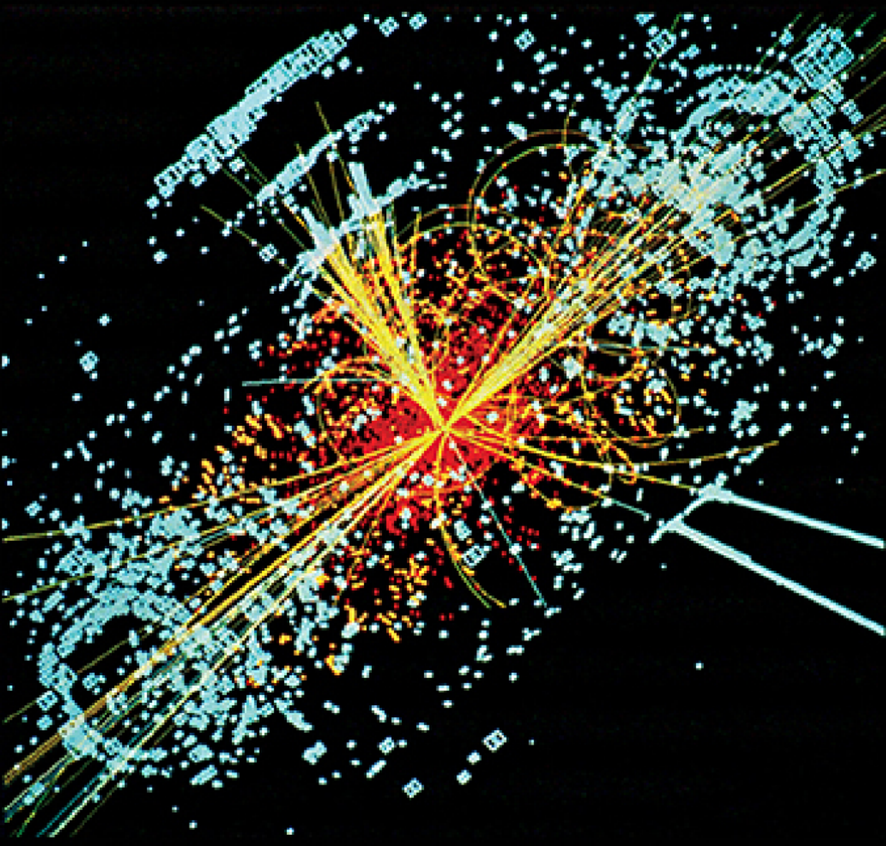
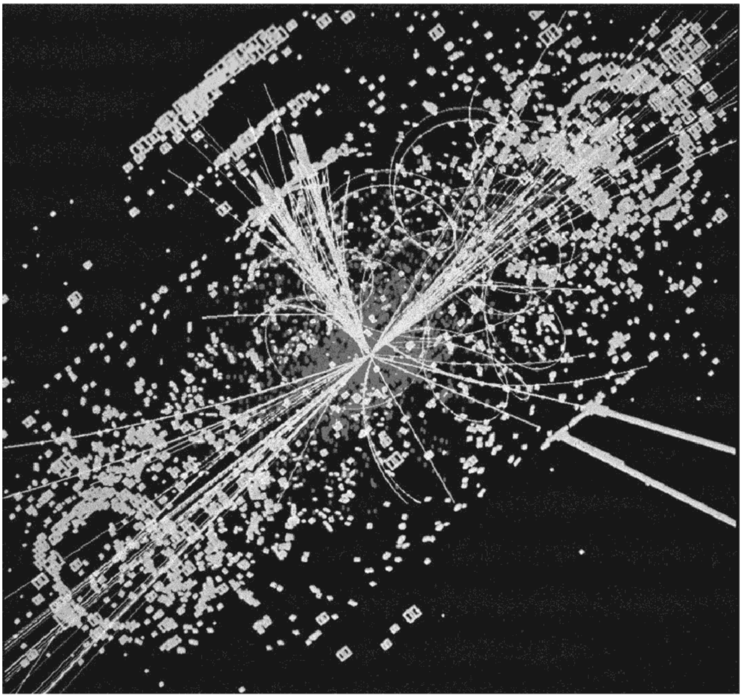
EFFECTIVE FIELD THEORIES



Alexey A. Petrov • Andrew E. Blechman

World Scientific

EFFECTIVE FIELD THEORIES



Alexey A. Petrov • Andrew E. Blechman
Wayne State University, USA

World Scientific
NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Petrov, Alexey A., 1971– author.
Effective field theories / Alexey A. Petrov, Andrew E. Blechman, Wayne State University, USA.
pages cm
Includes bibliographical references and index.
ISBN 978-9814434928 (hard cover : alk. paper)
1. Field theory (Physics). 2. Quantum field theory. 3. Gravitational fields.
I. Blechman, Andrew E., author. II. Title.
QC173.7.P48 2015
530.14--dc23
2015006575

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Printed in Singapore

Contents

This page intentionally left blank

Chapter 1

Introduction

What's in a name? That which we call a rose
By any other name would smell as sweet.

W. Shakespeare

1.1 Wherefore EFT?

This book is about effective field theories (EFTs). What makes a field theory effective? Is it better or worse than a "regular" field theory? We shall argue in this book that the way calculations are set up in EFTs makes them the most natural and convenient tools to address multi scale problems. Problems with separated scales often appear in Nature, and we intuitively know that it is most convenient to only work with degrees of freedom that are relevant for a particular scale – otherwise the problem quickly becomes intractable! You never worry about physics of the atoms when designing bridges, nor try to track each and every molecule of a gas through phase space; you instead define some "macroscopic" variables, and once you know how to relate those variables to the more "fundamental" laws, you can stop thinking about these laws and focus only on the relevant large-scale physics. EFT techniques codify this principle when working with problems in quantum field theory.

It is interesting to note that scale separation is very natural in physics. In quantum mechanics, we are not concerned with the value of the top quark mass when we calculate the energy levels of a hydrogen atom. Of course, given certain precision of an experimental measurement, we might want to be concerned about that.[1] Having this in mind, however, we would still insist that only degrees of freedom relevant to the problem in hand are needed to perform the calculation. In the language of quantum field theory this implies that operators that are responsible for experimental observables only include fields describing light degrees of freedom. By doing so we effectively eliminate all heavy particles with masses well above the scale associated with the problem at hand (say, a hydrogen atom).[2] They have not disappeared completely: quantum theory allows the possibility of these particles to be created and destroyed on very short time scales, and this leads to coupling constants and other parameters of our theory changing with scale – and this change could be affected by the removed degrees of freedom. Also, we might improve the accuracy of our predictions by introducing more operators in our theory. This certainly happens when it is not forbidden, for example, by the symmetries of the system we are studying. In a sense, symmetry becomes the guiding principle for our construction of effective description of physical systems: we do not even need to know what heavy particles we integrated out! We can keep adding more operators, provided we know the way to assess the importance of those new contributions – or, colloquially speaking, *power count* them. The structure and coefficients in front of those operators, when fit to experimental data, might tell us something about the heavy particles that we integrated out. Thus, EFTs become a very convenient way of studying new, undiscovered physics.

As we shall see in this book, this is one way of using techniques of effective field theories. There are many others. In the previous paragraph we discussed removing some (heavy) degrees of freedom completely. Instead of doing that, we can remove only "parts" of the fields. That is, we can remove the known solution, say, for a static or fast-moving particle, only concentrating on the corrections to the known result. This might reveal new symmetries of the theory, simplifying the overall description. The EFT method will allow us to do this as well.

[1] One of the most precise measurements known so far is the measurement of anomalous magnetic moment of the muon. That measurement is sensitive to the effects of heavy quarks, but not the top quark.

[2] Thinking in terms of a path integral formulation of a field-theoretic problem, we *integrate out* all heavy degrees of freedom.

