UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1978

The Powers of Perturbation Theory

Loops and Gauge Invariance in Particle Physics

JOHAN LÖFGREN

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2020

ISSN 1651-6214
ISBN 978-91-513-1041-1
urn:nbn:se:uu:diva-422716

Dissertation presented at Uppsala University to be publicly examined in Polhemsalen, Ångströmlaboratoriet, Lägerhyddsvägen 1, Uppsala, Friday, 11 December 2020 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Kimmo Kainulainen (University of Jyväskylä, Department of Physics).

Online defence: https://uu-se.zoom.us/j/62746592034?pwd=V0VBTVhGWC96UEJOWWpOTmdNd0E2UT09

Passcode: 045337

Abstract

Löfgren, J. 2020. The Powers of Perturbation Theory. Loops and Gauge Invariance in Particle Physics. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1978. 58 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-1041-1.

The Standard Model is the best particle physics theory we have, but there are still phenomena that it cannot explain. In this thesis I have worked on two different projects that connect to two of the biggest unsolved questions of the Standard Model.

From observations of neutrino oscillations we know that at least one of the neutrinos has to be massive. But the neutrinos of the Standard Model are massless. The first paper in the thesis investigates a simple extension of the Standard Model that realizes a fifth force as a U(1) gauge group. In such models, extra care has to be taken to not introduce inconsistencies known as anomalies. It turns out that the simplest way to avoid these problems is to introduce three right-handed neutrinos. Such models can then incorporate neutrino masses in a convenient way. In the second paper we have investigated a twist on this model that does not have neutrino masses, but which makes other interesting models possible—such as a model with gauged lepton number.

The observed asymmetry between matter and antimatter cannot be explained by the Standard Model. One of the more popular of the possible explanations is known as electroweak baryogenesis. In this scenario the asymmetry is determined during the electroweak phase transition in the early universe. The second project—spanning the three final papers of the thesis—has aimed to improve the approximation methods we use to calculate features of this phase transition. Such calculations are plagued by two big problems that seem to compete with each other. On the one hand, gauge invariant results seem to demand that a strict loop counting must be enforced. On the other hand, the loop approximation does not work well close to the phase transition. We argue that the solution is to use a different power counting, but still be strict about sticking to it.

Keywords: Beyond the Standard Model, Electroweak phase transition, Perturbation theory

Johan Löfgren, Department of Physics and Astronomy, High Energy Physics, Box 516, Uppsala University, SE-751 20 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-1041-1
urn:nbn:se:uu:diva-422716 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-422716)

Dedicated to curious people

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I A. Ekstedt, R. Enberg, G. Ingelman, J. Löfgren, T. Mandal. *Constraining minimal anomaly free U(1) extensions of the Standard Model.* J. High Energ. Phys. **2016**, 71 (2016). `ArXiv:1605.04855`.

II A. Ekstedt, R. Enberg, G. Ingelman, J. Löfgren, T. Mandal. *Minimal anomalous U(1) theories and collider phenomenology.* J. High Energ. Phys. **2018**, 152 (2018). `ArXiv:1712.03410`.

III A. Ekstedt, J. Löfgren. *On the relationship between gauge dependence and IR divergences in the $\hbar$-expansion of the effective potential.* J. High Energ. Phys. **2019**, 226 (2019). `ArXiv:1810.01416`.

IV A. Ekstedt, J. Löfgren. *A Critical Look at the Electroweak Phase Transition.* `ArXiv:2006.12614`. Submitted to JHEP (2020).

V A. Ekstedt, J. Löfgren. *The High-Temperature Expansion of the Thermal Sunset.* SciPost Phys. Core 3, 008 (2020). `ArXiv:2006.02179`.

Reprints were made with permission from the publishers.

My contribution to the papers

Paper I: A. Ekstedt, T. Mandal, and I performed the original model-building calculations together. I wrote the Mathematica script which we used to perform the numerical calculations, while T. Mandal and A. Ekstedt performed the Madgraph simulations. R. Enberg and G. Ingelman supervised the project. We all contributed to writing the paper.

Paper II: A. Ekstedt had the original idea to implement the Green-Schwarz mechanism for the U(1) extended Standard Model. We performed this derivation together. Andreas took the lead on the 1-loop calculations necessary to test the theory, with me checking the results. The rest of the project proceeded as for paper I.

Paper III: Though I first had the idea that started this project, the project evolved and changed as A. Ekstedt and I worked on it together; in the end it is an amalgamation of our shared insights. We did almost all of the calculations together, and we both wrote the paper.

Paper IV: I had as a goal for a very long time to combine the $\hbar$-expansion with thermal resummations. Then A. Ekstedt had the realization that the problem laid in the power counting. We developed the formalism together and tested it by performing thermal 2-loop calculations. We wrote the paper together.

Paper V: The realization that the method of regions could be useful for thermal sum-integrals grew from our work on paper III. A. Ekstedt performed the original calculation and wrote a first draft. I double-checked the derivation and contributed to writing the paper.

Foreword

> "Let me share with you the terrible wonders I have come to know..."
> —The Narrator
>
> *Darkest Dungeon*

The goal of my research is to test, or enable others to test, particle physics models against reality. In my PhD studies this has taken the form of two different projects, with some common aspects. In my first two papers, papers I and II, I have explored the collider phenomenology of a U(1) extension of the Standard Model of particle physics. In such extensions, gauge anomalies must cancel such that the theory is consistent. This requires that particular care is applied during model building.

Papers III and IV concern the effective potential—a device used to understand spontaneous symmetry breaking and the Higgs mechanism when quantum and thermal fluctuations are included. Such calculations are relevant for the electroweak phase transition that took place in the early universe. The work behind these papers strives to improve the results of perturbation theory by taking certain consistency conditions very seriously.

Paper V is a derivation of sub-leading terms in an expansion of a thermal sum-integral called the sunset. The sunset shows up in 2-loop calculations of the effective potential; the sub-leading terms of the expansion are needed at higher orders in perturbation theory, as we discuss in paper IV.

A unifying principle of these two different projects is the quest to maintain gauge invariance. Or, more colloquially, to ensure proper accounting of the degrees of freedom. A gauge invariance represents a redundancy in the description; to not maintain it would mean wrongful accounting. This is especially troubling because we typically introduce artificial unphysical degrees of freedom to simplify calculations. These unphysical contributions must cancel in the end. If not, we have twisted our formalism such that predictions become inconsistent and untrustworthy.

In chapter 1 I give a brief overview of the Standard Model and why physicists believe it must be extended. Chapter 2 describes spontaneous symmetry breaking and the Higgs mechanism, which are relevant to modern model building and to understanding the electroweak phase transition. Chapter 3 concerns quantum corrections and how they relate to symmetries, and chapter 4 brings thermal corrections into the picture. Each of these chapters ends with an annotated bibliography of recommended readings. The final chapter, chapter 5, is a popular summary in Swedish.

Thank you for reading,

Johan Löfgren

Contents

1. Particle physics and you

> The story so far: In the beginning the Universe was created. This has made a lot of people very angry and been widely regarded as a bad move.
>
> Douglas Adams, *The Restaurant at the End of the Universe*

Physics is an ambitious endeavor that aims to consistently describe planets, solar-systems, galaxies, (and beyond), and everything as small as molecules, atoms, electrons, (and beyond). Though this project is far from complete, great progress was made in the last century. There were several paradigm shifts that forever changed our view of reality. With the introduction of relativity and quantum mechanics, the 20th century marked the advent of the field *particle physics*. We learned that to study particles we must study quantum fields—*quantum field theory* can combine the principles of special relativity and quantum mechanics.

The most well-tested quantum field theory we have is the Standard Model of particle physics, which I describe in the following section. This introductory chapter has a light tone, with the aim to place my research into a larger context. In future chapters I get into more technical details.

1.1 A brief introduction to the Standard Model

The Standard Model is a quantum field theory—a collection of quantum fields with different properties, linked together by interactions. Excitations of these fields appear to us as particles, and when the particles affect each other it is due to the fields interacting. The discovery of all the different building blocks, and how they fit into the Standard Model, is an impressively large scientific project that culminated during the previous century. The Higgs boson—the final piece—was discovered in 2012 at the Large Hadron Collider [1, 2, 3].

There are several different types of fields in the Standard Model. We think of the *fermionic* fields as matter because their particle excitations obey Fermi-Dirac statistics. Loosely speaking, they cannot be compressed together indefinitely—they take up space. This is in contrast to *bosonic* fields, such as electromagnetic waves, that can be layered on top of each other with little limitation. There are three generations of fermions in the Standard Model, where the first generation corresponds to particles like the electron, or the quarks that form atomic nuclei. That is, atoms consist of protons and neutrons—composite particles of up- and down-quarks—orbited by electrons.

Fermions from the other two generations are not as well-known, for good reasons: they are very heavy and decay quickly into other particles. But they are of course important to physics at short distances.

There is in addition another kind of matter particle that is abundant in the universe. Though we do not notice neutrinos in our daily life, they are ever-present. They are created in nuclear processes, but they hardly interact with other particles.

With the Standard Model's matter-content out of the way, we turn to its forces. The concept of a force is, in particle physics, a slippery one. In the physics of our everyday lives, we typically think of a force as something that pushes or pulls something else. These forces may be emergent effects that arise from microscopic interactions, such as how the ground pushes your feet up when you are standing on it. Or a force might be the manifestation of the presence of a fundamental force-field—such as the electromagnetic field.

In contrast, quantum fields can in one instance act like a particle (something that scatters) and in another instance like a field (something that mediates interaction). Particles interact through intermediate states—fields are excited and their excitations trigger the excitations of other fields through interactions—and really all fields are forces in this sense. An electron can scatter off a photon (the particle that mediates the electromagnetic force) by "exchanging" an intermediate electron.[1] In this sense we might talk of the "electron force."

But this is not how particle physicists usually use the word force. Typically, we only refer to forces if they are mediated by *gauge bosons*. In this sense there are four different forces that particles are subject to. There is the electromagnetic force mediated by the photon γ, the weak force mediated by the Z and W bosons, the strong force mediated by the gluon g, and gravity (presumably) mediated by gravitons. All of these forces, except gravity, are a part of the Standard Model—the first three correspond to the gauge group $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. In particular, $\mathrm{SU}(3)_c$ corresponds to the strong force, and $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ to the electroweak unification of the electromagnetic and the weak force. The electroweak gauge group is "broken" by the Higgs mechanism, leaving us with the two separate forces we usually speak of.

That said, there are physicists who think that we should really call *all* bosons forces. In particular, this would imply that the Higgs field mediates a fifth force—the Higgs force. After all, propagating massive particles can interact with each other by exchanging a Higgs boson. Just like how an electron and a proton might attract each

[1] Or a more complicated intermediate state, represented by a Feynman diagram with loops.

other by exchanging a photon. Most physicists do not subscribe to this nomenclature though, and I will not use it further in this thesis.

The Higgs field implements the Higgs mechanism. All the gauge bosons and all the fermions of the Standard Model would need to be massless if there was no such mechanism. This mechanism plays a key role in both of the research projects that this thesis is based on; I elaborate on it in chapter 2.

1.2 Beyond the Standard Model

Now that we are familiar with the Standard Model, let's scrutinize its imperfections. There are flaws, both theoretical and experimental.

The Standard Model does not contain gravity. General relativity, the theory of gravity at large distances, has technical issues when constructed as a quantum field theory. Constructing a well-functioning quantum theory of gravity is one of the biggest challenges of theoretical physics. There are a few different approaches, with string theory as the most popular candidate (for good reasons). The fact that the Standard Model does not deal with gravity at all means that the theory does not make accurate predictions for certain extreme cases—such as near black holes or in the very early universe, where the quantum nature of gravity is important. But the Standard Model together with general relativity works well for the energies that are available in our colliders.

Neutrinos are massive. Due to the observed phenomenon of neutrino mixing, we have known since the 90s that at least one neutrino must be massive. But the neutrinos in the Standard Model are massless. Their masses could in principle be added to the Standard Model, but to do it correctly we need to know more about neutrinos. Depending on if they behave as Dirac or Majorana fermions, the fine details of their mass terms and their interactions will be different. I discuss this issue more in subsection 1.3.1.

Dark matter is likely to have a particle nature. Yet the particles of the Standard Model cannot explain the various phenomena that we need dark matter for, such as the rotation curves of galaxies. This problem can be addressed by adding massive particles to the Standard Model that interact weakly with everything besides gravity. The problem is of course to know which model is correct, and there is a large industry dedicated to test such models in different ways.

Matter and antimatter do not exist in equal amounts. This fact is very puzzling from the Standard Model's point of view. Though matter and antimatter can be differentiated in the Standard Model, since the CP symmetry is slightly broken, it cannot explain the sheer difference we have observed. There are a few competing explanations, but they all require additional particles. I will discuss the prospect of electroweak baryogenesis in subsection 1.3.2.

Dark energy. Cosmological observations suggest that we live in an accelerating universe. The simplest explanation (there are other candidates) is that there exists a very small, but nonzero and positive, *cosmological constant*—a background energy density. The Standard Model can accommodate such a constant, but there would have to be some serious fine-tuning of the parameters in the Lagrangian to explain its small value. It is up to extensions of the Standard Model to explain why the value is what it is, or to propose another mechanism that explains the cosmological observations.

The hierarchy problem. This is one of the Standard Model's longest-standing theoretical problems. If there are heavier particles that couple to the Higgs, then the Higgs boson's mass should get large quantum corrections. But because we know that the Higgs mass is actually around the weak scale, then there must be minute cancellations between these corrections such that they do not contribute too much. This is a fine-tuning. These considerations lead particle physicists to believe that the LHC would reveal more particles than what we have seen so far—particles that would enact some mechanism that renders fine-tuning unnecessary. Models with supersymmetry, and models in which the Higgs is a composite particle, are examples of such models. There is no concrete evidence for such extensions of the Standard Model as of yet.

1.3 Particle physics and me

In this section I elaborate on two of the problems I mentioned above, because they are related to my own work. This section is necessarily more technical than the previous one.

1.3.1 The neutrino masses

The Standard Model does not contain neutrino masses. But even though we are sure that at least one of the neutrinos is massive, it is not clear how to add mass terms for neutrinos.[2] Indeed, the type of mass term to add depends on whether neutrinos are Majorana or Dirac particles. Consider a left-handed neutrino ν_L. A Dirac mass term needs an accompanying right-handed neutrino ν_R, and is written as

$$m_D(\overline{\nu}_L \nu_R + \overline{\nu}_R \nu_L). \tag{1.1}$$

If such right-handed neutrinos do exist (though we have not observed them), the coefficient m_D would have to be incredibly small to explain the observed neutrino masses—which are around a few eV. In the Standard Model, such mass terms

[2] This subsection is based on the paper [4], and the textbook [5].

would be generated by interactions between the neutrinos and the Higgs field. With additional right-handed neutrinos, the mass term would be of the form

$$y\bar{l}_L \nu_R \tilde{H} + \text{h.c.}, \tag{1.2}$$

where H is the Higgs doublet, $\tilde{H} = i\sigma_2 H^*$, and l is the left-handed fermion doublet. Here y would need to be very small compared to other Yukawa couplings in the Standard Model.

