Contents

List of Figures

List of Tables

Chapter 1

Goldstone Bosons

Goldstone bosons are weakly-coupled states which appear in the low-energy spectrum of any system for which a rigid (or global) symmetry is spontaneously broken (that is, the symmetry is not preserved by the system's ground state). A great deal is known about the properties of these bosons, since at low energies their properties are largely governed by the nature of the symmetries which are spontaneously broken, depending only weakly on the details of the system itself.

This review is devoted to explaining the modern *effective lagrangian* method for identifying Goldstone boson properties. These methods are much more efficient than are the older current-algebra techniques of yore.

1.1 Introduction

It is a common feature of many physical systems that their behaviour is relatively simple when examined only at low energies (or temperatures) compared to the system's own characteristic scales. It often happens that there are relatively few states which can participate in low-energy processes, and their interactions can sometimes become less and less important the lower the energies that are examined. Very general theoretical tools

exist to exploit this simplicity, when it arises.

One such tool is the technique of effective lagrangians. The guiding idea for this method is the belief, first clearly enunciated by Weinberg, that there is no loss of generality in using a field theory to capture the low-energy behaviour of any system. This is because field theory in itself contains very little content beyond ensuring the validity of general 'motherhood' properties like unitarity, cluster decomposition, and so on. According to this point of view, if a field theory is identified which is the most general consistent with the low-energy degrees of freedom and symmetries of any particular system (together with a few 'motherhood' properties) then this field theory *must* provide a good description of the system's low-energy limit.

This is a particularly useful observation when the low-energy degrees of freedom are weakly interacting (*regardless* of how strongly interacting their higher-energy counterparts might be), because then the resulting field theory may be simple enough to be used to predict explicitly the system's low-energy properties. This simplicity is somewhat paradoxical since, as we shall see, the low-energy effective lagrangians are typically very complicated, involving all possible powers of the various fields and their derivatives. Simplicity is achieved in spite of the complicated effective lagrangian because, for weakly-coupled theories, general power-counting arguments exist which permit an efficient identification of the comparatively few interactions which appear at any given order in a low-energy expansion.

Remarkably, there turns out to be a very important situation for which very general results are known concerning the existence of very light degrees of freedom whose low-energy interactions are weak. This occurs whenever a continuous global symmetry is spontaneously broken (*i.e.* which is a symmetry of the hamiltonian but *not* a symmetry

of the ground state), since when this happens Goldstone's theorem guarantees the existence of low-energy Goldstone bosons, as well as determining a great deal about their interactions. These theorems, and their description in terms of an effective lagrangian formulation, are the subject of this review.

1.1.1 A Road Map

This section outlines how the material covered in this review is organized.

1. **General Formalism:** All of the general results may be found in Chapter 1, starting with a statement of the key theorems — those of Noether and Goldstone — which underlie everything else. This is followed by a motivational discussion of the simplest example to which Goldstone's theorem applies. Although the properties of the Goldstone bosons are guaranteed by general theorems, the moral of the example is that these properties are generally not manifest in a low-energy effective lagrangian unless a special choice of variables is used. These variables are identified and exploited first for spontaneously-broken abelian internal symmetries, and then the process is repeated for nonabelian internal symmetries. Both lorentz-invariant and nonrelativistic systems are considered. For the nonrelativistic case, special attention given to the breaking of time reversal, since this qualitatively affects the nature of the low-energy effective lagrangian. The spontaneous breaking of spacetime symmetries, like rotations, translations and lorentz transformations, is not discussed in this review.

2. **Applications:** Chapters 2 through 4 are devoted to specific applications of the methods of Chapter 1 to examples in high-energy/nuclear and condensed-matter

physics. Chapter 2 starts with the classic relativistic example of pions as pseudo-Goldstone bosons, whose study gave birth to many of the techniques described in Chapter 1. (A pseudo-Goldstone boson is the Goldstone boson for an *approximate* symmetry, as opposed to an exact symmetry.) This is followed in Chapter 3 by a study of spin waves (magnons) in both ferromagnets and antiferromagnets. Chapter 4 then closes with a recent, more speculative, application of these ideas to the $SO(5)$ proposal for the high-temperature superconductors.

3. **Bibliography:** Finally, Chapter 5 contains a brief bibiography. It is not meant to be exhaustive, as a great many articles have emerged over the past decades of applications of these methods. I therefore restrict myself to listing those papers and reviews of which I am most familiar. I apologize in advance to the authors of the many excellent articles I have omitted.

The review is aimed at upper-year graduate students, or practicing researchers, since it presupposes a familiarity with quantum field theory. It was written with an audience of high-energy and nuclear phycisists in mind, and so for the most part units are used for which $\hbar = c = 1$. However, I hope it will prove useful to condensed-matter physicists as well. Enjoy!

1.2 Noether's Theorem

We start with a statement of Noether's theorem, since this plays a role in the statement of Goldstone's theorem, which is the main topic of this chapter.

For a field theory Noether's theorem guarantees the existence of a conserved current, j^μ, for every global continuous symmetry of the action. To low orders in the derivative

expansion it is usually enough to work with actions which depend only on the fields and their first derivatives, so we restrict our statement of the theorem to this case.

Consider therefore a system governed by an action $S = \int d^4x\ \mathcal{L}(\phi, \partial_\mu \phi)$, where $\phi(x)$ generically denotes the fields relevant to the problem. We imagine that S is invariant under a set of transformations of these fields, $\delta\phi = \xi_a(\phi)\,\omega^a$, where ω^a denote a collection of independent, spatially constant symmetry parameters. Invariance of S implies that the lagrangian density, $\mathcal{L}$, must vary at most into a total derivative:

$$\delta\mathcal{L} \equiv \partial_\mu \left(\omega^a\, V_a^\mu\right), \tag{1.2.1}$$

for some quantities $V_a^\mu(\phi)$. This equation is meant to hold as an identity, for arbitrary field configurations, ϕ, and for arbitrary constant parameters, ω^a. Rewriting the variation of $\mathcal{L}$ directly in terms of the variations of the fields, and equating to zero the coefficient of the arbitrary constant ω^a in the result then gives:

$$\begin{aligned} \partial_\mu\, V_a^\mu &= \frac{\partial\mathcal{L}}{\partial\phi}\,\xi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\,\partial_\mu\xi_a \\ &= \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right)\right]\ \xi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\,\xi_a\right). \end{aligned} \tag{1.2.2}$$

The statement of the theorem follows from this last equation. It states that the quantities

$$j_a^\mu \equiv -\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\,\xi_a + V_a^\mu, \tag{1.2.3}$$

satisfy the special property

$$\partial_\mu j_a^\mu = 0, \tag{1.2.4}$$

when they are evaluated at a solution to the equations of motion for S — *i.e.* on field configurations for which the square bracket on the right-hand-side of eq. (1.2.2) vanishes.

Even though we have used relativistic notation in this argument, the conclusion, eq. (1.2.4), is equally valid for nonrelativistic systems. For these systems, if we write $\rho_a = j_a^0$ for the temporal component of j_a^μ, and denote its spatial components by the three-vector $\mathbf{j}_a$, then current conservation (eq. (1.2.4)) is equivalent to the familiar continuity equation

$$\frac{\partial \rho_a}{\partial t} + \nabla \cdot \mathbf{j}_a = 0. \tag{1.2.5}$$

Eq. (1.2.4) or eq. (1.2.5), are called conservation laws because they guarantee that the *charges*, Q_a, defined by

$$Q_a(t) = \int_{\text{fixed } t} d^3\mathbf{r}\ \rho_a(\mathbf{r}, t) = \int d^3\mathbf{r}\ j_a^0(x), \tag{1.2.6}$$

are conserved in the sense that they are independent of t. These charges are sometimes called the generators of the symmetry because their commutator with the fields give the symmetry transformations themselves

$$i\omega^a\ [Q_a, \phi(x)] = \omega^a\, \xi_a = \delta\phi. \tag{1.2.7}$$

The existence of such a conserved current carries special information if the symmetry involved should be spontaneously broken, as we now describe.

1.3 Goldstone's Theorem

Whenever the ground state of a system does not respect one of the system's global continuous symmetries, there are very general implications for the low-energy theory. This is the content of Goldstone's theorem, which we now present. This theorem is central to the purpose of this chapter, which is devoted to making its implications manifest in a low-energy effective theory.

Goldstone's theorem states that any system for which a continuous, global symmetry is spontaneously broken, must contain in its spectrum a state, $|G\rangle$ — called a *Goldstone mode*, or *Goldstone boson* since it must be a boson[1] — which has the defining property that it is created from the ground state by performing a spacetime-dependent symmetry transformation. In equations, $|G\rangle$ is defined by the condition that the following matrix element cannot vanish:[2]

$$\langle G|\rho(\mathbf{r},t)|\Omega\rangle \neq 0. \tag{1.3.1}$$

Here, $|\Omega\rangle$ represents the ground state of the system, and $\rho = j^0$ is the density for the conserved charge — guaranteed to exist by Noether's theorem — for the spontaneously broken symmetry.

Before turning to its implications, we outline the proof of this result. The starting point is the assumption of the existence of a *local order parameter*. This can be defined to be a field, $\phi(x)$, in the problem which satisfies two defining conditions. Firstly, it transforms nontrivially under the symmetry in question: *i.e.* there is another field, $\psi(x)$, for which:

$$\delta\psi \equiv i[Q,\psi(x)] = \phi(x). \tag{1.3.2}$$

Q is the conserved charge defined by integrating the density $\rho(\mathbf{r},t)$ throughout all of space. Secondly, the field ϕ must have a nonzero expectation in the ground state:

$$\langle\phi\rangle \equiv \langle\Omega|\phi(x)|\Omega\rangle \equiv v \neq 0. \tag{1.3.3}$$

This last condition would be inconsistent with eq. (1.3.2) if the ground state were invariant

[1] Supersymmetry is an exception to this statement, since spontaneously broken global supersymmetry ensures the existence of a Goldstone fermion, the goldstino.

[2] We use nonrelativistic notation here to emphasize that the conclusions are not specific to relativistic systems. This will prove useful when nonrelativistic applications are considered in later sections.

under the symmetry of interest, since this would mean $Q|\Omega\rangle = 0$, implying the right-hand-side of eq. (1.3.3) must vanish.

The proof of the theorem now proceeds from the following steps. (*i*) Substitute eq. (1.3.2) into eq. (1.3.3); (*ii*) Use $Q = \int \rho \; d^3\mathbf{r}$ in the result, as is guaranteed to be possible by Noether's theorem; (*iii*) Insert a partition of unity as a sum over energy eigenstates, $1 = \sum_n |n\rangle\langle n|$, on either side of the operator ρ. The resulting expression shows that if no energy eigenstate exists which satisfies the defining condition, eq. (1.3.1), then the right-hand-side of eq. (1.3.3) must vanish, in contradiction with the starting assumptions. This proves the theorem. We next elaborate on its consequences.

The defining matrix element, eq. (1.3.1), and the conservation law, eq. (1.2.5), together imply that Goldstone bosons must have a number of important properties. Besides determining their spin and their statistics, it implies two properties which are of particular importance:

1. The Goldstone boson must be *gapless*, in that its energy must vanish in the limit that its (three-) momentum vanishes. That is:

$$\lim_{p \to 0} E(p) = 0. \tag{1.3.4}$$

   To see why this follows from eq. (1.3.1), it is helpful to make the dependence on position and time in this equation explicit by using the identities $\rho_a(\mathbf{r}, t) = e^{-iHt}\, \rho_a(\mathbf{r}, 0)\, e^{iHt}$ and $\mathbf{j}_a(\mathbf{r}, t) = e^{i\mathbf{P}\cdot\mathbf{r}}\, \mathbf{j}_a(0, t)\, e^{-i\mathbf{P}\cdot\mathbf{r}}$, together with the energy- and momentum-eigenstate conditions: $H|\Omega\rangle = \mathbf{P}|\Omega\rangle = 0$, $\mathbf{P}|G(p)\rangle = \mathbf{p}|G(p)\rangle$ and $H|G(p)\rangle = E_p|G(p)\rangle$. Then, differentiation of eq. (1.3.1) with respect to t, and use of the continuity equation, eq. (1.2.5), gives:

$$-iE_p\, e^{-iE_p\, t}\, \langle G|\rho_a(\mathbf{r}, 0)|\Omega\rangle \;\; = \;\; \langle G|\frac{\partial \rho_a}{\partial t}(\mathbf{r}, t)|\Omega\rangle$$

$$= \quad - \langle G|\nabla \cdot \mathbf{j}_a(\mathbf{r},t)|\Omega\rangle \qquad (1.3.5)$$

$$= \quad -i\,\mathbf{p}\cdot\langle G|\mathbf{j}_a(\mathbf{r},t)|\Omega\rangle.$$

Eq. (1.3.4) follows from this last equality in the limit $\mathbf{p} \to 0$, given that the matrix element, $\langle G|\rho_a(\mathbf{r},0)|\Omega\rangle$, cannot vanish *by definition* for a Goldstone boson.

In relativistic systems, for which $E(p) = \sqrt{p^2 + m^2}$ where m is the particle mass, the gapless condition, eq. (1.3.4), is equivalent to the masslessness of the Goldstone particle.

2. More generally, the argument just made can be extended to more complicated matrix elements. One finds in this way that the Goldstone boson for any exact symmetry must completely decouple from all of its interactions in the limit that its momentum vanishes. Physically, this is because eq. (1.3.1) states that in the zero-momentum limit the Goldstone state literally is a symmetry transformation of the ground state. As a result it is *completely indistinguishable* from the vacuum in this limit.

These properties have a lot of implications for the low-energy behaviour of any system which satisfies the assumptions of the theorem. The first guarantees that the Goldstone boson must itself be one of the light states of the theory, and so it must be included in any effective lagrangian analysis of this low energy behaviour. The second property ensures that the Goldstone mode must be weakly coupled in the low-energy limit, and strongly limits the possible form its interactions can take.

The properties of gaplessness and low-energy decoupling can also be useful even if the spontaneously broken 'symmetry' in question is really not an exact symmetry. To the extent that the symmetry-breaking terms, $H_{\rm sb}$, of the system's Hamiltonian are small, the symmetry may be regarded as being approximate. In this case the violation of the gapless

and decoupling properties can usefully be treated perturbatively in $H_{\rm sb}$. The Goldstone particles for any such approximate symmetry — called *pseudo*-Goldstone bosons — are then systematically light and weakly coupled at low energies, instead of being strictly massless or exactly decoupled.

The purpose of the remainder of this chapter is to show in detail how these properties are encoded into the low-energy effective lagrangian. By considering simple examples we find that although these properties are always true, they need not be manifest in the lagrangian in an arbitrary theory. They can be made manifest, however, by performing an appropriate field redefinition to a standard set of field variables. We first identify these variables, and use them to extract the implications of Goldstone's theorem for the low-energy effective theory in the simplest case, for which the symmetry group of interest is abelian. The results are then generalized in subsequent sections to the nonabelian case.

1.4 Abelian Internal Symmetries

In order to see the issues which are involved, it is instructive to consider a simple field theory for which a symmetry is spontaneously broken. We therefore first consider a simple model involving a single complex scalar field, ϕ.

1.4.1 A Toy Example

The lagrangian density:

$$\begin{aligned} \mathcal{L} &= -\partial_\mu \phi^* \partial^\mu \phi - V(\phi^* \phi), \\ \text{with} \qquad V &= \frac{\lambda}{4}\left(\phi^* \phi - \frac{\mu^2}{\lambda}\right)^2, \end{aligned} \tag{1.4.1}$$

is invariant with respect to a $U(1)$ group of symmetries: $\phi \to e^{i\alpha}\ \phi$. This is a global symmetry because the term involving derivatives of ϕ is only invariant if the symmetry

parameter, α, is a constant throughout spacetime. It is called an internal symmetry since the symmetry acts only on fields and does not act at all on the spacetime coordinate, x^μ. For later reference, the Noether current for this symmetry is:

$$j_\mu = -i\left(\phi^* \partial_\mu \phi - \phi\, \partial_\mu \phi^*\right). \tag{1.4.2}$$

For small λ this system is well approximated by a semiclassical expansion, provided that the field ϕ is $O\left(\lambda^{-\frac{1}{2}}\right)$ in size. This may be seen by redefining $\phi = \tilde{\phi}/\sqrt{\lambda}$, and noticing that all of the λ-dependence then scales out of the lagrangian: $\mathcal{L}(\phi, \mu, \lambda) = \frac{1}{\lambda}\mathcal{L}(\tilde{\phi}, \mu, 1)$ — for which the limit $\lambda \to 0$ is seen to be equivalent to $\hbar \to 0$ in the semiclassical limit.

The vacuum of the theory is therefore well described, for small λ, by the classical configuration of minimum energy. Since the classical energy density is a sum of positive terms, $\mathcal{H} = \dot{\phi}^*\dot{\phi} + \nabla\phi^* \cdot \nabla\phi + V(\phi^*\phi)$, it is simple to minimize. The vacuum configuration is a constant throughout spacetime, $\dot{\phi} = \nabla\phi = 0$, and its constant value, $\phi = v$, must minimize the classical potential: $V(v^*v) = 0$. We may use the $U(1)$ symmetry to choose v to be real, and if μ^2 is positive then the solution becomes $v = \mu/\sqrt{\lambda}$. Happily this configuration lies within the conditions of validity of this semiclassical analysis.

Since the vacuum configuration, $\phi = v \neq 0$, is not invariant under the $U(1)$ transformations, $\phi \to e^{i\alpha}\phi$, the $U(1)$ symmetry is seen to be spontaneously broken. Goldstone's theorem should apply, and so we now identify the Goldstone degree of freedom.

The spectrum may be identified by changing variables to the real and imaginary parts of the deviation of the field ϕ from its vacuum configuration. Defining $\mathcal{R} \equiv \sqrt{2}\ \mathrm{Re}\,(\phi - v)$ and $\mathcal{I} \equiv \sqrt{2}\ \mathrm{Im}\,\phi$ diagonalizes the kinetic and mass terms, and the scalar potential in terms of these variables becomes:

$$V = \frac{m_R^2}{2}\mathcal{R}^2 + \frac{g_{30}}{3!}\mathcal{R}^3 + \frac{g_{12}}{2}\mathcal{R}\mathcal{I}^2 + \frac{g_{40}}{4!}\mathcal{R}^4 + \frac{g_{22}}{4}\mathcal{R}^2\mathcal{I}^2 + \frac{g_{04}}{4!}\mathcal{I}^4, \tag{1.4.3}$$

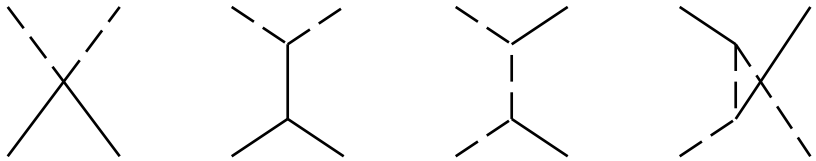

Figure 1.1: The Feynman graphs which describe $\mathcal{R}-\mathcal{I}$ scattering at tree level. Solid lines denote $\mathcal{R}$ and dashed lines represent $\mathcal{I}$.

where the couplings and masses in this potential are given in terms of the original parameters, λ and μ, by:

$$m_R^2 = \lambda\mu^2, \qquad \frac{g_{30}}{3!} = \frac{g_{12}}{2} = \frac{\lambda v}{2\sqrt{2}}, \qquad \frac{g_{40}}{4!} = \frac{g_{04}}{4!} = \frac{g_{22}}{4} = \frac{\lambda}{16}. \tag{1.4.4}$$

Notice the existence of a massless field, $\mathcal{I}$, as is required by Goldstone's theorem. We can verify that $\mathcal{I}$ really is the Goldstone boson by writing the Noether current, eq. (1.4.2), in terms of the mass eigenstates, $\mathcal{R}$ and $\mathcal{I}$:

$$j_\mu = v\sqrt{2}\,\partial_\mu\mathcal{I} + (\mathcal{R}\,\partial_\mu\mathcal{I} - \mathcal{I}\,\partial_\mu\mathcal{R})\,. \tag{1.4.5}$$

Clearly the matrix element:

$$\langle\mathcal{I}(p)|j^\mu(x)|0\rangle \propto v\sqrt{2}\,p^\mu\,e^{-ipx} \tag{1.4.6}$$

does not vanish (unless $v = 0$), as is required of a Goldstone boson.

A puzzle with the potential of eqs. (1.4.3) and (1.4.4) is that the Goldstone boson, $\mathcal{I}$, appears in the scalar potential, and so its couplings do not appear to vanish in the limit of vanishing momentum. This is only an appearance, however, and $\mathcal{I}$ really does decouple at low energies, as can be tested by computing Goldstone boson scattering in this limit. For example, the S-matrix at tree level for $\mathcal{I}-\mathcal{R}$ scattering may be computed

by evaluating the Feynman graphs of Fig. (1.1). The result is:

$$S[\mathcal{R}(r)+\mathcal{I}(s)\to\mathcal{R}(r')+\mathcal{I}(s')]=\frac{i\mathcal{A}\,\delta^4(r+s-r'-s')}{(2\pi)^2\sqrt{16s^0r^0s'^0r'^0}}\,, \tag{1.4.7}$$

with

$$\mathcal{A}=-g_{22}+\frac{g_{12}\,g_{30}}{(s+s')^2+m_R^2-i\epsilon}+g_{12}^2\left[\frac{1}{(s+r)^2-i\epsilon}+\frac{1}{(s-r')^2-i\epsilon}\right]. \tag{1.4.8}$$

In the limit $s^\mu, s'^\mu\to 0$ this becomes (using the condition $r^2=r'^2=-m_R^2$):

$$\begin{aligned}\mathcal{A} &\to -g_{22}+\frac{g_{12}\,g_{30}}{m_R^2}-\frac{2g_{12}^2}{m_R^2},\\ &= \lambda\left(-\frac{1}{2}+\frac{3}{2}-1\right)=0.\end{aligned} \tag{1.4.9}$$

The scattering amplitude indeed vanishes in the zero-momentum limit, as it must according to Goldstone's theorem. This vanishing is not manifest in the lagrangian, however, and is only accomplished through a nontrivial cancellation of terms in the S-matrix. For many purposes, not least when constructing an effective theory to describe the low energy interactions of the Goldstone bosons, it would be preferable to have this decoupling be manifest in the lagrangian. We will now do so, by making a field redefinition to a new set of variables for which decoupling becomes explicit.

1.4.2 A Better Choice of Variables

In order to identify which variables would make the decoupling of Goldstone bosons more explicit in the lagrangian, it is useful to recall the definition of what the Goldstone mode physically is. Its defining condition, eq. (1.3.1), can be interpreted to mean that the Goldstone modes are obtained from the ground state by performing a symmetry transformation, but with a spacetime dependent transformation parameter.

In the example considered in the previous section the ground state configuration is $\phi = v$, and so a local symmetry transformation of this ground-state would be $\phi = ve^{i\theta(x)}$. If this is substituted into the lagrangian of eq. (1.4.1), we find $\mathcal{L}\left(\phi = ve^{i\theta(x)}\right) = -v^2\partial_\mu\theta\partial^\mu\theta$. θ does not drop out of the problem because, although the lagrangian vanishes when it is evaluated at $\phi = v$, the configuration $ve^{i\theta(x)}$ is only related to $\phi = v$ by a symmetry when θ is a constant. This fact that θ parameterizes a symmetry direction when it is restricted to constant field configurations guarantees that any θ-dependence of $\mathcal{L}$ *must* involve at least one derivative of θ, thereby dropping out of the problem in the limit of small derivatives — *i.e.* small momenta, or long wavelengths.

All of this suggests that θ would make a good representation for the Goldstone mode, since this is precisely what a Goldstone mode is supposed to do: decouple from the problem in the limit of small momenta. We are led to the suggestion of using polar coordinates in field space,

$$\phi(x) = \chi(x)\, e^{i\theta(x)}, \tag{1.4.10}$$

in order to better exhibit the Goldstone boson properties. In this expression both θ and χ are defined to be real. Substituting this into the lagrangian gives:

$$\mathcal{L} = -\partial_\mu\chi\partial^\mu\chi - \chi^2\partial_\mu\theta\partial^\mu\theta - V(\chi^2). \tag{1.4.11}$$

It is clear that these variables do the trick, since the fact that θ appears in the definition, eq. (1.4.10), in the same way as does a symmetry parameter guarantees that it completely drops out of the scalar potential, as must a Goldstone boson if its low-energy decoupling is to be made explicit.

A price has been paid in exchange for making the low-energy decoupling of the Goldstone boson explicit, however. This price is most easily seen once the fields are

canonically normalized, which is acheived by writing $\chi = v + \frac{1}{\sqrt{2}}\,\chi'$ and $\theta = \frac{1}{v\sqrt{2}}\,\varphi$. With these variables the lagrangian is seen to have acquired nominally nonrenormalizable interactions:

$$\mathcal{L}_{\rm nr} = -\left[\frac{\chi'}{\sqrt{2}\,v^2} + \frac{\chi'^2}{4v^2}\right]\,\partial_\mu\varphi\partial^\mu\varphi. \tag{1.4.12}$$

Of course, the S-matrix for the theory in these variables is identical to that derived from the manifestly renormalizable lagrangian expressed in terms of the variables $\mathcal{R}$ and $\mathcal{I}$. So the S-matrix remains renormalizable even when computed using the variables χ' and φ. (The same is not true for *off-shell* quantities like Green's functions, however, since the renormalizability of these quantities need not survive a nonlinear field redefinition.)

In this toy model there is therefore a choice to be made between making the lagrangian manifestly display either the renormalizability of the theory, or the Goldstone boson nature of the massless particle. Which is best to keep explicit will depend on which is more convenient for the calculation that is of interest. Since, as we shall see, renormalizability is in any case given up when dealing with effective low-energy field theories, it is clear that the variables which keep the Goldstone boson properties explicit are the ones of choice in this case.

1.4.3 The General Formulation

The reason why the above redefinition works may be seen by asking how the $U(1)$ symmetry acts on the new variables. The key observation is that the symmetry transformation becomes *inhomogeneous*: $\theta \to \theta + \alpha$, where α is a constant. In terms of the canonically normalized field, φ, this transformation law becomes:

$$\varphi \to \varphi + \sqrt{2}\,v\,\alpha. \tag{1.4.13}$$

This kind of transformation rule is the hallmark of a Goldstone boson, since it enforces the explicit nature of all of the Goldstone boson properties in the lagrangian. In fact — as can be expected from the generality of Goldstone's theorem — they can all be derived purely on the grounds of this symmetry transformation, and do not rely at all on the details of the underlying model which motivated its consideration.