1.2 EFT vs MFT

In condensed matter physics, scientists have often employed the technique of "Mean Field Theory" (MFT), and for those that are familiar with it, you might think that we are simply re-casting an old idea in new clothes. It is certainly true that MFT and EFT have a common heritage, and they borrow a lot from each other; but there are a few philosophical differences between the two approaches that should be understood from the outset.

In MFT, the program is to try and calculate a background state of the dynamical degrees of freedom, which is an average, or "mean" field. This quantity is usually called the *order parameter*. You can set up your action (or in condensed matter systems, free energy) as a function of this order parameter, and use the result to make predictions. Like EFT, there are several ways to set up this free energy functional. One approach is to attempt to compute the order parameter directly and then re-express the free energy as a function of this mean field; this is the so-called *Bragg-Williams MFT*. Another thing you can try to do is identify the order-parameter and symmetries using physical arguments, and then write down the most general free energy as a function of this order parameter that is consistent with all the symmetries you identified; this is the so-called *Landau MFT*. There are other approaches to MFT that have been tried and tested over the last century as well. A nice review of these approaches to MFT is given in [Chaikin and Lubensky (1995)].

MFT has its advantages and disadvantages. It is simple and intuitive, and it often does a great job making *qualitative* predictions, such as the general structure of the phase diagram for a system. It also tends to do a good job making quantitative predictions far away from any phase transition or other breakdown of the assumptions that go into constructing it. However, MFT explicitly ignores fluctuations away from the mean field. For a system in equilibrium, this is not generally a problem, although there is a certain limit of accuracy; but when describing a phase transition, for example, it leads to nonsense! Phase transitions are precisely the point at which fluctuations can dominate the system, and as a result, MFT calculations of things like critical exponents are often far from the mark. At this point, physicists need to look elsewhere. There are many clever approaches that have been developed to correct for MFT's missing information. For a wonderful review of these methods, see [Zinn-Justin (2002); Parisi (1988)], for example.

EFT does not share the problems of MFT, at least not on the surface.

One can think of MFT as a "leading-order" EFT result. Many of the methods we alluded to in the above paragraph are actually built into EFT's general make-up. Furthermore, EFT takes many of the ideas from quantum field theory, such as Feynman diagrams, path integrals and renormalization, and builds it in directly to the theory, rather than as an add-on to correct for MFT's deficiencies. In short, EFT is a more general approach to a problem than MFT.

Of course, since both methods are applied to similar problems, they often borrow from each other, and so becoming an expert in EFT will help you to have a better understanding and appreciation for the MFT approach to condensed matter systems. As both of the authors of this book are trained as particle physicists, most of our examples are focussed in that direction. Nevertheless, we hope that seeing how the techniques are generally applied will help readers from many different fields have a better understanding of how to perform effective calculations, and the power you gain by casting your theory in this language.

1.3 An example from Newton

As a simple example that nevertheless shows many features of a real application of effective Lagrangians, let us consider Galileo Galilei's Leaning Tower of Pisa experiment. According to Galileo's student, Vincenzo Viviani, Galileo dropped balls of different mass m from the Leaning Tower of Pisa. We can write a Lagrangian for one of the balls,

$$L = \frac{mv^2}{2} - V(h) = \frac{mv^2}{2} - mgh, \tag{1.1}$$

where g is the free-fall acceleration, h is the ball's height, say, above the ground, and $v = \dot{h}$ is the velocity of the ball,.

We argue that Eq. (1.1) represents an *effective Lagrangian* to the full mechanical description of this problem. We are all taught that the zero level of potential energy $V(h)$ can be chosen arbitrarily, it is only the potential difference that is physical. Since constant terms can be dropped in the definition of a Lagrangian in Eq. (1.1), this fact is represented by a manifest shift symmetry, $h \to h + a$, where a is constant distance. Another way of saying this is that the force acting on the ball, $F = mg$, is independent of the height of the tower.

Now this is not, strictly speaking, correct! If the Leaning Tower of Pisa is moved to the top of Mount Everest, the force and the potential energy would change, simply because free-fall acceleration, $g(R) = GM/R^2$, depends on how high the object is located above the Earth.[3] In fact, one can show that $g(R)$ satisfies the following differential equation,

$$R\frac{\partial}{\partial R}g(R) = \gamma_g g(R), \tag{1.2}$$

with $\gamma_g = -2$. This equation looks very much like a renormalization group equation.[4]

We already know that Eq. (1.1) is not an exact expression, in fact, it is an approximation. One can try to make it better by observing that the radius of the Earth is much bigger than the height of the Leaning Tower of Pisa, i.e. $h/R \to 0$. A better approximation to a true potential can take a form of a power expansion,

$$V(h) = C_1(R)\ m\left(\frac{h}{R}\right) + C_2(R)\ m\left(\frac{h}{R}\right)^2 + \cdots . \tag{1.3}$$

Here $C_i(R)$ are unknown coefficients, which can be found if precise experimental data is available. In principle, an exact potential can be guessed if a sufficient number of terms in the expansion of Eq. (1.3) is determined.

In the case at hand, however, we can do better. Indeed, as Sir Isaac Newton tells us, the *full theory* is described by a Newtonian interaction potential between the Earth of mass M and a ball, separated by distance r,

$$V(h) = G\frac{Mm}{r} = G\frac{Mm}{R+h}, \tag{1.4}$$

which, however, does not have a manifest symmetry $h \to h + a$ that the *effective* Lagrangian of Eq. (1.1) possesses. Expanding $V(R)$ in h/R and *matching* Eqs. (1.3) and (1.4) we can determine the unknown coefficients,

$$C_1(R) = -C_2(R) = ... = \frac{GM}{R}, \tag{1.5}$$

which results in the expansion of the potential,

$$\begin{aligned} V(h) &= \frac{GM}{R}\ m\left(\frac{h}{R}\right) - \frac{GM}{R}\ m\left(\frac{h}{R}\right)^2 + \cdots \\ &= mgh - \frac{mg}{R}h^2 + \cdots , \end{aligned} \tag{1.6}$$

where we dropped the constant term. We indeed found that the leading term in Eq. (1.6) is the potential term of the original Lagrangian of Eq. (1.1). It is interesting that even in such a simple example we performed all the steps needed in a derivation of a classical effective Largangian! It simply shows that using effective Lagrangian techniques is very natural.

[3]Indeed, g is not the same at the top and at the bottom of the Leaning Tower of Pisa. In making this argument we neglected the variation of g along the height h.

[4]This curious observation was pointed out to us by T. Huber.

1.4 A theorem of Weinberg

Much more useful for practical calculations is the notion of *quantum* effective Lagrangians. A theoretical basis of any quantum effective field theory can be formulated in terms of a theorem, first given by S. Weinberg [Weinberg (1979b)]

Theorem 1.1. *To any given order in perturbation theory, and for a given set of asymptotic states, the most general possible Lagrangian containing all terms allowed by the assumed symmetries will yield the most general S-matrix elements consistent with analyticity, perturbative unitarity, cluster decomposition and assumed symmetry principles.*

Initially, this theorem was written to conjecture the equivalence of current algebra results and methods of effective Lagrangians in pion physics, where it was shown to work in all cases. Nothing in this theorem says that it is only applicable to pion physics, so it is expected that this theorem should work in any EFT. So far, there are no known counterexamples of this theorem.

Theorem 1.1 is very plausible, but not as trivial as it seems. As we know from our quantum field theory courses, an important part of quantum theory is renormalization, i.e. a correct treatment of the unknown behavior of that theory at ultra-small scales. According to the theorem, we need to write the most general set of operators in a Lagrangian consistent with given symmetries in order to get the most general S-matrix elements. That set would surely contain a very large number of operators; moreover, in general the number of such operators is infinite! How can one make any predictions when there are an infinite number of contributions to take into account?