On the other hand, a Majorana mass term can be written purely in terms of ν_L,

$$\frac{1}{2} m_M \overline{\nu^c}_L \nu_L. \tag{1.3}$$

However, there are additional complications with Majorana mass terms. Such terms are not possible with renormalizable operators in the Standard Model, because of gauge invariance. Though it is possible to induce such a mass term with non-renormalizable higher-dimensional operators like

$$\frac{c}{\Lambda}(\overline{l^c}_L H)(l_L H) + \text{h.c.}, \tag{1.4}$$

where Λ is a cut-off scale and c is a numerical coefficient. This explanation for the fermion mass would hence also require new physics (around the scale Λ). But an added bonus is that it could explain why neutrinos are very light. Because the mass terms are then suppressed by the high scale Λ.

To know which of these mechanisms is realized in nature, we need more experimental data about neutrino interactions. But such data is sparse because neutrinos interact weakly with other particles. Hence it is currently unknown wether neutrinos are Dirac or Majorana fermions. So far we have only observed the three left-handed neutrinos. A generic prediction of Majorana neutrinos is that of neutrinoless double-beta decay, where certain atomic decays produce two electrons (or positrons) and no neutrinos. This kind of decay cannot occur if the neutrinos are Dirac particles—it is forbidden by lepton-number conservation. No such decay has yet been observed.

Although there's a good case for believing in pure Majorana neutrinos, there are other compelling possibilities. If right-handed neutrinos ν_R exist, they can form both Majorana and Dirac mass terms simultaneously. And so a neutrino mass matrix M with off-diagonal elements is possible:

$$\mathcal{M} \sim \begin{pmatrix} 0 & m_D \\ m_D & m_M \end{pmatrix}. \tag{1.5}$$

If $M_M \gg m_D$, then the masses of the propagating states are approximately

$$m_1 \sim \frac{m_D^2}{m_M}, m_2 \sim m_M. \tag{1.6}$$

In other words, there are two classes of neutrinos—one that is very light, and one that is very heavy. This is known as a *seesaw-mechanism*.

In paper I we have considered a simple extension of the Standard Model, with an extra U(1) gauge symmetry. Such theories have certain consistency requirements—possible gauge anomalies must cancel, see section 3.2—and it turns out that the simplest way to resolve these is to add three fermions that are not charged under the Standard Model gauge group [4]. This precisely corresponds to three right-handed neutrinos.

The model also features a new complex scalar field φ which acquires a vacuum expectation value to break the new U(1) group and give mass to the new neutral gauge boson Z'. It is then possible to have Majorana mass terms for these three right-handed neutrinos, such as

$$\varphi^{\dagger}\overline{\nu^c}_R \nu_R + \text{h.c.} \tag{1.7}$$

The vacuum expectation value of φ is on the same order as the mass of Z', which should be around the TeV-scale to have evaded the experimental bounds so far. With $m_D \sim 1$ MeV (close to the electron mass), and $m_M \sim 1$ TeV, we find $m_1 \sim 1$ eV.

Such anomaly-free U(1) extensions of the Standard Model hence conveniently support the existence and smallness of the neutrino masses.

1.3.2 Matter-antimatter asymmetry

Though it is clear that there are more baryons than anti-baryons in our universe, it is possible to be more quantitative about it. By observing cosmic rays that reach the earth, and comparing the number of incoming anti-protons $\overline{p}$ to the number of incoming protons p, we have measured that [6]

$$\frac{\overline{p}}{p} \sim 10^{-4}. \tag{1.8}$$

This fraction is consistent with an abundance of *primordial* protons that have persisted since the early universe. The small fraction of antiprotons is just what you expect if they were created in the cosmic rays [7]—any primordial antiprotons have long since annihilated.

Just how the matter-antimatter asymmetry came about is an open question. There are a number of different proposed explanations, but none with much evidence in favour of it. One popular mechanism is known as *electroweak baryogenesis* [8]. Just before the electroweak phase transition—which broke the electroweak symmetry—a net baryon number can be generated by certain non-perturbative processes known as *sphalerons*. To prevent this number from being washed away, it is important that the phase transition is *first-order*. Such phase transitions occur through bubble nucleation, just like boiling water, and offer just the correct circumstances to make electroweak baryogenesis possible.

In the Standard Model, the phase transition is known to be *second-order* from lattice simulations [9, 10]. But because physics at the electroweak scale is difficult to test experimentally—especially the Higgs potential, which is relevant for the phase transition—it is possible that there is new physics lurking there. Such new physics might change the nature of the phase transition, and therefore there is much interest in extensions of the Standard Model that modify the Higgs potential.

But, as I describe further in chapter 4, it is tricky to perform perturbative calculations relating to the phase transition. In papers III and IV, a colleague and I have developed perturbative methods to improve such calculations. Paper V is our calculation of further terms in the high-temperature expansion of a 2-loop thermal sum-integral known as the sunset. As we discuss in paper IV, such terms are relevant at high orders of perturbation theory.

Recommended readings

Sean Carroll, *The Particle at the End of the Universe* [11]. This popular science book explores the discovery of the Higgs boson at the LHC. It is well written but still keeps a good level of accuracy.

A. Zee, *Quantum Field Theory in a Nutshell* [12]. This is a textbook. But it emphasizes the physical content of quantum field theory, not the technical details.

2. Spontaneous symmetry breaking

> Symmetry: you break it, you buy it.
>
> Label on coffee mug from University of Washington, Seattle.

The general idea of spontaneous symmetry breaking is one of the most far reaching and interesting ones in modern physics; see [13] for a recent and comprehensive review. In gauge theory it shows itself in the guise of the Higgs mechanism.

The Higgs mechanism, and also spontaneously broken global symmetries, play a key role in the Standard Model and in many of its extensions. In this chapter and those that follow, I will dig into some of the details these mechanisms. I start with a simple example and build on it sequentially. Though the concept of broken discrete symmetries is also interesting, I will focus on continuous symmetries for brevity.

2.1 Global symmetries, spontaneously broken

2.1.1 A simple example

Let's start by considering a 4D quantum field theory with a complex scalar field $\Phi(x)$, and a global U(1) symmetry. The Lagrangian is

$$\mathcal{L} = -\partial^\mu \Phi^\dagger \partial_\mu \Phi - V_0[\Phi, \Phi^\dagger], \tag{2.1}$$

$$V_0[\Phi, \Phi^\dagger] = m^2 \left(\Phi^\dagger \Phi\right) + \lambda \left(\Phi^\dagger \Phi\right)^2 . \tag{2.2}$$

These are all the possible terms with mass-dimension ≤ 4. Here V_0 is the *classical potential*, a functional that defines the interactions among the theory's scalar fields. As we shall see shortly, the form of the classical potential is crucial for determining the spectrum of the theory.

Though the complex representation used above realizes the U(1) transformation in a simple manner, $\Phi \to e^{-i\theta}\Phi$, the rest of the discussion will benefit from introducing the real and imaginary components of Φ,

$$\Phi = \frac{1}{\sqrt{2}} \left(\phi_1 + i\phi_2\right). \tag{2.3}$$

Introducing the real vector $\vec{\phi} = (\phi_1, \phi_2)^T$, we can write the potential as

$$V_0[\vec{\phi}] = \frac{1}{2}m^2\vec{\phi}^2 + \frac{1}{4}\lambda\vec{\phi}^4. \tag{2.4}$$

As this theory stands, a naive derivation of the spectrum would lead us to believe that there are two real scalar particles, both with mass m. Is this correct? Well, sometimes.

When we are constructing a quantum field theory, we should try to be as general as possible: write down all possible terms and consider all possible values of the parameters. In general, the Hamiltonian must be a Hermitian operator to produce a physical spectrum. In our case this tells us that the parameters must be real (though in more complicated theories they may be complex). From the argument above we might also expect that m should be positive, because it seems to correspond to the mass of a particle. But this assumption would be a mistake.

To see why, we need to consider the *ground state* of the theory.[1] We can start with the classical energy, and focus on quantum corrections later. The classical energy is given by the space-integral of the Hamiltonian density,

$$E = \int \mathrm{d}^3x\mathcal{H} = \int \mathrm{d}^3x\left(\frac{1}{2}\partial_0\vec{\phi}\cdot\partial^0\vec{\phi} + \frac{1}{2}\partial_i\vec{\phi}\cdot\partial^i\vec{\phi} + V_0[\vec{\phi}]\right), \tag{2.5}$$

where we can think of the first two terms as the kinetic energy of the scalar fields. The aptly named classical potential corresponds to the potential energy.

To find the ground state we should minimize this energy. Because the kinetic terms contribute as positive squares, their minimal contribution is zero. We then immediately draw the conclusion that the field with minimal energy, $\vec{\phi}_0$, should be static: $\partial_\mu\vec{\phi}_0 = 0$. All that remains is the potential energy,

$$E_{\mathrm{min}} = \int \mathrm{d}^3x V_0[\vec{\phi}_0] = \mathcal{V}\times\left(\frac{1}{2}m^2\vec{\phi}_0^2 + \frac{1}{4}\lambda\vec{\phi}_0^4\right), \tag{2.6}$$

where $\mathcal{V}$ is the volume of space and $\vec{\phi}_0$ minimizes the classical potential,

$$\partial_i V_0|_{\phi=\phi_0} \overset{!}{=} 0. \tag{2.7}$$

Here I used the shorthand $\partial_i \equiv \frac{\partial}{\partial\phi_i}$.

To solve this equation we should first consider the possible values of the parameters, and to do so we must let go of their physical interpretations—forget that m usually corresponds to a mass. If λ is negative then the energy is unbounded from below. This theory is *unstable* and is not of interest to us. If $\lambda = 0$, then the theory

[1]The following demonstration is based on Rubakov's textbook [14].

is unstable if $m^2 < 0$; if $m^2 \geq 0$ it describes two free scalar fields, both with mass m.

When $\lambda > 0$ the theory is bounded from below and the existence of a ground state is guaranteed. To find it, we solve equation (2.7) to find possible extrema,

$$0 \stackrel{!}{=} \left(m^2 + \lambda|\vec{\phi}_0|^2\right)\vec{\phi}_0 \implies \vec{\phi}_0 = 0 \quad \text{or} \quad |\vec{\phi}_0|^2 = -\frac{m^2}{\lambda}. \tag{2.8}$$

The natures of the extrema are found by computing the second derivative matrix (the Hessian). Due to the U(1) invariance, it is enough to focus on the absolute value of $\vec{\phi}_0$. Defining $\phi \equiv |\vec{\phi}|$, and $\partial \equiv \frac{\partial}{\partial \phi}$, we have that

$$\partial^2 V_0\Big|_{\vec{\phi}=0} = m^2, \tag{2.9}$$

$$\partial^2 V_0\Big|_{\phi^2=-\frac{m^2}{\lambda}} = -2m^2. \tag{2.10}$$

Now it is time to face to the possible values of m^2. If $m^2 = 0$, then the two extrema are both minima and located at the origin—the two scalars are massless.[2] If $m^2 > 0$, then the second extremum is not realized (there are no real values of $\vec{\phi}$ which satisfy that equation); the ground state is at the origin and the naive analysis holds: there are two real scalars with mass m.

With negative m^2, the extremum at the origin is a maximum, and the second one is a minimum. The ground state satisfies

$$|\vec{\phi}_0| = \sqrt{\frac{-m^2}{\lambda}}, \tag{2.11}$$

and is hence actually a continuum of states—all 2D vectors with this length.

The existence of this multitude of states is related to the symmetry of the original theory, which will be more apparent when we consider the general case in subsection 2.1.2. Because these states all have the same energy, we can pick one of them as a representative. Take

$$\vec{\phi}_0 = (\phi_0, 0)^T; \quad \phi_0 = \sqrt{\frac{-m^2}{\lambda}}. \tag{2.12}$$

Now that we have the ground state, we should expand around it to find the spectrum of the theory; in that vein,

$$\phi_1 = \phi_0 + H, \phi_2 = G, \tag{2.13}$$

[2]The second-derivative test actually fails in this case, but they can be confirmed to be minima by inspection.

where I gave the fluctuations around the ground state the names H and G—for reasons that shall soon become apparent. To find their masses, simply insert this expansion into the Lagrangian and look at the appropriate terms (H^2 and G^2),

$$m_H^2 = m^2 + 3\lambda\phi_0^2 = -2m^2, \tag{2.14}$$

$$m_G^2 = m^2 + \lambda\phi_0^2 = 0. \tag{2.15}$$

The spectrum hence consists of two scalars: H, which has squared mass $-2m^2$ (a positive number), and G, which is massless.

Though we started with a theory in which all fields were treated the same, where ϕ_1 and ϕ_2 were interchangeable—and could be intermingled with a U(1) rotation—we ended up with a spectrum in which the two "physical" fields are not interchangeable. This is the essence of spontaneous symmetry breaking; though the theory is invariant, the ground state breaks the symmetry—splitting the spectrum.

In this section we considered a simple Abelian symmetry, that was spontaneously broken. In the spectrum we found one massive scalar and one massless one. In the next section I will discuss the general case of a non-Abelian symmetry.

2.1.2 The general case (global)

The most general continuous symmetry that is of interest to us corresponds to a compact semi-simple Lie group $\mathcal{G}$ [14]. To specify how the various fields of our model transform under $\mathcal{G}$, we must specify in which representation they are. It is enough for our interests to consider the transformation of the scalars in the model.[3]

Because our end goal is to do perturbation theory, it is simplest to use a real representation of the scalars (this can be done without loss of generality). We collect the N real scalars in the vector $\vec{\phi}$, with components ϕ_i.

Let's denote the generators of this representation as T^a, with $a = 1, \ldots, D_{\mathcal{G}}$ and $D_{\mathcal{G}}$ is the dimension of the group $\mathcal{G}$. An infinitesimal transformation of ϕ_i is then given by

$$\phi_i' = (1 - i\theta^a T^a)_{ij}\phi_j, \tag{2.16}$$

where θ^a are arbitrary infinitesimal parameters. Note that this representation is real, and hence the generators T^a are antisymmetric Hermitian matrices.

For this to be a symmetry, we require that the action is unchanged under $\mathcal{G}$ transformations, $\delta S \overset{!}{=} 0$. For our purposes we can focus on the scalars, and in particular on static configurations. Then we just require that the classical potential is invariant, $\delta V_0 \overset{!}{=} 0$, that is

$$V_0(\vec{\phi}) \overset{!}{=} V_0'(\vec{\phi}') \implies (T^a\vec{\phi})_i \partial_i V_0 = 0. \tag{2.17}$$

[3]The following discussion is partially based on Srednicki's textbook [15], and partially on the two papers [16] and [17].