To show that this is true, imagine writing an arbitrary effective theory for a real scalar field, φ, subject only to the symmetry of eq. (1.4.13) (and, for simplicity, to Poincaré invariance). The most general lagrangian which is invariant under this transformation is an arbitrary function of the derivatives, $\partial_\mu\varphi$, of the field. An expansion in interactions of successively higher dimension then gives:

$$\mathcal{L}^{\text{eff}}(\varphi) = -\frac{1}{2}\,\partial_\mu\varphi\partial^\mu\varphi - \frac{a}{4\,v^4}\,\partial_\mu\varphi\partial^\mu\varphi\;\partial_\nu\varphi\partial^\nu\varphi + \cdots, \tag{1.4.14}$$

where we have inserted a power of v as appropriate to ensure that the parameter a is dimensionless. This accords with the expectation that it is the symmetry-breaking scale, v, which sets the natural scale relative to which the low energy limit is to be taken. In the toy model just considered, integrating out the heavy field, χ' produces these powers of v through the appearance of the inverse of the heavy mass, $m_{_R}$. The result must be an effective lagrangian of the form of eq. (1.4.14), but with a specific, calculable coefficient for the parameter a.

This, most general, lagrangian automatically ensures that φ has all of the Goldstone boson properties. For instance, since the symmetry implies that $\mathcal{L}^{\text{eff}}$ can only depend on derivatives of φ, it ensures that φ cannot appear at all in the scalar potential, and so in particular ensures that φ is massless. Similarly, applying Noether's theorem to the kinetic term for φ implies that there is a contribution to the Noether current, j^μ, which is linear

in φ:

$$j^\mu = \sqrt{2}\, v \, \left[\partial^\mu \varphi + \frac{a}{v} (\partial^\nu \varphi \, \partial_\nu \varphi) \, \partial^\mu \varphi + \cdots \right] . \tag{1.4.15}$$

The ellipses in this expression represent contributions to j^μ which come from other terms in the lagrangian besides the φ kinetic term. Clearly this ensures that the matrix element $\langle G | j^\mu | 0 \rangle \neq 0$ so long as $v \neq 0$.

Such an understanding of the Goldstone nature of a field, like φ, as an automatic consequence of a symmetry is clearly invaluable when constructing effective lagrangians for systems subject to spontaneous symmetry breaking. We next turn to the generalization of these results to the more general case of nonabelian internal symmetries.

1.5 Nonabelian Internal Symmetries

The lesson learned from the abelian example is that half of the art of constructing effective lagrangians for Goldstone bosons lies in the choice of a convenient set of variables in terms of which their properties are explicitly displayed in the lagrangian. In this section the above construction is generalized to the case of nonabelian, global, internal symmetries.

1.5.1 A Second Toy Model

As guidance towards an appropriate choice of field variables we once more start with a simple toy model for which the underlying theory is explicitly known. Consider, therefore, a system of N real scalar fields, $\phi^i, i = 1, \ldots, N$, which for simplicity of notation we arrange into an N-component column vector, denoted by ϕ with no superscript. Notice that there is no loss of generality in working with real fields, since any complex fields could always be decomposed into real and imaginary parts. We take as lagrangian density

$$\mathcal{L} = -\frac{1}{2} \, \partial_\mu \phi^T \partial^\mu \phi - V(\phi), \tag{1.5.1}$$

where the superscript 'T' denotes the transpose, and where $V(\phi)$ is a potential whose detailed form is not required. The kinetic term of this lagrangian is manifestly invariant under the Lie group, $O(N)$, of orthogonal rotations amongst the N real fields: $\phi \to O\phi$, where the O's are independent of spacetime position, $\partial_\mu O = 0$, and $O^T O = 1$. In general, the potential $V(\phi)$ need not be also invariant under these $O(N)$ transformations, but may only preserve some subgroup of these, $G \subset O(N)$. That is, if $g \in G$, then $V(g\phi) = V(\phi)$ for all fields ϕ.

Suppose now that for some regime of parameters this model is well described by the semiclassical approximation, and further that the potential, V, is minimized for some nonzero value for the fields: $\phi = v \neq 0$. If this is the case, then the symmetry group G may be spontaneously broken to some subgroup, $H \subset G$, which is defined by: $hv = v$, for all $h \in H$.

1.5.2 A Group-Theoretic Aside

It is important to notice that the current whose existence is guaranteed by Noether's theorem — and so which plays the central role in Goldstone's theorem — arises only if the symmetry of interest is *continuous*. Continuous here means that the group elements may be parameterized by a continuous parameter (like a rotation angle), as opposed to a discrete label. Groups with continuous labels are called Lie groups provided their labels are sufficiently smooth. Before proceeding it is useful to pause to record some mathematical properties of such Lie groups, and their associated Lie algebras.

1. Typically the continuous symmetry groups which arise in physical applications do so as explicit finite-dimensional unitary matrices. As a result a special role is played by compact groups, for which the parameter space of the group is a compact set.

Compact groups are of such special interest since it is only for compact groups that finite-dimensional, unitary and faithful matrix representations exist.[3] We assume compact groups throughout what follows, and we work explicitly with representations involving finite-dimensional and unitary matrices, $g^\dagger = g^{-1}$.

2. There is also no loss of generality in assuming our representation matrices, g, to be real: $g = g^*$. This is because any complex representation may always be decomposed into its real and imaginary parts. This convention is ensured in the scalar-field example we are considering by choosing to employ only real fields. We do *not* assume these matrices to be irreducible. Recall that if the matrices are reducible, then there is a basis in which they can be written in a block-diagonal form:

$$g = \begin{pmatrix} g_{(1)} & & \\ & \ddots & \\ & & g_{(n)} \end{pmatrix}. \tag{1.5.2}$$

3. It is useful to phrase much of what follows in terms of the Lie algebra of G and H rather than in terms of the Lie groups themselves. That is, we take advantage of the fact that any group element which is connected to the identity element, $g = 1$, may be written as a matrix exponential: $g = \exp[i\alpha^a T_a]$, of a linear combination of a collection of basis matrices, or generators, T_a, $a = 1, ..., d$ where d is called the dimension of the group. The T_a's lie inside what is called the Lie algebra of G. The unitarity and reality of the group elements, g, imply the matrices T_a to be hermitian and imaginary:

$$T_a = T_a^\dagger = -T_a^* = -T_a^T. \tag{1.5.3}$$

[3] A representation is faithful if there is a one-to-one correspondence between the group elements and the matrices which represent them. Since the groups of interest are usually *defined* by a finite-dimensional and unitary representation, this representation is, by definition, faithful.

4. Since the generators, T_a, are finite dimensional and hermitian, it follows that the matrix $N_{ab} = \mathrm{Tr}(T_a T_b)$ is positive definite. As a result we are free to redefine the generators to ensure that $N_{ab} = \delta_{ab}$. With this choice there is no distinction to be made between indices a and b which are superscripts and subscripts. We assume this convenient choice to have been made in what follows.

5. Closure of the group multiplication law — *i.e.* the statement that $g_1, g_2 \in G$ implies $g_1 g_2 \in G$ —implies commutation relations for the T_a's: $T_a\, T_b - T_b\, T_a = i\, c_{abd}\, T_d$ where the c_{abd}'s are a set of constant coefficients which are characteristic of the group involved. From its definition it is clear that c_{abd} is antisymmetric under the interchange of the indices a and b. Whenever the generators are chosen so that $N_{ab} = \delta_{ab}$ it also turns out that c_{abd} is completely antisymmetric under the interchange of *any* two indices.

6. For the present purposes it is convenient to choose a basis of generators which includes the generators of the subgroup H as a subset. That is, choose $\{T_a\} = \{t_i, X_\alpha\}$, where the t_i's generate the Lie algebra of H, and the X_α's constitute the rest. Since H is defined as the group which preserves the vacuum configuration, its generators must satisfy $t_i v = 0$. The closure of the subgroup, H, under multiplication ensures that $t_i\, t_j - t_j\, t_i = i\, c_{ijk}\; t_k$ (with no X_α's on the right-hand-side). Schematically this can be written $c_{ij\alpha} = 0$.

7. The X_α's do not lie within the Lie algebra of H, and so satisfy $X_\alpha v \neq 0$. They are said to generate the space, G/H, of *cosets*. A coset is an equivalence class which is defined to contain all of the elements of G that are related by the multiplication by an element of H. Physically, the X_α's represent those generators of the symmetry

group, G, which are spontaneously broken.

8. The group property of H described above, together with the complete antisymmetry of the c_{abd}'s implies a further condition: $c_{i\alpha j} = 0$. This states that $t_i\, X_\alpha - X_\alpha\, t_i = i\, c_{i\alpha\beta} X_\beta$ (with no t_j's on the right-hand-side). This states that the X_α's fall into a (possibly reducible) representation of H. Once exponentiated into a statement about group multiplication, the condition $tX - Xt \propto X$ implies, for any $h \in H$, that $hX_\alpha h^{-1} = L_\alpha{}^\beta X_\beta$ for some coefficients, $L_\alpha{}^\beta$.

   By contrast, $X_\alpha\, X_\beta - X_\beta\, X_\alpha$ need not have a particularly simple form, and can be proportional to both X_γ's and t_i's.[4]

1.5.3 The Toy Model Revisited

Returning to the toy model defined by the lagrangian, eq. (1.5.1), we know Goldstone's theorem implies that the assumed symmetry-breaking pattern must give rise to a collection of massless Goldstone bosons, whose interactions we wish to exhibit explicitly. The Goldstone modes are, intuitively, obtained by performing symmetry transformations on the ground state. Since an infinitesimal symmetry transformation on the ground state corresponds to the directions $X_\alpha v$ in field space, we expect the components of ϕ in this direction, $v^T X_\alpha \phi$, to be the Goldstone bosons. It is indeed straightforward to verify that the G-invariance of the lagrangian ensures the masslessness of these modes. There is consequently precisely one Goldstone degree of freedom for each generator of G/H.

More generally, in order to make low-energy decoupling of these Goldstone bosons manifest we require that they do not appear at all in the scalar potential. Following the

[4] If the right-hand-side of $X_\alpha\, X_\beta - X_\beta\, X_\alpha$ were assumed not to contain any X_γ's, then the coset G/H would be called a *symmetric* space, but we do not make this assumption here.

example taken for the case of the abelian symmetry, we therefore change variables from $\phi(x) = \{\phi^i\}$ to $\chi(x) = \{\chi^n\}$ and $\theta(x) = \{\theta^\alpha\}$, where

$$\phi = U(\theta)\ \chi, \tag{1.5.4}$$

and $U(\theta) = \exp[i\theta^\alpha(x)X_\alpha]$ is a spacetime-dependent symmetry transformation in the direction of the broken generators, X_α.

In order for eq. (1.5.4) to provide a well defined change of variables, χ must satisfy some kind of constraint. We therefore require that χ be perpendicular (in field space) to the Goldstone directions, $X_\alpha v$. That is:

$$v^T X_\alpha \chi = 0, \qquad \text{for all } x^\mu \text{ and } X_\alpha. \tag{1.5.5}$$

Notice that this constraint — together with the identity $v^T X_\alpha v = 0$, which follows from the antisymmetry of the X_α's — is precisely what is required to ensure the vanishing of the cross terms, proportional to $\partial_\mu\theta^\alpha\partial^\mu\chi'^n$, in the quadratic part of the expansion of the kinetic terms about the ground state configuration: $\chi = v + \chi'$.

Since $U(\theta)$ is an element of G, the variable θ is guaranteed to drop out of the scalar potential. Of course, this is the point of this change of variables, and it happens because G-invariance requires the potential to satisfy $V(U\chi) = V(\chi)$. As a result, all of the terms in $\mathcal{L}$ which involve the Goldstone bosons, θ, vanish when $\partial_\mu\theta^\alpha = 0$, and eqs. (1.5.4) and (1.5.5) define the change of variables which makes explicit the low-energy Goldstone-boson decoupling.

We pause now to briefly argue that it is always possible to satisfy eq. (1.5.5) starting from an arbitary smooth field configuration, $\phi(x)$. That is, we argue that it is always possible to find a spacetime dependent group element, $U(\theta) \in G$, for which $\chi = U^{-1}\phi$ satisfies eq. (1.5.5). To this end consider the following function, $\mathcal{F}[g(\alpha)] \equiv v^T g(\alpha)\phi$, where

$g(\alpha)$ is an arbitrary, spacetime-dependent element of G. Focus, for a moment, on $\mathcal{F}$ as a function of the parameters, α^a, of the group for a fixed spacetime position, x^μ. Since all of the variables, ϕ, v and g have been chosen to be real, and since the group, G, is compact, $\mathcal{F}(\alpha)$ defines a real-valued function having a compact range. It is a theorem that any such function must have a maximum and a minimum, and so there exist group elements, $\overline{g} = g(\overline{\alpha})$, for which $(\partial\mathcal{F}/\partial\alpha^a)|_{\alpha=\overline{\alpha}}$ vanishes. Repeating this condition for each point in spacetime defines functions $\alpha^a(x)$ whose smoothness follows from the assumed smoothness of $\phi(x)$.

The final step in the argument is to show that the existence of these stationary points of $\mathcal{F}$ also give solutions to the problem of finding a U for which $\chi = U^{-1}\phi$ satisfies eq. (1.5.5). This last step follows by explicitly taking the derivative of g with respect to α^a: and using parameters α such that $(\partial g/\partial\alpha^a)g^{-1} = T_a$. In this case the vanishing of $\partial\mathcal{F}/\partial\alpha^a$, when evaluated at $g = \overline{g} = g(\overline{\alpha})$, implies $v^T T_a \overline{g}\phi = 0$. We see that the choice $\chi = U^{-1}\phi$, with $U = \overline{g}^{-1}$, therefore satisfies eq. (1.5.5), and the existence of such a solution follows from the existence of a maximum and minimum of $\mathcal{F}$.

This concludes the argument. We next explore the properties of the new variables.

1.5.4 The Nonlinear Realization

Having motivated this choice of variables, we now determine how they transform under the group G of symmetry transformations. The transformation rules which we obtain — and which we show to be unique, up to field redefinitions, below — carry all of the information concerning Goldstone boson properties, and so requiring low-energy lagrangians to be invariant under these transformations automatically encodes these properties into the low-energy theory.

The starting point is the transformation rule for ϕ: $\phi \to \tilde{\phi} = g\phi$, where $g = \exp[i\alpha^a T_a]$. The transformation rule for the new variables is then $\theta \to \tilde{\theta}$ and $\chi \to \tilde{\chi}$, where $\phi = U(\theta)\chi$ and $\tilde{\phi} = U(\tilde{\theta})\tilde{\chi}$. That is, under any transformation, $g \in G$, $\theta, \chi, \tilde{\theta}$ and $\tilde{\chi}$ are related by:

$$gU(\theta)\chi = U(\tilde{\theta})\tilde{\chi}. \tag{1.5.6}$$

This last equation states that the matrix $\gamma \equiv \tilde{U}^{-1}gU$ (where $\tilde{U}$ denotes $U(\tilde{\theta})$) has the property that $\gamma\chi = \tilde{\chi}$. The central result to be now proven is that this condition implies that γ must lie within the subgroup, H of unbroken transformations, and so may be written $\gamma = \exp(iu^i t_i)$, for some function $u^i(\theta, g)$. Once this has been demonstrated, the transformation law therefore becomes:

$$\theta^\alpha \to \tilde{\theta}^\alpha(\theta, g) \qquad \text{and} \qquad \chi \to \tilde{\chi}(\theta, g, \chi), \tag{1.5.7}$$

where

$$\begin{aligned} g\, e^{i\theta^\alpha X_\alpha} &= e^{i\tilde{\theta}^\alpha X_\alpha}\, e^{iu^i t_i}, \\ \tilde{\chi} &= e^{iu^i t_i}\, \chi. \end{aligned} \tag{1.5.8}$$

The first of the eqs. (1.5.8) should be read as defining the nonlinear functions $\tilde{\theta}^\alpha(\theta, g)$ and $u^i(\theta, g)$. They are defined by finding the element, $ge^{i\theta\cdot X} \in G$, and then decomposing this matrix into the product of a factor, $e^{\tilde{\theta}\cdot X}$, lying in G/H times an element, $e^{iu\cdot t}$, in H. The second line of eqs. (1.5.8) then defines the transformation rule for the non-Goldstone fields, χ.

These rules generally define transformation laws which are nonlinear in the Goldstone fields, θ^α. They furnish, nonetheless, a faithful realization of the symmetry group G, in that $\tilde{\theta}(\theta, g_1 g_2) = \widetilde{\tilde{\theta}}(\tilde{\theta}(\theta, g_2), g_1)$ *etc.*. This may either be directly verified using the defi-

nitions of eqs. (1.5.8), or by noticing that this property is inherited from the faithfulness of the original linear representation of G on ϕ.

There is a particularly interesting special case for which eqs. (1.5.8) can be explicitly solved for $\gamma = e^{iu\cdot t}$ and $\tilde{U} = e^{i\tilde{\theta}\cdot X}$. This is when $g = h$ lies in H. In this case, the solution is easily seen to be: $\gamma = h$ and $\tilde{U} = hUh^{-1}$. Both χ and θ therefore transform *linearly* under the unbroken symmetry transformations, H. That is:

$$\begin{aligned} \theta^\alpha X_\alpha &\rightarrow \tilde{\theta}^\alpha X_\alpha = h\,\theta^\alpha X_\alpha\, h^{-1}, \\ \chi &\rightarrow \tilde{\chi} = h\chi, \end{aligned} \tag{1.5.9}$$

for all $h \in H$.

For the broken symmetries, $g \in G/H$, it is useful to specialize to an infinitesimal transformation, $g = 1 + i\omega^\alpha X_\alpha + \cdots$. In this case we have $\gamma = 1 + iu^i(\theta,\omega)t_i + \cdots$, and $U(\tilde{\theta}) = U(\theta)[1 + i\Delta^\alpha(\theta,\omega)X_\alpha + \cdots]$, where $u^i(\theta,\omega)$ and $\Delta(\theta,\omega)$ must both also be infinitesimal quantities. Eq. (1.5.8) gives them explicitly to be:

$$\begin{aligned} \Delta_\alpha &= \mathrm{Tr}\Big[X_\alpha e^{-i\theta\cdot X}(\omega\cdot X)e^{i\theta\cdot X}\Big], \\ &\approx \omega_\alpha - c_{\alpha\beta\gamma}\omega^\beta\theta^\gamma + O(\theta^2); \end{aligned} \tag{1.5.10}$$

$$\begin{aligned} u_i &= \mathrm{Tr}\Big[t_i e^{-i\theta\cdot X}(\omega\cdot X)e^{i\theta\cdot X}\Big], \\ &\approx -c_{i\alpha\beta}\omega^\alpha\theta^\beta + O(\theta^2) \end{aligned} \tag{1.5.11}$$

where we used: $\mathrm{Tr}(X_\alpha X_\beta) = \delta_{\alpha\beta}$, $\mathrm{Tr}(t_i t_j) = \delta_{ij}$ and $\mathrm{Tr}(t_i X_\alpha) = 0$.

This last expression for $\Delta_\alpha(\theta,\omega)$ can be re-expressed in terms of the change, $\delta\theta^\alpha \equiv \xi^\alpha(\theta,\omega) \equiv \tilde{\theta}^\alpha - \theta^\alpha$, of the Goldstone-boson fields. The relation between Δ_α and ξ^α is linear: $\Delta_\alpha = M_{\alpha\beta}(\theta)\,\xi^\beta$, where the matrix, $M_{\alpha\beta}$, of coefficients may be computed using

the following useful identity, which holds for any two square matrices, A and B:

$$\begin{aligned} e^{-iA}\, e^{i(A+B)} &= 1 + i\int_0^1 ds\; e^{-isA} B e^{is(A+B)}, \\ &= 1 + i\int_0^1 ds\; e^{-isA} B e^{isA} + O(B^2). \end{aligned} \tag{1.5.12}$$

Using this with $A = \theta \cdot X$ and $B = \xi \cdot X$ gives:

$$M_{\alpha\beta} = \int_0^1 ds\;\; \mathrm{Tr}\Big[X_\alpha e^{-is\theta\cdot X} X_\beta e^{is\theta\cdot X}\Big]. \tag{1.5.13}$$

The transformation rules for the θ^α with respect to the broken symmetries in G/H have two important properties. The first crucial property is that the transformation law is *inhomogeneous* in the broken symmetry parameters, since

$$\delta\theta^\alpha = \omega^\alpha - c^\alpha{}_{\beta\gamma}\omega^\beta\theta^\gamma + O(\theta^2). \tag{1.5.14}$$

As was observed earlier for the abelian example, it is this property which enforces the decoupling of the Goldstone bosons at low energies.

The second important property is that, for a nonabelian group, the symmetries in G/H act *nonlinearly* on the fields θ^α. This property is also significant since it ruins many of the consequences which would otherwise hold true for symmetries which are linearly realized. For example, the masses of those particles whose fields lie in a linear representation of a symmetry group necessarily have equal masses, *etc.*. The same is not true for particles whose fields are related by nonlinear transformations.

There is a corollary which follows from the nonlinearity of the realization of the symmetries in G/H. The fact that the transformation of θ^α and χ^n are both field dependent implies that the action of these symmetries are spacetime dependent. For example, even though the transformation parameters themselves, ω^α, are constants — since G is a global

symmetry — the transformation matrix $\gamma = e^{iu\cdot t}$ which appears in the χ transformation law is not a constant, $\partial_\mu \gamma \neq 0$. This fact complicates the construction of lagrangians which are invariant with respect to these symmetries.

1.6 Invariant Lagrangians

With the transformation rules for the Goldstone boson fields in hand we may now turn to the construction of invariant Lagrangians which can describe their low-energy interactions. The main complication here is in the construction of the kinetic terms, since the transformation rules for the fields are spacetime dependent due to their complicated dependence on the fields.

A clue as to how to proceed can be found by reconsidering the toy model of scalar fields ϕ. In this case the kinetic term, proportional to $\partial_\mu \phi^T \partial^\mu \phi$ is manifestly G invariant. It must therefore remain so after performing the change of variables to θ and χ. To see how this is comes about, we notice that after the replacement $\phi = U\chi$ we have: $\partial_\mu \phi = U(\partial_\mu \chi + U^{-1}\partial_\mu U \chi)$. In terms of the new variables, the kinetic term is invariant because the combination $\mathcal{D}_\mu \chi = \partial_\mu \chi + U^{-1}\partial_\mu U \chi$ transforms covariantly: $\mathcal{D}_\mu \chi \to \gamma \mathcal{D}_\mu \chi$. It does so because $U^{-1}\partial_\mu U$ transforms like a gauge-potential:

$$\begin{aligned} U^{-1}\partial_\mu U &\to \tilde{U}^{-1}\partial_\mu \tilde{U}, \\ &= \gamma \left(U^{-1}\partial_\mu U\right) \gamma^{-1} - \partial_\mu \gamma \, \gamma^{-1}. \end{aligned} \tag{1.6.1}$$

More information emerges if we separate out the parts of $U^{-1}\partial_\mu U$ which are proportional to X_α from those which are proportional to t_i since the inhomogeneous term, $\partial_\mu \gamma \, \gamma^{-1}$, is purely proportional to t_i. That is, if we define:

$$U^{-1}\partial_\mu U = -i\mathcal{A}^i_\mu t_i + i e^\alpha_\mu X_\alpha, \tag{1.6.2}$$

then each of these terms transforms separately under G transformations:

$$\begin{aligned} -i\mathcal{A}^i_\mu(\theta)t_i &\rightarrow -i\mathcal{A}^i_\mu(\tilde{\theta})t_i = \gamma\,[-i\mathcal{A}^i_\mu(\theta)t_i]\,\gamma^{-1} - \partial_\mu\gamma\,\gamma^{-1}, \\ ie^\alpha_\mu(\theta)X_\alpha &\rightarrow ie^\alpha_\mu(\tilde{\theta})X_\alpha = \gamma\,[ie^\alpha_\mu(\theta)X_\alpha]\,\gamma^{-1}. \end{aligned} \tag{1.6.3}$$

We see that the quantity $\mathcal{A}^i_\mu$ transforms like a gauge potential. For infinitesimal transformations, $g \approx 1 + i\omega^\alpha X_\alpha$ and $\gamma(\theta, g) \approx 1 + iu^i(\theta,\omega)t_i$, we have:

$$\delta\mathcal{A}^i_\mu(\theta) = \partial_\mu u^i(\theta,\omega) - c^i{}_{jk}u^j(\theta,\omega)\mathcal{A}^k_\mu(\theta). \tag{1.6.4}$$

Similarly, $e^\alpha_\mu(\theta)$ transforms covariantly, with

$$\delta e^\alpha_\mu(\theta) = -c^\alpha{}_{i\beta}u^i(\theta,\omega)\; e^\beta_\mu(\theta). \tag{1.6.5}$$

In this last expression, the structure constants define representation matrices, $(\mathcal{T}_i)_{\alpha\beta} = c_{\alpha i\beta}$, of the Lie algebra of H. These are the same matrices which define the representation of H that the generators X_α form, and it is important for later purposes to recall that these representation matrices need *not* be irreducible. If this representation *is* reducible then it is possible to define more G-invariant quantities with which to build the low energy lagrangian than would otherwise be possible.

If we extract the overall factor of $\partial_\mu\theta^\alpha$, so that $\mathcal{A}^i_\mu = \mathcal{A}^i_\alpha(\theta)\;\partial_\mu\theta^\alpha$ and $e^\alpha_\mu = e^\alpha{}_\beta(\theta)\;\partial_\mu\theta^\beta$, then the identity, eq. (1.5.12), gives the following expressions for the coefficients:

$$\begin{aligned} \mathcal{A}^i_\alpha(\theta) &= -\int_0^1 ds\;\mathrm{Tr}\Big[t_i e^{-is\theta\cdot X}X_\alpha e^{is\theta\cdot X}\Big], \\ &\approx \frac{1}{2}c_{i\alpha\beta}\theta^\beta + O(\theta^2), \end{aligned} \tag{1.6.6}$$

and

$$e^\alpha{}_\beta(\theta) = \int_0^1 ds\;\mathrm{Tr}\Big[X_\alpha e^{-is\theta\cdot X}X_\beta e^{is\theta\cdot X}\Big],$$

$$\approx \quad \delta_{\alpha\beta} - \frac{1}{2} c_{\alpha\beta\gamma}\theta^\gamma + O(\theta^2). \tag{1.6.7}$$

With these tools it is now clear how to build G-invariant couplings among the θ^α, and between the θ^α's and other fields, such as χ from the scalar-field example.

