation values of the Wilson loops by numerical methods and in this way we will extract what is called the *bare* string tension $\sigma_0(\beta_G)$ (assuming that $\langle W(C) \rangle$ really falls of as the area). This bare string tension is just a number and will, for dimensional reasons, be related to the *physical* string tension σ_{ph} by

$$\sigma_0(\beta_G) = \sigma_{ph} a^2(\beta_G) \tag{9.125}$$

where $a(\beta_G)$ denotes the lattice spacing which (in four dimensions) is given as a function of β_G by (9.112). Only if $\sigma_0(\beta_G)/a^2(\beta_G)$ has a finite limit for $\beta_G \to \infty$ can we say that the non-abelian gauge theories have a linear confining heavy $q - \bar{q}$ potential.

By now there seems to be substantial evidence, collected from years of numerical simulations on the fastest computers, that $\sigma_{ph} > 0$. If correct, this means that QCD indeed is a theory which confines quarks. At this point a few words should be said about the numerical simulations (we use here the string tension as an example, but similar remarks are valid for all other non-perturbative mass parameters extracted from numerical simulations). The extraction of σ_{ph} from the raw numerical data, i.e. $\langle W_{R\times T}\rangle$, is not straight forward. For small R we are clearly probing the perturbative part of $V(R)$, not the linear part. Where is the cross over, if there *is* a sharp cross over? And even in the region where the string tension dominates, the dynamics of the string might be important: Long strings will vibrate and such vibrations actually gives corrections to the linear potential. Apart from these problems $\langle W_{R\times T}\rangle$ contains terms which, although formally subdominant for large values of R and T, can be very important. In reality we have to fit to a formula like

$$\langle W_{R\times T}\rangle = \exp(-\sigma_0(\beta_G)RT - c_1(\beta_G)(R+T) + c_0 + \cdots) \tag{9.126}$$

where the problem is that the constant $c_1(\beta_G)$ does not scale to zero, unlike $\sigma_0(\beta_G)$ which falls off exponentially with β_G. This non-scaling is understood as follows: the area law is (in a perturbative language) due to exchange of gluons between distant (opposite) parts of the Wilson loop, but there are short distance contributions, essentially between neighbouring links in the lattice and it is precisely the same singular perimeter term we encountered in ex. 9.3. These contributions do not reflect any continuum physics and should in a theory with dynamical quarks be absorbed in mass and wave function renormalizations of the quark Lagrangian. Here they will just appear as constants. However, since $\sigma(\beta_G)$ is exponentially small for large β_G it means that $R \times T$ must be exponentially larger than $R+T$ in terms of lattice distances if the area term shall dominate. This in turn implies that $\langle W_{R\times T}\rangle$ is exponentially small for large β_G values if we at the same time require that the distance R is larger than the distances where we expect perturbative calculations to be reliable. The same conclusion is reached if we just look at the basic problem of determining σ_{ph} from (9.125). We want to take the limit $\beta_G \to \infty$ while keeping a fixed *physical length* ξ_L unchanged. Clearly this requires exponentially growing distances R measured in lattice units since $a(\beta_G)$ goes exponentially fast to zero according to (9.112). Since the interesting values of $\langle W_{R\times T}\rangle$ will be so small we get in addition problems with errorbars, statistical independence of configurations on the very large lattices needed etc..

In summary the situation resembles a lot the situation encountered in "real" experimental physics, where there sometimes is a long way from the raw data to the physical quantities extracted. Nevertheless it has been possible due to joint efforts of international collaborations by extensive numerical simulations to construct the potential $V(R)$ out to the physical distances of 1-2 fermi. The picture is a potential which at small distances agrees with perturbation theory and at larger distances change to a linear potential. Before this can be considered a proof that QCD is a confining theory it would be preferable to be able to construct the potential out to larger distances, but it requires a huge increase in computer power.