We can get a lot of mileage from this equation. If we differentiate it with respect to ϕ_i, we find

$$\partial^2_{ij} V_0 \, (T^a \vec{\phi})_j = T^a_{ij} \partial_j V_0, \tag{2.18}$$

where $\partial^2_{ij} V_0$ is the scalar mass-matrix as a function of the background field ϕ. The equations above must hold for any static scalar field configuration, even if it does not minimize the potential. But let's get more specific and consider *vacuum solutions* that do extremize the potential.

First off, if we consider the standard non-breaking type of solutions with $\vec{\phi}_0 = 0$, then the equations above trivially hold—all generators annihilate the vacuum: $T^a \vec{\phi}_0 = 0$. This equation tells us that $\vec{\phi}_0$ does not transform under the group $\mathcal{G}$. In other words, the vacuum does not break any symmetries; we say that all generators that annihilate the vacuum are *unbroken*.

But if we consider vacua located away from the origin, then the situation is more interesting. We can split the generators into two families, $\{T^a\} = \{\mathcal{T}^\alpha, t^b\}$, where t^b annihilate the vacuum and $\mathcal{T}^\alpha$ do not. So, if $\mathcal{T}^\alpha \vec{\phi}_0 \neq 0$ then the vacuum transforms under this "part" of the group , and we say that the generators $\mathcal{T}^\alpha$ are *broken*.

Using group theory, it is possible to show that the $D_\mathcal{H}$ unbroken generators form a subgroup, $\mathcal{H}$, that we call the *unbroken subgroup*. The broken generators $\mathcal{T}^\alpha, \alpha = 1, \ldots, D_\mathcal{G} - D_\mathcal{H}$ do not in general form a subgroup, but they are interesting in their own right.

Returning to equation (2.18), but focusing on a vacuum and its broken generators, we find

$$\partial^2_{ij} V_0 \Big|_{\vec{\phi}_0} (\mathcal{T}^\alpha \vec{\phi}_0)_j = 0. \tag{2.19}$$

We hence conclude that $\mathcal{T}^\alpha \vec{\phi}_0$ is an eigenvector of the mass-matrix with eigenvalue zero. In other words, there are $(D_\mathcal{G} - D_\mathcal{H})$ massless scalar bosons—one for each broken generator. This is Goldstone's theorem, and the massless scalar bosons are known as Goldstone bosons.

In the context of the Abelian U(1) example above, there is only one generator to begin with; in the real 2-dimensional representation the generator is proportional to the antisymmetric symbol ϵ_{ij}. The non-trivial vacuum breaks this generator, and the resulting unbroken subgroup is trivial. There is $D_\mathcal{G} - D_\mathcal{H} = 1 - 0 = 1$ Goldstone boson.

Furthermore, we chose a representative vacuum $\vec{\phi}_0 = (\phi_0, 0)^T$, and did not worry about a loss of generality. With this representative we found that ϕ_2 corresponded to the Goldstone boson G, and we let H refer to the massive fluctuation along the direction of the vacuum expectation value (vev) $\vec{\phi}_0$.

The Goldstone field corresponds to fluctuations perpendicular to the vev direction. This is because any static field configuration $\vec{\phi}$ is orthogonal to the corresponding Goldstone direction $\mathcal{T}^\alpha \vec{\phi}$,

$$\vec{\phi}_i(\mathcal{T}^\alpha \vec{\phi})_i = \vec{\phi}_i \mathcal{T}^\alpha_{ij} \vec{\phi}_j = 0, \tag{2.20}$$

due to the antisymmetry of $\mathcal{T}^\alpha$. In the Abelian case, if we instead would have considered a generic background field $\vec{\tilde{\phi}} = (\tilde{\phi}_1, \tilde{\phi}_2)^T$, we would have found the Goldstone direction to be $\epsilon \vec{\tilde{\phi}} = (\tilde{\phi}_2, -\tilde{\phi}_1)^T$—which is orthogonal to $\vec{\tilde{\phi}}$.

Of course, because the potential is invariant under the group $\mathcal{G}$, we are free to transform $\vec{\phi}$ into the simplest form possible. That is why we can simply choose to work with the vev in the ϕ_1 direction in Abelian case. But I think it helps to see that it can be done with general fields as well.

To summarize, for a particular vacuum $\vec{\phi}$ that breaks the generators $\mathcal{T}^\alpha$, we are free to choose a form of $\vec{\phi}$ that makes our lives simpler—as long as that form is reachable by a $\mathcal{G}$-transformation.

2.2 Gauge symmetries and the Higgs mechanism

We are now ready to tackle the Higgs mechanism.

2.2.1 A (kind-of) simple example

Let's return to the Abelian U(1) example, but now consider a *gauge symmetry*, $\Phi(x) \to e^{-ig\theta(x)}\Phi(x)$.[4] Such an invariance can be dealt with by introducing a companion gauge field A^μ with appropriate interactions to the fields, such that the full action is gauge invariant. Assuming that A^μ transforms as $A^\mu \to A^\mu - \partial^\mu \theta$, the form of the Lagrangian can be found by a simple construction, and involves the covariant derivative $D^\mu = \partial^\mu - igA^\mu$ and the field-strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$; the Lagrangian is

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - (D^\mu \Phi)^\dagger D_\mu \Phi - V_0[\Phi^\dagger, \Phi], \tag{2.21}$$

with V_0 as in subsection 2.1.1.

Again, let's find the vacua of this theory. We consider the energy,

$$E = \int \mathrm{d}^3x \mathcal{H} = \int \mathrm{d}^3x \left(\frac{1}{2} F_{0i}^2 + \frac{1}{4} F_{ij}^2 + \frac{1}{2} D_0 \vec{\phi} \cdot D^0 \vec{\phi} + \frac{1}{2} D_i \vec{\phi} \cdot D^i \vec{\phi} + V_0[\vec{\phi}] \right), \tag{2.22}$$

and note that the first four "kinetic" terms are positive squares—their minimum contribution is zero.

In the global symmetry case we could conclude that minimizing the energy required a static configuration. When there are gauge fields involved, we then

[4] This subsection is based on Rubakov [14] and Srednicki [15].

instead have covariant derivatives, and all we can say is that the vacua are static up to a gauge transformation. That is, we can in principle remove all the spatial dependence by performing a transformation; the vacua take the form

$$A_\mu(x) = -\partial_\mu \theta, \tag{2.23}$$
$$\vec{\varphi}(x) = e^{g\theta(x)\epsilon} \vec{\phi}, \tag{2.24}$$

where $i\epsilon$ is the generator of the real 2-dimensional representation of U(1), and $\vec{\phi}$ is a static configuration.

Though this technicality will become relevant later when we want to perform a gauge-fixing, we can forget about it for now. The remaining part of the energy resides in the potential V_0, and due to its gauge invariance we can again focus on a static solution. Hence the minimum energy is

$$E_{\text{min}} = \int \mathrm{d}^3x V_0[\vec{\phi}_0] = \mathcal{V} \times V_0(\vec{\phi}_0), \tag{2.25}$$

where again $\mathcal{V}$ is the volume of space and $\vec{\phi}_0$ is a static field that minimizes V_0.

From here the analysis proceeds just as for the global symmetry; we must consider the form of V_0 and the values of the parameters. The case analysis is the same for each case where there is no expectation value for $\vec{\phi}$, with the addition of the massless vector boson.[5] There are four physical degrees of freedom, two modes of the massless vector boson and two massive scalars.

We can focus on the interesting case of $\lambda > 0, m^2 < 0$, for which the ground state fulfills $|\vec{\phi}_0| = \sqrt{-m^2/\lambda}$. Again, there is one massive and one massless scalar, but now the vector field has also acquired a mass:

$$m_A^2 = g^2 \phi_0^2. \tag{2.26}$$

Counting the degrees of freedom again, there appears to be two scalar modes and three modes of the vector boson. It seems as if the number of degrees of freedom has increased from four to five.

To resolve this issue we need to reconsider the role of the massless scalar, the Goldstone boson. By examining the square terms in the Lagrangian, we find

$$-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}m_H^2 H^2 - \frac{1}{2}m_A^2 \left(A^\mu + \frac{1}{m_A}\partial^\mu G\right)^2. \tag{2.27}$$

We can perform a change of variables by introducing $B^\mu \equiv A^\mu + \frac{1}{m_A}\partial^\mu G$, which completely removes G as a propagating field. The conclusion is that there in fact are

[5] The $m^2 = 0$ case is special, as I discuss in subsection 3.3.4.

only four degrees of freedom, three modes of the massive vector and one massive scalar.

This is the Higgs mechanism. We started with a theory that is fully gauge invariant, and we spontaneously "broke" the gauge symmetry analogously to how we would a global symmetry. In the end we found a spectrum of a massive scalar—the Higgs boson H—and a massive vector boson. The *would-be Goldstone boson* is "eaten" (technical term) by the gauge boson, which gives it the missing degree of freedom it needed to obtain a mass.

Before moving on to a general model, I want to discuss the terminology a bit. We say that these gauge theories have a "local symmetry," and we speak of breaking them with the Higgs mechanism. But the gauge transformations are not symmetries in the usual sense, they are redundancies in our description of the physical world. We can usually forget about this difference, but we might naively be lead to believe false statements if we are not careful. Above, we first figured that an additional degree of freedom had surfaced in the non-trivial vacuum. But we had not considered that there was a redundancy and that one of the degrees of freedom was not actually free. Later in this thesis this kind of reasoning will become important. For now I will note that saying a gauge invariance is spontaneously broken is a bit of a misnomer, and some physicists would rather not use this phrasing.[6] But the phrase "spontaneously broken" brings to mind the correct analogy of the global case, and I will hence continue using this phrasing in this thesis. But keep in mind that, really, *gauge symmetries are not symmetries and they cannot break* [13].

2.2.2 The general case

Consider again a compact semi-simple Lie group $\mathcal{G}$ that acts on a real vector of the theory's scalars $\vec{\phi}$ with generators $T^a, a = 1, \dots, D_\mathcal{G}$. Now there are also $D_\mathcal{G}$ vector bosons A^a_μ that transform under the adjoint representation.[7]

The same equations as in the global case now hold, but before discussing these again I want to focus on the object ${F^a}_i \equiv (T^a\vec{\phi})_i$. As we have seen, ${F^a}_i$ will play the role of the Goldstone directions in the minimum. And in general ${F^a}_i$ is always perpendicular to the corresponding field value $\vec{\phi}$. It can be thought of as a rectangular matrix with $D_\mathcal{G}$ rows and N columns, and as such it can be rewritten using a singular value decomposition,

$$ {F^a}_i = S^{ab} {M^b}_j R^T_{ij}, \tag{2.28} $$

[6]Some physicists do not like to use the phrase spontaneously broken at all, even for global symmetries. Instead they might say that the symmetries are *hidden*, because the underlying equations respect the symmetry—only the ground state breaks it.

[7]Also this subsection is based on Srednicki [15] and the two papers [16] and [17].

where S and R are orthogonal $D_{\mathcal{G}} \times D_{\mathcal{G}}$ and $N \times N$ matrices respectively; M is a $D_{\mathcal{G}} \times N$ rectangular matrix with the singular values on the diagonal: $M^b{}_j = M^b \delta^b_j$ (no summation over b). All of these matrices inherit the $\vec{\phi}$ dependence from the definition of $F^a{}_i$.

From this object we can construct two square matrices that will be relevant later, by contracting the indices of two Fs in two different ways. The matrices are

$$(m_A^2)^{ab} = F^a{}_i (F^T)_i{}^b = S^{ac} M^c{}_j (M^T)_j{}^d (S^T)^{bd}, \tag{2.29}$$

$$(m_A^2)_{ij} = (F^T)_i{}^a F^a{}_j = R_{ik} (M^T)_k{}^a M^a{}_l (R^T)_{jl}. \tag{2.30}$$

These two square matrices are, perhaps confusingly, given the same name. This is because they have the same non-zero eigenvalues, though the multiplicities of the zero eigenvalues may be different. The matrices will relate the Goldstone bosons to the appropriate gauge bosons, as we shall see. The singular value decomposition above suggests the existence of bases (one in the a space and one in the i space) that diagonalize these mass-matrices.

We will return to the matrix $(m_A^2)_{ij}$ later when we discuss the details of gauge-fixing. For now I will simply state that the matrix $(m_A^2)^{ab}$ is what shows up in the Lagrangian as a mass-matrix for the gauge bosons. We are well advised to use the basis $\tilde{A}^a_\mu \equiv S^{ab} A^b_\mu$, which diagonalizes the gauge boson mass-matrix.

We can say something further about this matrix by considering its form in a nontrivial vacuum $\vec{\phi}_0$. Then we again split up the generators into two families, the $D_{\mathcal{H}}$ unbroken t^b, and the $(D_{\mathcal{G}} - D_{\mathcal{H}})$ broken $\mathcal{T}^\alpha$. Hence we know, in the vacuum, that the zero eigenvalues of $(m_A^2)^{ab}$ must lie in the block specified by the $D_{\mathcal{H}}$ unbroken generators. The $(D_{\mathcal{G}} - D_{\mathcal{H}})$ non-zero eigenvalues form their own block given by the broken generators.

The general conclusion is that the unbroken subgroup $\mathcal{H}$ still comes with $D_{\mathcal{H}}$ massless gauge bosons. And the rest of the group $\mathcal{G}$ that is broken by the Higgs mechanism comes with $(D_{\mathcal{G}} - D_{\mathcal{H}})$ massive vector bosons. From the analysis of the global symmetry case we know that this is the same number as the number of would-be Goldstone bosons.

The counting of degrees of freedom hence still works out, but the argument is somewhat more complicated than in the simple Abelian case. I will return to it when I discuss gauge-fixing in chapter 3.

Recommended readings

Valerij Rubakov, *Classical Theory of Gauge Fields* [14]. This textbook does an amazing job balancing mathematical rigor and physical motivations. It treats spontaneous symmetry breaking and the Higgs mechanism in great detail.

3. Symmetries and quantum effects

> "You have to be realistic about these things."
> —Logen Ninefingers
>
> Joe Abercrombie, *The First Law*

The calculations and arguments in chapter 2 all related to classical symmetries. In this chapter I will explore what can happen when quantum effects are included.

3.1 The path integral

The transition from a classical theory to a quantum one can most intuitively be understood from the path integral formalism.

In classical physics we write down an action in terms of our degrees of freedom, following certain rules and regulations. Then we minimize the action to find physical solutions. This is the principle of least action, and has been used to great success to derive field equations for electromagnetism and general relativity, and for deriving classical trajectories for particles present in such force fields.

The principle of least action, useful as it is, might seem a bit arbitrary. To clarify what I mean by that, let's consider the *path integral* in field theory. The path integral of a field Φ in the presence of a classical source J is, schematically,

$$Z[J] = \int \mathrm{D}\Phi e^{i\frac{1}{\hbar}\left(S[\Phi]+\int J\Phi\right)}. \tag{3.1}$$

This is a *functional integral*, with a *functional integration measure* $\mathrm{D}\Phi$. From this definition we can now "derive" the principle of least action. If we treat this path integral just like any other complex contour integral, then we can approximate it using the stationary phase approximation. For the following demonstration, I will set the classical source to zero, $J = 0$. But a similar analysis is applicable if J is nonzero.