It is simplest to build self-interactions for the Goldstone bosons. An invariant Lagrangian density may be built by combining the covariant quantity, $e^\alpha_\mu = e^\alpha{}_\beta \partial_\mu \theta^\beta$ in all possible H-invariant ways. This is simple to do since this quantity transforms very simply under G: $e_\mu \cdot X \to \gamma e_\mu \cdot X \gamma^{-1}$.

Derivatives of e^α_μ can also be included by differentiating using the covariant derivative constructed from $\mathcal{A}^i_\mu t_i$:

$$(D_\mu e_\nu)^\alpha = \partial_\mu e^\alpha_\nu + c^\alpha{}_{i\beta} \mathcal{A}^i_\mu \, e^\beta_\nu, \tag{1.6.8}$$

which transforms in the same way as does e^α_μ: $\delta(D_\mu e_\nu)^\alpha = -c^\alpha{}_{i\beta} u^i (D_\mu e_\nu)^\beta$.

The lagrangian is $\mathcal{L}(e_\mu, D_\mu e_\nu, \ldots)$, where the ellipses denote terms involving higher covariant derivatives. Provided only that this lagrangian is constrained to be globally H invariant:

$$\mathcal{L}(h e_\mu h^{-1}, h D_\mu e_\nu h^{-1}, \ldots) \equiv \mathcal{L}(e_\mu, D_\mu e_\nu, \ldots), \tag{1.6.9}$$

the result is guaranteed to be *automatically* globally G invariant, as required. For a Poincaré invariant system, the term involving the fewest derivatives therefore becomes:

$$\mathcal{L}_{GB} = -\frac{1}{2} f_{\alpha\beta} \, \eta^{\mu\nu} \, e^\alpha_\mu \, e^\beta_\nu + (\text{higher-derivative terms}). \tag{1.6.10}$$

In this expression, positivity of the kinetic energy implies that the matrix $f_{\alpha\beta}$ must be positive definite. G-invariance dictates that it must also satisfy $f_{\lambda\beta} c^\lambda{}_{i\alpha} + f_{\alpha\lambda} c^\lambda{}_{i\beta} = 0$. It was remarked earlier that on general grounds the matrices, X_α, fill out a representation, $\mathcal{R}$, of the unbroken subgroup H with representation matrices given by $(\mathcal{T}_i)_{\alpha\beta} = c_{\alpha i\beta}$, and

in terms of these matrices G-invariance requires the vanishing of all of the commutator, $[\mathcal{T}_i, f]$, for all of the generators, $\mathcal{T}_i$. If this representation, $\mathcal{R}$, of H is irreducible then, by Schur's lemma, $f_{\alpha\beta}$ must be proportional to the unit matrix, with positive coefficient: $f_{\alpha\beta} = F^2\delta_{\alpha\beta}$. Otherwise, if $\mathcal{R}$ is reducible into n irreducible diagonal blocks, then $f_{\alpha\beta}$ need only be block diagonal, with each diagonal element being proportional to a unit matrix:

$$f_{\alpha\beta} = \begin{pmatrix} F_1^2\delta_{\alpha_1\beta_1} & & \\ & \ddots & \\ & & F_n^2\delta_{\alpha_n\beta_n} \end{pmatrix}, \tag{1.6.11}$$

for n independent positive constants, F_n^2. We see that the lowest-dimension terms in the most general low-energy Goldstone-boson self-coupling lagrangian is parameterizable in terms of these n constants.

If other fields — denoted here collectively by χ — also appear in the low-energy theory then, since the symmetry H is not broken by the ground state, the fields χ must also transform linearly under H: $\chi \to h\chi$, where the matrices $\{h\}$ form a (possibly reducible) representation of H. In this case the starting point for inferring the coupling of the Goldstone bosons is an arbitrary, globally H-invariant lagrangian: $\mathcal{L}(\chi, \partial_\mu\chi, \ldots)$ with $\mathcal{L}(h\chi, h\partial_\mu\chi, \ldots) \equiv \mathcal{L}(\chi, \partial_\mu\chi, \ldots)$, for $\partial_\mu h = 0$. This lagrangian will be automatically promoted to become G-invariant by appropriately coupling the Goldstone bosons.

The promotion to G invariance proceeds by assigning to χ the nonlinear G-transformation rule: $\chi \to \gamma\chi$, where $\gamma = \gamma(\theta, g) \in H$ is the field-dependent H matrix which is defined by the nonlinear realization, eq. (1.5.8). An arbitrary globally H-invariant χ-lagrangian then becomes G invariant if all derivatives are replaced by the θ-dependent covariant derivative: $\partial_\mu\chi \to D_\mu\chi = \partial_\mu\chi - i\mathcal{A}^i t_i\chi$, for which $D_\mu\chi \to \gamma D_\mu\chi$.

The general lagrangian therefore becomes: $\mathcal{L}(e_\mu, \chi, D_\mu e_\nu, D_\mu\chi, \ldots)$, where G-invariance

is ensured provided only that $\mathcal{L}$ is constrained by global H invariance:

$$\mathcal{L}(he_\mu h^{-1}, h\,\chi, hD_\mu e_\nu h^{-1}, h\,D_\mu\chi, \ldots) \equiv \mathcal{L}(e_\mu, \chi, D_\mu e_\nu, D_\mu\chi, \ldots). \tag{1.6.12}$$

1.7 Uniqueness

The previous construction certainly defines a G-invariant lagrangian for the interactions of the Goldstone bosons which arise from the symmetry-breaking pattern $G \to H$, given the transformation rules which were derived in earlier sections. Our goal in the present section is to show that this construction gives the most general such invariant lagrangian. That is, we wish to show that the most general lagrangian density which is invariant under the transformation rules of eqs. (1.5.8) may be constructed using only the quantities $e_\mu^\alpha(\theta)$ and $\mathcal{A}_\mu^i(\theta)$ in addition to any other fields, χ.

We start with a general lagrangian density, $\mathcal{L}(\theta, \partial_\mu\theta, \chi, \partial_\mu\chi)$, involving the fields θ^α, χ^n and their derivatives. We do not include a dependence on second and higher derivatives of these fields, but this extension is straightforward to make along the lines that are described in this section. It is more convenient in what follows to trade the assumed dependence of $\mathcal{L}$ on $\partial_\mu\theta$ for a dependence on the combinations $e_\mu^\alpha = e^\alpha{}_\beta(\theta)\;\partial_\mu\theta^\beta$ and $\mathcal{A}_\mu^i = \mathcal{A}_\alpha^i(\theta)\;\partial_\mu\theta^\alpha$. There is no loss of generality in doing so, since any function of θ and $\partial_\mu\theta$ can always be written as a function of θ, e_μ^α and $\mathcal{A}_\mu^i$. This equivalence is most easily seen in terms of the matrix variable $U(\theta) = e^{i\theta\cdot X}$. Any function of θ and $\partial_\mu\theta$ can equally well be written as a function of U and $\partial_\mu U$, or equivalently as a function of U and $U^{-1}\partial_\mu U$. But an arbitrary function of $U^{-1}\partial_\mu U$ is equivalent to a general function of e_μ^α and $\mathcal{A}_\mu^i$, as may be seen from expression (1.6.2).

The condition that a general function, $\mathcal{L}(\theta^\alpha, e_\mu^\alpha, \mathcal{A}_\mu^i, \chi, \partial_\mu\chi)$, be invariant with re-

spect to G transformations is:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\theta^\alpha}\,\delta\theta^\alpha + \frac{\partial\mathcal{L}}{\partial e^\alpha_\mu}\,\delta e^\alpha_\mu + \frac{\partial\mathcal{L}}{\partial\mathcal{A}^i_\mu}\,\delta\mathcal{A}^i_\alpha + \frac{\partial\mathcal{L}}{\partial\chi^n}\,\delta\chi^n + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\chi^n)}\,\delta\partial_\mu\chi^n = 0. \tag{1.7.1}$$

We first specialize to the special case where the symmetry transformation lies in H: $g = e^{i\eta\cdot t} \in H$. In this case we must use, in eq. (1.7.1), the transformations:

$$\begin{aligned} &\delta\theta^\alpha = -c^\alpha{}_{i\beta}\eta^i\theta^\beta, \qquad \delta e^\alpha_\mu = -c^\alpha{}_{i\beta}\eta^i e^\beta_\mu, \qquad \delta\mathcal{A}^i_\mu = -c^i{}_{jk}\eta^i\mathcal{A}^k_\mu, \\ &\text{and} \qquad \delta\chi^n = i\eta^i(t_i\chi)^n, \qquad \delta\partial_\mu\chi^n = i\eta^i(t_i\partial_\mu\chi)^n. \end{aligned} \tag{1.7.2}$$

Requiring $\delta\mathcal{L} = 0$ for all possible transformation parameters, η^i, then implies the following identity:

$$\frac{\partial\mathcal{L}}{\partial\theta^\alpha}\,c^\alpha{}_{i\beta}\theta^\beta + \frac{\partial\mathcal{L}}{\partial e^\alpha_\mu}\,c^\alpha{}_{i\beta}e^\beta_\mu + \frac{\partial\mathcal{L}}{\partial\mathcal{A}^j_\mu}\,c^j{}_{ik}\mathcal{A}^k_\mu - \frac{\partial\mathcal{L}}{\partial\chi^n}\,i(t_i\chi)^n - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\chi^n)}\,i(t_i\partial_\mu\chi)^n = 0. \tag{1.7.3}$$

This identity simply states that $\mathcal{L}$ must be constructed to be an H-invariant function of its arguments, all of which transform linearly with respect to H transformations.

For the remaining symmetry transformations which do not lie in H, $g = e^{i\omega\cdot X} \in G/H$, we instead evaluate eq. (1.7.1) using the following transformations:

$$\begin{aligned} &\delta\theta^\alpha = \xi^\alpha{}_\beta\omega^\beta, \qquad \delta e^\alpha_\mu = -c^\alpha{}_{i\beta}u^i e^\beta_\mu, \qquad \delta\mathcal{A}^i_\mu = \partial_\mu u^i - c^i{}_{jk}u^j\mathcal{A}^k_\mu, \\ &\text{and} \qquad \delta\chi^n = iu^i(t_i\chi)^n \qquad \delta\partial_\mu\chi^n = iu^i(t_i\partial_\mu\chi)^n, \end{aligned} \tag{1.7.4}$$

where $\xi^\alpha = \xi^\alpha{}_\beta(\theta)\,\omega^\beta$ and $u^i = u^i_\alpha(\theta)\,\omega^\alpha$ are the nonlinear functions of θ that are defined by eq. (1.5.8). Using these in eq. (1.7.1), and simplifying the resulting expression using eq. (1.7.3), leads to the remaining condition for G invariance:

$$\frac{\partial\mathcal{L}}{\partial\theta^\alpha}\left(\xi^\alpha{}_\beta + c^\alpha{}_{i\gamma}u^i_\beta\theta^\gamma\right) + \frac{\partial\mathcal{L}}{\partial\mathcal{A}^j_\mu}\,\partial_\mu u^i_\beta + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\chi^n)}\,i\partial_\mu u^i_\beta(t_i\chi)^n = 0. \tag{1.7.5}$$

This last identity contains two separate pieces of information. The first piece can be extracted by specializing to $\theta^\alpha = 0$. In this case, since $\partial_\mu u^i_\beta = \partial_\alpha u^i_\beta \, \partial_\mu \theta^\alpha$ vanishes when $\theta^\alpha = 0$, and since eq. (1.5.14) implies $\xi^\alpha{}_\beta(\theta = 0) = \delta^\alpha_\beta$, we find:

$$\left.\frac{\partial \mathcal{L}}{\partial \theta^\alpha}\right|_{\theta=0} = 0. \tag{1.7.6}$$

But, since the group transformation law for θ^α is inhomogeneous, we may always perform a symmetry transformation to ensure that $\theta^\alpha = 0$ for *any* point $p \in G/H$. As a result, eq. (1.7.6) implies the more general statement:

$$\frac{\partial \mathcal{L}}{\partial \theta^\alpha} \equiv 0 \qquad \text{throughout} \qquad G/H. \tag{1.7.7}$$

The rest of the information contained in eq. (1.7.5) may now be extracted by using $\partial \mathcal{L}/\partial \theta^\alpha = 0$ to eliminate the first term. One finds the remaining condition:

$$\left(\frac{\partial \mathcal{L}}{\partial \mathcal{A}^j_\mu} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \chi^n)}\, i(t_j \chi)^n\right) \partial_\mu u^j_\beta = 0. \tag{1.7.8}$$

This equation has a very simple meaning. It states that $\mathcal{L}$ can depend on the two variables, $\mathcal{A}^j_\mu$ and $\partial_\mu \chi^n$, only through the one combination: $(D_\mu \chi)^n \equiv \partial_\mu \chi^n - i\mathcal{A}^j_\mu (t_j \chi)^n$. That is, χ can appear differentiated in $\mathcal{L}$ only through its covariant derivative, $D_\mu \chi$.

We see from these arguments that the G-invariance of $\mathcal{L}$ is equivalent to the statement that $\mathcal{L}$ must be an H-invariant function constructed from the covariantly-transforming variables e^α_μ, χ and $D_\mu \chi$. If higher derivatives of θ had been considered, then the vanishing of the terms in $\delta\mathcal{L}$ that are proportional to more than one derivative of u^i would similarly imply that derivatives of e^α_μ must also only appear through its covariant derivative, $(D_\mu e_\nu)^\alpha$, defined by eq. (1.6.8). This proves the uniqueness of the construction of invariant lagrangians using these covariant quantities.

1.8 The Geometric Picture

There is an appealing geometric description of the resulting effective lagrangian, which makes available many powerful techniques from differential geometry to effective lagrangian methods. We pause here to briefly outline this connection.

Consider, for simplicity, only the self-interactions of the Goldstone bosons: $\mathcal{L}_{GB}(\theta)$. This can be expanded into terms having increasing numbers of derivatives acting on θ. The first few terms that are consistent with Poincaré invariance are:

$$\mathcal{L}_{GB}(\theta) = -V(\theta) - \frac{1}{2} g_{\alpha\beta}(\theta)\, \partial_\mu \theta^\alpha \partial^\mu \theta^\beta + \cdots . \tag{1.8.1}$$

Positivity of the kinetic energy for fluctuations about any configuration, θ^α, requires the symmetric matrix $g_{\alpha\beta}$ to be positive definite for all θ.

The geometrical interpretation arises once we recall that the fields θ take values in the coset space G/H, and so each Goldstone boson field configuration can be considered to be a map from spacetime into G/H. The function, V, can then be considered to be a real-valued function which is defined on the space G/H. Similarly, the positive symmetric matrix, $g_{\alpha\beta}$, defines a metric tensor on G/H. These identifications of V and $g_{\alpha\beta}$ with geometrical objects on G/H are consistent with their transformation properties under field redefinitions, $\delta\theta^\alpha = \xi^\alpha(\theta)$, which are the analogues of coordinate transformations on G/H. To see this, perform this transformation in the lagrangian of eq. (1.8.1). The result is to replace V and $g_{\alpha\beta}$ by the quantities $V + \pounds_\xi V$ and $g_{\alpha\beta} + \pounds_\xi g_{\alpha\beta}$, where the linear operator, $\pounds_\xi$, is known as the Lie derivative in the direction specified by ξ^α, and is given explicitly for a scalar and a covariant tensor by:

$$\begin{aligned} \pounds_\xi V &= \xi^\alpha \partial_\alpha V, \\ \text{and} \qquad \pounds_\xi g_{\alpha\beta} &= \xi^\lambda \partial_\lambda g_{\alpha\beta} + g_{\lambda\beta}\partial_\alpha \xi^\lambda + g_{\alpha\lambda}\partial_\beta \xi^\lambda . \end{aligned} \tag{1.8.2}$$

In these expressions derivatives, like $\partial_\alpha V$, $\partial_\lambda g_{\alpha\beta}$ or $\partial_\alpha \xi^\lambda$, all represent differentiation with respect to θ^α, and not with respect to spacetime position, x^μ.

Clearly, the G invariance of the first few terms of $\mathcal{L}_{GB}$ therefore becomes equivalent to the problem of finding a scalar and a metric for which $\pounds_\xi V = \pounds_\xi g_{\alpha\beta} = 0$ for each of the $\xi^\alpha(\theta)$'s which describe the action of G on G/H. Since every point on G/H can be reached from any other by performing such a G transformation — *i.e.* G/H is a *homogeneous space* — it follows that the only possible invariant function, V, is a constant which is independent of θ. This expresses the low-energy decoupling of the Goldstone bosons since it shows that G invariance completely forbids their appearance in the scalar potential.

Similarly, the condition for the G invariance of the kinetic terms is that the metric $g_{\alpha\beta}$ must also be invariant under the action of all of the $\xi^\alpha(\theta)$'s which generate G on G/H. That is, all of these ξ's must generate *isometries* of the metric $g_{\alpha\beta}$. The problem of finding the most general invariant kinetic term is therefore equivalent to constructing the most general G-invariant metric on G/H. A comparison of the lagrangian of eq. (1.8.1), with our earlier result, eq. (1.6.10), gives a representation of the metric $g_{\alpha\beta}$ in terms of the quantities $e^\alpha{}_\beta(\theta)$. We have

$$\begin{aligned} g_{\alpha\beta} &= f_{\gamma\lambda}\, e^\gamma{}_\alpha e^\lambda{}_\beta, \\ &= \sum_{r=1}^{n} F_r^2\, \delta_{\gamma_r\lambda_r}\, e^{\gamma_r}{}_\alpha e^{\lambda_r}{}_\beta, \\ &\approx \sum_{r=1}^{n} F_r^2 \Big[\delta_{\alpha_r\beta_r} - c_{\alpha_r\beta_r\gamma}\theta^\gamma + O(\theta^2)\Big]. \end{aligned} \tag{1.8.3}$$

Eq. (1.8.3) also has a geometric interpretation. It shows that the object $e^\alpha{}_\beta$ can be interpreted as a G-covariant *vielbein* for the space G/H. In Reimannian geometry a vielbein is the name given to a set of N linearly independent vectors, $e^a{}_\beta$, $a = 1, \ldots, N$, which are tangent to an N-dimensional space. (The name is German for 'many legs', with

viel meaning 'many', and so such vectors are also called *zweibein* in two dimensions — with *zwei* = 'two' — or *vierbein* in four dimensions — with *vier* = 'four'.) Part of the utility of identifying such a set of vectors is that it is always possible to reconstruct from them the space's metric, using: $g_{\alpha\beta} = \delta_{ab}\, e^a{}_\alpha\, e^b{}_\beta$.

The geometrical interpretation of $e^\alpha{}_\beta$ as a vielbein, and the uniqueness of the construction of invariant lagrangians proven in the previous section, gives the general solution to the geometrical problem of constructing G-invariant metrics on the space G/H. We see that there is an n-parameter family of such metrics, where n counts the number of irreducible representations of H which are formed by the generators, X_α, of G/H. The n parameters are given explicitly by the constants $F_r^2, r = 1, \ldots, n$.

For many physical applications the representation of H that is furnished by the X_α is irreducible, and in this case the G-invariant metric is uniquely determined up to its overall normalization: $g_{\alpha\beta} = F^2 \hat{g}_{\alpha\beta}$, with $\hat{g}_{\alpha\beta} = \delta_{\alpha\beta} + O(\theta)$. For such systems there is precisely one constant in the effective lagrangian for Goldstone bosons which is undetermined by the symmetries of the problem, if we include only the fewest possible (two) derivatives:

$$\mathcal{L}_{GB} = -\frac{F^2}{2}\, \hat{g}_{\alpha\beta}(\theta)\, \partial_\mu \theta^\alpha \partial^\mu \theta^\beta + (\text{higher derivative terms}). \tag{1.8.4}$$

Once the one constant, F^2, is determined, either by calculation from an underlying theory, or by appeal to experiment, the lowest-order form for all of the Goldstone boson interactions are completely determined by the symmetry breaking pattern. The resulting model-independent predictive power has wide applications throughout physics, as we shall see when we consider examples in subsequent chapters.

1.9 Nonrelativistic Lagrangians

Before turning to examples, we pause to outline the changes in the above analysis which become necessary when it is applied to nonrelativistic systems, for which the spacetime symmetry is not Poincaré invariance. This is what is appropriate, for example, in condensed-matter applications for which there is a preferred frame, defined by the centre-of-mass of the medium of interest.

So long as our attention is restricted to internal symmetries, most of the considerations of this chapter apply just as well to nonrelativistic problems as they do to relativistic ones. In particular, the expressions obtained for the nonlinear realization of broken symmetries on the Goldstone boson fields does not depend at all on the spacetime symmetries which are assumed.

We assume for simplicity here that the system remains invariant with respect to translations and rotations, although the generalization to different spacetime groups is straightforward in principle. In practice, the results we obtain also apply to some lattice systems, for which translations and rotations are not symmetries. This is because, at least for the first few derivatives in a derivative expansion of the lagrangian, the lattice group for some lattices, such as a cubic lattice for example, implies the same restrictions as do translation and rotation invariance.

There are two cases, which we consider separately below, depending on whether or not time reversal is a good symmetry of the system.

1.9.1 When Time-Reversal is a Good Symmetry

For unbroken time reversal, the main difference from the relativistic situation lies in the fact that the time- and space-derivatives become independent since they are unrelated

by any symmetries. For instance, assuming unbroken translation and rotation invariance, the most general self-couplings amongst the Goldstone bosons for the symmetry-breaking pattern $G \to H$ are:

$$\begin{aligned}\mathcal{L}_{GB} &= -\frac{1}{2}\sum_{r=1}^{n}\Big[F^2_{(r),t}\,\hat{g}^{(r)}_{\alpha\beta}(\theta)\,\dot{\theta}^\alpha\dot{\theta}^\beta + F^2_{(r),s}\,\hat{g}^{(r)}_{\alpha\beta}(\theta)\,\nabla\theta^\alpha\cdot\nabla\theta^\beta\Big] \\ &\qquad +(\text{higher derivative terms}).\end{aligned} \tag{1.9.1}$$

Compared to the relativistic case, twice as many — that is, $2n$ — constants, $F^2_{(r),t}$ and $F^2_{(r),s}$, are required to parameterize the terms containing the fewest number of derivatives. Once these two constants are determined, all other interactions at this order of the derivative expansion are clearly completely determined.

1.9.2 When Time-Reversal is Broken

New possibilities arise for the lagrangian when time-reversal symmetry is broken, as is the case for a ferromagnet, for example. In this case, it is possible to write down terms in $\mathcal{L}_{GB}$ which involve an odd number of time derivatives. In particular, the dominant, lowest-dimension term involving time derivatives involves only one:

$$\Delta\mathcal{L}_{GB} = -A_\alpha(\theta)\,\dot{\theta}^\alpha. \tag{1.9.2}$$

The coefficient function, $A_\alpha(\theta)$, can be considered to define a vector field on the coset space G/H. It is to be chosen to ensure the G invariance of the low-energy theory.

If $\Delta\mathcal{L}_{GB}$ is required to be invariant under G transformations, then it must be built using the quantity $e^\alpha{}_\mu(\theta) = e^\alpha{}_\beta(\theta)\,\partial_\mu\theta^\beta$ and its covariant derivatives. If it is to involve only a single time derivative, then it must be proportional only to $e_0(\theta)$. But the only such G-invariant quantity is: $k_\alpha e^\alpha{}_\beta\dot{\theta}^\beta$, where the constants, k_α, must satisfy $k_\alpha c^\alpha{}_{i\beta} = 0$ for all indices i and β. Such a k_α exists only if the corresponding generator, X_α, commutes with

all of the generators of the unbroken subgroup, H. This quite restrictive condition is not ever satisfied in many situations of physical interest, and for these systems it appears — at least superficially — that no terms involving only a single time derivative are consistent with G-invariance.

This conclusion would be too strong, however, because it is too restrictive a condition to demand the G-invariance of the lagrangian density, $\mathcal{L}$. We are only required by G symmetry to demand the invariance of the action. The lagrangian density need not be invariant, provided that its variation is a total derivative. It is therefore worth re-examining the time-reversal-violating term of eq. (1.9.2) in this light.

Once we drop total derivatives in $\Delta\mathcal{L}_{GB}$, it is clear that the coefficient $A_\alpha(\theta)$ is only defined up to the addition of a gradient. That is, any two choices A_α and $\tilde{A}_\alpha \equiv A_\alpha + \partial_\alpha\Omega(\theta)$, differ in their contribution to $\Delta\mathcal{L}_{GB}$ only by the total derivative:

$$\Delta\mathcal{L}_{GB}(A + \partial\Omega) - \Delta\mathcal{L}_{GB}(A) = -\partial_\alpha\Omega(\theta)\,\dot{\theta} = -\,\dot{\Omega}(\theta). \tag{1.9.3}$$

In geometrical terms we may therefore regard the coefficient function, $A_\alpha(\theta)$, as defining a gauge potential on the coset space G/H.

The condition that $\Delta\mathcal{L}_{GB}$ contribute a G-invariant term to the action therefore only requires the coefficient $A_\alpha(\theta)$ to be G-invariant *up to a gauge transformation*. In equations, G-invariance of the action only requires:

$$\pounds_\xi A_\alpha \equiv \xi^\beta\partial_\beta A_\alpha + A_\beta\partial_\alpha\xi^\beta = \partial_\alpha\Omega_\xi, \tag{1.9.4}$$

for each generator $\delta\theta^\alpha = \xi^\alpha$ of the action of G on G/H, and for some scalar functions, $\Omega_\xi(\theta)$, on G/H. This last condition is equivalent to demanding the invariance of the gauge-invariant quantity: $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$. That is,

$$\pounds_\xi F_{\alpha\beta} \equiv \xi^\lambda\partial_\lambda F_{\alpha\beta} + F_{\lambda\beta}\partial_\alpha\xi^\lambda + F_{\alpha\lambda}\partial_\beta\xi^\lambda = 0. \tag{1.9.5}$$

We shall find that this condition *does* admit solutions in many cases of interest — most notably for the example of a ferromagnet.