We will see that the situation is not hopeless: for a given precision of measurements, effective field theories will provide consistent and testable predictions even if they are not renormalizable in a "classical" sense. We shall show that computing quantum loops in effective theories is exactly the same as it is in (conventionally) renormalizable field theories: we would still integrate over *all* values of momentum, even if our EFT is only valid to some momentum scale μ.[5] Also, even in quantum loops, we would only deal with the degrees of freedom given in the effective Lagrangian, i.e the ones with which we started our calculations. While this might appear strange, as the integration over momenta greater that μ certainly misses

[5]Since we are using a "natural" system of units where we set $\hbar = c = 1$, we shall use the notions of energy and momentum interchangeably.

some physics, it is actually quite natural – and not different from what is done in conventional QFTs. After all, if structures of the terms generated by loop effects are the same as the structures already present in the original local Lagrangian, those "incorrect" loop contributions would be corrected by shifting the parameters of the Lagrangian, i.e. by renormalization! Thus, in order to execute the EFT program, we will need to find a way to assess the importance of loop-induced contributions in our calculations, and, if needed, introduce proper counterterms to render the result finite. That is, we would have to find a proper *power counting* scheme.

1.5 Organization of the book

As mentioned above, there are many ways to construct an effective field theory, depending on what it is you are trying to describe. We find it convenient to group different EFTs into one of three categories, based on what degrees of freedom they include.[6]

The first kind of EFT we discuss, which we call "Type-I," refers to the famous classic example of an effective field theory, where the list of degrees of freedom only includes those fields that can contribute at the energy and momentum scale of the interaction. For example, beta decay is described by a theory that never makes any mention of the W-boson; atomic and molecular physics does not make use of quarks; standard-model particle physics processes make no mention of any super-heavy particle such as GUT remnants; etc. These EFTs are the most straightforward examples, and we will use them to start the discussion.

The second kind of EFT, which we call "Type-II," refers to problems where some fields no longer participate in *dynamics*, but they are still a part of the Fock space. These objects are taken as (nearly) infinitely heavy, and simply sit there while other lighter degrees of freedom are bouncing off of it in totally elastic collisions. One example of a Type-II theory is the Newtonian example we considered earlier. In that problem, the Earth was taken as infinitely massive, and therefore does not recoil when objects hit it, but it would be wrong to say that we have "integrated out the Earth!" The function of the Earth was to provide the gravitational field (if you wish, it is a massive, static reservoir of gravitons!) that was represented by the potential in the Lagrangian. Other examples of Type-II EFT are when you

[6]Similar suggestions for the breakdown into EFT-types were made in a talk by Michael Luke at SCET2007 conference at MIT.

have a bound state between a heavy particle and a light particle, like the Hydrogen atom. In that case, the proton is a source of photons (generating an electrostatic field), but we ignore its recoil as it interacts with the lighter electron. We can then reincorporate that recoil as subleading terms in our effective Lagrangian (just like we included higher-order terms in h/R with the Earth example.

The third kind of EFT, called "Type-III," is the newest and also the most controversial construction. This is an attempt to describe objects that have large energy-momentum transfers, but *only in a given, fixed direction.* This implies that we need to attempt a Type-II construction, but *only on the components of the field that can create the large momentum.* This means that, roughly speaking, we should integrate out the part of the field with momentum in the z-direction, but leave the parts of the field that create particles moving in the x, y-directions in the dynamics. This is a strange theory, to say the least, and it is still not clear whether such a construction is even self-consistent. From the point of view of studying EFT for its own sake, Type-III is definitely "where the action is!" But whatever its philosophical and technical problems might be, its prime example of Soft-Collinear Effective Theory (SCET) has proven to be incredibly useful in helping us understand heavy-to-light particle decays; parton showering; event-shape distributions; IR factorization theorems; the list goes on!

While it is important to have a basic knowledge of quantum field theory techniques, we made an attempt to make this book self-consistent by providing all of the needed background and introductory material. Group theory techniques are reviewed in Appendix A; QED and QCD are briefly discussed in Appendix B; and the ideas and some more advanced uses of Dimensional Regularization, one of the most popular regulators in EFT, are reviewed in Appendix C; so if the readers feel like some brush-up is needed, those resources are there for their convenience.

In the following chapters we shall provide needed background, discuss the "mechanics" of EFT building and then consider several classes of EFTs. We suggest the readers try and solve problems at the end of the chapters. After each chapter we also provided references for further studies of the topics discussed in the chapter.

We employ the Minkowski metric with the mostly-minus sign convention $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ throughout this book.

Chapter 2

Symmetries

2.1 Introduction

The most vital part of effective field theory is knowing what symmetries apply to your system. This knowledge can get you very far in describing the nature of the problem, constructing model-independent equations to describe dynamics, and put constraints on matrix elements. This chapter will be a review of some of the more important results that follow from symmetries. We will discuss Noether's and Goldstone's theorem, and the consequences that arise from them. We will discuss various examples, but for simplicity we will stick mostly with scalar fields wherever we can, avoiding fermions until we need them for anomalies.

2.2 Noether's Theorem

The chief reason why symmetries are important is due to a theorem in Lagrangian mechanics known as *Noether's Theorem*:

Theorem 2.1 (Noether). *Every continuous symmetry of the action (and path integral measure) implies a conservation law.*

The caveat about the measure being invariant is important to handle the possibility of quantum anomalies, as we will see at the end of this chapter.

Proof. Consider a Quantum Field Theory (QFT) with fields ϕ^a and action $S[\phi]$ that is presumed to be invariant under a *global* transformation of the form:

$$\phi^a \longrightarrow \phi^a + \epsilon \Delta \phi^a \tag{2.1}$$

where ϵ is an infinitesimal parameter. Although the action is only supposed to be invariant under these transformations for constant ϵ, let us consider the behavior of the action under the above field transformation when ϵ is allowed to vary with space-time:

$$\begin{aligned} S[\phi+\epsilon(x)\Delta\phi] &= \int d^dx\ \mathcal{L}(\phi+\epsilon(x)\Delta\phi, \partial_\mu\phi+\epsilon(x)\Delta\partial_\mu\phi) \\ &= \int d^dx\ \mathcal{L}(\phi,\partial_\mu\phi) + \int d^4x \left\{\epsilon\Delta\phi^a \frac{\partial\mathcal{L}}{\partial\phi^a} + \partial_\mu(\epsilon\Delta\phi^a)\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a}\right\} \\ &= S[\phi] + \int d^dx(\partial_\mu\epsilon)\Delta\phi^a\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a} \\ &\quad + \int d^dx\ \epsilon\left\{\frac{\partial\mathcal{L}}{\partial\phi^a}\Delta\phi^a + \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a}\Delta\partial_\mu\phi^a\right\} \end{aligned} \tag{2.2}$$

The last term is there regardless of whether ϵ is constant or a function of space-time, so if $S[\phi]$ is to be invariant under the global transformation in Eq. (2.1) this term must take the form $\int d^dx\epsilon\partial_\mu\mathcal{J}^\mu$. The quantity $\mathcal{J}^\mu$ will generally be nonzero if $\Delta\phi$ explicitly involves changes in space-time coordinates, such as translations and rotations.