If we have a solution Φ_0 that extremizes the action, $\delta S|_{\Phi_0} = 0$, then we can expand the action around this solution,

$$S[\Phi] = S[\Phi_0] + \cancel{\delta S|_{\Phi_0}(\Phi-\Phi_0)} + \frac{1}{2}\,\delta^2 S\Big|_{\Phi_0}(\Phi-\Phi_0)^2 + \dots \tag{3.2}$$

Using this expansion inside the path integral—the stationary phase approximation—we get that the logarithm of the path integral schematically obeys [18]

$$\log Z[0] = i\frac{1}{\hbar}S[\Phi_0] - \frac{1}{2}\log\det\left[\delta^2 S\right]\Big|_{\Phi_0} + \mathcal{O}(\hbar). \tag{3.3}$$

The classical solution Φ_0 dominates the path integral, and then there is an array of quantum corrections. We can imagine taking the limit $\hbar \to 0$, picking out the classical result. In this formalism, the principle of least action is just the statement that the classical physics dominate the path integral.

Observables can then be derived from the path integral, generally by taking functional derivatives with respect to the classical sources. For a typical theory with interactions, we need to approximate the path integral in a fashion analogous to the expansion above. Perturbation theory is typically done with Feynman diagrams, ordered by the number of loops. But sometimes this expansion breaks down, and it is more useful to use another expansion parameter. I will return to this later.

3.1.1 The path integral and symmetries

If we have a symmetry of the classical action, is it guaranteed to be a symmetry of the path integral? In other words, do quantum effects preserve symmetries? Not always.

Again, we can gain som intuition by thinking of the path integral as just another integral.[1] If we consider an ordinary integral,

$$\int \mathrm{d}x f(x), \tag{3.4}$$

and we know that the function $f(x)$ is invariant under some transformation $x \to x'$, that is $f(x) = f'(x')$, then the only way for the whole integral to be invariant is if the integration measure is invariant as well: $\mathrm{d}x' = \mathrm{d}x$.

Hence if the classical theory is invariant under some group $\mathcal{G}$, but the path integral integration measure $\mathrm{D}\Phi$ transforms under $\mathcal{G}$, then we say that the symmetry is *anomalous*.

A global anomalous symmetry is not a problem. But, as I emphasized before, gauge symmetries are not really symmetries—they are redundancies in our description. We are now ready to consider this statement in further detail.

Imagine that the theory specified by $\mathcal{L}$ has a local U(1) symmetry with an accompanying gauge boson A^μ. This vector field has four degrees of freedom, but we know that any physical manifestation of this field only has two. Somewhere along the way we must compensate for the extra redundancy that we have introduced.

[1] This subsection is based on Srednicki [15].

Think of the path integral. A part of the integration measure will now be $\mathrm{D}A \sim \mathrm{D}A_0\mathrm{D}A_1\mathrm{D}A_2\mathrm{D}A_3$, which shows our overcounting. To compensate for this overcounting we perform a trick that is attributed to Faddeev and Popov [19]. The delicate part in our treatment of this overcounting is to ensure that we extract the physical contributions of A^μ, and quotient away the rest. But this is hard because of the redundancy; which parts that are extraneous depends on which gauge we are in.

Again, the intuition comes from a regular integral. Imagine that we are integrating over x and y, but we know that the integrand does not depend on y,

$$\int \mathrm{d}x\mathrm{d}y f(x). \tag{3.5}$$

We recognize that the integral over y is redundant, and we can divide the expression above by $V = \int \mathrm{d}y$, which in effect drops the integral over y. If the situation is more complicated and we are integrating over some 2-dimensional space in which we know there is a redundancy, then the process is not as simple. Instead we can imagine starting with $\int \mathrm{d}x f(x)$, dividing by V, inserting an integral over y and a delta function that extracts the correct contribution,

$$\frac{1}{V}\int \mathrm{d}x\mathrm{d}y \det\frac{\partial G}{\partial y}\delta(G)f(x), \tag{3.6}$$

where $G(x,y)$ is a cleverly chosen function.

In the path integral formalism, we start with an integral over $\mathrm{D}A$ and then insert a delta functional and a functional determinant,

$$Z[J] \propto \int \mathrm{D}A \det\left(\frac{\delta G}{\delta\theta}\right)\delta(G)e^{iS}, \tag{3.7}$$

and we call G the *gauge-fixing* function. Here θ refers to a redundant degree of freedom. The trick now is to rewrite this functional determinant and the delta functional to something which we can use to calculate observables.

The functional determinant and the gauge-fixing delta functional are incorporated into our theory by introducing unphysical degrees of freedom, *Faddeev-Popov ghosts* η and $\overline{\eta}$, and by including new parts in the action,

$$Z[J] \propto \int \mathrm{D}A\mathrm{D}\eta\mathrm{D}\overline{\eta} e^{iS+iS_{\mathrm{gh}}+iS_{\mathrm{g.f.}}}, \tag{3.8}$$

where S_{gh} is the ghost action and $S_{\mathrm{g.f.}}$ is the gauge-fixing action.

Loosely speaking, we can think of ghosts as anti degrees-of-freedom; we can use them to cancel other unphysical contributions. Because they must cancel modes of the gauge boson, they also need to obey Bose-Einstein statistics. But, at the same time, for the construction in equation (3.7) to work out, they also need to anticommute. This means that they appear to violate the spin-statistics theorem. Even though this is unintuitive, it is not really an issue. After all, the ghosts are not physical. I dig into this statement, and show some possible forms of the new actions S_{gh} and $S_{\text{g.f.}}$, in the next subsection.

3.1.2 BRST symmetry and gauge-fixing

After appending extra degrees of freedom to our theory, and extra pieces to the Lagrangian, we might be concerned that we are distorting the physical picture. Though we are right to be worried, it is possible to show that any unphysical degrees of freedom cannot contribute to a physical observable.[2] Our original action S is invariant under gauge transformations. But the gauge-fixing action $S_{\text{g.f.}}$ and ghost action S_{gh} are cooked up to explicitly fix a particular gauge, and they must break gauge invariance. But any theory that has gauge invariance also has an additional global symmetry, called BRST [20].

The BRST symmetry is analogous to the original gauge symmetry, but with Grassman (anticommuting) valued parameters; call such a transformation $\mathbf{s}$. This transformation is by construction nilpotent, $\mathbf{s}^2 = 0$.

With this extra symmetry in mind, it is possible to assign transformation rules for the ghosts and simultaneously construct $S_{\text{g.f.}}$ and S_{gh} such that the complete action is BRST invariant in the end. The trick is to construct the new parts as a BRST transformation of some operator $\mathcal{O}$, $S_{\text{g.f.}} + S_{\text{gh}} = \mathbf{s}\mathcal{O}$. The nilpotence of $\mathbf{s}$ then assures the full theory is still BRST invariant.

The reason that we care about BRST invariance is that it protects us from accounting mistakes. We can assuage the worries we had about the unphysical states and interactions that now populate our theory. The precise statement is that *physical states are in the cohomology of* $\mathbf{s}$. What this means is that if $|\psi\rangle$ is a physical state, then $\mathbf{s}|\psi\rangle = 0$ and $|\psi\rangle$ is physically equivalent to some state $|\psi\rangle + \mathbf{s}|\chi\rangle$. Any state that can be written as a BRST transformation of some other state, such as $\mathbf{s}|\chi\rangle$, can only contain fluctuations of unphysical degrees of freedom.

These statements are robust under time-evolution. Meaning, if we start with a physical state, then we must end with a physical state. For calculations in quantum field theory, this means that although we use propagators that include all modes of the gauge bosons, and loops including ghosts, in the end all unphysical contributions will cancel.

At this stage it is good if the niggling worry starts creeping back up again. What if we make a mistake in the calculation and the ghosts somehow actually contribute? How would we know? One way is to generalize the operator $\mathbf{s}\mathcal{O}$ from which we constructed the new parts of the action. In this way we can introduce some parameters, let's say ξ^a, that in the end cannot affect our results. Any physical observable

[2] This subsection is also based on Srednicki [15].

should be independent of ξ^a. If we find that it is not, this means that we have made a mistake somewhere.

A simple example

Let's consider how this could be done in the Abelian Higgs model. I won't go through the details of the BRST transformation; a suggestion for gauge-fixing and ghost Lagrangians are

$$\mathcal{L}_{\text{g.f.}} = -\frac{1}{2\xi}\left(\partial^\mu A_\mu - \tilde{\xi} g \phi_i \epsilon_{ij} R_j\right)^2, \tag{3.9}$$

$$\mathcal{L}_{\text{ghost}} = -\partial_\mu \overline{\eta} \partial^\mu \eta - \tilde{\xi} g^2 \phi_i (\phi_i + R_i) \overline{\eta} \eta \tag{3.10}$$

where ξ and $\tilde{\xi}$ are gauge-fixing parameters. Here I use the notation of [17]. To translate to the basis I used in chapter 2, use $R_1 = H, R_2 = G$ and $\phi_1 = \phi_0, \phi_2 = 0$.

Though it is important to recognize that the above gauge-fixing procedure has two parameters ξ and $\tilde{\xi}$, calculations in this theory are much more tractable when $\tilde{\xi} = \xi$. This special case is known as the background-field R_ξ gauge.[3]

In this gauge the Goldstone fields get a contribution to its propagator and the mass is now $m_G^2 + \xi m_A^2$. The ghosts interact with the scalars, and they also have a mass ξm_A^2. The longitudinal mode of the gauge boson has propagator $\xi \frac{1}{p^2 + \xi m_A^2} \frac{p^\mu p^\nu}{p^2}$. When we calculate an observable, ξ will enter our expressions through these terms. In the end all such contributions must cancel.

The general case

Though the gauge-fixing can be done in many ways, I suspect it is hard to do better than the generalized $R_{\xi,\tilde{\xi}}$ gauges [17].[4]

In principle, the requirements for constructing the gauge-fixing terms are not very strict—there are many different possible choices. We know that the terms in the end must be BRST invariant. But from a practical point of view, we want the new terms to simplify our calculations as much as possible.

Borrowing from [17], generalized slightly, we can use

$$\mathcal{L}_{\text{g.f.}} = -\frac{1}{2}\left(\partial^\mu A^a_\mu + i\tilde{\xi}_a \phi_i g^a T^a_{ij} R_j\right)\left(\xi^{-1}\right)_{ab}\left(\partial^\mu A^b_\mu + i\tilde{\xi}_b \phi_i g^b T^b_{ij} R_j\right), \tag{3.11}$$

$$\mathcal{L}_{\text{ghost}} = -\partial_\mu \overline{\eta}^a \partial^\mu \eta^a + g_a f^{abc} \partial^\mu \overline{\eta}^a \eta^b A^c_\mu + \tilde{\xi}_a g_a i T^a_{ij} \phi_j g_b i T^b_{ik} (\phi_k + R_k) \overline{\eta}^a \eta^b. \tag{3.12}$$

Here ξ is a symmetric $D_\mathcal{G} \times D_\mathcal{G}$ matrix, with $D_\mathcal{G}(D_\mathcal{G}+1)/2$ independent components.

This gauge-fixing has several advantageous properties that I will list here. (1), it is renormalization invariant if the running of both ξ and $\tilde{\xi}$ is taken into account [17].

[3] As Srednicki points out, "the R stands for renormalizable. The ξ stands for ξ."

[4] The $\tilde{\xi}$ stands for $\tilde{\xi}$.

(2), it is what Fukuda and Kugo [21] dubs a *good gauge*, meaning that it does not induce any spurious minima. (3), it is invariant under the global remnant of the gauge group $\mathcal{G}$. This is good because it allows us to consistently change bases.

To see further advantages, we first need to discuss a few things. We know that the non-trivial vacua of the Higgs mechanism are related by gauge transformations, $\vec{\phi}(x) = e^{-ig\theta^\alpha(x)\mathcal{T}^\alpha}\vec{\phi}_0$, where $\mathcal{T}$ are the broken generators. We can choose a representative of this vacuum such that the rest of the calculations are simpler. The global invariance of the Lagrangian and the specific form of the gauge-fixing function above ensures that we can do this without loss of generality.

When we pick such a representative, we also automatically pick out the corresponding perpendicular Goldstone directions $\mathcal{T}^\alpha\vec{\phi}$. Any basis change we perform will mix fields in such a way that the Goldstone field afterwards still are perpendicular to the vacuum representative.

This property, together with the global invariance of the full Lagrangian, implies that we can specify a basis without losing any information. In particular, we can choose the basis suggested by the singular value decomposition of $F^a{}_i$.

At this point it is helpful to change notation to that introduced in [17]. In the mass basis of the vector fields we use the bold indices $\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots$. In the subspace of the massive vectors we use $A, B, C, \ldots$; in the massless subspace we use $a, b, c, \ldots$ Similarly for the scalar bosons; in this basis we use $\mathbf{j}, \mathbf{k}, \mathbf{l}, \ldots$ for the scalars $R_\mathbf{j}$. In the "non-Goldstone" subspace we use the indices $i, j, k, \ldots$, and in the Goldstone subspace we use $A, B, C, \ldots$ and represent the fields as G_A. That is, we have the fields

$$\{A^\mathbf{a}_\mu\} = \{Z^A_\mu, A^a_\mu\}, \qquad \{R_\mathbf{j}\} = \{G_A, R_j\}. \tag{3.13}$$

We should now rethink the matrix ξ_{ab}. It is helpful if ξ_{ab} is diagonal in the mass basis, that is $\xi_{\mathbf{ab}} = \mathrm{diag}\{\xi_A, \xi_a\}$. With this in mind, it is in general not necessary to specify the whole matrix ξ_{ab}, and just use the eigenvalues in the end. But note that if one wishes to compare between different bases, the full form including the matrix must be used.

In the end, the gauge-fixing Lagrangian becomes

$$\mathcal{L}_{\text{g.f.}} = -\frac{1}{2\xi_A}\left(\partial^\mu Z^A_\mu - \tilde{\xi}_A M_A G_A\right)^2 - \frac{1}{2\xi_a}\left(\partial^\mu A^a_\mu\right)^2. \tag{3.14}$$

Specifying to $\tilde{\xi}_A = \xi_A$ leaves us with $D_\mathcal{G}$ gauge-fixing parameters, one for each real vector boson. These are the background-field R_ξ gauges for a general model; they are free from kinetic mixing between the Goldstone bosons and the longitudinal gauge bosons.