1.10 Power Counting

Before proceeding to real-life applications, a final important issue must be addressed. Since the lagrangians expressed using Goldstone boson variables is typically nonrenormalizable, it is necessary to know how to use nonrenormalizable lagrangians when making quantitative calculations.

The key to doing so is to consider the Goldstone Boson lagrangians to which we have been led in previous sections to be 'effective theories' which describe only the low-energy behaviour of the system of interest. For instance, in our toy models the Goldstone bosons (θ^α) are massless while the other degrees of freedom (χ) are not. (Although the pseudo-Goldstone bosons for an approximate symmetry are not exactly massless, they may nonetheless appear in the low-energy theory so long as their mass, m, is sufficiently small.) A lagrangian involving only Goldstone bosons or pseudo-Goldstone bosons can only hope to describe physics at energies, q, below the mass threshhold, M, for producing the heavier particles. (It is often the case that M is proportional to the symmetry-breaking scale(s), v). The predictions of such a lagrangian are to be regarded as reproducing, in powers of q/M, whatever the 'underlying' (or 'microscopic') theory — *i.e.* the theory involving the heavy χ states — might be.

In order to make this concrete, consider one such a Lagrangian, having the form:

$$\mathcal{L}^{\text{eff}} = f^4 \sum_n \frac{c_n}{M^{d_n}} \mathcal{O}_n \left(\frac{\phi}{v}\right), \tag{1.10.1}$$

where ϕ denotes a generic boson field, c_n are a set of dimensionless coupling constants

which we imagine to be at most $O(1)$, and f, M and v are mass scales of the underlying problem. (For example, in the application to pions which follows we will have $f = \sqrt{F_\pi \Lambda_\chi}$, $M = \Lambda_\chi$ and $v = F_\pi$, where F_π and Λ_χ are scales which characterize the strength of the appropriate symmetry breaking in the strong interactions.) d_n is the dimension of the operator $\mathcal{O}_n$, in powers of mass, as computed by counting only the dimensions of the field, ϕ, and derivatives, $\partial\phi$.

Imagine using this lagrangian to compute a scattering amplitude, $\mathcal{A}_E(q)$, involving the scattering of E particles whose four-momenta are collectively denoted by q. We wish to focus on the contribution to $\mathcal{A}$ due to a Feynman graph having I internal lines and V_{ik} vertices. The labels i and k here indicate two characteristics of the vertices: i counts the number of lines which converge at the vertex, and k counts the power of momentum which appears in the vertex. Equivalently, i counts the number of powers of the fields, ϕ, which appear in the corresponding interaction term in the lagrangian, and k counts the number of derivatives of these fields which appear there.

1.10.1 Some Useful Identities

The positive integers, I, E and V_{ik}, which characterize the Feynman graph in question are not all independent since they are related by the rules for constructing graphs from lines and vertices. This relation can be obtained by equating two equivalent methods of counting the number of ways that internal and external lines can end in a graph. On the one hand, since all lines end at a vertex, the number of ends is given by summing over all of the ends which appear in all of the vertices: $\sum_{ik} i\, V_{ik}$. On the other hand, there are two ends for each internal line, and one end for each external line in the graph: $2I + E$.

Equating these gives the identity which expresses the 'conservation of ends':

$$2I + E = \sum_{ik} i\, V_{ik}, \qquad \text{(Conservation of Ends)}. \tag{1.10.2}$$

A second useful identity *defines* of the number of loops, L, for each (connected) graph:

$$L = 1 + I - \sum_{ik} V_{ik}, \qquad \text{(Definition of } L\text{)}. \tag{1.10.3}$$

For simple planar graphs, this definition agrees with the intuitive notion what the number of loops in a graph means.

1.10.2 Dimensional Estimates

We now collect the dependence of $\mathcal{A}_{E}(a)$ on the parameters in $\mathcal{L}^{\text{eff}}$.

Reading the Feynman rules from the lagrangian of eq. (1.10.1) shows that the vertices in the Feynman graph contribute the following factor:

$$\text{(Vertex)} = \prod_{ik} \left[i(2\pi)^4 \delta^4(p) \, \left(\frac{p}{M}\right)^k \, \left(\frac{f^4}{v^i}\right) \right]^{V_{ik}}, \tag{1.10.4}$$

where p generically denotes the various momenta running through the vertex.

Similarly, each internal line in the graph contributes the additional factors:

$$\text{(Internal Line)} = \left[-i \int \frac{d^4p}{(2\pi)^4} \left(\frac{M^2 v^2}{f^4}\right) \frac{1}{p^2 + m^2} \right]^I, \tag{1.10.5}$$

where, again, p denotes the generic momentum flowing through the line. m denotes the mass of the light particles which appear in the effective theory, and it is assumed that the kinetic terms which define their propagation are those terms in $\mathcal{L}^{\text{eff}}$ involving two derivatives and two powers of the fields, ϕ.

As usual for a connected graph, all but one of the momentum-conserving delta functions in eq. (1.10.4) can be used to perform one of the momentum integrals in eq. (1.10.5).

The one remaining delta function which is left after doing so depends only on the external momenta, $\delta^4(q)$, and expresses the overall conservation of four-momentum for the process. Future formulae are less cluttered if this factor is extracted once and for all, by defining the reduced amplitude, $\tilde{\mathcal{A}}$, by

$$\mathcal{A}_{E}(q) = i(2\pi)^4 \delta(q)\, \tilde{\mathcal{A}}_{E}(q). \tag{1.10.6}$$

The number of four-momentum integrations which are left after having used all of the momentum-conserving delta functions is then $I - \sum_{ik} V_{ik} + 1 = L$. This last equality uses the definition, eq. (1.10.3), of the number of loops, L.

We now wish to estimate the result of performing the integration over the internal momenta. In general these are complicated integrals for which a simple result is not always possible to give. Considerable simplifications arise, however, if all of the masses and energies of the particles in the low-energy theory are of the same order of magnitude, since in this case much can be said about the order of magnitude of the momentum integrals purely on dimensional grounds. (Although this is often the situation of interest when employing effective theories, it must be borne in mind that it does not always apply. For instance it excludes a situation of considerable practical interest, where the low-energy theory includes very massive but slowly-moving, nonrelativistic particles. Power counting for such systems is beyond the scope of this review.)

In order to proceed with a dimensional argument it is most convenient to regulate the ultraviolet divergences which arise in the momentum integrals using dimensional regularization. For dimensionally-regularized integrals, the key observation is that the size of the result is set on dimensional grounds by the light masses or external momenta of the theory. That is, if all external energies, q, are comparable to (or larger than) the

masses, m, of the light particles whose scattering is being calculated, then q is the light scale controlling the size of the momentum integrations, so dimensional analysis implies that an estimate of the size of the momentum integrations is:

$$\int \cdots \int \left(\frac{d^n p}{(2\pi)^n}\right)^A \frac{p^B}{(p^2+q^2)^C} \sim \left(\frac{1}{4\pi}\right)^{2A} q^{nA+B-2C}, \tag{1.10.7}$$

with a dimensionless prefactor which carries all of the complicated dependence on dimensionless ratios like q/m. The prefactor also depends on the dimension, n, of spacetime, and may be singular in the limit that $n \to 4$.[5]

With this estimate for the size of the momentum integrations, we find the following contribution to the amplitude $\tilde{\mathcal{A}}_E(q)$:

$$\int \cdots \int \left(\frac{d^4 p}{(2\pi)^4}\right)^L \frac{p^{\sum_{ik} kV_{ik}}}{(p^2+q^2)^I} \sim \left(\frac{1}{4\pi}\right)^{2L} q^{4L-2I+\sum_{ik} kV_{ik}}, \tag{1.10.8}$$

which, with liberal use of the identities (1.10.2) and (1.10.3), gives as estimate for $\tilde{\mathcal{A}}_E(q)$:

$$\tilde{\mathcal{A}}_E(q) \sim f^4 \left(\frac{1}{v}\right)^E \left(\frac{M^2}{4\pi f^2}\right)^{2L} \left(\frac{q}{M}\right)^{2+2L+\sum_{ik}(k-2)V_{ik}}. \tag{1.10.9}$$

This expression is the principal result of this section. Its utility lies in the fact that it links the contributions of the various effective interactions in the effective lagrangian, (1.10.1), with the dependence of observables on small mass ratios such as q/M. As a result it permits the determination of which interactions in the effective lagrangian are required to reproduce any given order in q/M in physical observables.

Most importantly, eq. (1.10.9) shows how to calculate using nonrenormalizable theories. It implies that even though the lagrangian can contain arbitrarily many terms, and so potentially arbitrarily many coupling constants, it is nonetheless predictive *so long as its predictions are only made for low-energy processes, for which* $q/M \ll 1$. (Notice also

[5] We ignore here any logarithmic infrared mass singularities which may arise in this limit.

that the factor $(M/f)^{4L}$ in eq. (1.10.9) implies, all other things being equal, the scale f cannot be taken to be systematically smaller than M without ruining the validity of the loop expansion in the effective low-energy theory.)

Before stating more explicitly the effective-lagrangian logic, which eq. (1.10.9) summarizes, we pause to generalize it to include fermions in the low-energy effective theory.

1.10.3 Including Fermions

It is straightforward to extend these results to include light fermions in the effective theory, although once again subject to the important assumption that all masses and energies are small in the effective theory. To this end, first generalize the starting form assumed for the lagrangian to include fermion fields, ψ, in addition to boson fields, ϕ:

$$\mathcal{L}^{\text{eff}} = f^4 \sum_n \frac{c_n}{M^{d_n}} \mathcal{O}_n \left(\frac{\phi}{v_B}, \frac{\psi}{v_F^{3/2}} \right). \tag{1.10.10}$$

An important difference between fermion and boson propagators lies in the way each falls off for large momenta. Whereas a boson propagator varies like $1/p^2$ for large p, a fermion propagator goes only like $1/p$. This leads to a difference in their contribution to the power counting of a Feynman graph. It is therefore important to keep separate track of the number of fermion and boson lines, and we therefore choose to now label vertices using *three* indices: k, i_B and i_F. As before, k labels the numbers of derivatives in the corresponding interaction, but now i_B and i_F separately count the number of bose and fermi lines which terminate at the vertex of interest. The number of vertices in a graph which carry a given value for k, i_B and i_F we now label by $V_{i_B i_F k}$.

Consider now computing an amplitude which has E_B external bosonic lines, E_F external fermion lines, and I_B and I_F internal bose and fermi lines. Repeating, with the lagrangian of eq. (1.10.10), the power counting argument which led (using dimensional

regularization) to eq. (1.10.9) now gives instead the following result:

$$\tilde{\mathcal{A}}_{E_B,E_F}(q) \sim f^4 \left(\frac{1}{v_B}\right)^{E_B} \left(\frac{1}{v_F}\right)^{3E_F/2} \left(\frac{M^2}{4\pi f^2}\right)^{2L} \left(\frac{q}{M}\right)^P, \tag{1.10.11}$$

where the power P can be written in either of the following two equivalent ways:

$$\begin{aligned} P &= 4 - E_B - \frac{3}{2}E_F + \sum_{i_B,i_F,k} \left(k + i_B + \frac{3}{2}i_F - 4\right) V_{i_B i_F k}, \\ &= 2 + 2L - \frac{1}{2}E_F + \sum_{i_B,i_F,k} \left(k + \frac{1}{2}i_F - 2\right) V_{i_B i_F k}. \end{aligned} \tag{1.10.12}$$

1.11 The Effective-Lagrangian Logic

The powercounting estimates just performed show how to organize calculations using nominally nonrenormalizable theories, considering them as effective field theories. They suggest the following general logic concerning their use.

Step I Choose the accuracy (*e.g.* one part per mille) with which observables, such as $\mathcal{A}_E(q)$, are to be computed.

Step II Determine the order in the small mass ratios q/M or m/M that must be required in order to acheive the desired accuracy.

Step III Use the power counting result, eq. (1.10.9), to find which terms in the effective lagrangian are needed in order to compute to the desired order in q/M. Eq. (1.10.9) also determines which order in the loop expansion is required for each effective interaction of interest.

Step IVa Compute the couplings of the required effective interactions using the full underlying theory. If this step should prove to be impossible, due either to ignorance

of the underlying theory or to the intractability of the required calculation, then it may be replaced by the following alternative:

Step IVb If the coefficients of the required terms in the effective lagrangian cannot be computed then they may instead be regarded as unknown parameters which are to be taken from experiment. Once a sufficient number of observables are used to determine these parameters, all other observables may be unambiguously predicted using the effective theory.

A number of points bear emphasizing at this point.

1. The possibility of treating the effective lagrangian phenomenologically, as in Step IVb above, immeasurably broadens the utility of effective lagrangian techniques, since they need not be restricted to situations for which the underlying theory is both known and calculationally simple. Implicit in such a program is the underlying assumption that there is no loss of generality in working with a local field theory. This assumption has been borne out in all known examples of physical systems. It is based on the conviction that the restrictions which are implicit in working with local field theories are simply those that follow from general physical principles, such as unitarity and cluster decomposition.

2. Since eq. (1.10.9) — or eqs. (1.10.11) and (1.10.12) — states that only a finite number of terms in $\mathcal{L}^{\text{eff}}$ contribute to any fixed order in q/M, and these terms need appear in only a finite number of loops, it follows that only a finite amount of labour is required to obtain a fixed accuracy in observables.

   Renormalizable theories represent the special case for which it suffices to work only to zeroeth order in the ratio q/M. This can be expected to eventually dominate

at sufficiently low energies compared to M, which is the reason why renormalizable theories play such an important role throughout physics.

3. An interesting corollary of the above observations is the fact that only a finite number of renormalizations are required in the low-energy theory in order to make finite the predictions for observables to any fixed order in q/M. Thus, although an effective lagrangian is not renormalizable in the traditional sense, it nevertheless *is* predictive in the same way that a renormalizable theory is.

Chapter 2

Pions: A Relativistic Application

We now present a relativistic application of these techniques to the low-energy interactions of pions and nucleons. This example provides a very useful, and experimentally successful, description of the low-energy limit of the strong interactions, and so illustrates how effective lagrangians can remain predictive even if it is impossible to predict their effective couplings from an underlying theory. This example is also of historical interest, since the study of low-energy pion scattering comprises the context within which the above Goldstone-boson formalism initially arose.

2.1 The Chiral Symmetries of QCD

The modern understanding of the strong interactions is based on the theory of mutually interacting spin-half quarks and spin-one gluons that is called Quantum Chromodynamics (QCD). This theory is described by the lagrangian density

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^a_{\mu\nu} G_a^{\mu\nu} - \sum_n \overline{q}_n (\not{D} + m_{q_n}) q_n, \tag{2.1.1}$$

where $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g f^a{}_{bc} G^b_\mu g^c_\nu$ is the field strength tensor for the gluon fields, G^a_μ. Here $a = 1, \ldots, 8$ labels the generators of the gauge symmetry group of the theory,

$SU_c(3)$, for which the $f^a{}_{bc}$ are the structure constants. The subscript 'c' of $SU_c(3)$ stands for 'colour', which is the name given to the strong charge.

The quarks are represented by Dirac spinors, q_n, where $n = 1, \cdots, 6$ counts the six kinds of quarks. In order of increasing mass, these are: u, d, s, c, b and t. For the purposes of later comparison we list here the quark masses, m_{q_n}, in GeV: $m_u = 0.0015 - 0.005$, $m_d = 0.003 - 0.009$, $m_s = 0.06 - 0.17$, $m_c = 1.1 - 1.4$, $m_b = 4.1 - 4.4$ and $m_t = 173.8 \pm 5.2$. All of these quarks are assumed to transform in the three-dimensional representation of the colour symmetry group, $SU_c(3)$, and so their covariant derivative (which appears in the combination $\not{D} = \gamma^\mu D_\mu$ in the lagrangian) is: $D_\mu q_n = \partial_\mu q_n - \frac{i}{2} g\, G^a_\mu \lambda_a q_n$. The eight matrices, λ_a, denote the three-by-three Gell-Mann matrices, which act on the (unwritten) colour index of each of the quarks. The explicit form for these matrices is not required in what follows. In all of these expressions g represents the coupling constant whose value controls the strength of the quark-gluon and gluon-gluon couplings.

The strong interactions as given by the above lagrangian density are believed to bind the quarks and gluons into bound states, which correspond to the observed strongly interacting particles (or, *hadrons*), such as protons (p), neutrons (n), pions (π), kaons (K), *etc.*. Table 2.1 lists the masses and some of the quantum numbers for all of the hadrons whose masses are less than 1 GeV. Considerably more states have masses above 1 GeV.

For the present purposes the most significant feature about this particle spectrum is that the lightest two quarks, u and d, have masses which are much smaller than all of the masses of the states which make up the particle spectrum. This suggests that the QCD dynamics may be well approximated by taking $m_u, m_d \approx 0$, and working perturbatively in these masses divided by a scale, Λ, which is typical of the strong interactions. From

Particle	Quark Content	Mass (GeV)	Spin	Isospin
$\pi^-(\pi^+)[\pi^0]$	$d\overline{u}(u\overline{d})[u\overline{u}, d\overline{d}]$	0.140 [0.135]	0	1
π^0	$u\overline{u}, d\overline{d}$	0.135	0	1
$K^+(K^0)$	$u\overline{s}(d\overline{s})$	0.494 (0.498)	0	$\frac{1}{2}$
$K^-(\overline{K}^0)$	$s\overline{u}(s\overline{d})$	0.494 (0.498)	0	$\frac{1}{2}$
η	$u\overline{u}, d\overline{d}, s\overline{s}$	0.547	0	0
$\rho^-(\rho^+)[\rho^0]$	$d\overline{u}(u\overline{d})[u\overline{u}, d\overline{d}]$	0.770	1	1
ω	$u\overline{u}, d\overline{d}, s\overline{s}$	0.782	1	0
$K^{*+}(K^{*0})$	$u\overline{s}(d\overline{s})$	0.892 (0.896)	1	$\frac{1}{2}$
$K^{*-}(\overline{K}^{*0})$	$s\overline{u}(s\overline{d})$	0.892 (0.896)	1	$\frac{1}{2}$
η'	$u\overline{u}, d\overline{d}, s\overline{s}$	0.958	0	0
f_0	$u\overline{u}, d\overline{d}, s\overline{s}$	0.980	0	0
a_0	$u\overline{u}, d\overline{d}, s\overline{s}$	0.980	0	1
$p(n)$	$uud(ddu)$	0.938 (0.940)	$\frac{1}{2}$	$\frac{1}{2}$

Table 2.1: Masses and Quantum Numbers of the Lightest Hadrons

the observed bound-state spectrum we expect Λ to be roughly 1 GeV.

The approximation for which m_u and m_d vanish turns out to be a very useful one. This is because the QCD lagrangian acquires the very useful symmetries

$$\begin{pmatrix} u \\ d \end{pmatrix} \rightarrow \left(U_{_L}\, \gamma_{_L} + U_{_R}\, \gamma_{_R} \right) \begin{pmatrix} u \\ d \end{pmatrix}, \tag{2.1.2}$$

in this limit, where $U_{_L}$ and $U_{_R}$ are arbitrary two-by-two unitary matrices having unit determinant. The Dirac matrix $\gamma_{_L} = \frac{1}{2}(1+\gamma_5)$ projects onto the left-handed part of each of the quarks, u and d, while $\gamma_{_R} = \frac{1}{2}(1-\gamma_5)$ projects onto their right-handed part. The group of symmetries which is obtained in this way is $G = SU_{_L}(2) \times SU_{_R}(2)$, with the subscripts '$L$' and '$R$' indicating the handedness of the quarks on which the corresponding factor of the group acts. A symmetry such as this which treats left- and right-handed fermions differently is called a *chiral* symmetry. These transformations are exact symmetries of QCD in the limit of vanishing m_u and m_d, but are only approximate symmetries when

these masses take their real values. Because the approximate symmetry involved is chiral, the technique of expanding quantities in powers of the light-quark masses is called *Chiral Perturbation Theory*.

If this symmetry, G, were not spontaneously broken by the QCD ground state, $|\Omega\rangle$, then we would expect all of the observed hadrons to fall into representations of G consisting of particles having approximately equal masses. This is not seen in the spectrum of observed hadrons, although the known particles *do* organize themselves into roughly degenerate representations of the approximate symmetry of *isospin*: $SU_I(2)$. The isospin quantum number, I, for the observed $SU_I(2)$ representations of the lightest hadrons are listed in Table 2.1. (The dimension of the corresponding representation is $2I+1$.) Isospin symmetry can be understood at the quark level to consist of the diagonal subgroup of G, for which $U_L = U_R$. That is, the approximate symmetry group which is seen to act on the particle states is that for which the left- and right-handed components of the quarks u and d rotate equally.

This suggests that if QCD is to describe the experimentally observed hadron spectrum, then its ground state must spontaneously break the approximate symmetry group G down to the subgroup $H = SU_I(2)$, for which:

$$\begin{pmatrix} u \\ d \end{pmatrix} \to U \begin{pmatrix} u \\ d \end{pmatrix}. \tag{2.1.3}$$

There is indeed good theoretical evidence, such as from numerical calculations, that the ground state of QCD really does behave in this way.

Given this symmetry-breaking pattern, we know that the low-energy spectrum of the theory must include the corresponding Goldstone bosons. If G had been an exact symmetry, then the corresponding Goldstone bosons would be exactly massless. But

since G is only a real symmetry in the limit that m_u and m_d both vanish, it follows that the Goldstone bosons for spontaneous G breaking need only vanish with these quark masses. So long as the u and d quarks are much lighter than the natural scale — $\Lambda \approx 1$ GeV — of the strong interactions, so must be these Goldstone bosons. Indeed, the lightest hadrons in the spectrum, $\pi^{\pm}$ and π^0, have precisely the quantum numbers which are required for them to be the Goldstone bosons for the symmetry-breaking pattern $SU_{L}(2) \times SU_{R}(2) \to SU_{I}(2)$.[1] Particles such as these which are light, but not massless, because they are the Goldstone bosons only of an approximate symmetry of a problem are called *pseudo-Goldstone bosons*.

Since the low-energy interactions of Golstone bosons are strongly restricted by the symmetry-breaking pattern which guarantees their existence, it is possible to experimentally test this picture of pions as pseudo-Goldstone bosons. The remainder of this chapter is devoted to extracting some of the simplest predictions for pion interactions which can be obtained in this way. The fact that these predictions successfully describe the low-energy interactions of real pions gives support to the assumed symmetry-breaking pattern for the ground state of the strong interactions.

2.2 The Low-Energy Variables

In order to proceed, we must first construct the nonlinear realization for the case $G = SU_{L}(2) \times SU_{R}(2)$ and $H = SU_{I}(2)$. To do so, we first write out the representation we shall use for the elements of each of these groups. Denoting the Pauli matrices, $\vec{\tau} = \{\tau_n\}$,

[1] In fact, the next-lightest particles, K and η, together with the pions have the quantum numbers to be the Goldstone bosons for the pattern $SU_{L}(3) \times SU_{R}(3) \to SU_{V}(3)$, which would be appropriate in the limit that the lightest three quarks, u, d and s, were *all* massless. The unbroken subgroup here, $SU_{V}(3)$, again is the diagonal, handedness-independent, subgroup.

by $\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and $\tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, we write:

$$g = \begin{pmatrix} e^{\frac{i}{2}\vec{\omega}_L \cdot \vec{\tau}} & 0 \\ 0 & e^{\frac{i}{2}\vec{\omega}_R \cdot \vec{\tau}} \end{pmatrix} \in SU_L(2) \times SU_R(2), \tag{2.2.1}$$

and

$$h = \begin{pmatrix} e^{\frac{i}{2}\vec{\omega}_I \cdot \vec{\tau}} & 0 \\ 0 & e^{\frac{i}{2}\vec{\omega}_I \cdot \vec{\tau}} \end{pmatrix} \in SU_I(2). \tag{2.2.2}$$

We adopt here, and throughout the remainder of the chapter, an obvious vector notation for the three-component quantities ω^n, θ^n, u^n, *etc.*.

In this representation, the Goldtone boson field becomes:

$$U(\vec{\theta}) = \begin{pmatrix} e^{\frac{i}{2}\vec{\theta} \cdot \vec{\tau}} & 0 \\ 0 & e^{-\frac{i}{2}\vec{\theta} \cdot \vec{\tau}} \end{pmatrix}, \tag{2.2.3}$$

and the nonlinear H transformation, γ, is:

$$\gamma(\vec{\theta}, g) = \begin{pmatrix} e^{\frac{i}{2}\vec{u} \cdot \vec{\tau}} & 0 \\ 0 & e^{\frac{i}{2}\vec{u} \cdot \vec{\tau}} \end{pmatrix}. \tag{2.2.4}$$

2.2.1 A Notational Aside

Before passing to the nonlinear realization, we briefly pause to make contact between the variables as defined here, and those that are often used in the literature. We have defined the elements, $g \in G$, the matrices $U(\vec{\theta})$, and $\gamma(\vec{\theta}, g)$ in a block-diagonal form which emphasizes the left- and right-handed parts of the transformations:

$$g = \begin{pmatrix} g_L & 0 \\ 0 & g_R \end{pmatrix}, \quad U(\vec{\theta}) = \begin{pmatrix} u_L(\vec{\theta}) & 0 \\ 0 & u_R(\vec{\theta}) \end{pmatrix}, \quad \gamma(\vec{\theta}, g) = \begin{pmatrix} h(\vec{\theta}, g) & 0 \\ 0 & h(\vec{\theta}, g) \end{pmatrix}. \tag{2.2.5}$$

In terms of these quantities, the transformation law $U \to \tilde{U} = gU\gamma^\dagger$ becomes $u_L \to \tilde{u}_L = g_L u_L h^\dagger$ and $u_R \to \tilde{u}_R = g_R u_R h^\dagger$. It is common practice to work with the composite quantity, Ξ, for which the transformation rule does not depend on the implicitly-defined matrix h. That is, if $\Xi \equiv u_L u_R^\dagger$, then $\Xi \to \tilde{\Xi} = g_L \Xi g_R^\dagger$. This transformation law has the

advantage of involving only explicit, constant matrices. In terms of the Goldstone boson fields, $\vec{\theta}$, we have $u_{_L} = e^{\frac{i}{2}\vec{\theta}\cdot\vec{\tau}} = u_{_R}^\dagger$, so $\Xi = u_{_L} u_{_R}^\dagger = e^{i\vec{\theta}\cdot\vec{\tau}}$.