9.5.3 Inclusion of matter fields

The lattice approach can be used to address a number of other questions in QCD and the standard model, for which perturbation theory is inappropriate. We have already mentioned the wish to verify that the theory correctly gives the hadron masses. The question of chiral symmetry breaking in QCD, the question about the behaviour of quarks and gluons at high temperatures and high densities and the question of phase transitions in the standard model at high temperature. To address some of the mentioned questions we have to couple gauge fields to matter fields in a gauge invariant way on the lattice. For scalar fields there are no problems. The scalar fields $\phi(x)$ are placed on the lattice sites i. Assume that $\phi(x) \to V(x)\phi(x)$ under a local gauge transformation (we consider here the simplest case where $\phi(x)$ transforms in the fundamental representation, generalization to other cases is obvious). The lattice version is $\phi_i \to V_i\phi_i$, and the gauge invariant continuum observables

$$\phi^\dagger(y)\,\left[Pe^{i\int_{C:x\to y} dx_\mu A_\mu}\right]\,\phi(x) \tag{9.127}$$

are replaced by

$$\phi_i^\dagger\, U_{\{C:j\to i\}}\,\phi_j \tag{9.128}$$

The link variables act as parallel transport of ϕ_j to site i and this ensures that ϕ_i and $U_{\{C:j\to i\}}\phi_j$ transforms in the same way, precisely as in the continuum. This is especially true for neighbouring variables ϕ_i and ϕ_j connected by the link variable U_{ij} and we can introduce the lattice covariant derivative in analogue with the introduction of the ordinary lattice derivative:

$$\begin{aligned} \partial_\mu\phi(x) &\to \phi_{j+\hat\mu} - \phi_j \\ D_\mu\phi(x) &\to \phi_{j+\hat\mu} - U_{j+\hat\mu,j}\phi_j. \end{aligned} \tag{9.129}$$

It is now trivial to "latticize" the kinetic term $|D_\mu\phi|^2$. The final action for a gauge-Higgs system on the lattice becomes:

$$S[\phi,U] = -\kappa\sum_{(ij)}\mathrm{Re}\,\phi_i^\dagger U_{ij}\phi_j + \mu\sum_i\phi_i^\dagger\phi_i + \lambda\sum_i\left(\phi_i^\dagger\phi_i\right)^2 - \beta_G\sum_p\mathrm{Re}\,\mathrm{tr}\,U_p \tag{9.130}$$

where i denote the sites, (ij) the links and p the plaquettes. As in the case of the pure gauge field there is of course a considerable freedom in the choice of lattice actions which reduce to the continuum version in the naive scaling limit.

In principle the fermions can be introduced in the same way. We have no space to discuss this in detail. Let us only mention one point: there *are* problems with fermions on the lattice: It is not known how to introduce *chiral fermions* on the lattice. It turns out that the spectrum of fermions is 2^d times degenerate on a d-dimensional lattice. It is not difficult to remove the additional artificial massless (lattice) excitations by adding mass counter terms to the Lagrangian. However, if we insists on dealing with chiral fermions this is unfortunate, since we break explicitly the chiral symmetry by adding such mass terms. In a theory like the electroweak theory chiral symmetry is very important in the fermionic sector and it would be preferable not to start by breaking it explicitly, since we then have to recover it again in some way in the continuum limit. This problem is intimately related to the question of chiral anomalies discussed in a continuum context earlier. The point is that the lattice provides a gauge invariant regularization of the theory. If we had no problems with chiral invariance we would have a theory which at the quantum level is both gauge- and chiral invariant. This would violate the anomaly equations of the continuum (recall that the standard model was considered a consistent theory because the fermion content is such that the anomalies cancel between the different species). A convincing way of introducing chiral fermions on the lattice is still not known, but strictly speaking it could be argued that the same is true in the continuum formulation. As an example we can consider dimensional regularization. It is gauge invariant, but it has not been fully proven that the perturbative chiral sector is well defined to all orders in perturbation theory. The problem is that γ_5 has no natural definition away from even dimensions. From this point of view one can say that the lattice regularization is no worse than other types of gauge invariant regularization. It just makes the problem more visible.

9.5.4 Numerical simulations

We now turn to point (3) mentioned in the introduction to this chapter. As already mentioned a number of times the fast modern computers offer the possibility of performing large scale simulations of lattice gauge theories. One can say that the usefulness of lattice gauge theories to some extend is linked to this possibility of performing computer simulations since it is difficult to perform analytic calculations within the formalism. Let us therefore briefly describe the simplest approach to such simulations, just to give the reader a feeling of the principles involved. The topic of numerical simulations of large statistical systems is by now a vast one, and it would take many chapters to cover the different methods which can be used. Since the region of coupling constant space which is interesting is the one close the phase transitions where the correlation length diverges, and since one in addition often has to extract subleading behaviour from the raw data, extraction of reliable information from critical systems by numerical methods is an art, more than a question of trivial technique.