In this simplified gauge-fixing, the scalars will get a contribution to their masses as $\xi(m_A^2)_{ij}$, where $(m_A^2)_{ij}$ is the matrix introduced in equation (2.30). To find out what this means for the scalars, we can multiply equation (2.18) by $F^a{}_k$ (summing over a), giving

$$(m_A^2)_{ij}\partial^2_{jk}V_0\Big|_{\vec{\phi}} = T^a_{ij}\phi_j T^a_{kl}\partial_l V_0. \tag{3.15}$$

In particular, at the vacuum we have that

$$(m_A^2)_{ij}\partial^2_{jk}V_0\Big|_{\vec{\phi}_0} = 0. \tag{3.16}$$

This equation tells us that the non-zero eigenvalues of $\xi(m_A)^2_{ij}$ (the gauge-fixing induced scalar mass matrix) and those of the scalar mass matrix $\partial^2_{ij}V_0$ live in different subspaces [15, 16, 17]. In other words, only the Goldstone bosons get ξ-dependent masses.

In the simple cases, G_A can also be a mass eigenstate on its own. But this only occurs when the real scalars transform under an irreducible representation of the gauge group. In the more complicated cases (such as in the Two Higgs Doublet Model), the non-Goldstone scalars and Goldstone scalars actually mix outside the vacua. At the vacua, equation (3.16) applies and the scalar subspaces separate again.

I will return to discussing gauge-fixing in subsection 3.3.3.

3.2 More on anomalies

As I mentioned briefly in subsection 3.1.1, an anomaly arises when the path integral does not respect the same symmetries as the classical Lagrangian. In the path integral language this means that the integration measure is not invariant under the symmetry transformation. In practice, anomalies of internal symmetries arise in theories with chiral fermions: if a 2-component fermion ψ transforms under some group, $\psi \to \psi'$, then the measure transforms as $\mathrm{D}\psi \to \mathrm{D}\psi'$.

Non-chiral transformations then always cancel—leaving the measure invariant. But if the symmetry is chiral, then there might be a trace left of the transformation. For global symmetries this is not an issue.

For gauge symmetries, on the other hand, anomalies are disastrous. Gauge symmetries are not symmetries in the usual sense, they are more like redundancies. As I have discussed before, the distinction is important. For such gauge anomalies to exist would break our counting of degrees of freedom, and invalidate our calculations.

At the perturbative level, anomalous gauge symmetries reveal themselves in loop diagrams with external gauge bosons and chiral fermions running in the loops. For Abelian symmetries it is enough to consider *triangle diagrams*—-diagrams with one chiral fermion loop, and three external gauge bosons. Such diagrams carry anomalous factors that depend on the charges of the fermions in the loops. As an example, consider a theory with a U(1) symmetry, an accompanying Z', and some chiral fermions with U(1) charges z_i. Summing all triangle diagrams with external Z' bosons gives a result proportional to

$$\mathcal{A}^3 \equiv \sum_i z_i^3. \tag{3.17}$$

If this factor is non-zero, then we say that there is a $[\mathrm{U}(1)]^3$ anomaly. On the other hand, if the charges line up in such a way that $\mathcal{A}^3 = 0$, then the theory is anomaly free.

In a gauge theory with chiral fermions, such as the Standard Model, there are a number of different possible anomalies. For the theory to be consistent all of them must exactly cancel. This does happen in the Standard Model, with no great explanation as to why. It is a fact that invites speculation: perhaps there is a larger symmetry group that is broken at some higher scale? The answer as to why the Standard Model is anomaly free is not settled.

Because the Standard Model contains chiral fermions, we need to be careful when considering extensions of the Standard Model. For the theory to make sense we need all gauge anomalies to cancel, and care must be taken to ensure that this is the case. In paper I we considered a simple extension of the Standard Model with an extra U(1) gauge group, taking care to cancel anomalies. The simplest way to cancel the anomalies is by introducing three right-handed fermions neutral under the Standard Model gauge groups [4]—three right-handed neutrinos. As I discussed in subsection 1.3.1, it is also possible to give the neutrinos Majorana masses. This minimal model hence contains a convenient solution to the neutrino mass problem.

Cancellation Mechanisms

Having established the necessity of cancelling anomalies in theories with chiral fermions, some questions may arise.

As an example, if we did not know of the top-quark's existence, we would think our current theory to be anomalous. This is not a completely implausible scenario, because the top quark is heavier than all the other particles in the Standard Model.

But how can this be? From the effective field theory perspective we would expect the theory to make sense when we integrate out the heaviest particles. This tells us that some extra care has to be taken when we construct such effective theories.

To resolve this question we can again turn to the path integral. We think of the anomaly as the measure transforming under the gauge transformation. The problem is that the Lagrangian is invariant by construction, and hence there is an uncancelled transformation left in the end.

This perspective also suggests an alternative scenario, where the Lagrangian *does* transform under such transformations, in just the way to cancel the transformation of the measure! The path integral is then invariant again. In reality the story is a bit more complicated, but by following this line of argument one can carefully construct an effective theory that is not gauge-invariant at the Lagrangian level, but still produces gauge-invariant observables.

This is known as the Green-Schwarz mechanism, and was first used to cancel anomalies in string theory [22]. But the construction works equally well for quantum field theory, though there the terms can be due to heavy fermions instead of string theory effects. In paper II we explore the collider phenomenology of such a theory. In particular, we consider the same theory as described above—the Standard Model with an extra U(1) gauge symmetry—but without any neutrinos. We worked out the details of this theory, based on the formalism developed in [23].

3.3 The effective potential

We have seen that the classical potential V_0 plays a key role in understanding spontaneous symmetry breaking and the Higgs mechanism. But how is this picture changed by quantum effects?

To find the classical minimum and the corresponding spectrum, we had to first consider general static field-values $\vec{\phi}$. We considered the classical potential as a function of these real numbers and then minimized it to find the physical solution. In other words, we had to go "off-shell," outside the solution to the classical equations of motion, to actually find the solution in the first place.

For quantum effects there is a similar picture. It is possible to formulate an *effective potential*, which has the quantum corrected static background field as its minimum. See [24] for a detailed analysis and review of the effective potential.

We obtain the effective action $\Gamma[\phi]$ by performing a Legendre transform of the functional $W[J] = i \log Z[J]$, trading the classical source J for a static field ϕ. It can be shown that a minimum of the effective action corresponds to $J = 0$ [18].

To find the pattern of symmetry breaking in these quantum theories, we can restrict ourselves to static fields (up to gauge transformations, as usual). Then we find that the effective action is proportional to the effective potential $V(\phi)$,

$$\Gamma[\phi] = -\mathcal{V}\mathcal{T} \times V(\phi), \tag{3.18}$$

with $\mathcal{V}\mathcal{T}$ the spacetime volume. We can then minimize the effective potential to find the true quantum corrected minimum $\vec{\phi}_{\text{min}}$, and the accompanying minimal energy density $V_{\text{min}} \equiv V(\vec{\phi}_{\text{min}})$.

At this stage we have to consider how to calculate $V(\phi)$. In general, exact results are not available and we must turn to perturbation theory.[5] From the sketch in

[5] Or lattice calculations.

equation (3.3), we infer that the solution will be something like

$$V(\phi) = V_0(\phi) + V_1(\phi) + V_2(\phi) + \dots, \qquad V_1(\phi) \sim -i\frac{1}{2}\log\left[\det D^{-1}\big|_{\phi}\right], \quad (3.19)$$

where D represents a propagator, and det is a functional determinant. The higher order corrections $V_2(\phi), \dots$ are better represented as Feynman diagrams [24].

3.3.1 Symmetries and the effective potential

Global symmetries

If the Lagrangian and the path-integral measure are invariant under some internal and global group $\mathcal{G}$, then we expect the effective potential to also be invariant under such transformations. Much of the analysis performed in the classical case in chapter 2 now carries over to the effective potential. In particular, we can upgrade equation (2.18) to a quantum corrected version,

$$\partial^2_{ij} V \; (T^a \vec{\phi})_j = T^a_{ij} \partial_j V. \tag{3.20}$$

Evaluating this at the vacuum $\vec{\phi}_{\text{min}}$, we find a quantum version of Goldstone's theorem,

$$\partial^2_{ij} V\Big|_{\vec{\phi}_{\text{min}}} (\mathcal{T}^\alpha \vec{\phi}_{\text{min}})_j = 0. \tag{3.21}$$

In the completely general case, one should consider that the effective potential might have different broken generators than the classical potential. That is, the quantum effects might change the picture of symmetry breaking. In practice this only happens for very particular theories where regular perturbation theory breaks down—see subsection 3.3.4 for an example. In the standard cases, we can relax and consider the same directions in field space as for the classical potential.

Local symmetries

For local symmetries the situation is more tricky. Remember that to perform the path integral we needed to fix a gauge, spoiling the gauge invariance of the theory.

This gauge-fixing does not have to respect the global part of the gauge symmetry in general. If it does not, then there are complications for general off-shell values of $\vec{\phi}$. As an example, we are not free to upgrade equation (2.18) to the effective potential for general field values. Even though on-shell, in the minimum, the Goldstone theorem will always apply.

But since we have to go off-shell to find the minimum in the first place, it is more convenient to work with a gauge-fixing that is invariant under the global transformations. This is another reason why the gauge-fixing described in subsection 3.1.2 is highly recommended.

3.3.2 The 1-loop potential

At one loop, the effective potential can be calculated either by a functional logarithmic determinant [25], or by a particular perturbative expansion that involves infinite "prototype" Feynman diagrams [26]. Both derivations offer their own insights, but I will not detail either of them here.

Using $\overline{\text{MS}}$ and dimensional regularisation with $d = 4 - 2\epsilon$, μ as the $\overline{\text{MS}}$ renormalization scale, and the short-hand $\int_p \equiv \left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon \int \frac{\mathrm{d}^d p}{(2\pi)^d}$, the contribution for a degree of freedom with squared mass x to the effective potential is

$$\begin{aligned} \mathbf{f}(x) &\equiv \frac{1}{2}\int_p \log\left[p^2 + x\right] = -\frac{1}{2}\left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon \frac{\Gamma(-\frac{d}{2})}{(4\pi)^{d/2}} x^{d/2} \\ &= \frac{1}{(16\pi)^2}\left(-\frac{x^2}{4\epsilon} + \frac{x^2}{4}\left(\log\left[\frac{x}{\mu^2}\right] - \frac{3}{2}\right) + \mathcal{O}(\epsilon)\right). \end{aligned} \tag{3.22}$$

The 1-loop basis function $\mathbf{f}(x)$ in this equation is not renormalized; the $1/\epsilon$ terms can be removed with local counterterms.

When the propagators of a theory is diagonal, then each mode with squared mass X will contribute simply as $\mathbf{f}(X)$. But, when the propagators of particles are not diagonal—which they are not in the general case—the logarithmic determinant has to be separated as a trace over logarithms of eigenvalues: $\log\det D^{-1} \sim \mathrm{Tr}\log D^{-1}$.

This seems like it might make an analysis of possible ξ-dependence of a 1-loop calculation difficult. But, as is shown in [16], it is still possible to make a straightforward analysis in background-field R_ξ-gauges. Here I repeat their argument. The general structure of the 1-loop potential, in a generic non-Abelian gauge theory with this gauge-fixing, looks like

$$\begin{aligned} \frac{1}{2}\int_p \Big(&\mathrm{Tr}\log\left[p^2 + M^2_{ij} + \xi(m^2_A)_{ij}\right] + (d-1)\,\mathrm{Tr}\log\left[p^2 + (m^2_A)^{ab}\right] \\ &+ \mathrm{Tr}\log\left[p^2 + \xi(m^2_A)^{ab}\right] - 2\,\mathrm{Tr}\log\left[p^2 + \xi(m^2_A)^{ab}\right]\Big). \end{aligned} \tag{3.23}$$

Here, the first log corresponds to the scalars, the second to the transverse gauge bosons, the third to longitudinal gauge bosons, and the fourth to ghosts. The three ξ-dependent logarithms include unphysical degrees of freedom. For general off-shell values of ϕ, the logarithms do not separate and there is a ξ dependence. But, at the broken minimum ϕ_0, the matrices $M^2_{ij} \equiv \partial^2_{ij} V_0$ and $(m^2_A)_{ij}$ are mutually diagonalizable—their non-zero eigenvalues live in different subspaces. The logarithm hence separates at the minimum,

$$\mathrm{Tr}\log\left[p^2 + M^2_{ij} + \xi(m^2_A)_{ij}\right]\Big|_{\phi_0} = \mathrm{Tr}\log\left[p^2 + M^2_{ij}\right]\Big|_{\phi_0} + \mathrm{Tr}\log\left[p^2 + \xi(m^2_A)_{ij}\right]\Big|_{\phi_0}. \tag{3.24}$$

A cancellation now occurs in equation (3.23), because the matrices $(m_A^2)_{ij}$ and $(m_A^2)^{ab}$ have the same non-zero eigenvalues. All ξ-dependence cancels, in a conceited effort between the Goldstones, the longitudinal modes of the gauge bosons, and the ghosts.

3.3.3 Gauge dependence and Nielsen identities

As noted above, the effective potential is gauge-dependent for general field values ϕ. This can make it tricky to extract observables from it. It can help to understand in what way the effective potential is gauge-dependent, which can be understood by the help of *Nielsen identities* [27]. These non-perturbative identities can be thought of as Ward identities for the effective potential (they are derived in a similar manner), and they take the form

$$\partial_\xi V(\phi,\xi) = -C(\phi,\xi)\partial_\phi V(\phi,\xi). \tag{3.25}$$

Here $C(\phi,\xi)$ is a *Nielsen coefficient* that is calculable in perturbation theory. Se the textbook by Das [28] for a pedagogical derivation of the identity, including a calculation of $C(\phi,\xi)$ in a simple model.[6]

We expect that the energy density of any particular vacuum is an observable—independent of ξ—because it can in principle be measured. The Nielsen identities confirm this view, as can be seen by evaluating the equation in the minimum $\vec{\phi}_{\text{min}}$:

$$\partial_\xi V(\vec{\phi}_{\text{min}},\xi) = 0. \tag{3.26}$$

This is a neat story that seems to suggest that we can simply find V_{min} by minimizing V and then evaluating it at the minimum. But here we actually have to be a bit careful, due to perturbation theory. This is emphasized in [16], which I base the following analysis on—but see also the earlier papers by Fukuda and Kugo [21] and Laine [29].

In practice, V is calculated in some perturbative expansion that we truncate at some order, say one loop: $V = V_0 + \hbar V_1(\phi) + \mathcal{O}\left(\hbar^2\right)$.[7] If we would minimize this function numerically, then we would find some minimum $\phi'(\xi)$ that depends on ξ. And evaluating $V'_{\text{min}} = V(\phi') = V_0(\phi') + \hbar V_1(\phi')$ we would find a residual gauge dependence. This seems to contradict the Nielsen identity. So what went wrong?

The issue is that we are not taking the powers of perturbation theory seriously. We are treating $V_1(\phi)$ and $V_0(\phi)$ on the same footing, even though they clearly are not—they are separated in size by the power counting. Indeed, by expanding the

[6] Das' textbook really covers finite temperature field theory. But, as I mention in chapter 4, the formalism is essentially the same.