It is possible, and often convenient, to reformulate all of the Goldstone boson self-couplings that are obtained elsewhere in this chapter in terms of this field Ξ. It is *not* possible to express the Goldstone-boson couplings to other fields, χ, in this way since the matrix γ cannot be removed from the transformation law for these other fields.

2.2.2 The Nonlinear Realization

The nonlinear realization is now obtained by constructing both $\vec{\tilde{\theta}}(\vec{\theta}, g)$ and $\vec{u} = \vec{u}(\vec{\theta}, g)$, using the condition $g\, U(\vec{\theta}) = U(\vec{\tilde{\theta}})\, \gamma$. For the groups under consideration this construction may be performed in closed form by using the identity:

$$\exp\Big[i\vec{\alpha}\cdot\vec{\tau}\Big] = \cos\alpha + i\hat{\alpha}\cdot\vec{\tau}\ \sin\alpha, \tag{2.2.6}$$

where $\alpha = \sqrt{\vec{\alpha}\cdot\vec{\alpha}}$, and $\hat{\alpha} = \vec{\alpha}/\alpha$.

Using this identity to multiply out both sides of the defining equation $g\ U(\vec{\theta}) = U(\vec{\tilde{\theta}})\ \gamma$, and equating the coefficients of 1 and $\vec{\tau}$, separately for the left- and right-handed parts of the matrices, gives explicit expressions for $\delta\vec{\theta} = \vec{\xi}$ and $\vec{u}$. If $g_{_{L,R}} = \exp\left[\frac{i}{2}\vec{\omega}_{_{L,R}}\cdot\vec{\tau}\right]$, and defining $\vec{\omega}_{_I} \equiv \frac{1}{2}(\vec{\omega}_{_L} + \vec{\omega}_{_R})$ and $\vec{\omega}_{_A} \equiv \frac{1}{2}(\vec{\omega}_{_L} - \vec{\omega}_{_R})$, then:

$$\begin{aligned}
\vec{\xi} &= \vec{\theta}\times\vec{\omega}_{_V} + \frac{\theta}{2}\left(\tan\frac{\theta}{2} + \cot\frac{\theta}{2}\right)[\vec{\omega}_{_A} - \hat{\theta}(\hat{\theta}\cdot\vec{\omega}_{_A})] + \hat{\theta}(\hat{\theta}\cdot\vec{\omega}_{_A}), \\
&= \vec{\omega}_{_A} + \vec{\theta}\times\omega_{_V} + O(\theta^2); \qquad (2.2.7) \\
\vec{u} &= \vec{\omega}_{_V} + (\hat{\theta}\times\vec{\omega}_{_A})\ \tan\frac{\theta}{2} \\
&= \vec{\omega}_{_V} + \frac{\vec{\theta}\times\vec{\omega}_{_A}}{2} + O(\theta^2). \qquad (2.2.8)
\end{aligned}$$

For future reference we notice that the transformation law for $\vec{\theta}$ implies that the

three broken generators of $G = SU_L(2) \times SU_R(2)$ form an irreducible, three-dimensional representation of the unbroken subgroup, $H = SU_I(2)$.

Similarly evaluating the combination

$$U^\dagger \partial_\mu U = \frac{i}{2}\,\vec{e}_\mu \cdot \begin{pmatrix} \vec{\tau} & \\ & -\vec{\tau} \end{pmatrix} + \frac{i}{2}\,\vec{\mathcal{A}}_\mu \cdot \begin{pmatrix} \vec{\tau} & \\ & \vec{\tau} \end{pmatrix}, \tag{2.2.9}$$

gives the quantities with which the invariant lagrangian is built:

$$\begin{aligned} \vec{e}_\mu &= \left(\frac{\sin\theta}{\theta}\right)\,\partial_\mu\vec{\theta} - \left(\frac{\sin\theta - \theta}{\theta^3}\right)\,(\vec{\theta}\cdot\partial_\mu\vec{\theta})\,\vec{\theta}, \\ &= \partial_\mu\vec{\theta}\left(1 - \frac{1}{6}\,\theta^2\right) + \frac{1}{6}\,(\vec{\theta}\cdot\partial_\mu\vec{\theta})\,\vec{\theta} + O(\theta^5); \qquad (2.2.10) \\ \vec{\mathcal{A}}_\mu &= -2\,\left(\frac{\sin^2\frac{\theta}{2}}{\theta^2}\right)\,(\vec{\theta}\times\partial_\mu\vec{\theta}) \\ &= -\frac{1}{2}\,\vec{\theta}\times\partial_\mu\vec{\theta} + O(\theta^4). \qquad (2.2.11) \end{aligned}$$

Notice that $\vec{e}_\mu$ is odd, and $\vec{\mathcal{A}}_\mu$ is even, under the interchange $\vec{\theta} \to -\vec{\theta}$. The low-energy Goldstone boson lagrangian will be required to be invariant under such an inversion of $\vec{\theta}$, since this is a consequence of the parity invariance of the underlying QCD theory.

It is useful to also record here the G-transformation rules for the other fields which can appear in the low-energy theory. Of particular interest are the nucleons — neutrons and protons — since low-energy pion-nucleon interactions are amenable to experimental study. The nucleon transformation rules under $G = SU_L(2) \times SU_R(2)$ are completely dictated by their transformations under the unbroken isospin subgroup, $H = SU_I(2)$. Since the nucleons form an isodoublet, $N = \begin{pmatrix} p \\ n \end{pmatrix}$, they transform under isospin according to $\delta N = \frac{i}{2}\,\vec{\omega}_I \cdot \vec{\tau}\,N$. The rule for the complete G tranformations is therefore simply

$$\delta N = \frac{i}{2}\,\vec{u}\cdot\vec{\tau}\,N. \tag{2.2.12}$$

We therefore see that the appropriate covariant derivative for nucleons is:

$$D_\mu N = \partial_\mu N - \frac{i}{2}\,\vec{\mathcal{A}}_\mu(\theta)\cdot\vec{\tau}\,N. \tag{2.2.13}$$

2.3 Invariant Lagrangians

We may now turn to the construction of the invariant lagrangian which governs the low-energy form for pion interactions. The lagrangian describing pion self-interactions involving the fewest derivatives is uniquely determined up to an overall normalizing constant. As was discussed in detail in the previous chapter, this is because of the irreducibility of the transformation rules of the broken generators, $\vec{X} = \frac{1}{2}\,\vec{\tau}\,\gamma_5$, under the unbroken isospin transformations. The most general G-invariant lagrangian density involving only two derivatives is

$$\mathcal{L}_{\pi\pi} = -\frac{F^2}{2}\,\hat{g}_{mn}(\vec{\theta})\,\partial_\mu\theta^m\partial^\mu\theta^n + (\text{higher-derivative terms}), \tag{2.3.1}$$

where the metric, $\hat{g}_{mn}$, on G/H is:

$$\begin{aligned} g_{mn}(\theta) &= \delta_{rs}\,e^r{}_m\,e^s{}_n = \delta_{mn}\,\left(\frac{\sin^2\theta}{\theta^2}\right) + \theta_m\theta_n\,\left(\frac{\theta^2-\sin^2\theta}{\theta^4}\right), \\ &= \delta_{mn}\,\left(1-\frac{1}{3}\,\theta^2\right) + \frac{1}{3}\,\theta_m\theta_n + O(\theta^4). \end{aligned} \tag{2.3.2}$$

For applications to pion scattering it is useful to canonically normalize the pion fields, that is, to ensure that their kinetic terms take the form: $-\frac{1}{2}\,\partial_\mu\vec{\pi}\cdot\partial^\mu\vec{\pi}$. This requires the rescaling: $\vec{\theta} = \vec{\pi}/F$. With this choice we have:

$$\begin{aligned} \mathcal{L}_{\pi\pi} &= -\frac{1}{2}\,\left[\frac{F^2\sin^2\left(\frac{\pi}{F}\right)}{\pi^2}\right]\partial_\mu\vec{\pi}\cdot\partial^\mu\vec{\pi} + \left[\frac{\pi^2 - F^2\sin^2\left(\frac{\pi}{F}\right)}{\pi^4}\right](\vec{\pi}\cdot\partial_\mu\vec{\pi})(\vec{\pi}\cdot\partial^\mu\vec{\pi}) \\ &\qquad +(\text{higher-derivative terms}). \end{aligned} \tag{2.3.3}$$

$$= -\frac{1}{2}\,\partial_\mu\vec{\pi}\cdot\partial^\mu\vec{\pi} - \frac{1}{2F^2}\,(\vec{\pi}\cdot\partial_\mu\vec{\pi})\,(\vec{\pi}\cdot\partial^\mu\vec{\pi}) + O(\pi^6) + \cdots.$$

An integration by parts has been performed in writing the π^4 term of the expansion of the lagrangian.

The couplings between nucleons and pions to lowest order in the derivative expansion involve only one derivative. The most general form for these that is consistent with the nonlinearly-realized G-invariance, and with parity invariance, is:

$$\begin{aligned}\mathcal{L}_{\pi N} &= -\overline{N}\left(\partial\!\!\!/ - \frac{i}{2}\vec{\mathcal{A}\!\!\!/}(\vec{\theta})\cdot\vec{\tau} + m_{_N}\right)N - \frac{ig}{2}\left(\overline{N}\gamma^{\mu}\gamma_5\vec{\tau}N\right)\ \vec{e}_{\mu}(\vec{\theta}), \\ &= -\overline{N}\left(\partial\!\!\!/ + m_{_N}\right)N - \frac{ig}{2F}\left(\overline{N}\gamma^{\mu}\gamma_5\vec{\tau}N\right)\cdot\partial_{\mu}\vec{\pi} \\ &\qquad -\frac{i}{2F^2}\left(\overline{N}\gamma^{\mu}\vec{\tau}N\right)\cdot(\vec{\pi}\times\partial_{\mu}\vec{\pi}) + \cdots .\end{aligned} \tag{2.3.4}$$

The ellipses here represent terms which involve either three or more powers of the pion field, more than two powers of the nucleon field, or involve more than one derivative.

Clearly, only the one constant F need be determined in order to completely fix the dominant low-energy pion self-interactions, and a second constant, g, is also required to determine the lowest-order pion-nucleon couplings.

2.3.1 Conserved Currents

For future reference it is instructive to compute the Noether currents for the symmetry group $G = SU_{_L}(2)\times SU_{_R}(2)$ in both the underlying theory (*i.e.* QCD), and in the effective low-energy pion-nucleon theory.

In QCD, the symmetry transformation under G is given by $\delta q = \frac{i}{2}\left(\vec{\omega}_{_L}\gamma_{_L} + \vec{\omega}_{_R}\gamma_{_R}\right)\cdot\vec{\tau}\ q$, where $q = \begin{pmatrix} u \\ d \end{pmatrix}$ denotes the two-component quantity containing the two lightest quarks. The corresponding Noether currents that are obtained from the QCD lagrangian, eq. (2.1.1) are:

$$\vec{\jmath}^{\,\mu}_{_L} = \frac{i}{2}\,\overline{q}\gamma^{\mu}\gamma_{_L}\vec{\tau}\ q, \qquad \text{and} \qquad \vec{\jmath}^{\,\mu}_{_R} = \frac{i}{2}\,\overline{q}\gamma^{\mu}\gamma_{_R}\vec{\tau}\ q. \tag{2.3.5}$$

The current, $\vec{\jmath}^{\,\mu}_I$, which corresponds to the unbroken $SU_I(2)$ isospin symmetry is therefore: $\vec{\jmath}^{\,\mu}_I = \vec{\jmath}^{\,\mu}_L + \vec{\jmath}^{\,\mu}_R = \frac{i}{2}\,\overline{q}\gamma^\mu\vec{\tau}\,q$. The current for the broken symmetry is similarly: $\vec{\jmath}^{\,\mu}_A = \vec{\jmath}^{\,\mu}_L - \vec{\jmath}^{\,\mu}_R = \frac{i}{2}\,\overline{q}\gamma^\mu\gamma_5\vec{\tau}\,q$.

In the effective pion-nucleon theory the corresponding current may also be constructed using the known action of G on $\vec{\pi}$ and N, and using the lagrangian, whose lowest-derivative terms are given by eqs. (2.3.3) and (2.3.4). Keeping only the terms involving a single pion or only two nucleons, at the lowest order in the derivative expansion, then gives:

$$\begin{aligned} \vec{\jmath}^{\,\mu}_I &= -\Big(\vec{\pi}\times\partial^\mu\vec{\pi}\Big) + \frac{i}{2}\,\overline{N}\gamma^\mu\vec{\tau}\,N + \cdots, \\ \vec{\jmath}^{\,\mu}_A &= F\,\partial^\mu\vec{\pi} + \frac{ig}{2}\,\overline{N}\gamma^\mu\gamma_5\vec{\tau}\,N + \cdots. \end{aligned} \tag{2.3.6}$$

There are an infinite number of higher-order terms in these currents corresponding to the infinite number of interactions in the effective pion-nucleon lagrangian. All of the terms not written explicitly above involve additional factors of the fields $\vec{\pi}$ or N, or involve more derivatives of these fields than do the terms displayed.

2.3.2 Determining F and g

These expressions for the Noether currents for G turn out to furnish a handle for experimentally determining the constants F and g. This is because, as is made explicit in the following section, experimental information exists concerning the value of some of the matrix elements of the broken current $\vec{\jmath}^{\,\mu}_A$.

This experimental information exists because it is precisely the current $\vec{\jmath}^{\,\mu}_A$ which appears in that part of the weak-interaction lagrangian which describes transitions from d quarks to u quarks. Since these transitions are responsible for many reactions, including

all nuclear β-decays, free-neutron decay, and $\pi^{\pm}$ decay, the corresponding matrix elements of this current can be measured.

The terms in the underlying lagrangian which describe these decays are obtained by supplementing the QCD interactions of eq. (2.1.1) with the weak-interaction term:

$$\mathcal{L}_{\text{weak}} = \frac{G_{_F} \cos\theta_{_C}}{\sqrt{2}}\, \overline{u}\gamma^{\nu}(1+\gamma_5)d\; \overline{\nu}_{\ell}\gamma_{\nu}(1+\gamma_5)\ell + \text{h.c.}. \tag{2.3.7}$$

Here the Dirac spinor field ℓ and the Majorana field ν_{ℓ} respectively represent a charged lepton — in practice, the electron and muon — and the corresponding neutrino. $G_{_F}$ is the *Fermi coupling constant*, which is determined from the muon decay rate to be $G_{_F} = 1.16649(2) \times 10^{-5}$ GeV^{-2}. The angle $\theta_{_C}$ is called the *Cabbibo angle*, and it parameterizes the fact that the size of the coupling constant, $G_{_F}\cos\theta_{_C}$, as seen in superallowed nuclear β-decays is smaller than $G_{_F}$ as is measured in muon decay. Numerically, $\cos\theta_{_C} = 0.9753(6)$.

The main feature to be noticed from eq. (2.3.7) is that the quark combination which appears is a linear combination of the conserved $SU_{_L}(2) \times SU_{_R}(2)$ currents:

$$\begin{aligned} i\overline{u}\gamma^{\mu}(1+\gamma_5)d &= i\overline{q}\gamma^{\mu}\gamma_{_L}(\tau_1 + i\tau_2)\; q \\ &= \Big[(j_{_I})^{\mu}_1 + i(j_{_I})^{\mu}_2\Big] + \Big[(j_{_A})^{\mu}_1 + i(j_{_A})^{\mu}_2\Big]. \end{aligned} \tag{2.3.8}$$

In preparation for using eqs. (2.3.6) we have re-expressed the left-handed currents which appear in the weak interactions in favour of the axial and vector currents using: $\vec{\jmath}^{\,\mu}_{_L} = \frac{1}{2}\,(\vec{\jmath}^{\,\mu}_{_I} + \vec{\jmath}^{\,\mu}_{_A})$.

To compute the decay rate for the reaction $\pi^+ \to \mu^+\nu_{\mu}$ we require the following matrix element: $\langle \mu^+, \nu_{\mu}|\mathcal{L}_{\text{weak}}|\pi^+\rangle$. The part of this matrix element which involves strongly-interacting particles is $\langle\Omega|\vec{\jmath}^{\,\mu}_{_A}|\pi^+\rangle$, where $|\Omega\rangle$ is the QCD ground state. The isospin current, $\vec{\jmath}^{\,\mu}_{_I}$, does not appear in π^+ decay because its matrix element vanishes due to the parity

invariance of the strong interactions. The most general form for this matrix element which is consistent with Poincaré and isospin invariance is given by:

$$\langle \Omega | (j_A)^\mu_n(x) | \pi_m(q) \rangle = \frac{i F_\pi \, q^\mu \, e^{iqx}}{\sqrt{(2\pi)^3 \, 2q^0}} \, \delta_{mn}, \tag{2.3.9}$$

where it is conventional to extract the numerical factor $1/\sqrt{(2\pi)^3 2}$, and the pion states are labelled here as members, $|\pi_n\rangle$ $(n = 1, 2, 3)$, of an isotriplet. These are related to the physical states, having definite electric charge, by: $|\pi^\pm\rangle = \frac{1}{\sqrt{2}} \, (|\pi_1\rangle \mp i|\pi_2\rangle)$, and $|\pi^0\rangle = |\pi_3\rangle$.

The only unknown quantity in this matrix element is the constant F_π, which is inferred to be $F_\pi = 92\,\text{MeV}$ by comparing the prediction, $1/\tau_{\text{th}} = (G_F^2 \cos^2\theta_C F_\pi^2 m_\mu^2 m_\pi / 4\pi) \, (1 - m_\mu^2/m_\pi^2)^2$, with the observed mean lifetime, $\tau_{\text{exp}} = 2.6030(24) \times 10^{-8}$ s, for the decay $\pi^+ \to \mu^+ \nu_\mu$.

Now, to lowest order in the derivative expansion, the matrix element of eq. (2.3.9) can be directly evaluated as a function of the parameter F using the second of eq. (2.3.6). Comparing these results permits the inference

$$F = F_\pi = 92 \;\; \text{MeV}. \tag{2.3.10}$$

With this constant in hand, we may now use the low-energy effective lagrangian to predict the low-energy pion self-interactions.

Before proceeding to these predictions, we first repeat these steps for another matrix element in order to infer the value of the constant, g, which governs the size of the pion-nucleon coupling. We once again consider the weak interaction, eq. (2.3.7), but this time consider its prediction for the decay rate of a free neutron into a proton, an electron and an antineutrino: $n \to pe\overline{\nu}_e$. In this case, the most general Poincaré-, parity-, time-reversal-

and isospin-invariant form for the desired matrix element is:

$$\begin{aligned}
\langle N(k,\sigma)|\vec{\jmath}_I^{\,\mu}(x)|N(l,\zeta)\rangle &= \frac{ie^{iqx}}{2(2\pi)^3}\ \overline{u}(k,\sigma)\Big[F_1(q^2)\gamma^\mu + F_2(q^2)\gamma^{\mu\nu}q_\nu\Big]\,\vec{\tau}\ u(l,\zeta), \\
\langle N(k,\sigma)|\vec{\jmath}_A^{\,\mu}(x)|N(l,\zeta)\rangle &= \frac{ie^{iqx}}{2(2\pi)^3}\ \overline{u}(k,\sigma)\Big[G_1(q^2)\gamma^\mu\gamma_5 + G_2(q^2)\gamma_5 q^\mu\Big]\,\vec{\tau}\ u(l,\zeta).
\end{aligned} \tag{2.3.11}$$

Here, l^μ and k^μ are the four-momenta of the initial and final nucleons, and $q^\mu = (l-k)^\mu$ is their difference. ζ and σ similarly represent the spins of the initial and final nucleons. $u(k,\sigma)$ is the Dirac spinor for a free particle having momentum k^μ, with $k^2 + m_{_N}^2 = 0$, and spin σ. (Our normalization is: $\overline{u}(k,\sigma)u(k,\sigma') = (m_{_N}/k^0)\ \delta_{\sigma\sigma'}$.) Finally, $\gamma^{\mu\nu}$ stands for the commutator $\frac{1}{2}\,[\gamma^\mu,\gamma^\nu]$.

The unknowns in this matrix element are the four Lorentz-invariant functions, F_1, F_2, G_1 and G_2, of the invariant momentum transfer, q^2. These functions are not completely arbitrary, however, since they must encode the fact that we are working in a limit where $G = SU_{_L}(2) \times SU_{_R}(2)$ is taken to be a symmetry of the QCD lagrangian. The implications of G-invariance are easily extracted by demanding current conservation, $\partial_\mu \vec{\jmath}_I^{\,\mu} = \partial_\mu \vec{\jmath}^{\,\mu} = 0$, for all of the currents. Keeping in mind that the nucleons have equal mass in the G-invariant limit in which we are working, this implies no conditions on the functions F_1 and F_2, but implies for the others:

$$2im_{_N}\ G_1(q^2) = q^2\ G_2(q^2). \tag{2.3.12}$$

In the rest frame of the decaying neutron, the components of the momentum transfer, q^μ, are at most of order 1 MeV. Since this is much smaller than the typical strong-interaction scale, $\Lambda \sim 1$ GeV, which characterizes the matrix element, for neutron decay it suffices to simplify eqs. (2.3.11) using $q^\mu \approx 0$. In this approximation the neutron decay rate depends only on the two unknown constants, $F_1(0)$ and $G_1(0)$.

Figure 2.1: The Feynman graphs which give the dominant nucleon matrix elements of the Noether currents in the low-energy effective theory. Solid lines represent nucleons, and dashed lines represent pions.

Since the constants $F_1(0)$ and $G_1(0)$ correspond to the low-energy limit of these current matrix elements, they may be related to the constants which appear in the dominant terms of the low-energy effective lagrangian. This may be done by using eqs. (2.3.6) to directly evaluate the matrix elements of eqs. (2.3.11). Doing so, we find contributions from the two Feynman graphs of Fig. 2.1. The first of these gives the direct matrix element of eqs. (2.3.6), and contributes to the form factors F_1 and G_1. The second graph uses the $NN\pi$ interaction of the effective lagrangian, eq. (2.3.4), together with the vacuum-pion matrix element of eq. (2.3.9). It contributes only to the form factor G_2. Evaluating these graphs, we find:

$$F_1 = 1, \qquad G_1 = g, \qquad \text{and} \qquad G_2 = \frac{2igm_N}{q^2}, \tag{2.3.13}$$

from which we see $F_1(0) = 1$ and $G_1(0) = g$. The factor $1/q^2$ in G_2 comes from the massless pion propagator in the second of Figs. 2.1. Notice that this result for G_2 is precisely what is required to satisfy the current-conservation condition of eq. (2.3.12).

The finding that $F_1(0) = 1$ states that this part of the matrix element is not renormalized by the strong interactions, since this value for $F_1(0)$ is the same as would have been obtained if the matrix elements of $\vec{J}^\mu_I$ were taken using the underlying quark states rather than with the composite nucleon states. $F_1(0)$ is the same for both quarks and nucleons because $F_1(0)$ is the quantity which determines the matrix elements in these

states of the conserved isospin charges, $\vec{\mathcal{I}} = \int d^3\mathbf{r}\ \vec{\jmath}_I^0$. But these have matrix elements which depend only on the $SU_I(2)$ transformation properties of the states whose matrix elements are taken. Since both quarks and nucleons are isodoublets, and since inspection of eq. (2.3.5) shows that quarks have $F_1(0) = 1$, the same must be true for nucleons.

The same argument does not hold for the axial current because this is a current for a symmetry which is spontaneously broken. This turns out to imply that the corresponding conserved charge is not well defined when acting on particle states, and so $G_1(0)$ need not be unity.

We finally arrive at the desired conclusion: the neutron decay rate, which is completely determined by the constants $F_1(0) = 1$ and $G_1(0) = g$, can be used to experimentally infer the numerical value taken by the remaining constant, g, of the effective lagrangian. The measured neutron mean life — which is $\tau_{\text{exp}} = 887(2)$ s — then implies $g = 1.26$.

Having determined from experiment the values taken by F and g, we are now in a position to use the effective pion-nucleon lagrangian to predict the low-energy properties of pion-pion and pion-nucleon interactions.

2.3.3 The Goldberger-Treiman Relation

Historically the trilinear $N - N - \pi$ interaction has been written with no derivatives, as a Yukawa coupling:

$$\mathcal{L}_{NN\pi} = i g_{NN\pi}\ (\overline{N}\,\gamma_5 \vec{\tau}\, N) \cdot \vec{\pi}, \tag{2.3.14}$$

with the constant $g_{NN\pi}$ found from phenomenological studies to be close to 14. But the value of this constant can be predicted in terms of the constant g, and this prediction serves as the first test of the low-energy pion-nucleon lagrangian.

The prediction starts with the trilinear $N-N-\pi$ interaction of eq. (2.3.4):

$$\mathcal{L}_{NN\pi} = -\frac{ig}{2F_\pi}\left(\overline{N}\gamma^\mu\gamma_5\vec{\tau}N\right)\cdot\partial_\mu\vec{\pi}, \tag{2.3.15}$$

and performs an integration by parts to move the derivative to the nucleon fields. One then uses the lowest-order equations of motion for N: *i.e.* $(\partial\!\!\!/ + m_N)\,N = 0$, to simplify the result. One obtains a result of the form of eq. (2.3.14), but with

$$g_{NN\pi} = \frac{gm_N}{F_\pi}. \tag{2.3.16}$$

Using the experimental values: $g = 1.26$, $m_N = 940$ MeV and $F_\pi = 92$ MeV gives the prediction $g_{NN\pi} = 12.8$, which agrees well with the phenomenologically inferred value. This prediction, eq. (2.3.16), is known as the *Goldberger-Treiman relation.*

We turn now to one last dangling issue which remains to be addressed before we can compute low-energy pion-pion and pion-nucleon scattering.

2.4 Explicit Symmetry-Breaking

Notice that the effective lagrangian, eqs. (2.3.3) and (2.3.4), has very definite implications for the masses of the pions and nucleons. It states that the pion multiplet must be exactly massless, and that the nucleon masses must be equal. Since these predictions rely only on the assumption of unbroken G invariance, and since G-invariance only holds for QCD in the limit that m_u and m_d vanish, corrections to the pion and nucleon mass predictions can only be inferred by including the effects of the symmetry-breaking quark mass terms for the low energy effective theory. We do so, in this section, to lowest order in the light-quark masses.