Integrating the second term by parts and dropping a surface term, we have:

$$S[\phi+\epsilon(x)\Delta\phi] = S[\phi] - \int d^dx\ \epsilon(x)\ \partial_\mu\left[\Delta\phi^a\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a} - \mathcal{J}^\mu\right] \tag{2.3}$$

If we insert this into the path integral and use the invariance of the measure we get:

$$\begin{aligned} Z &= \int \mathcal{D}\phi\ e^{iS[\phi]} = \int \mathcal{D}(\phi+\epsilon\Delta\phi)\ e^{iS[\phi+\epsilon\Delta\phi]} \\ &= \int \mathcal{D}\phi\ e^{iS[\phi]}\ e^{-i\int d^dx\epsilon\partial_\mu j^\mu} \end{aligned} \tag{2.4}$$

where in the first line we shifted the dummy functional integration variable $\phi(x)$ and in the second step we have defined

$$j^\mu = \Delta\phi^a\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a} - \mathcal{J}^\mu\ . \tag{2.5}$$

Since ϕ is a dummy field variable, Z must be independent of ϵ and so we can write:

$$0 = \frac{i}{Z}\frac{\delta Z}{\delta\epsilon}\bigg|_{\epsilon=0} = \frac{\int\mathcal{D}\phi\ (\partial_\mu j^\mu)e^{iS}}{\int\mathcal{D}\phi\ e^{iS}} = \langle\partial_\mu j^\mu\rangle\ . \tag{2.6}$$

This is just the correspondence principle at work: j^μ is the classically conserved Noether current, and it is conserved quantum mechanically in all correlation functions. □

Eq. (2.5) is the Noether current corresponding to the symmetry transformation in Eq. (2.1). Throughout most of this book we will be concerned with internal symmetries, where you can set $\mathcal{J}^\mu = 0$. In that case, the Noether current only involves kinetic (gradient) terms in the Lagrangian. This tells us that Noether currents are *universal* in the sense that they are independent of the potential (non derivative) interactions in the action. Thus, even if a symmetry is violated by potential terms, one can still construct a Noether current: *the symmetry structure of any quantum field theory is dictated by the kinetic terms alone!* Although contact interactions can (and often do) break symmetries, we can still identify a Noether current and treat the symmetry breaking effects as corrections to the theory.

Noether currents are not uniquely defined: we could construct a new current from the old one:

$$j_2^\mu = j_1^\mu + \partial_\nu K^{\mu\nu} \tag{2.7}$$

where j_1^μ is the current in Eq. (2.5) and $K^{\mu\nu} = -K^{\nu\mu}$ is an arbitrary antisymmetric two-indexed tensor. Both currents are locally conserved, as can be easily checked. What is the meaning of j_2^μ? To answer that, consider the charge that is conserved:

$$Q_1 = \int d^{d-1}\vec{x}\ j_1^0 \tag{2.8}$$

$$\begin{aligned} Q_2 &= \int d^{d-1}\vec{x}\ j_2^0 = \int d^{d-1}\vec{x}\ (j_1^0 + \partial_\nu K^{0\nu}) = Q_1 + \int d^{d-1}\vec{x}\ \partial_i K^{0i} \\ &= Q_1 + \int_\infty d\mathbf{S}\cdot\mathbf{K} \end{aligned} \tag{2.9}$$

So the extra term represents a surface charge at infinity. So long as we are considering the case that all charges are confined to the interior of our space-time, this extra term has no physical effects. However, in cosmology or condensed matter systems, where boundary effects might be important, you must remember to take this term into account.

2.3 Examples of Noether currents

Let us consider one of the most important starting actions for an effective field theory of any kind: that of a (complex) scalar field with arbitrary potential that only depends on the modulus of the field:

$$S[\phi] = \int d^dx\ \{(\partial_\mu\phi^*)(\partial^\mu\phi) - \mathcal{V}(|\phi|) \tag{2.10}$$

This action is invariant under a constant rephasing of the scalar field, which in infinitesimal form is

$$\phi \longrightarrow \phi + i\epsilon\phi \tag{2.11}$$

In the notation of the previous section, $\Delta\phi = i\phi$ and so the Noether current is

$$\begin{aligned} j^\mu &= \Delta\phi^a \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi^a} \\ &= (i\phi)(\partial^\mu\phi^*) + (-i\phi^*)(\partial^\mu\phi) \\ &= -i(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*) \end{aligned} \tag{2.12}$$

Notice that we have to consider the transformation of every field – in this case, that means *both* ϕ and ϕ^*. Also note that this result is independent of the choice of $\mathcal{V}(\phi)$ as we previously advertised.

The above calculation generalizes to the case of N scalar fields[1] that can transform into each other and still leave the action invariant:

$$S[\phi^a] = \int d^dx\ \frac{1}{2}g_{ab}(\partial_\mu\phi^a)(\partial^\mu\phi^b) \tag{2.13}$$

where g_{ab} is a real, nonsingular $N \times N$ matrix, and we're considering real fields for concreteness, although you can generalize to complex fields if you like. This action will be invariant under the transformation

$$\phi^a \longrightarrow \phi^a + i\epsilon^A[T^A]^a_b\phi^b \tag{2.14}$$

where ϵ^A is a vector of infinitesimal parameters, and iT^A are a set of real, $N \times N$ matrices that satisfy the rule:

$$g_{ac}g^{db}[T^A]^c_d = -[T^A]^b_a \tag{2.15}$$

Here we use the usual Einstein convention that g^{ab} is the inverse matrix of g_{ab}, and repeated upper-lower indices are summed. We will never lower the capital letter indices, however.

Since there is one symmetry transformation for each A, there is a corresponding current. We can group them together from the arguments above to find

$$j^{\mu A} = \phi_a[T^A]^a_b\partial^\mu\phi^b \tag{2.16}$$

[1] We will index the fields by lowercase letters from the front of the alphabet. Upper case letters will be used to label the matrices; note that a set of $N \times N$ matrices are spanned by N^2 matrices, although depending on the details of the action, the precise form of g_{ab}, and any additional terms, not all of these matrices need lead to symmetries.

where $\phi_a \equiv g_{ab}\phi^b$. Usually we will suppress the lowercase indices as understood, but we include them above so you can see how they contract. In an index-free notation:

$$j^{\mu A} = \overline{\phi} T^A \partial^\mu \phi \tag{2.17}$$

Let us consider one more example: the real (free) scalar field

$$S[\theta] = \int d^d x \ \frac{1}{2}(\partial_\mu \theta)(\partial^\mu \theta) \tag{2.18}$$

This action is invariant under the shift symmetry:

$$\theta(x) \longrightarrow \theta(x) + \epsilon v \tag{2.19}$$

where v is a constant with the same units as θ. Then $\Delta\theta = v$ and the Noether current is

$$j^\mu = v\partial^\mu \theta \tag{2.20}$$

The conserved current is just the gradient of the scalar field! We will see how this current appears in applications shortly, as well as the meaning of the number v introduced in the transformation rule.

2.4 Gauged symmetries and Noether's procedure

When a transformation like Eq. (2.1) remains a symmetry of the action even when the parameter ϵ is a function of space-time, then the symmetry is said to be a *local* or *gauged* symmetry. There are a few general points about gauging a symmetry that should be pointed out:

(1) Gauge symmetries imply global symmetries, since we can always take $\epsilon(x)$ to be a constant. Thus everything mentioned previously about Noether's theorem still holds.

(2) The "gauge" part of the symmetry is strictly speaking not a *symmetry* but a *redundancy* coming from the fact that (in $d = 4$) a massless spin-1 field has two degrees of freedom, but is described by a four-vector potential that allows you to maintain Lorentz covariance. "Gauge invariance" is simply the statement that the unphysical degrees of freedom do not contribute to any physical amplitude.

(3) For a global symmetry, Noether's theorem is only valid *on-shell* – the quantum version of the theorem we proved above only had the current being conserved within correlation functions, not as an operator equation. However, when the symmetries are gauged, Noether currents are conserved *as operators!* Thus people sometimes say that gauged symmetries are more "fundamental" than global symmetries.