The one most important feature when we try the calculate the functional integral on a large lattice is the huge number of variables. On a 16^4 lattice we have $16^4 \cdot 4 \cdot 8 \approx 2 \cdot 10^7$ variables for a pure $SU(3)$ theory. Let us in the following consider lattice gauge theories and an observable $\mathcal{O}(U)$. A typical observable could be a Wilson loop of a given size $R \times T$. We want to calculate the expectation value of $\mathcal{O}(U)$ defined by:

$$\langle \mathcal{O}(U) \rangle = \frac{\int \prod_l dU_l \; \mathcal{O}(U) \, e^{-S[U]}}{\int \prod_l dU_l \; e^{-S[U]}}. \tag{9.131}$$

A direct attempt to calculate such high dimensional integrals would be doomed to fall, since most configurations contribute an exponentially small amount to the integral. As soon as we are far away from the minimum of $S[U]$ the configurations play little role in the integral. To get an efficient approximation to the integral we need an *importance sampling* of configurations U_l which we want to use in our evaluation of (9.131). Suppose we select a sequence of lattice configurations $\{U_l\}_i$, $i = 1, \ldots, n$ according to a given probability distribution $P[U]$. We can then approximate the expectation value (9.131) by

$$\mathcal{O}_n = \frac{\sum_{i=1}^n \mathcal{O}(U(i)) e^{-S[U(i)]} P^{-1}[U(i)]}{\sum_{i=1}^n e^{-S[U(i)]} P^{-1}[U(i)]} \tag{9.132}$$

and it follows from the general identity valid for any probability distribution $P(U)$ and and function $f(U)$

$$\frac{1}{n} \sum_{i=1}^n f(U(i)) P(U(i)) \rightarrow \int dU f(U) P(U) \quad \text{for } \; n \rightarrow \infty, \tag{9.133}$$

that

$$\lim_{n\to\infty} \mathcal{O}_n \to \langle \mathcal{O} \rangle \,. \tag{9.134}$$

It is not surprising that the optimal choice of $P[U]$ is the Boltzmann weight $e^{-S[U]}$ itself and we have the formula:

$$\langle \mathcal{O} \rangle = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathcal{O}(U(i)) \tag{9.135}$$

where the set of independent lattice configurations $\{U_l(i)\}$ are chosen according to the probability distribution

$$P[U(i)] \propto e^{-S[U(i)]}. \tag{9.136}$$

The problem of evaluating $\langle \mathcal{O} \rangle$ is now reduced to the problem of generating an ensemble of independent lattice configurations $\{U_l(i)\}$, $i = 1, \ldots, n$ according to (9.136). A number of computer algorithms exists for doing this. Let us only mention a very general (but for the same reason not alway the most efficient) one called the *Metropolis Algorithm.*

Given a lattice configuration U_l, where l runs over all links, we want to generate a new lattice configuration U_l'. We do this by changing a U_l at a link according to a procedure to be described shortly. This induces a change in the action :ΔS. If $\Delta S < 0$ we accept the change and if $\Delta S > 0$ the change is accepted with the conditional probability $e^{-\Delta S}$. In practise this means that one picks a random number $x \in]0,1[$, selected with uniform probability and accept the change if $e^{-\Delta S} > x$. All this takes place at a given link and one says that an attempt to *update* the link has been made. After this one moves to another link and repeats the process. By a so-called *sweep* one means that an attempt to update each link on the lattice has been made. If a lattice configuration $\{U_l(2)\}$ is the results of a sweep starting from the configuration $\{U_l(1)\}$ the two configurations are of course not independent. After a sufficient number of sweeps we will however have obtained a configuration which is independent of the first one and if we assume the first one was chosen with the correct Boltzmann distribution the same will be the case for the second one. They will therefore qualify as configurations in the sum (9.135). Usually it is *not* easy to determine when two configurations are independent and often one will have to analyse the behaviour of $\mathcal{O}_n$ from (9.132) as a function of n in order to determine how many sweeps are needed in order to generate independent configurations.