The quark mass terms in the QCD lagrangian are proportional to $\overline{q}\ M\gamma_L\ q + \text{h.c.}$, where $M = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix}$ is the light-quark mass matrix. Under the $G = SU_L(2)\times SU_R(2)$

symmetry, $q \to (g_L\gamma_L + g_R\gamma_R)\, q$, this transforms into:

$$\overline{q}\; M\gamma_L\; q \to \overline{q}\; g_R^\dagger M g_L\gamma_L\; q + \text{h.c.}. \tag{2.4.1}$$

Although this is not invariant, it *would* have been invariant if the mass matrix had been a field which had also transformed under G according to: $M \to g_R M g_L^\dagger$.

We imagine the effective pion-nucleon theory having an expansion in the light quark masses, M: $\mathcal{L}_{\text{eff}} = \mathcal{L}_0 + \mathcal{L}_1 + \cdots$, where the subscript indicates the power of M it contains. Each of these terms may be separately expanded in powers of the derivatives, and of the fields π and N. The construction to this point has given the lowest-derivative terms which can appear in $\mathcal{L}_0$. Our goal now is to determine the most general form which may be taken by $\mathcal{L}_1$, and which contains no derivatives of any fields. This will give the dominant symmetry-breaking contribution at low energies.

2.4.1 Pions Only: Vacuum Alignment

We start by focussing on the part of $\mathcal{L}_1$ which depends only on the pion fields. The form taken by $\mathcal{L}_1$ may be obtained from the following argument. We require that $\mathcal{L}_1$ be G-invariant, but only if we take $M \to g_R M g_L^\dagger$ in addition to transforming the fields π in their usual way.

It is straightforward to construct one such a term involving only the pion fields. The simplest construction is to use the quantity $\Xi \equiv u_L u_R^\dagger = e^{i\vec{\theta}\cdot\vec{\tau}} = \cos\theta + i\hat{\theta}\cdot\vec{\tau}\sin\theta$, defined in section (8.2.1), which transforms according to $\Xi \to \tilde{\Xi} = g_L \Xi g_R^\dagger$. (Recall here that θ and $\hat{\theta}$ are defined by $\theta = \sqrt{\vec{\theta}\cdot\vec{\theta}}$ and $\hat{\theta} = \vec{\theta}/\theta$.)

A possible lagrangian therefore is:

$$\mathcal{L}_{1,\pi\pi} \;\; = \;\; -A\ \text{Re}\ \text{Tr}\,[M\ \Xi] - B\ \text{Im}\ \text{Tr}\,[M\ \Xi] \tag{2.4.2}$$

$$= \quad -A\,(m_u + m_d)\,\cos\theta - B\,(m_u - m_d)\,\theta_3\,\frac{\sin\theta}{\theta}. \tag{2.4.3}$$

Clearly this generates a potential energy which is a function of $\vec{\theta}$, as is possible because of the explicit breaking of the $SU_{_L}(2) \times SU_{_R}(2)$ symmetry by the quark masses. As a result, all values for $\vec{\theta}$ are not equally good descriptions of the vacuum, and it is necessary to minimize the potential in order to determine the vacuum value for $\vec{\theta}$. This choosing of the vacuum value for the pseudo-Goldstone fields after the introduction of explicit symmetry-breaking is a process known as *vacuum alignment*.

In the present instance the potential is minimized by $\theta_1 = \theta_2 = 0$, and has the schematic form $V(\theta) = -\mathcal{L}_1(\theta) = A\cos\theta + B\sin\theta = -|A|\cos(\theta - \theta_0)$, for $\theta = \theta_3$ and A and B (or, equivalently, A and θ_0) constants. This, once minimized (giving $\theta_{\min} = \theta_0$) and expanded about the minimum (with $\theta = \theta_{\min} + \theta'$), the potential becomes $V(\theta') = -|A|\cos\theta'$, leaving our lagrangian density of the form:

$$\begin{aligned}
\mathcal{L}_{1,\pi\pi} &= \frac{\mathcal{M}^3}{2}\,\mathrm{Tr}\,[M\,(\Xi + \Xi^\dagger)], \\
&= (m_u + m_d)\,\mathcal{M}^3\,\cos\theta, \\
&= m_\pi^2\,\left[F_\pi^2 - \frac{1}{2}\,\vec{\pi}\cdot\vec{\pi} - \frac{1}{4!\,F_\pi^2}\,(\vec{\pi}\cdot\vec{\pi})^2 + O(\pi^6)\right],
\end{aligned} \tag{2.4.4}$$

with the constant $\mathcal{M}^3$ positive.

Eq. (2.4.4) gives the required symmetry-breaking interaction, where the $\vec{\pi}$'s are chosen so that the vacuum is at $\vec{\pi} = 0$, and, in the last line, we have also eliminated the arbitrary parameter, $\mathcal{M}$, which has the dimensions of mass, in terms of the common mass, m_π, we find for all three pions:

$$m_\pi^2 = (m_u + m_d)\,\frac{\mathcal{M}^3}{F_\pi^2}. \tag{2.4.5}$$

There are several features here worth highlighting. Firstly, notice that all of the pion

self-interactions necessarily preserve isospin to this order in the derivative and quark-mass expansions. This implies, among other things, degenerate masses for all three pions. This preservation of isospin does *not* rely on the isospin-breaking difference, $m_u - m_d$, being small in comparison with m_u or m_d. Rather, it relies only on m_u and m_d both being small compared to the characteristic scale of QCD. We must look elsewhere for an understanding of the observed mass difference between the charged and neutral pions, such as to the isospin-breaking electromagnetic interactions.

Secondly, the lagrangian of eq. (2.4.4) necessarily implies a quark-mass-dependent contribution to the vacuum-energy density, $-\rho_{_V} = \mathcal{L}_{1,\pi\pi}(\vec{\pi} = 0) = m_\pi^2 F_\pi^2 = (m_u + m_d)\,\mathcal{M}^3$. This contribution permits a physical interpretation for the parameter $\mathcal{M}$, as follows. In the underlying theory the derivative of the total vacuum energy, $\rho_{_V}$, with respect to any quark mass is given by:

$$\frac{\partial \rho_{_V}}{\partial m_q} = \langle \Omega | \overline{q}\, q | \Omega \rangle, \tag{2.4.6}$$

where $|\Omega\rangle$ is the QCD ground state. We see, by comparison with the pion scalar potential, eq. (2.4.4), that

$$\langle \Omega | \overline{u}\, u | \Omega \rangle = \langle \Omega | \overline{d}\, d | \Omega \rangle = -\mathcal{M}^3 + \cdots, \tag{2.4.7}$$

where the ellipses here denote the contributions due to quantum effects in the low-energy pion-nucleon theory. Evidently $\mathcal{M} \approx [m_\pi^2 F_\pi^2/(m_u + m_d)]^{1/3}$ gives the size of the expectation value which is responsible for the spontaneous breaking of the chiral $SU_{_L}(2) \times SU_{_R}(2)$ symmetry. Using the values $m_\pi = 140$ MeV, $F_\pi = 92$ MeV and 4.5 MeV $< m_u + m_d <$ 14 MeV then gives 230 MeV $< \mathcal{M} <$ 330 MeV for this scale.

Next, we remark that since none of the terms which appear in $\mathcal{L}_{1,\pi\pi}$ depend on derivatives of $\vec{\pi}$, they do not at all affect expressions (2.3.6) for the conserved Noether

currents of the theory. We therefore need not at all change the above analysis which determined the experimental values for the constants F and g from pion and nucleon weak decays.

Finally, we note that the mass term given in eq. (2.4.4) is the only possible term (up to normalization) which is linear in M and depends only on θ, and not on its derivatives. This uniqueness follows from the impossibility of building a G-invariant scalar potential. To see this, we write $\mathcal{L}_1 = \mathrm{Tr}\,(M\,\mathcal{O}) + \mathrm{h.c.}$, for $\mathcal{O}(\vec{\theta})$ a two-by-two matrix function of the Goldstone boson fields. $\mathcal{O}$ must transform under G according to $\mathcal{O} \to g_{_L}\mathcal{O}g_{_R}^\dagger$. $\mathcal{O}_1 = \Xi$ satisfies these conditions, but suppose $\mathcal{O} = \mathcal{O}_2$ were a second, independent solution. In this case, the combination $V(\vec{\theta}) = \mathrm{Tr}\,[\mathcal{O}_1\mathcal{O}_2^\dagger]$ or $V(\vec{\theta}) = \det\,[\mathcal{O}_1\mathcal{O}_2^\dagger]$ would be a G-invariant scalar potential, as would any of the eigenvalues of the matrix $\mathcal{O}_1\mathcal{O}_2^\dagger$. Since we know from the previous chapter no such potential is possible, it follows that an independent quantity, $\mathcal{O}_2$, also cannot exist.

2.4.2 Including Nucleons

We next consider the part of $\mathcal{L}_1$ which involves precisely two factors of the nucleon field, and no derivatives. That is: $\mathcal{L}_{1,N\pi} = -\overline{N}\;f(\vec{\theta}, M)\;\gamma_{_L}N + \mathrm{h.c.}$, where the transformation laws: $\vec{\theta} \to \tilde{\vec{\theta}}$ and $N \to \tilde{N} = h(\vec{\theta}, g)\,N$ imply that the matrix-valued function, $f(\vec{\theta}, M)$, must satisfy: $f(\tilde{\vec{\theta}}, g_{_R}Mg_{_L}^\dagger) = h\;f(\vec{\theta}, M)\;h^\dagger$.

The solution, unique up to normalization, to this condition is: $f = u_{_R}^\dagger M u_{_L}$, where $u_{_L} = u_{_R}^\dagger = \exp\left[\frac{i}{2}\,\vec{\theta}\cdot\vec{\tau}\right]$. We therefore find:

$$\begin{aligned}\mathcal{L}_{1,N\pi} &= -\lambda\,\overline{N}\,\left[e^{\frac{i}{2}\vec{\theta}\cdot\vec{\tau}}\,M\,e^{\frac{i}{2}\vec{\theta}\cdot\vec{\tau}}\right]N + \mathrm{h.c.}\\ &= -\lambda\,\overline{N}\,M\,N - i\lambda\,\left(\frac{\sin\theta}{2\theta}\right)\,\overline{N}\Big\{\vec{\theta}\cdot\vec{\tau}, M\Big\}\gamma_5\,N\end{aligned}$$

$$+\lambda\left(\frac{\sin^2\frac{\theta}{2}}{\theta^2}\right)\ \overline{N}\Big[\vec{\theta}\cdot\vec{\tau}M\vec{\theta}\cdot\vec{\tau}+\theta^2 M\Big]\ N \tag{2.4.8}$$

$$\begin{aligned} = \ & -\lambda\,\overline{N}\,M\,N-\frac{i\lambda}{2F_\pi}\,\overline{N}\Big\{\vec{\pi}\cdot\vec{\tau},M\Big\}\,N \\ & +\frac{\lambda}{4F_\pi^2}\,\overline{N}\Big[\vec{\pi}\cdot\vec{\tau}M\vec{\pi}\cdot\vec{\tau}+\vec{\pi}\cdot\vec{\pi}\,M\Big]\ N+\cdots. \end{aligned}$$

As usual, curly braces denote the anticommutator of the corresponding matrices: $\{A,B\}=AB+BA$.

Notice that besides providing nonderivative pion-nucleon couplings, this term also splits the neutron and proton masses by an amount:

$$\delta_\lambda m_N=\lambda(m_d-m_u). \tag{2.4.9}$$

Even though they do not contribute to the pion mass splittings, the differing u and d quark masses do act to split the masses of the nucleon isodoublet. Now, m_u and m_d may be determined by repeating the above analysis for the masses of the lightest eight mesons, π, K, η, under the assumption that these are all pseudo-Goldstone bosons for the symmetry group $SU_L(3)\times SU_R(3)$ which is appropriate when the s quark is assumed to be light in addition to the u and d quarks. In principle, once this has been done, eq. (2.4.9) permits the constant λ to be extracted from the experimental difference, $m_n-m_p=1.293318(9)$ MeV. It is important in so doing to include also the contributions of the electromagnetic interactions to this mass difference, since these are similar in size to eq. (2.4.9).[2]

2.5 Soft Pion Theorems

We may now proceed to work out some of the implications of the effective lagrangian for low-energy pion-pion scattering. As usual, the first question must be to ask which

[2]I thank John Donoghue for reminding me of the importance of this electromagnetic contribution.

interactions need be considered in which Feynman graphs in order to properly mimic the low-energy expansion of the underlying QCD theory. To this end we use the powercounting results of Chapter 1.

2.5.1 Power Counting

For simplicity we consider here only the case where there are no nucleons in the low-energy theory, since only in this case we can directly use the power-counting results obtained in Chapter 1. As has already been emphasized, these results cannot be directly applied to nucleons, because they were derived under the assumption of very light, relativistic fermions, and the nucleons in the low-energy pion-nucleon theory are very massive and nonrelativistic at the energies of interest.

Power Counting in the Symmetric Limit

We start by omitting all symmetry-breaking terms of the pion lagrangian which are proportional to the quark masses. These are considered in the next section. In this case the pions are massless, and their Goldstone-boson lagrangian has the form given in eq. (2.3.1):

$$\mathcal{L}_{\pi\pi} = -F_\pi^2 \left[\frac{1}{2}\,\hat{g}_{mn}(\vec{\theta})\,\partial_\mu\theta^m\partial^\mu\theta^n + \frac{c}{\Lambda_\chi^2}\,h_{mnpq}(\vec{\theta})\,\partial_\mu\theta^m\partial^\mu\theta^n\,\partial_\nu\theta^p\partial^\nu\theta^q + \cdots\right], \qquad (2.5.1)$$

where c is a dimensionless number which is, in principle, calculable from QCD. The ellipses represent an infinite sequence of additional terms, including several others which also have four derivatives, as does the displayed term proportional to c.

Eq. (2.5.1) has the form of eq. (1.10.1), with: $f = \sqrt{F_\pi\Lambda_\chi}$, $v = F_\pi$, and $M = \Lambda_\chi \sim 4\pi F_\pi \sim 1$ GeV. The powercounting estimate of eq. (1.10.9) for the scattering of pions becomes:

$$\tilde{\mathcal{A}}_{_E}(q) \sim F_\pi^2\Lambda_\chi^2\,\left(\frac{1}{F_\pi}\right)^E\,\left(\frac{\Lambda_\chi}{4\pi F_\pi}\right)^{2L}\,\left(\frac{q}{\Lambda_\chi}\right)^{2+2L+\sum_{ik}(k-2)V_{ik}}, \qquad (2.5.2)$$

which is a famous result, due first to Weinberg.

For a given observable, $\mathcal{A}_E(q)$, the number, E, of external particles is fixed. In this case it is only the last two factors of eq. (2.5.2) which differentiate different types of contributions. We remark that in practical applications for pion scattering, it happens that $\Lambda_\chi \sim 4\pi F_\pi \sim 1$ GeV. As a result, the second-last factor, $(\Lambda_\chi/4\pi F_\pi)^{2L}$, turns out to be $O(1)$ for realistic pion scattering. This means that it is only the last factor which controls the importance of various interactions.

According to eq. (2.5.2), the contribution of higher-derivative interactions is clearly only suppressed by the ratio q/Λ_χ, which limits us to considering only low-energy pion dynamics near threshhold. The dominant term in the expansion in powers of q/Λ_χ corresponds to choosing the smallest possible value for the quantity $P = 2+2L+\sum_{ik}(k-2)V_{ik}$. It is noteworthy, when using this expression, to remark that all of the interactions have at least two derivatives (we temporarily ignore pion masses *etc.*), and so $k \geq 2$. Furthermore, it is only the first term in the derivative expansion which has $k = 2$, and so the $k = 2$ interaction is unique.

As a result the lowest value possible for P is $P = 2$, and this is only possible if $L = 0$ and if $V_{ik} = 0$ for all $k > 2$. This implies that the dominant contribution to pion scattering is computed by using only the first term in the effective lagrangian, eq. (2.5.1), and working only to tree level with these interactions.

The next-to-leading terms in q/Λ_χ then have $P = 4$, which can arise in either one of two ways. (i) We can have $L = 1$ and $V_{ik} = 0$ for all $k > 2$; or (ii) we can have $L = 0$ and $V_{i4} = 1$ for some i, while V_{i2} can take any value. This states that the subleading, $O(q^4)$, contribution is obtained by either working to one loop order using only the interactions of the first term of the lagrangian of eq. (2.5.1), or using tree graphs having exactly one

vertex taken from the four-derivative interactions in the lagrangian, as well as any number of interactions from the first term in the lagrangian.

In this way it is clear how to compute any given order in the expansion in powers of q/Λ_χ.

Symmetry-Breaking Terms

Before proceeding to calculations, we must also include one other feature. We must track the appearance of the explicit symmetry-breaking terms, of which we only keep those which are proportional to a single power of the light-quark masses, $m_q \sim m_u, m_d$. These vertices, which come from the symmetry-breaking term $\mathcal{L}_1$, can be very simply included into the power-counting results of Chapter 1 by considering all of the non-derivative interactions to be suppressed by a dimensionless coupling, $c_{k=0} \sim (m_q/\Lambda_\chi) \sim (m_\pi^2/\Lambda_\chi^2)$.

With these points in mind, eq. (2.5.2) becomes:

$$\tilde{\mathcal{A}}_E(q) \sim \Lambda_\chi^2 F_\pi^2 \left(\frac{1}{F_\pi}\right)^E \left(\frac{q}{\Lambda}\right)^P \left(\frac{m_q}{\Lambda_\chi}\right)^{\sum_{k=0} V_{ik}}, \tag{2.5.3}$$

where P can be written in either of two equivalent ways:

$$\begin{aligned} P &= 4 - E + \sum_{ik} (k + i - 4)\, V_{ik}, \\ &= 2 + 2L + \sum_{ik} (k - 2)\, V_{ik}. \end{aligned} \tag{2.5.4}$$

Using the first of these forms for P we see that the contribution to the powercounting estimate due to the insertion of the symmetry-breaking terms with $k = 0$ is:

$$\prod_{k=0} \left(\frac{q}{\Lambda}\right)^{(i-4)V_{ik}} \left(\frac{m_q}{\Lambda_\chi}\right)^{V_{ik}}. \tag{2.5.5}$$

The dangerous interactions are clearly those for which $i < 4$. For example, the pion mass

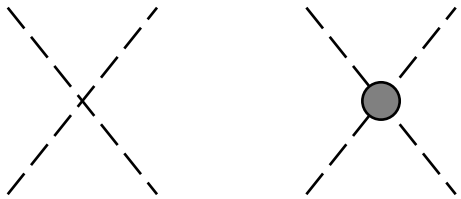

Figure 2.2: The Feynman graphs which give the dominant contributions to pion-pion scattering in the low-energy pion-nucleon theory. The first graph uses a vertex involving two derivatives. The second involves the pion mass, but no derivatives.

term, $\sim m_\pi^2 \pi^2$, has $i = 2$, and so contributes the factor

$$\prod_{k=0} \left(\frac{m_q \Lambda_\chi}{q^2} \right)^{V_{ik}} . \tag{2.5.6}$$

Now, we are interested in applications for which the external momenta, q, are of order several hundred MeV, and so $q \sim m_\pi \sim \sqrt{m_q \Lambda_\chi}$. For these momenta the factor $(m_q \Lambda_\chi / q^2) \sim (m_\pi^2 / q^2) \sim O(1)$. It follows that it is *not* a good approximation to perturb in the pion mass term, and so we should include this term in the unperturbed lagrangian. That is, we should include the pion mass explicitly into the pion propagator so that $G_\pi(q) = -i/(q^2 + m_\pi^2 - i\epsilon)$. It is legitimate to perturb in all of the other symmetry-breaking interactions of the scalar potential, however, since for these $k = 0$ and $i \geq 4$.

We now turn to specific interactions, starting with pion-pion scattering, for which $E = 4$. In this case the above powercounting shows that there are precisely two dominant contributions. The first of these consists of the tree graph of Fig. 2.2, using the four-point vertex from the G-invariant term which involves two derivatives, eq. (2.3.3). The second contribution is also obtained using the graph of Fig. 2.2, but this time takes the four-point pion self-interaction from the symmetry-breaking scalar potential of eq. (2.4.4). Although the first term is unsuppressed by the light-quark masses, it gives a contribution which is down relative to the second term by two powers of external momenta, q. Both are therefore comparable in size for pions near threshhold, $q^2 \sim m_\pi^2$. All other graphs are

smaller than these two by powers of either m_q or q.

2.5.2 Pion-Pion Scattering

We now compute pion-pion scattering by evaluating the graphs of Fig. 2.2 using the effective pion self-couplings of eqs. (2.3.3) and 2.4.4. A straightforward calculation gives the following S matrix element for the scattering $\pi_a\pi_b \to \pi_c\pi_d$:

$$S(\pi_a\pi_b \to \pi_c\pi_d) = \frac{i\delta^4(p_a + p_b - p_c - p_d)}{(2\pi)^2\sqrt{p_a^0 p_b^0 p_c^0 p_d^0}}\,\mathcal{A}_{ab,cd}, \tag{2.5.7}$$

with

$$\mathcal{A}_{ab,cd} = \frac{1}{F_\pi^2}\Big[\delta_{ab}\delta_{cd}\,(s - m_\pi^2) + \delta_{ac}\delta_{bd}\,(t - m_\pi^2) + \delta_{ad}\delta_{bc}\,(u - m_\pi^2)\Big], \tag{2.5.8}$$

where the Lorentz-invariant Mandelstam variables, $s = -(p_a + p_b)^2$, $t = -(p_a - p_c)^2$ and $u = -(p_a - p_d)^2$ are related by the identity: $s + t + u = 4m_\pi^2$. In the CM frame s, t and u have simple expressions in terms of the pion energy, E, and three-momentum, q: $s = 4E^2$, $t = -2E^2 + 2q^2\cos\vartheta$ and $u = -2E^2 - 2q^2\cos\vartheta$. Here ϑ denotes the scattering angle, also in the CM frame.

Comparison with the data is made using channels having definite angular momentum and isospin. If we decompose $\mathcal{A}_{ab,cd}$ into combinations, $\mathcal{A}^{(I)}$, having definite initial isospin:

$$\mathcal{A}_{ab,cd} = \mathcal{A}^{(0)}\,\frac{1}{3}\delta_{ab}\delta_{cd} + \mathcal{A}^{(1)}\,\frac{1}{2}(\delta_{ac}\delta_{bd} - \delta_{ad}\delta_{bc}) + \mathcal{A}^{(2)}\,\left[\frac{1}{2}(\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}) - \frac{1}{3}\delta_{ab}\delta_{cd}\right], \tag{2.5.9}$$

then

$$\mathcal{A}^{(0)} = \frac{2s - m_\pi^2}{F_\pi^2}, \qquad \mathcal{A}^{(1)} = \frac{t - u}{F_\pi^2}, \qquad \mathcal{A}^{(2)}_{cd} = -\,\frac{s - 2m_\pi^2}{F_\pi^2}. \tag{2.5.10}$$

The next step is to resolve these amplitudes into partial waves:

$$\mathcal{A}_\ell^{(I)} \equiv \frac{1}{64\pi}\int_{-1}^{1} d\cos\vartheta\; P_\ell(\cos\vartheta)\,\mathcal{A}^{(I)} \tag{2.5.11}$$

Parameter	Leading Order		Next Order	Experiment
a_0^0	$7m_\pi^2/32\pi F_\pi^2$	0.16	0.20	0.26(5)
b_0^0	$m_\pi^2/4\pi F_\pi^2$	0.18	0.26	0.25(3)
a_1^1	$m_\pi^2/24\pi F_\pi^2$	0.030	0.036	0.038(2)
a_0^2	$-m_\pi^2/16\pi F_\pi^2$	-0.044	- 0.041	- 0.028(12)
b_0^2	$-m_\pi^2/8\pi F_\pi^2$	-0.089	-0.070	-0.082(8)

Table 2.2: Theory *vs* Experiment for Low-Energy Pion Scattering

where $P_\ell(\cos\vartheta)$, as usual, denote the Legendre polynomials (so $P_0(x) = 1$ and $P_1(x) = x$). Since all of the dependence on ϑ appears through the variables t and u, and since eqs. (2.5.10) give $\mathcal{A}^{(0)}$ and $\mathcal{A}^{(2)}$ as functions of s only, it is clear that only the partial wave $\ell = 0$ is predicted at lowest order for the even isospin configurations. Also, since $\mathcal{A}^{(1)}$ is strictly linear in $\cos\vartheta$, it only involves the partial wave $\ell = 1$.

The actual comparison with the data is made by expanding the (real part of) $\mathcal{A}_\ell^{(I)}$ in powers of the squared pion momentum: $q^2/m_\pi^2 = E^2/m_\pi^2 - 1 = (s - 4m_\pi^2)/4m_\pi^2$. That is, writing

$$\mathcal{A}_\ell^{(I)} = \left(\frac{q^2}{m_\pi^2}\right)^\ell \left(a_\ell^I + b_\ell^I \frac{q^2}{m_\pi^2} + \cdots\right), \tag{2.5.12}$$

defines the pion scattering lengths, a_ℓ^I, and slopes, b_ℓ^I. Applying these definitions to eqs. (2.5.10) gives the predictions of the second and third columns of Table 2.2. Column three gives the numerical value corresponding to the analytic expression which is given in column two. The predictions including the next-order terms in the q^2/Λ_χ^2 expansion have also been worked out, and are given in the fourth column of this Table.[3]

Comparison of these predictions with experiment is not straightforward, since it is not feasible to directly perform pion-pion scattering experiments. Instead, the pion-pion

[3]I have taken these values from the excellent book *Dynamics of the Standard Model* by Donoghue, Golowich and Holstein (see bibliography).

scattering amplitudes at low energies are inferred from their influence on the final state in other processes, such as $K \to \pi\pi e\nu_e$ or $\pi N \to \pi\pi N$. The experimental results, as obtained from kaon decays, for those quantities which are predicted to be nonzero at lowest order are listed in the right-hand-most column of Table 2.2. Data also exist for other partial waves which are predicted to vanish at lowest order, such as $I = 0, \ell = 2$, and these are found to be in good agreement with the nonzero predictions which arise at next-to-leading order in the low-energy expansion.