Recall that when we promoted a global transformation to a local one we generated an extra term that was proportional to the Noether current:

$$S_0[\phi + \epsilon(x)\Delta\phi] = S_0[\phi] + \int d^d x \; (\partial_\mu \epsilon) j^\mu \tag{2.21}$$

For the action to remain invariant under this local symmetry we must include a term that cancels this; we accomplish this by introducing a four-vector potential A_μ that transforms as

$$A_\mu \longrightarrow A_\mu + \partial_\mu \epsilon \tag{2.22}$$

and couple it to the Noether current:

$$S_1[\phi, A_\mu] = -\int d^d x \left[A_\mu j^\mu + \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right] \tag{2.23}$$

The first term cancels the shift in S_0 upon transforming ϕ, while the second term is the usual kinetic energy operator for A_μ. In this simple case, $F_{\mu\nu}$ and j^μ are already invariant under the combined transformation of Eqs. (2.1) and (2.22) and so the total action $S_0 + S_1$ is the minimum extension of the globally invariant action that is also invariant under the gauge symmetry.

This procedure of correcting the global action step by step by adding terms designed to cancel changes from the modified transformation law is known as the *Noether's procedure* and is a very powerful way of building gauge invariant actions from knowledge of globally invariant ones. This case we have considered is an example of an Abelian gauge theory like Maxwell's electromagnetism. Let us see how to extend this idea to more complicated gauge symmetries.

Consider a collection of vector fields A^a_μ $a = 1, \ldots, N$ which transform in the adjoint representation of some Lie algebra[2]:

$$A^a_\mu \longrightarrow A^a_\mu + g c^{abc} \epsilon^b A^c_\mu \quad \Rightarrow \quad \Delta^b A^a_\mu = g c^{abc} A^c_\mu \tag{2.24}$$

where c^{abc} are the structure constants from Eq. (A.7) and g is a small constant (the gauge coupling). The kinetic term for a collection of spin-1 vector bosons,

$$-\int d^d x \; \frac{1}{4} f^a_{\mu\nu} f^{\mu\nu a}, \qquad f^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu \tag{2.25}$$

is invariant under the global transformation in Eq. (2.24) and generates a Noether current:

$$j^{\mu a} = (g c^{abc} A^c_\nu) \frac{\partial \mathcal{L}}{\partial \partial_\nu A^a_\mu} = g c^{abc} f^{\mu\nu b} A^c_\nu \tag{2.26}$$

[2]See Appendix A.2 for details on algebra theory.

Notice that this current is *not* invariant under the transformation!

Noether's procedure then says that to make this theory invariant under a local transformation, we must add a new term to the action

$$S_1 = -\int d^d x j^{\mu a} A_\mu^a \tag{2.27}$$

and also modify the transformation law to

$$A_\mu^a \longrightarrow A_\mu^a + g c^{abc} \epsilon^b A_\mu^c + \partial_\mu \epsilon^a \tag{2.28}$$

$S_0 + S_1$ is now invariant under the transformation of Eq. (2.28) up to first order in g, but not at order g^2:

$$\delta S_0 + \delta S_1 = -\int d^d x \; g^2 c_{abc} c_{bde} A_\mu^a A_\nu^c A^{\nu e} \partial^\mu \epsilon^d \tag{2.29}$$

To cancel this term, one must add another term to the action

$$S_2 = +\int d^d x \; \frac{g^2}{4} c_{abc} c_{bde} A_\mu^a A_\nu^c A^{\nu e} A^{\mu d} \tag{2.30}$$

Now $S_0 + S_1 + S_2$ is invariant under Eq. (2.28) to order g^2, and in fact to all orders in g, so we may stop here. Our final result is in fact the usual gauge invariant action of Yang-Mills:

$$S = -\int d^d x \; \frac{1}{4} F_{\mu\nu}^a F^{\mu\nu a} \,, \qquad F_{\mu\nu}^a = f_{\mu\nu}^a - g c^{abc} A_\mu^b A_\nu^c \tag{2.31}$$

$$A_\mu^a \longrightarrow A_\mu^a + D_\mu \epsilon^a \tag{2.32}$$

as it should be.

In this case, Noether's procedure terminated at $\mathcal{O}(g^2)$, but that does not have to be the case. The process can go much further, often never terminating at all, and you must modify both the action and the field transformation rules at each order in the coupling. In practice, this is a very powerful way to derive the action of a theory based on knowing the symmetries of the problem. We will see how this works in some exercises.

2.5 Broken symmetries and Goldstone's Theorem

Consider a complex scalar field with a potential of the form

$$\mathcal{V}(\phi) = m^2 |\phi|^2 + \lambda |\phi|^4 \tag{2.33}$$

This potential has a symmetry

$$\phi \longrightarrow e^{i\alpha} \phi \tag{2.34}$$

regardless of the value of (m^2, λ). However, the spectrum of the theory is very sensitive to these parameters. In particular, the ground state of the theory where $\mathcal{V}$ is minimized occurs when

$$\frac{\partial \mathcal{V}}{\partial \phi} = \phi^*(m^2 + 2\lambda|\phi|^2) = 0 \Rightarrow |\phi_0|^2 = 0, \ -\frac{m^2}{2\lambda} \tag{2.35}$$

We will assume that $\lambda > 0$ so that $\mathcal{V}$ is bounded from below. Then if $m^2 > 0$, the only minimum is at $\phi_0 = 0$ which is clearly still invariant under Eq. (2.34). On the other hand, if $m^2 < 0$ the minimum actually occurs at the nonzero value, and so the ground state is *not* invariant under the symmetry. The field forms a background *condensate* in the ground state, and we can no longer treat the other fields in the theory as living in a vacuum. When the action of a QFT has a symmetry while the ground state forms a condensate that breaks that symmetry, we say that the symmetry is *spontaneously broken*.

A word of warning: the name is unfortunate, as the symmetry is *not* really broken! The action is still symmetric, and therefore there is still a conserved Noether current (remember that the current was independent of the potential).

We will define the nonzero value of ϕ_0 in the vacuum state as

$$v^2 \equiv -\frac{m^2}{\lambda} \tag{2.36}$$

and write the field as[3]

$$\phi(x) = \frac{1}{\sqrt{2}}\left(v + \varphi(x)\right) \tag{2.37}$$

If we plug this back into our potential we find

$$\mathcal{V}(\varphi) = \left[\frac{1}{2}m^2v^2 + \frac{1}{4}\lambda v^4\right] + \frac{1}{2}\left(m^2 + \lambda v^2\right)\left[v(\varphi + \varphi^*) + |\varphi|^2\right] + \frac{\lambda}{4}\left[v^2(\varphi + \varphi^*)^2 + 2v|\varphi|^2(\varphi + \varphi^*) + |\varphi|^4\right] \tag{2.38}$$

The first term in brackets ($\mathcal{V}(\varphi = 0)$) is a constant and corresponds to the value of the potential in the ground state. It gives us information about the condensate, and represents a cosmological constant, but has no other meaning. This term is called the *Mean Field Theory* result in condensed matter systems, and v is the *order parameter*. The second term vanishes for

[3]The factor of $1/\sqrt{2}$ is conventional, so $|\phi_0| = v/\sqrt{2}$, but not every author uses this convention – make sure you know what convention is being used to avoid frustrating mistakes!

the value of v in Eq. (2.36), leaving just the third term. We can understand its meaning by decomposing our field even further into real and imaginary components:

$$\varphi(x) = \phi_1(x) + i\phi_2(x) \tag{2.39}$$

Then our potential is

$$\mathcal{V}(\phi_1, \phi_2) = V_0(v) + \lambda v^2 \phi_1^2 + \lambda v \phi_1(\phi_1^2 + \phi_2^2) + \frac{\lambda}{4}(\phi_1^2 + \phi_2^2)^2 \tag{2.40}$$

So upon spontaneous symmetry breaking our theory has three points of interest:

(1) A condensate described by a number v.
(2) A (real) massive scalar field (ϕ_1) with mass-squared equal to $2\lambda v^2$.
(3) A (real) *massless* scalar field (ϕ_2).

The presence of a massless mode is not a coincidence of the problem, nor is it only true classically or at leading order in perturbation theory. It turns out that whenever a symmetry is spontaneously broken, we expect to see a massless particle appear. This is a consequence of *Goldstone's Theorem* which we turn to now.