[7] Here $\hbar$ just counts loops—it is in principle not related to the reduced Planck constant.

Nielsen identity in perturbative powers, and equating order by order, we find that

$$\mathcal{O}\left(\hbar^0\right): \qquad \partial_\xi V_0 = -C_0 \partial_\phi V_0 \qquad \Longrightarrow C_0 = 0, \tag{3.27}$$

$$\mathcal{O}\left(\hbar^1\right): \qquad \partial_\xi V_1 = -C_0 \partial_\phi V_1 - C_1 \partial_\phi V_0 \qquad \Longrightarrow \partial_\xi V_1 = -C_1 \partial_\phi V_0, \tag{3.28}$$

$$\vdots$$

which suggests that the 1-loop potential is actually gauge invariant when evaluated in the tree-level extrema—and not in the minimum found from minimizing $V_0 + \hbar V_1$.

Looking back at equation 3.24 and the reasoning around it, this fact should not surprise us. The contributions from the unphysical degrees of freedoms—Goldstones, longitudinal modes, and ghosts—only cancel in two specific field points. At the origin we have that $m_A^2 = 0$, which removes the ξ-dependence. At the tree-level minimum ϕ_0 the Goldstone masses are zero and the logarithms separate such that the ξ-dependent contributions cancel.

I think of this fact in the following way. A gauge field A_μ has four degrees of freedom, the real scalar G has one, and the ghosts $\eta, \overline{\eta}$ have two. We can now consider going on-shell in two different ways. At the origin in field space (which is a maximum, and hence unstable fluctuations occur) there is one massless vector with two degrees of freedom, one "massive" scalar (the square mass is negative—hence the quotes), and two massless ghosts. The accounting of degrees of freedom reads

$$\phi = 0: \qquad 4 + 1 - 2 = 2 + 1 + 0. \tag{3.29}$$

At the broken minimum, the massive scalar has three degrees of freedom, and the G scalar must join the ghosts as unphysical,

$$\phi = \phi_0: \qquad 4 + 1 - 2 = 3 + 0 + 0. \tag{3.30}$$

But anywhere outside these extrema we cannot really assign a physicality to the different modes, and anything goes. There are $4 + 1 + 2 = 7$ degrees of freedom, and they all contribute to the effective potential. When we actually do approach a vacuum of the tree-level potential, then the accounting has to resolve into the usual one.

This distinction is relevant because we have to go off-shell to find a quantum corrected vacuum expectation value $\vec{\phi}_{\min}$, and we must then take care with our accounting. In other words, we must always expand around the tree-level vacua, where the accounting makes sense. That is, we expand

$$\vec{\phi}_{\min} = \phi_0 + \hbar\phi_1 + \hbar\phi_2 + \ldots, \tag{3.31}$$

and solve the minimum condition order-by-order in perturbation theory,

$$0 \stackrel{!}{=} \partial V(\vec{\phi}_{\text{min}}) = \partial V_0|_{\phi_0} + \hbar\,(\partial V_1 + \phi_1 \partial^2 V_0)\Big|_{\phi_0} + \mathcal{O}\left(\hbar^2\right), \tag{3.32}$$

$$\mathcal{O}\left(\hbar^0\right): \quad \Longrightarrow \partial V_0|_{\phi_0} \stackrel{!}{=} 0, \tag{3.33}$$

$$\mathcal{O}(\hbar): \quad \Longrightarrow \phi_1 \quad \stackrel{!}{=} - \left.\frac{\partial V_1}{\partial^2 V_0}\right|_{\phi_0}, \tag{3.34}$$

$$\vdots \tag{3.35}$$

And this expansion should then be used when evaluating V_{min},

$$V_{\text{min}} = V(\vec{\phi}_{\text{min}}) = V_0(\phi_0) + \hbar V_1(\phi_0) + \mathcal{O}\left(\hbar^2\right). \tag{3.36}$$

This expression is now gauge invariant, order-by-order in $\hbar$.

There is another, slightly more technical, interpretation as to why this careful expansion is necessary. The gauge-dependence of the effective potential can in the diagrammatic expansion be traced back to the lack of 1-particle-reducible diagrams. As an example, in scalar electrodynamics such diagrams are necessary to cancel the gauge dependence of the four-point function: the gauge dependence of the loop corrections of the external legs cancel that of the loop corrected vertex.

The construction of the effective potential removes all such 1-particle-reducible diagrams [25]. A particular class of such diagrams have *tadpoles* inserted, which are subdiagrams with one external leg. Fukuda and Kugo [21] showed that the $\hbar$-expansion detailed above corresponds to putting a subset of all tadpoles back into the diagrams (even in gauges such as the R_ξ-gauge). Then ϕ_1 corresponds to the 1-loop tadpoles, ϕ_2 to 2-loop, and so on. With the tadpoles reinserted, the end-result is again gauge invariant.[8]

With all of these considerations in mind, let's take an additional look at the Nielsen identity in equation (3.25). We can think of this as a first-order partial differential equation analogous to that which describes advection, and we can find *characteristics* in the ξ–ϕ plane which keep the effective potential constant. Any such characteristic $\phi(\xi)$ must satisfy

$$\partial_\xi \phi(\xi) = C(\phi, \xi). \tag{3.37}$$

In particular, the minimum $\vec{\phi}_{\text{min}}$ must be such a characteristic. This tells us that the vacuum expectation value of the field is a gauge dependent quantity—and hence

[8]The diagrams on the cover of this thesis are those of the effective potential evaluated to two loops—with the tadpoles reinserted.

not an observable. In terms of our perturbative expansion, we would find that

$$\mathcal{O}\left(\hbar^0\right): \quad \partial_\xi \phi_0 = 0, \tag{3.38}$$

$$\mathcal{O}\left(\hbar^1\right): \quad \partial_\xi \phi_1 = C_1(\phi_0, \xi), \tag{3.39}$$

$$\vdots \tag{3.40}$$

This tells us that the quantum correction ϕ_1 is ξ-dependent. This dependence exactly cancels the dependence of other terms, at higher orders in the expansion of V_{min}. The vacuum energy density is finite and gauge invariant, in the end.

There is an unfortunate confusion in the literature, that seems to have persisted since the original and early papers on the effective potential. Many authors claim that the background-field R_ξ gauges are not suitable for studies of the effective potential, because it introduces a fictitious dependence on the background field ϕ—see for example [30, 31, 32, 33, 34, 35, 24]. But I think this statement is *too* cautious. Fukuda and Kugo [21] showed that any spurious ϕ-dependence that is introduced in this gauge-fixing does not matter in the end—as long as one consistently reinserts the tadpoles in the calculation of the energy density. In other words, physical observables are gauge invariant in the $\hbar$-expansion as detailed above. Though, as explained in [17], there is a price to be paid: the background-field R_ξ gauges are not renormalizable at general field values ϕ. But this is also a non-issue once the physical limit is taken with the $\hbar$-expansion. Background-field R_ξ gauge works well as long as one calculates physical observables.

As a final comment, the perturbatively calculated Nielsen coefficient, $C_0(\vec{\phi}_{\text{min}}, \xi)$, $C_1(\vec{\phi}_{\text{min}}, \xi), \ldots$, and the perturbative vacuum expectation value $\phi_0, \phi_1, \ldots$, can become infinite due to infrared divergences. But again we must remember that these are not physical observables. Even though these divergences are inconvenient, they cannot contribute to the result in the end. In paper III, a colleague and I have argued that the $\hbar$-expansion detailed above is well-behaved in this sense, and does not in general need to be fixed with extra techniques.

But there are exceptions to this, as I will discuss in the next subsection.

3.3.4 Symmetry breaking by quantum effects

In the previous subsection I argued that quantum corrections to spontaneous symmetry breaking must be organized as a perturbative expansion around some tree-level effect. This implies that the quantum corrections cannot change the overall picture.

But this point of view is actually too strict. It *is* possible to have quantum corrections affect spontaneous symmetry breaking—but it is still necessary to apply proper care.

The first application of the effective potential, and the first derivation of its 1-loop contributions, was due to S. Coleman and E. Weinberg [26]—the 1-loop potential is hence often called the Coleman-Weinberg potential. What they studied was the idea of quantum-generated spontaneous symmetry breaking, also known as radiative symmetry breaking.

Consider the Abelian Higgs theory. If the classical potential would have $m^2 = 0$, then there would only be a minimum at the origin. The field-dependent squared masses are

$$m_Z^2 = e^2\phi^2, \tag{3.41}$$

$$m_H^2 = 3m_G^2 = 3\lambda\phi^2, \tag{3.42}$$

and these masses are zero when evaluated at the tree-level minimum $\phi = 0$. The effective potential to one loop accuracy is

$$V(\phi) = \frac{\lambda}{4}\phi^4 + \hbar\frac{3}{4}(m_Z^2)^2\left(\log\left[\frac{m_Z^2}{\mu^2}\right] - \frac{5}{6}\right). \tag{3.43}$$

Here it seems as if the 1-loop terms *could* induce a non-zero value for ϕ. But—with our strict loop counting discussed above—we are not allowed to have the 1-loop terms mix with the tree-level terms, rendering this "quantum-breaking" impossible. Or does it?

It all comes down to which power counting we should use. In the standard loop expansion of perturbative quantum field theory, one would count $\lambda \sim e^2$. This would make all 1-loop diagrams equally important, all 2-loop diagrams equally important, and so on. From equation (3.43) we see that if we would instead count $\lambda \sim e^4$, then the 1-loop effects would be just as important as the tree-level effects.

This modified power counting (which was recognized as important already by Coleman and Weinberg [26]) will change the perturbative expansion. What then happens with gauge invariance, that seemed to need a fixed loop counting? This question has been resolved by Andreassen, Frost, and Schwartz [24].

In this modified power counting, it turns out that a resummation is necessary. In Fermi gauges, there is an infinite amount of diagrams that contribute at next-to-leading order. These can be resummed by shifting the masses of the scalars according to the mass one would derive from the leading effective potential,

$$H \to \overline{H} = \partial^2(V_0 + V_1) = H + \frac{3}{2}e^2 m_Z^2\left(\log\left[\frac{m_Z^2}{\mu^2}\right] - \frac{1}{3}\right), \tag{3.44}$$

and similarly for G. With these new masses, the accounting of degrees of freedom works out just as in the regular loop expansion. An evaluation of V_{min} order-by-order in powers of e results in a finite and gauge-invariant quantity.

We can hence modify the previous verdict I made regarding the calculation of observables. The point is not that the expansion has to be a proper loop expansion centered on a classic effect. But, whatever power-counting is used, it has to be respected and centered around some *leading* effect. In the example above we can count powers of e instead of loops, with $\lambda \sim e^4$, and expand around the leading-order vacuum expectation value found from equation (3.43). This is a proper accounting.

This kind of reasoning is useful in the next chapter, which looks at how thermal fluctuations can modify the effective potential.

Recommended readings

Mark Srednicki, *Quantum Field Theory* [15]. This is my favourite quantum field theory textbook. The exposition is technical but clear, and the chapters are very self-contained—it's great for looking things up.

Andreassen, Frost, and Schwartz, *Consistent use of Effective Potentials* [24]. This paper is a great review of the different issues that arise in effective potential calculations, with references to the relevant original papers. The authors take the issues of gauge invariance and infrared divergences seriously.

Martin and Patel, *Two-loop effective potential for generalized gauge fixing* [17]. This paper details the calculation of the 2-loop effective potential in a generic gauge theory, using a generalized gauge-fixing. Even though there are around 500 equations, it is straight-forward to follow the logic—the structure of the paper is remarkable.

4. Phase transitions

“Time is a flat circle.”
—Rust Cohle

True Detective

As the early universe expanded and cooled down, there was a phase transition that broke the electroweak symmetry. The *symmetric vacuum* at the origin is lower in energy at high temperatures, but at some point the *broken vacuum* becomes energetically favorable, and the transition occurs.

The nature of this transition, if it is continuous or a sudden jump, is an important question. A *first-order transition* occurs via bubble nucleation, and can leave imprints for us to detect, such as a gravitational wave signal [36].

But how do we know how this transition occurs? From previous chapters we expect that there should be a potential for us to minimize. Indeed, in *finite temperature field theory*, the usual formalism of quantum field theory can be extended to include temperature effects [37]. There are new contributions to the effective potential that depend on the temperature, and this can in principle change the picture of spontaneous symmetry breaking dictated by the classical potential [38].

The finite temperature formalism is in certain ways easier to interpret than the quantum one. We can formulate thermodynamical observables in terms of the partition function, which encodes the thermal and quantum fluctuations. This partition function can in turn be formulated like a path integral, and observables can be derived from it.

In particular, the Helmholtz free energy of a theory at a particular temperature T is given, in units with $k_B = 1$, by

$$F = -T \log Z, \tag{4.1}$$

and represents the “useful” energy available in the system. The free energy as a function of the background field is proportional to the thermal effective potential.

Taking the limit $T \to 0$ recovers the quantum field theory effective potential. In a sense, the finite temperature formalism offers us a second interpretation of the zero temperature effective potential. It is just the free energy density of the system at zero temperature, where all fluctuations are quantum in nature.

4.1 Finite temperature effective potential

Finite temperature field theory can be formalized in several different ways. For our purposes, the imaginary time formulation will be most useful [39]—in this formalism time is imaginary and periodic. Operationally, we can think of it as a quantum field theory on $\mathbb{S}^1 \times \mathbb{E}^3$, where $\mathbb{S}^1$ is a circle with period $\beta \equiv 1/T$, and $\mathbb{E}^3$ is 3D Euclidean space,

$$Z[j] \propto \int \mathrm{D}\phi\, e^{\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3 x \mathcal{L}}. \tag{4.2}$$

For this path-integral to correspond to the partition function in our thermal theory, we need the bosonic fields to be periodic in τ, while the fermionic fields must be anti-periodic.

The Fourier composition of these fields then feature one sum (for the compact dimension) and three integrals. A bosonic propagator with square mass x, in momentum space, is then of the form

$$\SumInt_P \frac{1}{P^2 + x}, \text{ with}$$

$$P \equiv (2\pi n T)^2 + p^2, \qquad \SumInt_P = T \sum_{n=-\infty}^{\infty} \int_p, \qquad \int_p \equiv \left(\frac{Q^2 e^{\gamma_E}}{4\pi}\right)^{\epsilon} \int \frac{\mathrm{d}^{d-1}p}{(2\pi)^{d-1}},$$

and $d = 4 - 2\epsilon$, with μ as the $\overline{\text{MS}}$ scale. A fermionic propagator is similar, but then $P^2 = (\pi(2n+1)T)^2 + p^2$, and we use the notation $\SumInt_{\{P\}}$ to mark the difference.