This example nicely illustrates the predictive power which is possible with a low-energy effective lagrangian, even if it is impossible to predict the values for the couplings of this lagrangian in terms of an underlying theory. This predictive power arises because many observables — *e.g.* the pion scattering lengths and slopes — are all parameterized in terms of a single constant — the decay constant, F_π — which can be extracted directly from experiment. We emphasize that this predictive power holds regardless of the renormalizability of the effective theory. Computing to higher orders involves the introduction of more parameters, but predictions remain possible provided that more observables are computed than there are parameters to fix from experiment. The information underlying these predictions comes from the symmetries of the underlying theory, as well as the restrictions due to the comparatively small number of possible interactions which can appear at low orders of the low-energy expansion.

Chapter 3

Magnons: Nonrelativistic Applications

We now turn to a second application, this time to a nonrelativistic system. Besides once again illustrating the utility of the effective-lagrangian techniques, this example shows how the analysis can be applied to more complicated condensed-matter systems. It also illustrates how effective lagrangians permit the separation of the generic predictions which follow only from general properties such as the symmetry-breaking patterns, from the details of the models which may be used to establish these symmetry-breaking patterns from the underlying physics.

We take as our application the macroscopic behaviour of ferromagnets and antiferromagnets. These systems exhibit a transition at low temperatures to a phase which is characterized by a bulk order parameter, which we call **S** for the ferromagnet and **N** for the antiferromagnet, which transforms under rotations as a vector. For ferromagnets this order parameter can be taken to be the overall magnetization of the sample. Because this order parameter spontaneously breaks the rotational symmetry, Goldstone bosons must exist and so must appear in any low-energy (or long-wavelength) description of these systems. It is the low-energy interactions of these Goldstone bosons which is described in

this chapter.

The distinction between a ferromagnet and an antiferromagnet requires more information concerning the underlying material. As for most condensed-matter systems, the underlying microscopic system consists of an enormous number of electromagnetically interacting electrons and atomic nuclei. One picture of what is going on in a ferromagnet or antiferromagnet consists of imagining the electrons being reasonably localized to their corresponding atoms, with these atoms carrying a net magnetic moment due to its having a net electronic spin. The electron in each atom which carries the net spin interacts with its counterparts on neighbouring atoms, resulting in (among other things) an effective spin-spin interaction between these atoms. This spin-spin interaction can come about due to the exchange part of the Coulomb interaction, which arises due to the antisymmetrization of the wavefunction which is required because of the statistics of the electrons. Under varying circumstances one might suppose this spin-spin interaction to either favour the mutual alignment of neighbouring spins, or their *anti*alignment (where the spins line up to point in opposite directions).

The behaviour of such mutually interacting electronic or atomic spins may then be investigated by abstracting out just this spin dynamics into a simplifying model. For example, the interacting electrons can be replaced by a system of spins which are localized to each of the lattice sites which define the nuclear positions in the solid. The mutual interactions of the atoms can be reduced to a spin-spin coupling having a phenomenological sign and magnitude, according to whether it is energetically favourable for neighbouring spins to be aligned or antialigned. Such models show that at low temperatures macroscopic numbers of these spins tend to either align or antialign, according to which of these takes less energy. A system for which neighbouring spins tend to align, prefers to acquire

Figure 3.1: The Distinction Between Ferromagnets (F) and Antiferromagnets (AF).

a net magnetization since the magnetic moment of each atom adds to give a macroscopically large total. This is a ferromagnet. If neighbouring spins prefer to antialign, then the order of the ground state consists of spins which alternate in their alignment, with every other spin pointing in a fixed direction, and the others pointing in the opposite direction. Such an arrangement is called antiferromagnetic. These two alternative arrangements are pictured in Fig. 3.1.

The statistical mechanics of such spin models can successfully describe many features of real ferromagnets and antiferromagnets. So long as calculations are based on models, however, it is difficult to quantitatively assess their accuracy. One of the purposes of the present chapter is to show that some predictions for these systems are very robust, since they do not rely on more than the qualitative features of the models. The robust predictions are those which can be formulated completely within the framework of a low-energy effective theory, and which therefore rely only on the spectrum and symmetries which dominate at low energies. The accuracy of this kind of prediction *can* be quantitatively assessed since this accuracy is controlled by the domain of validity of the effective theory itself. The role played in this kind of calculations by the details of an underlying model is simply the prediction of the quantum numbers and symmetries of the low-energy degrees of freedom, and so the model need only get these qualitative features right in order to accurately reproduce the proper low-energy behaviour.

This line of reasoning, in which some quantitative predictions can be justified as general low-energy features of a given system can be of great practical importance. For example, some very high-precision measurements are now based on the macroscopic behaviour of complicated condensed matter systems. Examples are the Josephson effect, or the Integer Quantum Hall effect, both of which have been used to fix the best measured value for the electromagnetic fine structure constant, α. These determinations are accurate to within very small fractions of a percent. We should only believe such a determination of α if we can equally accurately justify the theoretical predictions of the effects on which the determination is based — a very tall order if the prediction is to be based on a model of the underlying system. Happily, such an accuracy is possible, and is one of the fruits of an effective lagrangian analysis of these systems.

We now turn to the application of these effective-lagrangian techniques to the description of the long-distance, low-energy behaviour of the ordered spin systems.

3.1 Antiferromagnetism: T Invariance

We start with applications to the low-energy properties of antiferromagnets. We do so because antiferromagnets preserve a type of time-reversal symmetry, which makes the analysis of its low energy behaviour fairly similar to what would apply for relativistic systems.

For an antiferromagnet, the order parameter, $\mathbf{N}$, can be taken to be the staggered sum of the spins, $\mathbf{s}_i$, for each lattice site, 'i':

$$\mathbf{N} = \sum_i (-)^i \, \mathbf{s}_i, \tag{3.1.1}$$

where the sign, $(-)^i$, is positive for one sublattice for which all spins are parallel in the

ordered state, and is negative for the other sublattice for which all spins are antiparallel with those of the first lattice. We shall refer to these sublattices in what follows as the 'even' and 'odd' sublattice respectively. By taking such an alternating sum we find the expectation $\langle \mathbf{N} \rangle \neq 0$ in the system's ground state, $|\Omega\rangle$.

The action of time-reversal invariance, T, is to reverse the sign of the spin of every site: $\mathbf{s}_i \to -\mathbf{s}_i$. Although this transformation also reverses the order parameter, $\mathbf{N} \to -\mathbf{N}$, it may be combined with another broken symmetry, S, to obtain a transformation, $\tilde{T} = TS$, which *is* a symmetry of $\mathbf{N}$. This other symmetry, S, consists of a translation (or shift) of the whole lattice by a single lattice site, taking the entire 'even' sublattice onto the 'odd' sublattice, and vice versa. Since both S and T act to reverse the direction of $\mathbf{N}$, they preserve $\mathbf{N}$ when they are performed together.

We next turn to the construction of the general low-energy lagrangian for the Goldstone bosons for the breaking of rotation invariance — called *magnons* — for these systems.

3.1.1 The Nonlinear Realization

The first step is to identify the symmetry breaking pattern, $G \to H$. At first it is tempting to assume that the role of G should be played by the spacetime symmetries, since these include rotations. This is not correct, however, for several reasons. Firstly, the spacetime symmetries of a lattice do not consist of the full group of translations and rotations since these symmetries are broken by the lattice itself. The unbroken subgroup consists only of the group of lattice symmetries: *i.e.* those translations and rotations which take the lattice to itself. There are indeed Goldstone bosons for the spontaneous breaking of translational and rotational symmetry down to this lattice group, but these are the phonons and are

not the focus of the present analysis.

In fact, the rotations of the spins on the lattice can be taken to be an internal $SU(2)$, or $SO(3)$, symmetry, rather than a spacetime symmetry. This is because the action of rotations on the intrinsic spin of a particle becomes an independent internal symmetry, separate from spacetime rotations, in the limit that the particle involved is nonrelativistic. This is because all of the interactions which couple the orbital angular momentum with the spin angular momentum vanish in the limit that the particle mass tends to infinity. Since the spins of interest for real systems are those for nonrelativistic electrons or atoms, we may consider the broken symmetry group to be an internal symmetry, $G = SU(2)$ (which equals $G = SO(3)$, locally). In real life, the electron mass is not infinite, so there are small 'spin-orbit' effects which really do break the internal spin symmetry. These introduce small corrections to predictions based on this symmetry, such as the exact gaplessness of the Goldstone mode. We ignore any such symmetry-breaking effects in what follows.

The order parameter for the symmetry breaking is the vector, $\mathbf{N}$, itself, and so the group of unbroken transformations is $H = U(1)$ (or, $SO(2)$), consisting of rotations about the axis defined by $\langle \mathbf{N} \rangle$. The coset space which is parameterized by the Goldstone bosons is therefore the space $G/H = SU(2)/U(1)$, or $SO(3)/SO(2)$. This last way of writing G/H identifies it as a two-sphere, S_2, since this describes the space swept out by the action of rotations on a vector, $\langle \mathbf{N} \rangle$, of fixed length.

We now have two ways to proceed. We could, on the one hand, follow the steps outlined in Chapter 1 to construct the nonlinear realization of G/H and its invariant lagrangian. Instead we choose here to take a simpler route. As discussed in Chapter 1, the most general possible low-energy Goldstone boson lagrangian must necessarily take

the form of eq. (1.9.1):

$$\mathcal{L}_{AF} = \frac{F_t^2}{2}\,\hat{g}_{\alpha\beta}(\theta)\,\dot{\theta}^\alpha\dot{\theta}^\beta - \frac{F_s^2}{2}\,\hat{g}_{\alpha\beta}(\theta)\nabla\theta^\alpha\cdot\nabla\theta^\beta + (\text{higher-derivative terms}), \tag{3.1.2}$$

where $\hat{g}_{\alpha\beta}(\theta)$ is an $SO(3)$-invariant metric on the two-sphere. This form is the most general consistent with the nonlinearly-realized $SO(3)$ invariance, as well as with invariance with respect to translations, rotations and the time-reversal-like symmetry, $\tilde{T}$, described above.[1] The $\tilde{T}$ invariance rules out interactions having an odd number of time derivatives, such as the term linear in time derivatives which was constructed in Chapter 1.

The main point to be made is that the lagrangian given in eq. (3.1.2) is unique, a result which follows from the uniqueness of the $SO(3)$-invariant metric on the two-sphere. The uniqueness of this metric is a consequence of the fact that the two broken generators of $SO(3)/SO(2)$ form an irreducible representation of the unbroken subgroup $SO(2)$ — a condition which was shown to imply a unique metric in Chapter 1. Since it is unique, any representation of it is equally good and we choose here to use the familiar polar coordinates, (θ,ϕ), for the two-sphere, in terms of which the invariant metric has the usual expression: $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$. With this choice the above lagrangian becomes:

$$\mathcal{L}_{AF} = \frac{F_t^2}{2}\left(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2\right) - \frac{F_s^2}{2}\left(\nabla\theta\cdot\nabla\theta + \sin^2\theta\,\nabla\phi\cdot\nabla\phi\right) + (\text{higher-derivative terms}). \tag{3.1.3}$$

Alternatively, we can equally well parameterize S_2 using a unit vector, $\vec{n}$, where

[1] We use translation and rotational invariance for simplicity, even though these are too restrictive for real solids, for which only the lattice symmetries should be imposed. For some lattices, such as cubic ones, the implications of the lattice group turn out to be the same as what is obtained using rotation and translation invariance, at least for those interactions involving the fewest derivatives which are studied here.

$n_x = \sin\theta\cos\phi$, $n_y = \sin\theta\sin\phi$ and $n_z = \cos\theta$, so $\vec{n}\cdot\vec{n} = 1$. Then

$$\begin{aligned}\mathcal{L}_{AF} &= \frac{F_t^2}{2}\,\dot{\vec{n}}\cdot\dot{\vec{n}} - \frac{F_s^2}{2}\,\nabla\vec{n}\cdot\nabla\vec{n} \\ &\qquad +(\text{higher-derivative terms}).\end{aligned} \tag{3.1.4}$$

This variable, $\vec{n}(\mathbf{r},t)$, makes most clear the physical interpretation of the Goldstone modes: they describe long-wavelength variations in the direction of the order parameter $\langle\mathbf{N}\rangle$. It has the drawback of hiding the self-interactions which are implied by $\mathcal{L}_{AF}$, since the lagrangian of eq. (3.1.4) is purely quadratic in $\vec{n}$. The self-interactions, which are manifest in expression (3.1.3), are nonetheless present, and are hidden in the constraint $\vec{n}\cdot\vec{n} = 1$.

The nonlinear realization of $SO(3)$ transformations on these variables is straightforward to work out, starting with the transformation rule for $\vec{s}$: $\delta\vec{n} = \vec{\omega}\times\vec{n}$, where the vector $\vec{\omega}$ represents the three $SO(3)$ transformation parameters. This implies the transformations:

$$\begin{aligned}\delta\theta &= \omega_y\cos\phi - \omega_x\sin\phi \\ \delta\phi &= \omega_z - \omega_x\cot\theta\cos\phi - \omega_y\cot\theta\sin\phi.\end{aligned} \tag{3.1.5}$$

With these transformation laws we may immediately write down the first terms in a derivative expansion of the Noether currents, $j^\mu = (\vec{\rho},\vec{\mathbf{j}})$, for the $SO(3)$ invariance in the low-energy effective theory. They may be most compactly written:

$$\vec{\rho} = F_t^2\,(\dot{\vec{n}}\times\vec{n}) + \cdots \qquad \text{and} \qquad \vec{\mathbf{j}} = -\,F_s^2\,(\nabla\vec{n}\times\vec{n}) + \cdots, \tag{3.1.6}$$

where the dots are a reminder of the unwritten higher-derivative contributions.

The other quantities which arise in the general nonlinear realization may also be constructed in terms of these variables. For example, the four independent components

of the covariant quantity (zweibein), $e^{\alpha}{}_{\beta}(\theta,\phi)$, are most easily constructed, following the geometrical picture of Chapter 1, as the components of any two orthogonal vectors which are tangent to the two sphere. These may be found by differentiating the unit vector, $\vec{n}$, because the identity $\vec{n}\cdot\vec{n}=1$ implies $\vec{n}\cdot\delta\vec{n}=0$, for any variation, $\delta\vec{n}$. Denoting these two vectors by $\vec{e}_\theta=\partial\vec{n}/\partial\theta$ and $\vec{e}_\phi=\partial\vec{n}/\partial\phi$, we have in cartesian components:

$$\begin{aligned}(\vec{e}_\theta)_x&=\cos\theta\cos\phi, \qquad (\vec{e}_\theta)_y=\cos\theta\sin\phi, \qquad (\vec{e}_\theta)_z=-\sin\theta;\\ (\vec{e}_\phi)_x&=-\sin\theta\sin\phi, \qquad (\vec{e}_\phi)_y=\sin\theta\cos\phi, \qquad (\vec{e}_\phi)_z=0.\end{aligned} \tag{3.1.7}$$

Clearly these vectors satisfy $\vec{e}_\theta\cdot\vec{e}_\phi=\vec{e}_\theta\cdot\vec{n}=\vec{e}_\phi\cdot\vec{n}=0$, and $\vec{e}_\theta\cdot\vec{e}_\theta=1$, $\vec{e}_\phi\cdot\vec{e}_\phi=\sin^2\theta$, so $\vec{e}_\alpha\cdot\vec{e}_\beta=\hat{g}_{\alpha\beta}$, as required. The two-by-two matrix, $e^{\alpha}{}_{\beta}$, of components may be found by expressing the two vectors, $\vec{e}_\beta$, defined by eqs. (3.1.7), as linear combinations of any two orthonormal basis vectors, $\vec{t}^{\alpha}$, which lie tangent to the sphere: *i.e.* $e^{\alpha}{}_{\beta}=\vec{t}^{\alpha}\cdot\vec{e}_\beta$. Using the basis vectors $\vec{e}_\beta$ themselves for this purpose leads to the result:

$$\begin{pmatrix} e^{\theta}{}_{\theta} & e^{\theta}{}_{\phi} \\ e^{\phi}{}_{\theta} & e^{\phi}{}_{\phi}\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & \sin\theta\end{pmatrix}. \tag{3.1.8}$$

3.1.2 Physical Applications

Any physical question that could be asked of the low-energy limit of the underlying theory can equally well be addressed using the low-energy effective lagrangian. In particular, the lagrangian just derived for the Goldstone modes for a ferromagnet may be used to describe the response (at zero or nonzero, but small, temperature) of the system to probes which couple to the spin degrees of freedom.

In order to interpret the constants F_s and F_t it is convenient to expand the field $\vec{n}$, or equivalently θ and ϕ, about its vacuum configuration, $\vec{n}_0=\langle\mathbf{N}\rangle$. We are free to perform an $SO(3)$ rotation to choose the direction of $\vec{n}_0$ arbitrarily, and so we use this

freedom to ensure that $\vec{n}_0$ points up the positive x-axis. This implies that θ and ϕ take the vacuum values, $\theta_0 = \frac{\pi}{2}$ and $\phi_0 = 0$. Writing the canonically normalized fluctuation fields by $\theta = \frac{\pi}{2} + \vartheta/F_t$ and $\phi = \varphi/F_t$, the lagrangian becomes:

$$\begin{aligned}
\mathcal{L}_{AF} &= \frac{1}{2}\left(\dot{\vartheta}^2 - v^2\,\nabla\vartheta\cdot\nabla\vartheta\right) + \frac{1}{2}\cos^2\left(\frac{\vartheta}{F_t}\right)\,\left(\dot{\varphi}^2 - v^2\,\nabla\varphi\cdot\nabla\varphi\right) \\
&\qquad\qquad +(\text{higher-derivative terms}), \qquad\qquad (3.1.9) \\
&= \frac{1}{2}\left(\dot{\vartheta}^2 - v^2\,\nabla\vartheta\cdot\nabla\vartheta + \dot{\varphi}^2 - v^2\,\nabla\varphi\cdot\nabla\varphi\right) \\
&\qquad\qquad -\frac{\vartheta^2}{2F_t^2}\left(\dot{\varphi}^2 - v^2\,\nabla\varphi\cdot\nabla\varphi\right) + \cdots .
\end{aligned}$$

The constant, v, here represents the ratio $v = F_s/F_t$. The ellipses denote terms that involve at least six powers of the fields, or which involve more than two derivatives with respect to either position or time.

The terms quadratic in the fields describe two real modes which propagate according to the linear dispersion law: $E(p) = vp$. These modes physically correspond to *spin waves*: small, long-wavelength precessions of the vector $\vec{n}$ about its vacuum value, $\vec{n}_0$. They carry ± 1 unit of the conserved $SO(2)$ spin in the direction parallel to $\langle \mathbf{N} \rangle$. This gives the physical interpretation of the parameter v to be the velocity of propagation of these modes. The condition that this velocity must be smaller than the velocity of light is $v \leq c = 1$ (in fundamental units), or, equivalently, $F_s \leq F_t$. F_t is similarly seen to govern the strength of the interaction terms in eq. (3.1.9).

These modes and their interactions are amenable to experimental study through their electromagnetic couplings. Although magnons carry no electric charge, they do couple to magnetic fields, $\mathbf{B}$, due to the interactions of the microscopic magnetic moments which participate in the long-wavelength spin waves. This gives a coupling of the magnetic field to the medium's spin density. This coupling can be probed, for example, by scattering

neutrons which are also electrically neutral but which carry an intrinsic magnetic moment which interacts with magnetic fields.

The interaction between magnons and electromagnetic fields is therefore given by a term of the form: $\mathcal{L}_{\rm em} = -\nu\ \vec{s}\cdot\mathbf{B}$, where $\vec{s}$ is the system's spin density. The lowest-dimension effective interaction between magnons and electromagnetic fields is now obtained by expressing the spin density in a derivative expansion, using Noether's result, eq. (3.1.6), for $\vec{s} = \vec{\rho}$. The result is:

$$\begin{aligned}
\mathcal{L}_{\rm em} &= -\nu\ \vec{s}\cdot\mathbf{B}, \\
&= -\nu F_t^2\Big[B_x\ (\dot\theta\ \sin\phi + \dot\phi\ \sin\theta\cos\theta\cos\phi) \\
&\qquad\qquad +B_y\ (-\dot\theta\ \cos\phi + \dot\phi\ \sin\theta\cos\theta\sin\phi) - B_z\ \dot\phi\ \sin^2\theta)\Big] \\
&= \nu F_t\Big(B_y\ \dot\vartheta + B_z\ \dot\varphi\Big) + \cdots,
\end{aligned} \tag{3.1.10}$$

where ν is an effective coupling parameter having the dimensions of magnetic moment (or: inverse mass, in fundamental units). Notice that the time derivative in this interaction ensures invariance with respect to $\tilde{T}$ transformations, under which $\vec{s}\to-\vec{s}$ and $\mathbf{B}\to-\mathbf{B}$.

A nonrelativistic neutron couples to the magnetic field with strength

$$\mathcal{L} = -\mu_{_N}\ n^\dagger\vec{\sigma}\,n\cdot\mathbf{B}, \tag{3.1.11}$$

where $n(x)$ denotes the two-component neutron field, and $\vec{\sigma}$ denotes the Pauli matrices acting in the two-component neutron spin space. The constant $\mu_{_N}$ is the neutron magnetic moment which is, in order of magnitude, $\mu_{_N}\sim e/m_{_N}$. As usual e is the electromagnetic coupling constant (*i.e.* the proton charge) and $m_{_N}$ is the neutron (or nucleon) mass.

Using these interactions, eqs. (3.1.10) and (3.1.11), the cross section per-unit-volume for neutron scattering from the medium can be computed. For slowly-moving neutrons,

and under the assumption that only the momentum, $\mathbf{p}'$, of the scattered neutron is measured we find:

$$\frac{d\sigma}{V d^2\mathbf{p}'} = \frac{\mu_N^2 \nu^2}{4\pi^3 v_N} V_{ij}(\mathbf{p}-\mathbf{p}')\ S_{ij}(E-E', \mathbf{p}-\mathbf{p}'). \tag{3.1.12}$$

Here $v_N \ll 1$ is the speed of the incoming, nonrelativistic, neutron, and V is the volume of the medium whose magnons are responsible for scattering the neutrons. E and E' are the energies of the initial and scattered neutrons, and $\mathbf{p}$ and $\mathbf{p}'$ are their momenta. The function, $V_{ij}(\mathbf{q})$, is the magnetic-moment interaction potential (in momentum space) which arises from the electromagnetic interaction between the neutron and magnon fields:

$$V_{ij}(\mathbf{q}) = \delta_{ij} - \frac{q_i q_j}{\mathbf{q}^2}. \tag{3.1.13}$$

The quantity $S_{ij}(\omega, \mathbf{q})$ is the spin correlation function, which contains all of the information about the scattering medium which is relevant for analyzing the neutron collision. It is defined by:

$$S_{ij}(\omega, \mathbf{q}) = \int dt\, d^3\mathbf{r}\ \langle s_i(\mathbf{r}, t)\, s_j(0)\rangle\ e^{i\omega t - i\mathbf{q}\cdot\mathbf{r}}. \tag{3.1.14}$$

The quantity $\langle s_i(\mathbf{r}, t)\, s_j(0)\rangle = \mathrm{Tr}[\rho\ s_i(\mathbf{r}, t)\, s_j(0)]$ appearing in here is the expectation defined by the density matrix, ρ, which characterizes the initial state of the medium with which the neutron scatters. Notice that it is *not* a time-ordered product of operators which appears in this expectation.

This correlation function may be explicitly computed at low energies using the previously-derived effective lagrangian which describes the low-energy magnon self-interactions. The simplest case is that for which the medium is initially in the no-magnon ground state, and where the energies involved are low enough to neglect the magnon self-interactions. In this case the spin density may be well-approximated by the first terms of eq. (3.1.6).

Then:

$$S_{ij}(\omega, \mathbf{q}) = \frac{\pi \omega^2 F_t^2}{v|\mathbf{q}|}\, \delta_{ij}\, \delta(\omega - v|\mathbf{q}|), \tag{3.1.15}$$

where $v = F_s/F_t$ is the speed of magnon propagation.

We see that the cross section has sharp peaks when the neutron energy and momentum transfers are related by the magnon dispersion relation: $E - E' = v|\mathbf{p} - \mathbf{p}'|$. This corresponds to inelastic scattering in which the neutron transfers its energy and momentum to the medium by creating a magnon. According to eq. (3.1.15) the resulting peaks in the cross section are infinitely sharp, but in real systems they have a finite width due to processes which cause the produced magnons to scatter or decay. If the lifetime for undisturbed magnon propagation, $\tau = 1/\Gamma$, is much longer than the other interaction times of interest in the neutron scattering, then the delta function in eq. (3.1.15) becomes replaced by the lineshape:

$$\delta(\omega - v|\mathbf{q}|) \to \frac{\Gamma}{2\pi} \frac{1}{(\omega - v|\mathbf{q}|)^2 + \frac{1}{4}\Gamma^2}. \tag{3.1.16}$$

Measurements of the positions and widths of these peaks as functions of the scattered neutron energy and momentum can be used to measure the magnon dispersion relation — and so the constant v — and its decay rate, Γ, for the scattering medium. The predicted linear spectrum is indeed found when neutrons are scattered from antiferromagnets. The situation when neutrons scatter from ferromagnets is different, as we shall now see.

3.2 Ferromagnetism: T Breaking

For ferromagnets, the order parameter is simply the total magnetization, or the total spin, of the system. Since this defines a vector in space, the spin symmetry, $G = SO(3)$, is spontaneously broken to $H = SO(2)$ just as for an antiferromagnet. The low energy

behaviour of ferromagnets and antiferromagnets are nevertheless quite different, and this difference is due to the fact that time-reversal symmetry is broken in a ferromagnet but not in an antiferromagnet.

We denote by $\mathbf{S}$ the order parameter for a ferromagnet, which is related to the spins, $\mathbf{s}_i$, of the underlying spin model by:

$$\mathbf{S} = \sum_i \mathbf{s}_i. \tag{3.2.1}$$

Ferromagnets are characterized by having ground states for which there is a nonzero expectation for this quantity: $\langle \mathbf{S} \rangle \neq 0$.

The action of time-reversal invariance, T, is to reverse the sign of the spin of every site, $\mathbf{s}_i \to -\mathbf{s}_i$, and so it does the same for the order parameter, $\mathbf{S} \to -\mathbf{S}$. The difference with the antiferromagnetic case arises because for a ferromagnet it is not possible to find another broken symmetry which combines with time reversal to preserve $\mathbf{S}$. The low-energy effective theory can therefore contain T-violating terms, and this changes the properties of its Goldstone bosons in an important way.