Theorem 2.2 (Goldstone). *For every spontaneously broken symmetry, there corresponds a massless particle whose quantum numbers are the same as that symmetry's Noether charge.*

The massless modes of the theorem are called **(Nambu-)Goldstone (NG) Fields**.

Proof. Consider a QFT with N fields whose action (and path integral measure) are invariant under a field transformation:

$$\phi^i \longrightarrow \phi^i + i\epsilon T^i_j \phi^j \tag{2.41}$$

where T is a finite, imaginary matrix. Rather than study the usual action, we will consider the effective action that generates the 1PI correlation functions:

$$\Gamma[\phi] = -\log\ Z[J] - \int d^d x\ J(x)\phi(x) \tag{2.42}$$

This will allow us to include quantum effects without any more work. In terms of the effective action, invariance under the transformation in Eq. (2.41) implies

$$\left(\frac{\delta\Gamma}{\delta\phi^n}\right) T^n_m \phi^m = 0 \tag{2.43}$$

Let us take the functional derivative of this equation and evaluate it at $\phi(x) = \phi_0$, the minimum of the full effective potential

$$\left(\frac{\delta^2\Gamma}{\delta\phi^l\delta\phi^n}\right)_{\phi=\phi_0} T^n_m\phi^m + \left(\frac{\delta\Gamma}{\delta\phi^n}\right)_{\phi=\phi_0} T^n_l = 0 \tag{2.44}$$

But since ϕ_0 is the vacuum solution, defined as the minimum of the effective action, the first derivative vanishes, and the second derivative is precisely the two-point 1PI correlation function evaluated with all external momenta equal to zero. This is the definition of the mass matrix M^2 of the theory, and so we have

$$\left[M^2\right]_{ln} T^n_m\phi^m_0 = 0 \tag{2.45}$$

Thus the vectors $T^n_m\phi^m_0$ are eigenvectors of the mass matrix with eigenvalue zero. If $\phi_0 = 0$, then there is nothing to be said; but if $\phi_0 \neq 0$, there must be a zero-eigenvalue of M^2 – that is, a massless mode. Furthermore, if $1 \leq k \leq N$ is the number of nonzero ϕ^m_0, there are k such zero-eigenvectors, and this is also the precise number of broken generators. Thus there is one massless mode for each broken generator, as we wished to show.

We will only sketch the proof of the statement that the quantum numbers of these massless modes must be the same as those of the broken charges. One can use the Källèn-Lehmann spectral techniques to show that

$$\langle B|j^0(x)|0\rangle \neq 0 \tag{2.46}$$

where B is one of the NG modes; see [Weinberg (1996)] for a nice explanation of this calculation. But this matrix element will certainly vanish if the B state has different parity, helicity or internal quantum numbers as the current. Therefore the NG fields must have the same quantum numbers as the broken generators. □

In most cases of interest, the space-time properties of the generators of symmetry groups are simple – they are space-time scalars. Therefore, the NG fields themselves must be spin-0. The famous exception to this is supersymmetry, where the current is a vector-spinor (so the SUSY charges are spinors) and the corresponding NG field is a spinor, affectionately known as a *goldstino*.

2.5.1 *Nonrelativistic NG-bosons*

Although Goldstone's theorem does not require Lorentz invariance to hold, the form of the theorem and the proof we sketched out is most appropriate for relativistic systems. It turns out that when the system is nonrelativistic, it is still true that each broken generator gives you a "flat direction" in the potential, but the corresponding *gapless excitation*,[4] which we can interpret as a massless (quasi-)particle, can be a bit more tricky to identify. In particular, it is no longer true that each broken generator corresponds to its own NG-boson.

The flaw in our proof comes from how we interpreted Eq. (2.45). We pointed out that if $\phi_0^m \neq 0$, then $T_m^n \phi_0^m$ is an eigenvector of the mass matrix with zero-eigenvalue. But it is also possible that a linear combination of the charges annihilates ϕ_0, which would mean that some of the purported massless modes are actually null-vectors! This happens if the symmetry-breaking field ϕ corresponds to one of the charges. This cannot occur in a Lorentz-invariant theory, but it is quite possible in condensed matter systems.

Perhaps the most famous example of this is the Heisenberg Ferromagnet. The model is a collection of spins on a lattice with Hamiltonian:

$$H = -J \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j \; ; \tag{2.47}$$

The notation $\langle ij \rangle$ refers to sums over neighboring spins. The model has a global $O(3)$ symmetry that rotates all the spins in the same way. J is a constant, taken to be positive for the ferromagnet. The ground state (at zero temperature) is the state where all the spins are aligned in one direction (the z direction, say). This state spontaneously breaks the $O(3)$ to an $O(2)$ symmetry, leaving only rotations about the z axis. The other two rotation generators represent broken symmetries, and therefore you might have expected to find two NG-bosons: gapless states that come from the broken generators acting on the spontaneous magnetization. This state is often called the *magnon* in the condensed matter literature. But there is something very strange: there is only one magnon state, while the Goldstone theorem seems to suggest that there should be two, since two symmetry generators are broken. The resolution to this apparent paradox is that the spontaneous magnetization (ϕ_0) is directly related to the symmetry generators, and so there is a null-vector. That leaves only one eigenvector left as a Goldstone boson.

This point was clarified further in a powerful theorem by [Nielsen and Chadha (1976)]. Their result states that as long as certain conditions are

[4]By this we mean a quantum excited state that can have infinitely long wavelength (or low momentum); that is, it obeys a dispersion relation such that $E(\vec{p} \to 0) = 0$.

satisfied by the theory, such as translational invariance and proper asymptotic behavior of the correlation functions, spontaneous symmetry breaking of N generators will always be accompanied by a collection of gapless modes that come in two types:

$$\begin{aligned} \text{Type}-1 \text{ modes}: &\ E(\vec{p}) \propto |\vec{p}|^{2l+1}, \\ \text{Type}-2 \text{ modes}: &\ E(\vec{p}) \propto |\vec{p}|^{2l}. \end{aligned}$$

Say that there are n Type-1 modes, and m Type-2 modes; then the Nielsen-Chadha theorem says that $n + 2m = N$. In other words, if the dispersion relation of the NG mode is even in (quasi-)momentum, then it counts twice! Indeed, if you work out the dispersion relation for a Heisenberg Ferromagnet, you will find that $E \to |\vec{p}|^2/2M$, where M is a nonuniversal constant that depends on the details of the model. As the magnon has a quadratic dispersion relation, we see that there is only one magnon in the spectrum, since it counts both of the broken generators.

On the other hand, the Heisenberg antiferromagnet (same as Eq. (2.47), but with $J < 0$) undergoes spontaneous symmetry breaking, where ϕ_0 is the *staggered magnetization*. This is not directly related to the generators and therefore we expect both generators to represent two independent NG modes. Sure enough, there are two gapless modes in the low-energy spectrum with a *linear* dispersion: $E \to \pm C|\vec{p}|$, and so there are two NG bosons, according to the Nielsen-Chadha theorem.

There are never any Type-2 modes in a relativistic theory; indeed the Ferromagnetic magnon dispersion relation violates our understanding of a relativistic dispersion relation, since it seems to represent a particle of effective mass M that is also gapless, violating Einstein's formula, $E = Mc^2$! Clearly, we can only have this sort of thing happen when dealing with nonrelativistic effective field theory.

We will see another example of nontrivial NG counting in the problems.

2.6 The BEHGHK Mechanism of Anderson

In 1964, a series of papers were published in the *Physical Review* in which an exception to Goldstone's theorem was noted.[5] These papers realized that if a continuous symmetry was *gauged*, then there need not be massless modes in the spectrum. To be more precise, the NG fields that you get are purely gauge artifacts and can be removed by a suitable choice of gauge.