Other than the differences noted above, the overall structure of the theory is the same as for a zero temperature quantum field theory—the calculations of the effective potential proceed exactly the same. In general it is possible to separate the contributions as

$$V(\phi; T) = V^{T=0}(\phi) + V^{T\neq 0}(\phi; T). \tag{4.3}$$

But in the approximation scheme that is relevant for phase transitions, this separation is not useful. Instead one performs a high-temperature expansion, where the involved square mass x is assumed to be much smaller than T^2, $x \ll T^2$.

The master-integral that corresponds to the 1-loop contribution of a boson to the effective potential reads

$$\mathbf{f}(x) \equiv \frac{1}{2} \SumInt_P \log\left[P^2 + x\right] \tag{4.4}$$

$$= -\frac{x^2}{4(16\pi^2)}\frac{1}{\epsilon} - \frac{\pi^2}{90}T^4 + \frac{T^2 x}{24} - \frac{T x^{3/2}}{12\pi} + \mathcal{O}(x, \epsilon), \tag{4.5}$$

in the high-temperature expansion. The corresponding fermionic master-integral reads

$$\mathbf{f}_F(x) \equiv \frac{1}{2}\sum\!\!\!\!\!\!\int_{\{P\}} \log\left[P^2 + x\right] \tag{4.6}$$

$$= -\frac{x^2}{4(16\pi^2)}\frac{1}{\epsilon} + \frac{7\pi^2}{720}T^4 - \frac{T^2 x}{48} + \mathcal{O}(x,\epsilon), \tag{4.7}$$

in the high-temperature expansion. A qualitative difference between $\mathbf{f}(x)$ and $\mathbf{f}_F(x)$ is the absence of a term linear in T in $\mathbf{f}_F(x)$.

At two loops there are additional master-integrals that are relevant. See appendix A in paper IV for a more exhaustive list. Paper V is a calculation of further terms in the high-temperature expansion of the 2-loop sunset sum-integral.

4.2 Phase transitions

We have seen that, perturbatively, the effective potential is the name of the game when it comes to symmetry breaking. This still applies at finite temperature, with the interesting complications that arise from temperature dependent coefficients [38].

As the temperature changes, so does the shape of the potential. To understand the behavior for very large temperatures, we can look at the leading ϕ-dependent term of the 1-loop potential. Imagine that a symmetry is broken at low temperatures, as specified by the classical potential. Adding the leading temperature corrections then gives

$$V(\phi) = V_0 + \hbar\frac{T^2}{24}\alpha\phi^2 + \mathcal{O}\left(T, \alpha^2\right), \tag{4.8}$$

where the collection of couplings α is determined by how the fields of the model couple to the Higgs field. In the limit $T \to \infty$, this potential is a second degree polynomial with its minimum at the origin. There is hence no symmetry breaking for large temperatures. But as the temperature decreases, the other terms of the potential become important again, and the shape of the potential goes through a change.

Depending on the particulars, this change can occur in two different ways, as illustrated in figure 4.1. There can be a continuous transition where the minimum at the origin bifurcates into a maximum at the origin and a broken vacuum. The minimizing background field $\phi(T)$—the order parameter of the transition—then continuously moves away from the origin as T decreases, and finally stopping at the zero temperature broken minimum.

But if a barrier develops, then the global minimum will jump from the origin to some non-zero value. This minimum then settles at the zero temperature broken minimum.

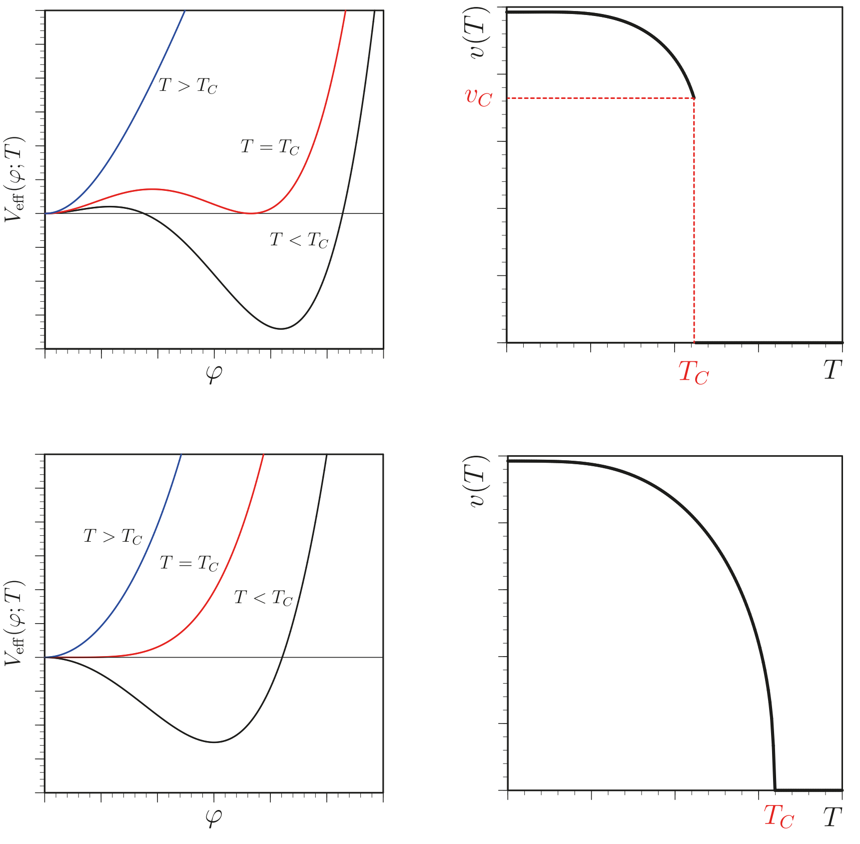

Figure 4.1. An illustration of two different kinds of phase transitions. The top row illustrates a first order phase transition, the bottom row a second order one. The left column shows the effective potential as a function of a background field φ, evaluated at three different temperatures. The right column shows the order parameter—the minimum of the potential—as a function of temperature. This figure is from the review by Senaha [40].

Personally, I really like this picture of the phase transition. It is such a natural extension of the classical results, with vivid imagery accompanying it. But, if I have done my job correctly earlier in the thesis, you should also be noticing some red flags right about now. How can a loop level effect—which is supposedly suppressed by powers of the couplings—actually be large enough to warp the classical potential? I dig into this question in the next section.

4.3 Perturbative problems

That finite temperature calculations suffer from perturbative problems is a well-established fact [37, 39]. Here I want to highlight just how connected this is to the phase transition. This analysis is based on the early work of Arnold and Espinosa [41], and our work in paper IV. It is an attempt to extend the gauge-invariant methods of [16], using the logic laid out at the end of subsection 3.3.4.

Starting with the 1-loop diagram of the effective potential, we can consider a class of higher-loop diagrams that arise from inserting loops on the propagator. The most important contributions occur when the inner loop has *soft* momentum $k^2 \ll T^2$, and the outer loops have hard momenta $k^2 \sim T^2$. An n-loop diagram d_n with $(n-1)$ insertions will then scale with the temperature as

$$d_n \sim T\left(e^2T^2\right)^{n-1} \implies \frac{d_n}{d_{n-1}} \sim e^2T^2, \tag{4.9}$$

which means that sufficiently large temperatures will break the perturbative expansion. When $T \sim 1/e$, the $(n-1)$-loop diagrams become as important as the n-loop diagrams.

This breakdown of perturbation theory signals that we are not treating our physics correctly. In this particular case, we can fix the problem by resumming all the problematic diagrams into one single contribution. In this contribution, the mass of the inner loop particle then gets shifted by $\sim e^2T^2$. The particle is screened by the thermal bath, and gets an effective mass [37].

Second order transitions

Returning to phase transitions, let's consider the leading contribution again. We can focus on the simple example of Abelian Higgs. Then the leading terms in the effective potential are schematically

$$V(\phi) \sim \frac{1}{2}(m^2+(e^2+\lambda)T^2)\phi^2 + \frac{1}{4}\lambda\phi^4, \tag{4.10}$$

where the T^2 term arises from the 1-loop effective potential.

If these terms are a good approximation for the potential, then there is a second order phase transition with a critical temperature $T_c \sim 1/e$. This is precisely the scale where perturbation theory breaks down, as seen in equation (4.9). Maybe it is not actually surprising. For the temperature fluctuations to affect the phase structure of the theory, we need 1-loop effects to be as large as tree-level effects. This means that we cannot use loops to order perturbation theory, just as for the Coleman-Weinberg mechanism explained in subsection 3.3.4. To reorder perturbation theory then requires a resummation, with the end result that $V_{\min}(T)$ is finite and gauge invariant near the phase transition. This enables us to study the phase transition in a consistent way.

First order transitions

There has to be a barrier in the potential for a first-order transition to occur. Operationally, with a positive quartic coupling we need a cubic term in the potential.

To find it we can look to the next term in the high-temperature expansion of the 1-loop potential—the term linear in T that arises from bosonic fluctuations. The potential is then schematically

$$V(\phi) \sim \frac{1}{2}(m^2 + (e^2 + \lambda)T^2)\phi^2 - Te^3\phi^3 + \frac{1}{4}\lambda\phi^4. \tag{4.11}$$

We should now perform a similar analysis as above. The analysis is a little more complicated, but Arnold and Espinosa showed [41] that we can balance these terms if we count the quartic coupling as $\lambda \sim e^3$.

In paper IV we put the two power-countings detailed above to the test by performing resummed 2-loop calculations in Abelian Higgs and the Standard Model, in Fermi gauges. We demonstrated that, if proper care is taken with the resummation and the power-counting—and the perturbative methods described in subsection 3.3.3 are used—then the result is finite and gauge-invariant.

Resummations and more complicated models

To find the thermal masses of the longitudinal vector bosons, one has to calculate their zero-momentum thermal self-energies $\Pi_A(T^2; p = 0)$ to $\mathcal{O}\left(e^2T^2\right)$. The corresponding scalars' self-energies can be found directly from the effective potential. In simple models like Abelian Higgs or the Standard Model, the resummed Higgs square mass $\overline{H}$, and the resummed Goldstone square mass $\overline{G}$, can be calculated as

$$\overline{H} = \partial^2 V_{\text{LO}}(\phi), \qquad \overline{G} = \frac{\partial V_{\text{LO}}(\phi)}{\phi}, \tag{4.12}$$

where V_{LO} is the leading-order potential. In the first-order phase transition power counting, V_{LO} includes terms like $T^2\phi^2$ and $T\phi^3$.

For more complicated theories, with more complicated field structure, the situation is more complicated. I would refer to the formula used to prove the Goldstone theorem of the effective potential,

$$\partial^2_{ij}V\ (\mathcal{T}^\alpha\vec{\phi})_j = \mathcal{T}^\alpha_{ij}\partial_j V. \tag{4.13}$$

If we use a gauge-fixing that preserves the global invariance of the Lagrangian (and hence the effective potential), then we can use this formula to connect the derivatives of the effective potential (with respect to whatever background-fields we activate), to the mass-matrix of the Goldstone fields.

Final remarks

In the Standard Model, we already know from lattice studies that the Higgs is too heavy to allow a first-order phase transition [9, 10]. With the power counting rules detailed above, this is reflected in the fact that the Higgs quartic coupling λ is too large to be able to count it as $\lambda \sim g^3, g'^3$, where g, g' are the gauge-couplings of the electroweak sector.

We can then make generic predictions based on this formalism. Any extension of the Standard Model that does not change its phase structure will need to fulfill at least one of the two following properties. Either, (1), there has to be new bosonic fields that couple to the Higgs with couplings parametrically larger than λ (generically, there needs to be new bosons with a mass just above the electroweak scale). Or, (2), there needs to be a mechanism that reduces the size of λ. The predictions of property (1) are consistent with the predictions of [42].

In models where the phase structure is different, for example in models with two-step phase transitions, the analysis is not as simple. But I expect that similar judgments based on how the different couplings compare in size are still relevant.

Recommended readings

Kapusta and Gale, *Finite-temperature field theory: principles and applications* [37]. I learned finite temperature field theory from this textbook. It is a book that rewards deep study.

Patel and Ramsey-Musolf, *Baryon Washout, Electroweak Phase Transition, and Perturbation Theory* [16]. This paper is an amazing review of the issues of gauge dependence in phase transition studies. It was a big influence on me.

5. Populärvetenskaplig sammanfattning (på svenska)

5.1 Introduktion

Standardmodellen, som är den bästa partikelfysik-teori vi har, kan förklara *nästan* alla fenomen vi ämnar beskriva inom partikelfysik. Men inte alla. Vi vet därmed att vi behöver förbättra våra teorier på olika sätt. Det övergripande målet med min forskning är att testa, eller att hjälpa andra att testa, partikelfysik-teorier mot data. Detta har tagit formen av två olika projekt som jag arbetat på.

Partikelfysik är relevant vid interaktioner som sker vid höga energier. Det betyder korta avstånd, där kvantmekanik beskriver verkligheten, och höga hastigheter, där speciell relativitetsteori är nödvändig. Dessa två fysikaliska principer kan kombineras med hjälp av *kvantfältteori*, ett teoretiskt ramverk som utvecklades under 1900-talet. I dessa teorier är de fundamentala objekten kvantfält som breder ut sig över rummet. När ett sådant fält exciteras ser vi det som en partikel. Två olika fält kan påverka varandra, och en resonans i ett fält kan därmed trigga resonanser i ett annat. På detta sätt kan partiklar interagera med varandra och skapas eller förstöras.

Kvantfältteori är på vissa sätt underbart, men på vissa andra sätt väldigt knepigt. Det är en lång väg från att skriva ner en teori till att beräkna någonting mätbart, en lång väg genom konceptuellt snårig skog: ofysikaliska frihetsgrader introduceras, oändligt många oändligheter tar ut varandra, med mera. Det kan vara svårt att hålla reda på vad som är matematiska knep och vad man skulle kalla för ”riktig fysik.”

Därför tycker jag att det är extra viktigt att vi tar vara på de få möjligheter vi har för att ha kontroll över våra teorier så att vi kan vara säkra på vad som kan förutsägas från dem. Detta är ett övergripande tema för de två projekt jag arbetat med.

5.2 Standardmodellen och dess problem

Standardmodellen etablerades under slutet på 1900-talet. Men det var först 2012 som den sista pusselbiten, Higgsbosonen, verifierades experimentellt. Det återstår dock många öppna frågor. I den här sektionen diskuterar jag två av dessa problem som relaterar till min forskning.

5.2.1 Neutriner och deras massor

Neutriner är ett flyktigt slag av partiklar som ingår i Standardmodellen—de är otroligt lätta jämfört med de andra partiklarna, samt så växelverkar de väldigt svagt med allt. Trots att det alltid finns många neutriner runt omkring oss så vet vi inte särskilt mycket om dem. När Standardmodellen formulerades så trodde man att neutriner var helt masslösa, men nu vet man att så inte är fallet på grund av att man observerat neutrino-oscillationer. Detta kvantmekaniska fenomen som beblandar de olika arterna av neutriner kan bara ske om åtminstone en av dem är massiv.