The final ingredient in the Metropolis scheme is the specification of the choice of transition $W(U_l \to U_l')$ for a given link. The transition W has to be chosen such that successive applications makes it possible to cover the whole gauge group G. It is also natural to choose it to cover the group space uniformly, since we have at this stage actually been a little sloppy with the measure factor dU. This is a uniform measure on the group and by choosing W uniform too we respect the Haar measure in the correct way. In practise one chooses W in the form:

$$U_l \to \Delta U \, U_l \tag{9.137}$$

where ΔU is selected among a set of random matrices chosen in a suitable neighbourhood of the identity $I \in G$. In this way one can monitor the change in the action ΔS and ensure that the acceptance rate in the test $e^{-\Delta S} > x$ is not too small. How to make the selection of the random matrices ΔU depends on the group and we have to refer to the literature. (For $SU(2)$ there is a simple way, since $SU(2)$ can be mapped to the 3-sphere S^3 by means of its representation by Pauli matrices, and the Haar measure in this representation just is the uniform measure on S^3.)

This description completes the ingredients needed for a numerical simulation (and we encourage the reader to write her/his own test program). Let us just end this chapter by explaining *why* the Metropolis algorithm will generate the correct probability distribution. Let us consider the relation between the n^{th} and the $(n+1)^{\text{th}}$ updating. The probability $P_{n+1}(U)$ of ending at a given configuration[5] U, if we start out with a probability distribution $P_n(U)$ and assume that the probability for a transition $U \to U'$ is $W(U \to U')$, is given by

$$\begin{aligned} P_{n+1}(U) &= \sum_{U'} W(U' \to U) P_n(U') \\ &= P_n(U) + \sum_{U'} \left(P_n(U') W(U' \to U) - P_n(U) W(U \to U') \right) \end{aligned} \qquad (9.138)$$

where we have used the normalization

$$\sum_{U'} W(U \to U') = 1. \qquad (9.139)$$

From (9.138) we see that a *sufficient* condition for a stationary probability distribution $P_n(U) = P(U)$ independent of n is that

$$P(U) W(U \to U') = P(U') W(U' \to U) \quad \forall\, U, U' \qquad (9.140)$$

If W is chosen such that it satisfies (9.140) we say that it fulfills *detailed balance.* By summing over U in (9.140) we get

$$\sum_{U} P(U) W(U \to U') = P(U') \qquad (9.141)$$

and this tells us that the distribution $W(U \to U')$, viewed as a matrix $W(U, U')$, in the space of configurations has $P(U)$ as eigenvector with eigenvalue 1. Due to (9.139) and the fact[6] that $W(U, U') > 0$ this is the maximal eigenvalue and $P(U)$ its unique eigenvector[7]. Since any probability distribution $P_0(U)$ will have a scalar product different from zero with $P(U)$ (assuming that $P(U) > 0\ \forall\, U$) we see that $P_n = W^n P_0$ will converges to $c \cdot P$, where c is the scalar product of P and P_0 considered as vectors. *Detailed balance is thus sufficient to ensure convergence to the correct probability distribution.*

Let us now check that the Metropolis algorithm satisfies detailed balance (9.140) with respect to the Boltzmann weight $P(U) \propto e^{-S(U)}$. The transition probability of Metropolis is

$$W(U \to U') = \begin{cases} 1 & \text{if } \ S(U) > S(U') \\ e^{-(S(U') - S(U))} & \text{if } \ S(U) < S(U') \end{cases} \qquad (9.142)$$

This means that

$$\frac{W(U \to U')}{W(U' \to U)} = e^{-(S(U') - S(U))} \quad \text{or} \quad e^{-S(U)} W(U \to U') = e^{-S(U')} W(U' \to U). \quad (9.143)$$

We conclude that (9.140) is satisfied with $P(U) \propto e^{-S(U)}$.

---

[5] We use the short hand notation U for a complete lattice configuration $\{U_l\}$.

[6] The condition $W(U, U') > 0$ expresses the fact that one should be able to reach any element from any other element, i.e. *ergodicity* of the *Markov process* W. Strictly speaking it need not be satisfied in a single step, as already mentioned above.

[7] It follows from the so-called Perron-Frobenius theorem of linear analysis.