[5] "BEHGHK" (pronounced "Beck") comes from a talk by Ben Kilminster.

2.6.1 *A little history*

This seems like a very good place to insert a bit of background on where the "Higgs Mechanism" came from. For those less interested in the history, feel free to skip this section.

After Yang and Mills published their theory of non-Abelian gauge invariance, it was hoped (and that hope was ultimately realized) that one can use this idea to explain the forces of nuclear and particle physics. The problem was that these theories required the existence of massless spin-1 particles, and there were no candidates (remember that gluons were not discovered until much later). When Nambu and Goldstone published their results on spontaneous symmetry breaking, it was found that in addition to the vector bosons, you also have a collection of spin-0 bosons that must be massless. The situation has gone from bad to worse!

It was Philip Anderson (motivated by a result from Julian Schwinger) who first seemed to realize that the situation was not as bad as it seemed. Schwinger had suggested that the requirement of massless gauge bosons was only a requirement in the weak coupling limit. When you have strongly interacting forces, he showed that you need not necessarily have massless gauge fields! Anderson was an expert in condensed matter physics, and he presented a physical example of Schwinger's ideas by studying the behavior of a non-relativistic free-electron gas, where it was known that transverse EM waves do not propagate below the plasma frequency, while above this frequency there are three propagating modes (one longitudinal, two transverse). He showed how this result followed from Schwinger's ideas, and by making the connection to superconductivity, he realized that there need be no massless modes in the system. He then made a very interesting point at the conclusion of his paper:

> It is noteworthy that in most of these cases, upon closer examination, the Goldstone bosons do indeed become tangled up with the Yang-Mills gauge bosons and, thus, do not in any true sense really have zero mass.... We conclude, then, that the Goldstone zero-mass difficulty is not a serious one, because we can probably cancel it off against an equal Yang-Mills zero-mass problem. [Anderson (1963)]

As an amusing afterthought, Anderson points out that when one takes gravity into account, that the breakdown of translational and rotational invariance leads to the presence of three phonons, which, when combined with the two graviton helicities, precisely gives you the right number of degrees of freedom for a massive, spin-2 field. Anderson was way ahead of his time!

The following year, three papers came out, all of which were published in *Physical Review*, Volume 13. All of these papers show, to various degrees, that Goldstone's theorem need not apply to gauge theories, and that when a gauge theory is spontaneously broken, there are no massless particles in the spectrum.

(1) The first paper was by Robert Brout and Francois Englert [Englert and Brout (1964)]. They showed that when a gauge theory is spontaneously broken, the vector boson acquires a mass in perturbation theory. They compute this mass, but admit: "We have not yet constructed a proof in arbitrary order..." although they point out that there should be no problems generalizing the result.

(2) The second paper was by Peter Higgs [Higgs (1964)]. He pointed out that, as Brout and Engert discovered, Goldstone's theorem need not hold when the symmetry is gauged, and furthermore, that when you send the gauge coupling to zero, the longitudinal modes of the gauge field manifest as the required NG fields. So it was Higgs who first makes the connection of "gauge fields eating the NG bosons". In addition, he proposes a model for how this might work; we will use this model in our example below. Higgs's approach is to treat the theory classically – he derives the field equations from a gauge invariant action, and shows that the result is a Proca Lagrangian with an extra (massive) scalar boson, forever known as the "Higgs boson".

(3) The third paper was by Gerald Guralnik, Carl Hagen and Tom Kibble [Guralnik *et al.* (1964)];. Their paper took a somewhat different approach. They pointed out that when a gauged symmetry is spontaneously broken, the corresponding (global) charge *is no longer conserved* despite the existence of a local conservation law. The problem comes from the gauge-fixing conditions: for example, in the radiation gauge, the theory is not manifestly covariant and you can get non-trivial surface terms at infinity. In a Lorentz-covariant gauge (such as Lorenz gauge), the authors point out that you *do* have massless modes as Goldstone's theorem requires, but that these are "gauge [artifacts] rather than physical particles." Indeed, this can be seen explicitly in terms of the later discovered R_ξ gauge with $\xi = 0$ where ghosts and NG bosons are realized as massless particles [Peskin and Schroeder (1995)]. In a physical gauge, however, these extra modes do not appear in the

spectrum, since Goldstone's theorem requires a global conservation law that is explicitly broken by gauge fixing terms!

We now know that the BEHGHK mechanism does seem to play a major role in the standard model of particle physics, with the residual boson mass at around 125 GeV. After this discovery was confirmed at the Large Hadron Collider, Englert and Higgs shared the 2013 Nobel Prize in Physics.

2.6.2 *An example*

We will start by studying the model of a complex, electrically charged scalar field. This is known as the *Landau-Ginzburg model* by condensed matter physicists; the *Coleman-Weinberg model* by particle physicists; and *scalar QED* among students:

$$\begin{aligned} S[\phi, A_\mu] &= \int d^dx \Big\{ |D_\mu\phi|^2 - m^2|\phi|^2 - \lambda|\phi|^4 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \Big\} \\ &= \int d^dx \Big\{ |\partial_\mu\phi|^2 - m^2|\phi|^2 - \lambda|\phi|^4 \\ &\qquad + eA_\mu j^\mu + e^2|\phi|^2 A_\mu A^\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \Big\} \end{aligned} \tag{2.48}$$

where $D_\mu\phi = \partial_\mu\phi + ieA_\mu\phi$, and j^μ is the current defined in Eq. (2.12). As we have seen in Section 2.5, when $m^2 < 0$ the phase symmetry is spontaneously broken and the field picks up a vacuum expectation value $v/\sqrt{2} = \sqrt{-m^2/2\lambda}$. We should therefore expect to see a massless field appear upon going to a physical basis for the fields. In this section we will decompose ϕ in terms of real fields with a different parametrization:

$$\phi(x) = \frac{1}{\sqrt{2}} \left(v + \varphi(x) \right) e^{i\theta(x)/v} \tag{2.49}$$

rather than splitting the field into real and imaginary parts. It is a straightforward exercise to relate our fields (φ, θ) to the previously chosen basis (ϕ_1, ϕ_2), and we leave it to you.

There is a good reason why we chose to decompose ϕ in this way: since the potential only depends on the modulus of ϕ, the dependence on θ is restricted to very few terms:

$$S[\varphi, \theta, A_\mu] = \int d^d x \Big\{ \frac{1}{2}(\partial_\mu \varphi)^2 + \frac{1}{2}(\partial_\mu \theta)^2 - \mathcal{V}(v+\varphi) + eA_\mu \tilde{j}^\mu + \frac{g^2}{2}(v+\varphi)^2 A_\mu A^\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \Big\}$$
$$= \int d^d x \Big\{ \frac{1}{2}(\partial_\mu \theta)^2 + \frac{1}{2} g^2 v^2 A_\mu A^\mu + g A_\mu \tilde{j}^\mu \Big\} + \cdots \tag{2.50}$$

where

$$\tilde{j}^\mu = v \partial^\mu \theta \tag{2.51}$$

is the current in terms of our physical fields, and the ellipses in Eq. (2.50) refer to θ-independent terms. Notice right away that φ does not appear in the current – this represents a *neutral* particle. Also notice that the phase field $\theta(x)$ is the massless NG field that we expect, but in this choice of parametrization it makes no appearance in the potential. This is why this choice of parametrization of real fields is so useful. You should also recognize this current as our Noether current that follows from a *shift symmetry* in the field from Eq. (2.20). This is precisely correct, since under the symmetry:

$$\phi \longrightarrow e^{i\alpha}\phi \quad \Rightarrow \quad \theta \longrightarrow \theta + \alpha v \tag{2.52}$$

while φ does not transform at all, and therefore does not appear in the current. However, the situation is much more interesting now that our symmetry is gauged, since now α can be a function of space-time. Therefore, why not choose $\alpha(x) = -\theta(x)/v$ everywhere? In that case, the field is completely cancelled by the gauge transformation and should not appear in the action at all!