Ett av de utstående problemen med Standardmodellen är därmed att den endast beskriver masslösa neutriner. Samtidigt är det inte så enkelt som att bara lägga till massor till modellen. Detta kan göras på olika sätt. Genom att testa neutriners egenskaper med hjälp av partikelfysikexperiment kan vi försöka avgöra vilket som är rätt.

Från den teoretiska sidan är en speciell klass av sådana förklaringar särskilt attraktiva. Om det finns tunga partners till de neutriner vi redan känner till så kan dessa tunga partiklar genom en balansmekanism förklara varför de neutrinerna vi observerat är så lätta. Detta kallas för en ”seesaw-mechanism” på engelska, eftersom det påminner om en tippande gungbräda.

I artikel I har jag och mina kollegor utforskat en enkel utvidgning av Standardmodellen där en ny kraft realiseras. Denna kraft motsvarar en lokal U(1)-symmetri (symmetrin av en cirkel). Sådana lokala symmetrier är väldigt speciella, för de antyder att det finns en redundans i vår beskrivning. Det betyder att några av våra frihetsgrader är ofysikaliska, och att vi måste vara extra försiktiga för att inte råka få med deras bidrag när vi beräknar någonting mätbart.

Mer specifikt så behöver man i sådana modeller se till att kvantfysikaliska effekter inte bryter denna symmetri. Sådana effekter kallas för anomalier. Lokala symmetrier måste vara anomalifria för att kunna beskriva verkligheten. Det visar sig att modeller med en extra U(1)-symmetri är lättast att göra anomalifria genom att lägga till tre nya partiklar som beter sig precis som neutriner. Dessa teorier kan därmed på ett elegant sätt realisera en gungbräde-mekanism. I vår artikel använder vi data från experimentet LHC för att begränsa de parametrar som finns i dessa teorier.

Artikel II frångår problemet med neutrinomassorna och fokuserar istället på anomalifria teorier. För att kancellera anomalierna kan man föreställa sig att det finns nya partiklar som är så tunga att de inte är relevanta för resten av modellen. Detta kallas för *Green-Schwarz* mekanismen, och härstammar från strängteori (ett teoretiskt ramverk som ämnar beskriva kvangravitation). I vår artikel översatte vi i detalj hur detta skulle fungera för U(1)-symmetrin vi studerar, och sedan gjorde vi förutsägelser för hur detta skulle manifesteras vid LHC. Dessa teorier är mindre begränsade än de vanliga anomalifria teorierna.

5.2.2 Materia och antimateria

En intressant förutsägelse av kvantfältteori är existensen av antimateria. För vissa partiklar finns det en motsvarande antipartikel, som har några av sina egenskaper spegelvända jämfört med sin partikelpartner. Ett exempel är elektronen, som vi känner till som den lätta och negativt laddade partikeln i våra atomer. Elektronen har en antipartikel, *positronen*, som har positiv elektrisk laddning men samma massa som elektronen. Materia och antimateria kan annihilera varandra: en positron kan annihilera en elektron och skapa två fotoner.

Eftersom materia och antimateria behandlas i stort sett likadant i Standardmodellen, så kan man fundera på varför det bara finns materia runt omkring oss. Antimateria bör vara ungefär lika vanligt som materia, vilket skulle innebära att nästan all materia annihilerats för länge sedan. Att materia dominerar universum just nu tyder på att någonting skedde tidigt i universums utveckling som skapade lite mer materia än man naivt förväntar sig.

Elektrosvag baryogenes är en möjlig förklaring som är populär bland fysiker, där materia-antimateria-asymmetrin utvecklades av tidiga termiska fluktuationer. Dessa blev sedan kvar efter den *elektrosvaga fasövergången*. Tidiga universum expanderade och svalnade av, och fasövergången skedde när den elektrosvaga symmetrin bröts.

Fasövergångar kan ske på flera olika sätt, men för att förenkla så kan vi dela upp dem i två olika sorter. Vatten som omvandlas till ånga när det kokar är ett exempel på en *första ordningens fasövergång*. Detta är en turbulent process där bubblor av ånga bildas och expanderar. En *andra ordningens fasövergång* sker mer kontinuerligt, och systemets egenskaper förändras under lugnare former. Ett exempel är hur ett block av ferromagnetiskt material kan magnetiseras när det placeras i ett magnetfält.

Det är ännu okänt hur den elektrosvaga fasövergången gick till, men enligt Standardmodellen var det en andra ordningens fasövergång. Vi är intresserade av fasövergångens natur eftersom den kan lämna avtryck på det vi kan observera idag. För att elektrosvag baryogenes ska kunna förklara materia-antimateria asymmetrin så måste det ha skett en första ordningens fasövergång.

För att studera fasövergången i en kvantfältteori så använder man sig vanligtvis av den *effektiva potentialen*, som beskriver den potentiella energin för olika möjliga grundtillstånd, vilka även kallas *vakuum*. Genom att minimera potentialen kan man hitta vakuumet när termiska och kvantmekaniska fluktuationer är med i bilden. Om man studerar hur vakuumet förändras när temperaturen ändras kan man avgöra hur fasövergången går till.

Men den här beräkningen har två motstridiga problem. För att kunna jämföra olika möjliga kandidater till vakuum så behöver man introducera ofysikaliska fluktuationer. Man måste vara väldigt noggrann för att inte råka få med dem i slutresultatet när man använder en approximationsmetod som kallas för *störningsteori*. Samtidigt så gör termiska fluktuationer nära fasövergången att den approximationen bryter ner. För att fixa det problemet så behöver man röra om i störningsteorin.

Det förstör den känsliga balansen som krävs för att inte få med de ofysikaliska bidragen till resultatet.

I artikel IV har jag och en kollega demonstrerat en metod som kan lösa båda dessa problem samtidigt. För att få beräkningen att gå ihop krävs det att man funderar extra noggrant på den approximationen man gör. Beroende på vilken sorts fasövergång som sker så måste man använda olika approximationer. I artikel III argumenterade vi att störningsteorin vanligtvis inte behöver repareras när man inte har med termiska fluktuationer. Artikel V är vår beräkning av en särskild slags termiska korrektioner som är relevanta om man vill studera fasövergången med väldigt noggrann störningsteori.

Acknowledgments

“Show me that solvable problem
We can get through this
I’ll do the hardest part with you”
—Steven Universe

Steven Universe the Movie

I am deeply grateful for being given the time to learn, read, think, talk physics. Forgive me if I get a bit sappy, for a moment.

I am grateful to my colleagues. I want to thank my advisors, Rikard and Gunnar, for giving me this opportunity—and for all the encouragement you have provided during this journey. Rikard, thank you for guiding me and letting me do my own thing at the same time. I have always felt better motivated after speaking to you. Thank you Gunnar for inspiring me to try to speak clearly about physics. You have a way with words that I admire.

Thank you Tanumoy and Stefan, for helping me get started with research. Tanumoy, I am glad you taught me that at some point you just have to start calculating. It was always fun to work with you. I also want to thank you Stefan, for believing in me, and for being an amazing teacher. Talking physics with you is always fun and always enlightening.

I am grateful to my friends. Thank you Andreas and Suvendu—lunch and coffee were great breaks from work with you two around. Andreas, I would not have been the same physicist if we had not studied together so many years. Thank you for teaching me so much, and thank you for sharing the frustrations and joys of our calculations. Your willpower is very inspiring. Suvendu, thanks for all the great conversations—about technology, design, food, physics, and everything. You are amazing at explaining so many different things, and your curiosity really shines through your explanations.

Tack alla i min familj som har inspirerat och stöttat mig, på era egna sätt. Tack pappa för att du alltid uppmuntrat mig att läsa. Tack mamma för att du lyssnat och förstått när livet är svårt. Tack syster för din humor och värme. Tack farmor och Pelle för ert stöd. Tack mostrar—ni har alltid inspirerat mig att vilja studera och lära mig mera.

Tack Anna för att du påminner mig om att njuta av det goda i livet. Jag älskar att vi ofta skrattar tillsammans. Tack Edith för din fantasi, din glädje, din koncentration, din nyfikenhet, och din leklust.

Thank you again, family, friends, and colleagues.

References

[1] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC." *Phys. Lett. B*, 2012. 716:1–29.

[2] S. Chatrchyan et al. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC." *Phys. Lett. B*, 2012. 716:30–61.

[3] S. Chatrchyan et al. "Observation of a New Boson with Mass Near 125 GeV in pp Collisions at $\sqrt{s} = 7$ and 8 TeV." *JHEP*, 2013. 06:081.

[4] T. Appelquist, B. A. Dobrescu, and A. R. Hopper. "Nonexotic Neutral Gauge Bosons." *Phys. Rev. D*, 2003. 68:035012.

[5] M. D. Schwartz. *Quantum field theory and the standard model*. Cambridge University Press, Cambridge, 2014. ISBN 9781107034730;1107034736;.

[6] J. M. Cline. "Baryogenesis." In "Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime," 9 2006 .

[7] M. Trodden. "Electroweak baryogenesis." *Rev. Mod. Phys.*, 1999. 71:1463–1500.

[8] D. E. Morrissey and M. J. Ramsey-Musolf. "Electroweak baryogenesis." *New J. Phys.*, 2012. 14:125003.

[9] K. Kajantie, M. Laine, K. Rummukainen, and M. E. Shaposhnikov. "Is there a hot electroweak phase transition at m(H) larger or equal to m(W)?" *Phys. Rev. Lett.*, 1996. 77:2887–2890.

[10] M. D'Onofrio and K. Rummukainen. "Standard model cross-over on the lattice." *Phys. Rev. D*, 2016. 93(2):025003.

[11] S. M. Carroll. *The particle at the end of the universe: how the hunt for the Higgs boson leads us to the edge of a new world*. Dutton, New York, 2012. ISBN 9780525953593;0525953590;.

[12] A. Zee. *Quantum field theory in a nutshell*. Princeton Univ. Press, Princeton, N.J, 2003. ISBN 9780691010199;0691010196;.

[13] A. J. Beekman, L. Rademaker, and J. van Wezel. "An Introduction to Spontaneous Symmetry Breaking." *SciPost Phys. Lect. Notes*, 2019. page 11.

[14] V. Rubakov. *Classical theory of gauge fields*. Princeton Univ. Press, Princeton, N.J., 2002. ISBN 0691059276.

[15] M. A. Srednicki. *Quantum field theory*. Cambridge University Press, Cambridge, 2007. ISBN 9780521864497;0521864496;.

[16] H. H. Patel and M. J. Ramsey-Musolf. "Baryon Washout, Electroweak Phase Transition, and Perturbation Theory." *JHEP*, 2011. 07:029.

[17] S. P. Martin and H. H. Patel. "Two-loop effective potential for generalized gauge fixing." *Phys. Rev. D*, 2018. 98(7):076008.

[18] M. E. Peskin and D. V. Schroeder. *An introduction to quantum field theory*. Addison-Wesley, Reading, Mass, 1995. ISBN 9780201503975;0201503972;.

[19] L. Faddeev and V. Popov. "Feynman Diagrams for the Yang-Mills Field." *Phys. Lett. B*, 1967. 25:29–30.

[20] C. Becchi, A. Rouet, and R. Stora. "Renormalization of Gauge Theories." *Annals Phys.*, 1976. 98:287–321.

[21] R. Fukuda and T. Kugo. "Gauge Invariance in the Effective Action and Potential." *Phys. Rev. D*, 1976. 13:3469.

[22] M. B. Green and J. H. Schwarz. "Anomaly Cancellation in Supersymmetric D=10 Gauge Theory and Superstring Theory." *Phys. Lett. B*, 1984. 149:117–122.

[23] P. Anastasopoulos, M. Bianchi, E. Dudas, and E. Kiritsis. "Anomalies, anomalous U(1)'s and generalized Chern-Simons terms." *JHEP*, 2006. 11:057.

[24] A. Andreassen, W. Frost, and M. D. Schwartz. "Consistent Use of Effective Potentials." *Phys. Rev. D*, 2015. 91(1):016009.

[25] R. Jackiw. "Functional evaluation of the effective potential." *Phys. Rev. D*, 1974. 9:1686.

[26] S. R. Coleman and E. J. Weinberg. "Radiative Corrections as the Origin of Spontaneous Symmetry Breaking." *Phys. Rev. D*, 1973. 7:1888–1910.

[27] N. Nielsen. "On the Gauge Dependence of Spontaneous Symmetry Breaking in Gauge Theories." *Nucl. Phys. B*, 1975. 101:173–188.

[28] A. Das. *Finite temperature field theory*. World Scientific, Singapore, 1997. ISBN 9789810228569;9810228562;.

[29] M. Laine. "Gauge dependence of the high temperature two loop effective potential for the Higgs field." *Phys. Rev. D*, 1995. 51:4525–4532.

[30] S. Weinberg. "Perturbative Calculations of Symmetry Breaking." *Phys. Rev. D*, 1973. 7:2887–2910.

[31] L. Dolan and R. Jackiw. "Gauge Invariant Signal for Gauge Symmetry Breaking." *Phys. Rev. D*, 1974. 9:2904.

[32] I. Aitchison and C. Fraser. "Gauge Invariance and the Effective Potential." *Annals Phys.*, 1984. 156:1.

[33] P. B. Arnold. "Phase transition temperatures at next-to-leading order." *Phys. Rev. D*, 1992. 46:2628–2635.

[34] M. Laine. "The Two loop effective potential of the 3-d SU(2) Higgs model in a general covariant gauge." *Phys. Lett. B*, 1994. 335:173–178.

[35] S. Ramaswamy. "Gauge invariance and the effective potential: The Abelian Higgs model." *Nucl. Phys. B*, 1995. 453:240–258.

[36] C. Caprini et al. "Detecting gravitational waves from cosmological phase transitions with LISA: an update." *JCAP*, 2020. 03:024.

[37] J. I. Kapusta and C. Gale. *Finite-temperature field theory: principles and applications*. Cambridge University Press, Cambridge, 2. edition, 2006. ISBN 0521173221;9780521173223;0521820820;9780521820820;.

[38] L. Dolan and R. Jackiw. "Symmetry Behavior at Finite Temperature." *Phys. Rev. D*, 1974. 9:3320–3341.

[39] M. Laine and A. Vuorinen. *Basics of Thermal Field Theory*, volume 925. Springer, 2016.

[40] E. Senaha. "Symmetry Restoration and Breaking at Finite Temperature: An Introductory Review." *Symmetry*, 2020. 12(5):733.

[41] P. B. Arnold and O. Espinosa. "The Effective potential and first order phase transitions: Beyond leading-order." *Phys. Rev. D*, 1993. 47:3546. [Erratum: Phys.Rev.D 50, 6662 (1994)].

[42] M. J. Ramsey-Musolf. "The electroweak phase transition: a collider target." *JHEP*, 2020. 09:179.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1978

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2020

Distribution: publications.uu.se
urn:nbn:se:uu:diva-422716