If we look at Eq. (2.50) and use Eq. (2.51) we find

$$S = \int d^d x \frac{1}{2} \Big\{ (\partial_\mu \theta)^2 + 2gvA_\mu \partial^\mu \theta + g^2 v^2 A_\mu A^\mu \Big\} + \cdots$$
$$= \int d^d x \frac{1}{2} g^2 v^2 \left[A_\mu + \partial_\mu \left(\frac{\theta}{gv} \right) \right]^2 + \cdots$$
$$= \int d^d x \frac{1}{2} g^2 v^2 A'_\mu A'^\mu + \cdots \tag{2.53}$$

where A'_μ is precisely the gauge field suitably transformed after the transformation in Eq. (2.52). So θ really has completely vanished from our action after a gauge transformation, and there is no NG field! The final result is a massive vector-boson, and a massive, neutral (real) scalar field.

It might bother you that a field has suddenly vanished from our theory – what about unitarity? How can we start with a theory with two degree of freedom (θ, φ) and lose one? The answer comes from keeping track of the degrees of freedom in the vector boson. In d space-time dimensions, a massless vector boson has $d-2$ degrees of freedom, while a massive vector boson has $d-1$ degrees of freedom. So where did that extra longitudinal degree of freedom come from? It is nothing more than θ! This has led to the idea that, "The massless vector boson *ate* the NG boson and became massive!"

2.6.3 *An interlude: superconductivity*

One of the amazing things about local spontaneous symmetry breaking is that it describes most of the phenomena of superconductivity, without the need to resort to a model! This was first noted in a paper by Steven Weinberg dedicated to Nambu [Weinberg (1986)], and is explained with even more detail in [Weinberg (1996)]. Although it is a bit of an aside from the main thrust of the book, we feel remiss if we do not give an explanation to how superconductivity follows from an EFT of a spontaneously broken U(1) gauge symmetry, since it is such a prime example of how far you can get without knowing "the man behind the curtain!"

The only assumptions we will make are that a U(1) gauge symmetry (E&M) is spontaneously broken by a charge-2 field. It could be a fundamental field (like a Higgs boson) or a composite object (like a Cooper pair of electrons) – those kinds of details are irrelevant for our purposes. We will parametrize this field as

$$\phi(x) = \frac{1}{\sqrt{2}}\,(v+h)\,e^{2ie\theta} \tag{2.54}$$

where θ is the (dimensionless) NG boson, and h is a massive excitation. Now the broken gauge invariance acts on θ through a shift symmetry as in Eq. (2.52), so a gauge transformation $\alpha(x)$ shifts $\theta \to \theta + \alpha$. But since this is a charge-2 object, there is a residual $\mathbb{Z}_2$ symmetry, so that the fields

$$\theta \cong \theta + \frac{\pi}{e} \tag{2.55}$$

must represent the same physics.

Specifying the action $S[\theta, A_\mu, h]$ would require us to pick a model for our theory, but whatever we choose, we know that it must be a function of $A_\mu - \partial_\mu\theta$ in order for the dynamics to remain gauge invariant: remember, only the ground state breaks gauge symmetry, not the action. Therefore, any charge densities and currents in our superconducting material are given by

$$J^0 = -\frac{\delta S}{\delta A^0} = -\frac{\delta S}{\delta \dot{\theta}}, \tag{2.56}$$

$$\vec{J} = \frac{\delta S}{\delta \vec{A}}. \tag{2.57}$$

This is enough to describe several phenomena relating to superconductivity. It is particularly nice that we have this kind of universality, since there is still a lot of debate as to the nature of high-temperature superconductors.

(1) **Meissner effect:** In the broken phase, the vector potential that describes the photon is described by a Proca Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A_\mu A^\mu, \tag{2.58}$$

with $m_A = 2ev$. The resulting field equation for the photon implies that there will be a penetration depth $\lambda \sim m^{-1}$, and the magnetic field inside $|\vec{B}| \sim e^{-\lambda r}$ as you penetrate into the superconductor.

(2) **Flux quantization:** The absence of magnetic fields deep inside the superconductor does not imply that the vector potential vanishes, but it does imply that $A_\mu = \partial_\mu \theta$, whose spacial components are $\vec{A} = \vec{\nabla}\theta$. Doing a line integral of this expression around a closed loop inside the superconductor gives:

$$\oint \vec{A} \cdot d\vec{x} = \oint \vec{\nabla}\theta \cdot d\vec{x}$$

$$\int (\vec{\nabla} \times \vec{A}) \cdot d\vec{S} = \Delta\theta.$$

The left-hand side of this equation is just the magnetic flux through the area bound by the integration loop, while the right-hand side is the change in the Goldstone field going all the way around the loop. This change can be any integer times π/e, since changes in θ by this amount do not affect any physical observable, according to Eq. (2.55). We have therefore proved the famous flux-quantization result:

$$\Phi_B = n\frac{\pi}{e}. \tag{2.59}$$

(3) **Infinite conductivity:** The fact that magnetic flux is quantized is enough to show that the conductivity is infinite deep inside the superconductor since the flux (and therefore the current, by Ampère's Law) cannot decay continuously, but only drop in discrete jumps, which is impossible without large changes in energy.

Another way to see how infinite conductivity arises is by using Eq. (2.56), which tells us that the charge density and $-\theta$ are canonically conjugate variables, so the Hamiltonian of the superconductor is a natural function of J^0 and θ. We therefore have

$$\dot{\theta} = -\frac{\partial \mathcal{H}}{\partial J^0} = -V, \tag{2.60}$$

where V is the energy per unit charge, which is simply the voltage. This tells us that a superconductor carrying a steady current with time-independent fields must have a vanishing potential difference. This is equivalent to the statement that the conductivity is infinite (zero voltage drop for a fixed current).

(4) **AC Josephson Junction:** Consider a gap between two superconducting slabs (labeled 1 and 2). The NG field is nonvanishing in either slab, but $\theta_1 \neq \theta_2$ in general. We can conclude that the dynamics in the gap will be determined by the difference between θ_1 and θ_2: $S_{\text{gap}} = \mathcal{A}F[\Delta\theta]$, where $\mathcal{A}$ is the area of the gap. Now thanks to Eq. (2.55), we know that both $\theta_{1,2}$ can change by $n\pi/e$ independently with no consequence to any observable, so it must be that F is periodic with period π/e. Now imagine there is a nonzero vector potential in the gap. By gauge invariance, we must replace $\Delta\theta \to \int(\vec{\nabla}\theta - \vec{A})\cdot \vec{dx}$, integrated across the gap. Then from Eq. (2.57), we have

$$\vec{J}_{\text{gap}} = \frac{\delta S_{\text{gap}}}{\delta \vec{A}} = \frac{\delta S_{\text{gap}}}{\delta \Delta\theta} \cdot \frac{\delta \Delta\theta}{\delta \vec{A}} = -\hat{n}F'[\Delta\theta], \tag{2.61}$$

where $\hat{n}$ is a unit vector pointing across the gap. So we find that a current flows across the gap. Furthermore, if we additionally impose a fixed potential difference (ΔV) between the two superconducting slabs, Eq. (2.60) gives us $\Delta\theta = -t\Delta V$. Since S_{gap}, and therefore $\vec{J}_{\text{gap}}$, is periodic, we have a an alternating current oscillating with frequency

$$\nu = \frac{e\Delta V}{\pi}. \tag{2.62}$$

This is the same expression Josephson derived using a specific model, but it is clear that the frequency is model independent.

2.7 CCWZ construction of EFT

In 1969, Callan, Coleman, Wess and Zumino (CCWZ) showed that physics truly is invariant to how you choose to parametrize your fields [Coleman *et al.* (1969); Callan *et al.* (1969)]. In particular, this means that we can *image not available*