3.2.1 The Nonlinear Realization

Since the symmetry-breaking pattern for both ferromagnets and antiferromagnets is $SO(3) \to SO(2)$, the nonlinear realization of this symmetry on the Goldstone bosons is identical for these two systems. We therefore use the same polar coordinates in this case, θ and ϕ, as in the previous sections. As before it is convenient to use these to define a unit vector, denoted by $\vec{s}$, with components $s_x = \sin\theta\cos\phi$, $s_y = \sin\theta\sin\phi$ and $s_z = \cos\theta$, so that $\vec{s}\cdot\vec{s} = 1$. The field, $\vec{s}(\mathbf{r}, t)$, again describes long-wavelength oscillations in the direction of $\langle \mathbf{S} \rangle$.

The action of $SO(3)$ on these variables is once more given by eq. (3.1.5), and the term in the effective lagrangian which involves the fewest spatial derivatives is again determined to be:

$$\mathcal{L}_{F,s} = -\frac{F_s^2}{2}\Big(\nabla\theta\cdot\nabla\theta + \sin^2\theta\,\nabla\phi\cdot\nabla\phi\Big). \tag{3.2.2}$$

The new features appear once the term with the fewest time derivatives is constructed. As is discussed in some detail in Chapter 1, this involves only a single time derivative because of the broken time-reversal symmetry. It has the form given by eq. (1.9.2):

$$\mathcal{L}_{F,t} = -A_\alpha(\theta)\,\dot\theta^\alpha, \tag{3.2.3}$$

where the coefficient function, $A_\alpha(\theta)$, may be considered to be a gauge field defined on the coset space G/H. In Chapter 1 it was determined that the condition that this term be G invariant is that A_α must only be G-invariant up to a gauge transformation, in the sense that:

$$\pounds_\xi A_\alpha \equiv \xi^\beta\partial_\beta A_\alpha + A_\beta\partial_\alpha\xi^\beta = \partial_\alpha\Omega_\xi, \tag{3.2.4}$$

for each generator $\delta\theta^\alpha = \xi^\alpha$ of G on G/H, where $\Omega_\xi(\theta)$ are a collection of scalar functions on G/H. This last condition is equivalent to the invariance of the field strength for A_α: $\pounds_\xi F_{\alpha\beta} = 0$. Our problem is to explicitly construct such a gauge potential for the example of interest, $G/H = SO(3)/SO(2) \equiv S_2$.

This construction is quite simple. Since our coset space is two dimensional, it is always possible to write the field strength in terms of a scalar field: $F_{\alpha\beta} = \mathcal{B}(\theta)\,\epsilon_{\alpha\beta}$, where $\epsilon_{\alpha\beta}$ is the antisymmetric tensor which is constructed using the coset's G-invariant metric. The condition that $F_{\alpha\beta}$ be G invariant is then equivalent to the invariance of $\mathcal{B}$.

That is:

$$\pounds_\xi \mathcal{B} \equiv \xi^\alpha \, \partial_\alpha \mathcal{B} = 0, \tag{3.2.5}$$

which is only possible for all G transformations if $\mathcal{B}$ is a constant, independent of θ^α.

Our solution for $\mathcal{L}_{F,t}$ for a ferromagnet therefore simply boils down to the construction of a gauge potential for which $F_{\alpha\beta} = \mathcal{B}\ \epsilon_{\alpha\beta}$ on the two-sphere, $S_2 = SO(3)/SO(2)$. But such a gauge potential is very familiar — it is the gauge potential for a magnetic monopole positioned at the centre of the two-sphere. The result may therefore be written (locally) as: $A_\alpha \, d\theta^\alpha = \mathcal{B}\ \cos\theta \, d\phi$, and so the corresponding lagrangian is given by

$$\mathcal{L}_{F,t} = -\,\mathcal{B}\ \cos\theta \, \dot{\phi}, \tag{3.2.6}$$

where $\mathcal{B}$ is a constant. In terms of the vectors $\vec{s}$, $\vec{e}_\theta = \partial\vec{s}/\partial\theta$ and $\vec{e}_\phi = \partial\vec{s}/\partial\phi$ this may be written:

$$\mathcal{L}_{F,t} = -\,\mathcal{B}\ \vec{s}\cdot\Big(\vec{e}_\theta \times \dot{\vec{e}}_\theta\Big). \tag{3.2.7}$$

The complete Goldstone boson lagrangian containing the fewest time and space derivatives is found by combining the contributions of eqs. (3.2.2) and (3.2.6), giving:

$$\mathcal{L}_F = -\,\mathcal{B}\ \cos\theta \, \dot{\phi} - \frac{F_s^2}{2}\Big(\nabla\theta\cdot\nabla\theta + \sin^2\theta\ \nabla\phi\cdot\nabla\phi\Big). \tag{3.2.8}$$

It is instructive to compute the Noether currents for the $SO(3)$ symmetry that is implied by this lagrangian density. The conserved current density is the same as was found for the antiferromagnet:

$$\vec{\mathbf{j}} = F_s^2\ (\vec{s}\times\nabla\vec{s}) + \cdots. \tag{3.2.9}$$

In computing the corresponding expression for the charge density, it is necessary to keep in mind that under these transformations $\mathcal{L}_F$ is not invariant, but instead transforms into

a total derivative:

$$\delta\mathcal{L}_{_F} = -\frac{d\Omega}{dt} = -\frac{\mathcal{B}}{\sin\theta}\Big(\omega_x\,\cos\phi + \omega_y\,\sin\phi\Big). \tag{3.2.10}$$

Using this in the general expression, eq. (1.2.3), for the Noether current gives the conserved charge density:

$$\vec{\rho} = \mathcal{B}\,\vec{s} + \cdots. \tag{3.2.11}$$

The ellipses in this equation, and in eq. (3.2.9), represent more complicated terms which are suppressed by additional derivatives. As is easily verified, the classical equations of motion for the lagrangian, (3.2.8), are equivalent to the conservation condition for this current:

$$\dot{\vec{s}} + k\,\Big(\vec{s}\times\nabla^2\vec{s}\Big) = 0. \tag{3.2.12}$$

This equation has long been known to describe long-wavelength spin waves in ferromagnets, and is called the *Landau-Lifshitz* equation. The constant, k, here is given in terms of F_s and $\mathcal{B}$ by

$$k = \frac{F_s^2}{\mathcal{B}}. \tag{3.2.13}$$

Equation (3.2.11) brings out a feature of time-reversal breaking systems which is qualitatively different from those which preserve time reversal. It states that it is the conserved charge density itself, $\vec{\rho}$, which acquires a vacuum expectation value and breaks the $SO(3)$ symmetry:

$$\langle\vec{\rho}\rangle = \mathcal{B}\;\langle\vec{s}\rangle = \mathcal{B}\;\vec{s}_0 \neq 0. \tag{3.2.14}$$

Clearly the breaking of time reversal (and lorentz invariance) are prerequisites for the acquisition of a nonzero ground-state expectation value for $\vec{\rho}$, which is the time component of a current.

3.2.2 Physical Applications

The propagation of small-amplitude, long-wavelength spin waves is therefore seen to be completely determined by the underlying pattern of spontaneous symmetry breaking: $SO(3) \to SO(2)$ together with T violation. Linearizing the Landau Lifshitz equation, eq. (3.2.12), shows the resulting propagating modes to have the quadratic dispersion relation:

$$E(p) = kp^2. \tag{3.2.15}$$

This dispersion relation, and the value of the constant k, can be measured by neutron scattering, in a manner that is similar to what was found for antiferromagnets. We highlight here only the differences which arise from the antiferromagnetic example.

The lowest-dimension effective interaction which couples the field $\vec{s}$ to electromagnetic fields in the ferromagnetic case is:

$$\begin{aligned} \mathcal{L}_{\rm em} &= -\mu\ \vec{s}\cdot\mathbf{B}, \\ &= -\mu\mathcal{B}\left(B_x\ \sin\theta\cos\phi + B_y\ \sin\theta\sin\phi + B_z\ \cos\theta\right), \\ &= -\mu\mathcal{B}\ B_x - \mu\mathcal{B}\left(B_y\ \delta\phi - B_z\ \delta\theta\right) + \cdots, \end{aligned} \tag{3.2.16}$$

where μ is an effective coupling parameter and we have taken the expectation value, $\langle\mathbf{S}\rangle$ to point in the positive x direction, so $\theta = \frac{\pi}{2} + \delta\theta$ and $\phi = \delta\phi$. The constant term, independent of $\delta\theta$ and $\delta\phi$, in the final line of eqs. (3.2.16) gives the interaction energy between the magnetic field and the expectation value, $\langle\mathbf{S}\rangle$. This permits the physical interpretation of the constant μ as the magnetic-moment density of the material.

This interaction between $\vec{s}$ and $\mathbf{B}$ breaks T invariance, and has the following puzzling feature. It does *not* involve any derivatives of the Goldstone boson fields, θ and ϕ, in apparent contradiction with the general results of Chapter 1. In fact, the absence of

derivatives in eq. (3.2.16) is very much like the absence of derivatives in the pion mass term. This is because the coupling between $\vec{s}$ and $\mathbf{B}$ relates the internal spin $SO(3)$ symmetry to ordinary rotations in space, and so destroys the freedom to consider both as separate symmetries. But because rotation invariance is a spacetime symmetry, the derivations of Chapter 1 do not directly apply, since these assumed the action of internal symmetries from the outset in the transformation rules of the fields.

Once more taking the neutron coupling to the magnetic field as in eq. (3.1.11), we may compute the cross section for inelastic neutron scattering. For slowly-moving neutrons, and under the assumption that only the momentum, $\mathbf{p}'$, of the scattered neutron is measured we find:

$$\frac{d\sigma}{V d^2\mathbf{p}'} = \frac{\mu_N^2 \mu^2}{4\pi^3 v_N} V_{ij}(\mathbf{p}-\mathbf{p}')\; S_{ij}(E-E', \mathbf{p}-\mathbf{p}'). \tag{3.2.17}$$

The variables are the same as for the antiferromagnetic example: v_N and μ_N are the speed and magnetic moment of the slow incoming neutron, V is the volume of the medium, E (E') is the energy of the initial (final) neutron, and $\mathbf{p}$ ($\mathbf{p}'$) are the corresponding momenta. The magnetic-moment interaction, $V_{ij}(\mathbf{q})$, is as given in eq. (3.1.13).

The medium-dependent quantity, $S_{ij}(\omega, \mathbf{q})$, once more represents the spin correlation function, defined by eq. (3.1.14). As was the case for the ferromagnet this is dominated by sharp peaks when the neutron scatters to produce a magnon, and so has an energy and momentum transfer related by the magnon dispersion relation. For a ferromagnet this is: $E - E' = k(\mathbf{p}-\mathbf{p}')^2$. Measurements of these peaks as functions of the scattered neutron energy and momentum indeed verifies the quadratic dispersion relation, and can be used to measure the constant k.

A second consequence of the quadratic magnon dispersion relation is the temperature

dependence of the magnetization, $M = |\mathbf{M}|$, of a sample at very low temperatures. Since the magnon field describes the long-wavelength deviations of the net magnetization from its ground state value, the net magnetization at very low temperatures is simply proportional to the average magnon occupation number. That is:

$$M(0) - M(T) \propto M(0) \int d^3\mathbf{p}\, n\left(\frac{E}{T}\right), \tag{3.2.18}$$

where $n(E/T) = (\exp[E/T] - 1)^{-1}$ is the Bose-Einstein distribution. The temperature dependence of this result can be determined by changing integration variables from p to the dimensionless quantity $x = E/T$. If $E(p) \propto p^z$, for some power, z, then:

$$p^2\, dp = p^2 \frac{dp}{dE}\, dE \propto E^{2/z}\; E^{-(z-1)/z}\; dE \propto T^{3/z}. \tag{3.2.19}$$

For $z = 2$ this predicts $[M(0) - M(T)]/M(0) \propto T^{3/2}$, in agreement with low-temperature observations.

Chapter 4

$SO(5)$-Invariance and Superconductors

The techniques described herein have recently proven useful to analyze the consequences of a remarkable proposal for the existence of an $SO(5)$ invariance amongst the cuprates which exhibit high-temperature superconductivity. This chapter presents a bare-bones outline of this proposal, together with a brief summary of the Goldstone-boson properties which emerge.

4.1 $SO(5)$ Symmetry

Although a proper presentation of the arguments for — and against, since the subject remains controversial — the $SO(5)$ proposal is beyond the scope of this review, the form of the proposed symmetry itself is easy to state. The starting point is the following experimental fact: by performing small adjustments to any of the high-T_c cuprates, it is possible to convert them from superconductors into antiferromagnets. This adjustment is typically accomplished in practice by altering the 'doping', which means that atoms having different valences are randomly substituted into a portion of the unit cells of the material of interest For example the element Sr might be substituted for the element La in

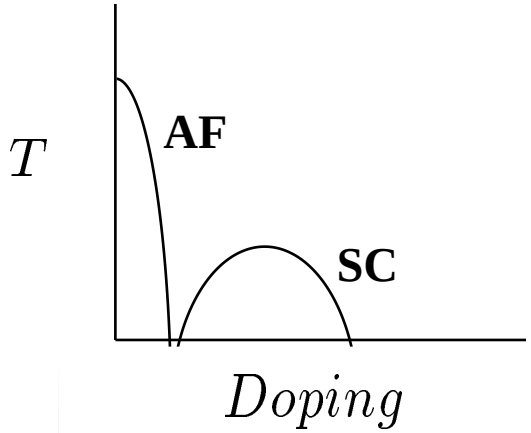

Figure 4.1: A typical Temperature–*vs*–Doping phase diagram for a high-T_c system.

some fraction, x, of the unit cells. Physically, this substitution has the effect of changing the number of charge carriers in the band from which the superconducting electrons are taken.

This basic association of superconductivity and antiferromagnetism suggests a fundamental connection between the two. Zhang's proposal is that — at least over part of the theory's parameter space — these two phases are related by an approximate $SO(5)$ symmetry. The action of the symmetry is simplest to state for the order parameters for the two phases.

As we have seen, in §3, the order parameter for the antiferromagnetic (AF) phase is simply the direction in space, $\vec{n}$, into which the alternating aligned spins point. A nonzero value of this order parameter spontaneously breaks the $SO(3) \simeq SU(2)$ symmetry of rotations amongst electron spins:

$$\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \to \mathcal{O}_3 \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}, \tag{4.1.1}$$

where $\mathcal{O}_n$ denotes an n-by-n real, orthogonal matrix.

For the superconducting (SC) phase, on the other hand, the order parameter, ψ, is a

quantity which carries the quantum numbers of a pair of electrons (or holes). Specifically, it has electric charge $q = \pm 2e$, if e is the proton charge, and is usually taken to have no spin. (The sign of this charge depends on whether the charge carriers are electrons or holes.) A nonzero value for ψ signals the breaking of the symmetry of electromagnetic phase rotations: $\psi \to e^{iq\omega}\,\psi$. This symmetry forms the group $SO(2) \simeq U(1)$, as is easily seen by grouping the real and imaginary parts of ψ into a two-component vector:

$$\begin{pmatrix} \psi_R \\ \psi_I \end{pmatrix} \to \mathcal{O}_2 \begin{pmatrix} \psi_R \\ \psi_I \end{pmatrix}, \tag{4.1.2}$$

where $\psi = \psi_R + i\,\psi_I$.

The $SO(5)$ proposal is that the system is approximately invariant under five-by-five rotations of all five components of the order parameter,

$$\begin{pmatrix} \psi_R \\ \psi_I \\ n_1 \\ n_2 \\ n_3 \end{pmatrix} \to \mathcal{O}_5 \begin{pmatrix} \psi_R \\ \psi_I \\ n_1 \\ n_2 \\ n_3 \end{pmatrix}. \tag{4.1.3}$$

This symmetry contains electron-spin rotations, eq. (4.1.1), and electromagnetic phase transformations, eq. (4.1.2), as the block-diagonal $SO(2) \times SO(3)$ subgroup which acts separately on the ψ_k and the n_a.

As of this writing, there is a controversy over whether such an approximate symmetry actually exists for high-T_c superconductors and, if so, to how large a portion of the phase diagram it might apply. Regardless of how this controversy ultimately becomes resolved, two points on which everyone must agree are:

1. For any part of the phase diagram for which approximate $SO(5)$ symmetry holds, and which lies within the ordered (AF or SC) phases, there must be a total of four Goldstone (or pseudo-Goldstone) bosons. This corresponds to one each for the four

$SO(5)$ generators which are broken by a nonzero value for one of the ψ_k or n_a. These bosons include the usual Goldstone bosons for $SO(3)$ or $SO(2)$ invariance (*e.g.*, the magnons), plus some new pseudo-Goldstone bosons which are consequences only of the assumed $SO(5)$ symmetry.

2. The low-energy properties of these Goldstone and pseudo-Goldstone bosons are completely dictated by the assumed symmetry-breaking pattern, and are independent of the details of whatever microscopic electron dynamics gives rise to the symmetry in the first place. These low-energy properties may be efficiently described using an effective lagrangian along the lines of those described in Chapter 1.

These two properties taken together make an unambiguous detection of the $SO(5)$ pseudo-Goldstone bosons a particularly attractive test of the $SO(5)$ proposal. Their detection would be a 'smoking gun' for the existence of an extended symmetry like $SO(5)$. Better yet, their properties are unambiguously predicted theoretically, without the the usual complications which arise when complicated electron dynamics is squeezed into a simple theoretical model. We now construct the low-energy effective lagrangian describing these pseudo-Goldstone bosons, following the general techniques of the previous sections.

4.2 The Effective Lagrangian in the Symmetry Limit

We start with the effective lagrangian in the (idealized) limit where $SO(5)$ is not just approximate, but is instead a *bona fide* symmetry of the system. In this case the lagrangian symmetry is $G = SO(5)$, and this is spontaneously broken (by the order parameters we are considering) to the subgroup, $H = SO(4)$. We require in this limit the nonlinear sigma model for the quotient space $G/H = SO(5)/SO(4)$. As discussed in general in §1, the

lowest terms in the derivative expansion of the Lagrangian for this system are therefore completely determined up to a small number of constants. (Precisely how many constants depends on how much symmetry – including crystallographic symmetries – the system has. As we describe shortly, more possibilities also arise once explicit $SO(5)$-breaking interactions are introduced.)

In this particular case the space $SO(5)/SO(4)$ is an old friend: the four-sphere, S_4, which is defined as the points swept out by an arbitrary five-dimensional vector, $\vec{N}$, which has unit length: $\vec{N} \cdot \vec{N} = \sum_{i=1}^{5} N_i^2 \equiv 1$. In terms of such a field, the invariant lagrangian obtained using the techniques of previous sections is

$$\mathcal{L}_{\rm inv} = \frac{f_t^2}{2}\, \partial_t \vec{N} \cdot \partial_t \vec{N} - \frac{f_s^2}{2}\, \nabla \vec{N} \cdot \nabla \vec{N} \tag{4.2.1}$$

(For simplicity of this and later expressions, eq. (4.2.1) assumes rotational invariance, which is *not* appropriate for real cuprates. For real systems, the electrons believed responsible for superconductivity and antiferromagnetism move preferentially along planes made up of copper and oxygen atoms. The low-energy lagrangian for such systems is better written either in two space dimensions (for Goldstone bosons confined to the planes), or in three dimensions with separate coefficients, f_a for each spatial direction, ∇_a. These complications are ignored here, but are discussed in more detail in the original references. As is also discussed in these referencse, in specific dimensions it is sometimes also possible to write more invariants than are considered here, such as those which depend on the completely antisymmetric tensor, $\epsilon_{ijk..}$.)

A convenient parameterization for the four-sphere, and so of our Goldstone bosons, is given by polar coordinates:

$$\vec{N} = \begin{pmatrix} n_Q \\ n_S \end{pmatrix}, \qquad \text{where} \qquad n_Q = \cos\theta \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}, \qquad n_S = \sin\theta \begin{pmatrix} \sin\alpha\cos\beta \\ \sin\alpha\sin\beta \\ \cos\alpha \end{pmatrix}. \tag{4.2.2}$$

(As usual, care is required to properly handle those points where these coordinates are singular).

Using the standard expression for the round metric on S_4 then gives the unique Goldstone-boson effective lagrangian for a rotational- and time-reversal-invariant system in the $SO(5)$ symmetry limit:

$$\begin{aligned}\mathcal{L} &= \frac{f_t^2}{2}\Big[(\partial_t\theta)^2 + \cos^2\theta(\partial_t\phi)^2 + \sin^2\theta\Big((\partial_t\alpha)^2 + \sin^2\alpha(\partial_t\beta)^2\Big)\Big] \\ &\qquad - \frac{f_s^2}{2}\left[(\nabla\theta)^2 + \cos^2\theta(\nabla\phi)^2 + \sin^2\theta\Big((\nabla\alpha)^2 + \sin^2\alpha(\nabla\beta)^2\Big)\right].\end{aligned}$$

4.3 Symmetry-Breaking Terms

Next consider how small $SO(5)$-breaking effects can change the low-energy lagrangian. We may do so by following the same steps as were taken in §2 to describe the implications of quark masses on low-energy pion properties. We do so here in two steps. We first classify the types of violation of $SO(5)$ symmetry which can arise in real systems. We then examine which consequences follow only from $SO(2) \times SO(3)$ invariance, in order to be able to disentangle these from predictions which are specific to $SO(5)$. Finally we perturb in the various $SO(5)$ symmetry-breaking parameters to obtain the predictions of approximate $SO(5)$ invariance. By contrasting what is obtained with the result assuming only $SO(2) \times SO(3)$ invariance, the consequences of approximate $SO(5)$ invariance may be found.

4.3.1 Kinds of Explicit Symmetry Breaking (Qualitative)

There are several kinds of $SO(5)$ symmetry breaking which are worth distinguishing from one another. These are:

1. *Electromagnetic Interactions:* One source of explicit $SO(5)$ symmetry breaking is any couplings to applied macroscopic electromagnetic fields. Electromagnetic gauge invariance requires these to be incorporated into the lagrangian using the usual precedure of minimal substitution:

$$\partial_t \psi \to (\partial_t - i\, q\, A_0)\psi, \qquad \nabla\psi \to (\nabla - i\, q\, \mathbf{A})\psi. \tag{4.3.1}$$

   Such couplings necessarily break $SO(5)$ because they treat the electrically-charged components of $\vec{N}$ differently than the electrically-neutral ones. Although such couplings do not pose any complications of principle, for simplicity's sake we imagine no macroscopic electromagnetic fields to be applied in what follows.

2. *Doping:* One of the physical variables on which the phase diagram of the cuprates crucially depends is the doping. Since changes in the doping correspond to changes in the density of charge carriers amongst the electrons which are relevant for both the antiferromagnetism and the superconductivity, it may be described within the effective theory by using a chemical potential, μ, coupled to electric charge. In this way adjustments in μ may be chosen to ensure that the system has any given experimental charge density. Mathematically a chemical potential is introduced by replacing the system Hamiltonian, H, with the quantity $H - \mu Q$, where Q is the electric charge. Within the Lagrangian formulation in which we are working, this amounts to simply making the replacement $A_0 \to A_0 + \mu$ in the electrostatic scalar potential, A_0.

3. *Intrinsic Breaking:* The third, and final, category of symmetry-breaking consists of everything apart from the previous two. It is known that $SO(5)$ is not an exact

symmetry, even with no chemical potential, and in the absence of any applied electromagnetic fields. It is the interactions (involving the fewest derivatives) in this last class of symmetry-breaking terms which we now wish to classify.

4.3.2 General $SO(2) \times SO(3)$-Invariant Interactions

If $SO(5)$ were not a symmetry at all, then there would be no guarantee that the low-energy spectrum should contain particles described by all of the fields α, β, θ and ϕ. For the purposes of later comparison, it is nevertheless useful to ask what kinds of low-energy interactions among such states are permitted by $SO(2) \times SO(3)$ invariance.

The most general such Lagrangian involving these four states may be written in terms of the fields n_S and n_Q, where these fields satisfy the constraint $n_S \cdot n_S + n_Q \cdot n_Q \equiv 1$, to the extent that we are interested in only those modes which would be Goldstone or pseudo-Goldstone modes in the $SO(5)$ limit. The most general such result, which involves at most two derivatives, supplements the invariant expression, eqs. (4.2.1) or (4.2.3), with the following terms:

$$\begin{aligned} \mathcal{L}_{\rm sb} &= -V + f_t^2 \Big[A\, \partial_t n_Q \cdot \partial_t n_Q + B\, \partial_t n_S \cdot \partial_t n_S + C\, (n_Q \cdot \partial_t n_Q)^2 \Big] \\ &\qquad - f_s^2 \Big[D \nabla_a n_Q \cdot \nabla_a n_Q + E \nabla_a n_S \cdot \nabla_a n_S + F (n_Q \cdot \nabla_a n_Q)^2 \Big] \end{aligned}$$

where f_t and f_s are the constants appearing in the invariant lagrangian. The quantities V, A, B, C, D, E and F are potentially arbitrary functions of the unique $SO(2) \times SO(3)$ invariant which involves no derivatives: $n_Q \cdot n_Q$. (Recall $n_S \cdot n_S$ is not independent due to the constraint $n_S \cdot n_S + n_Q \cdot n_Q = 1$).

In terms of polar coordinates, and inserting a chemical potential as just described,

the total effective lagrangian becomes:

$$\begin{aligned}\mathcal{L} \;=\; & -V+\frac{f_t^2}{2}\Big[\Big(1+2A\sin^2\theta+2B\cos^2\theta+2C\sin^2\theta\cos^2\theta\Big)(\partial_t\theta)^2 \\ & +(1+2A)\cos^2\theta(\partial_t\phi+q\mu)^2+(1+2B)\sin^2\theta\Big((\partial_t\alpha)^2+\sin^2\alpha(\partial_t\beta)^2\Big)\Big] \\ & -\frac{f_s^2}{2}\Big[\Big(1+2D\sin^2\theta+2E\cos^2\theta+2F\sin^2\theta\cos^2\theta\Big)(\nabla\theta)^2 \\ & +(1+2D)\cos^2\theta(\nabla\phi)^2+(1+2E)\sin^2\theta\Big((\nabla\alpha)^2+\sin^2\alpha(\nabla\beta)^2\Big)\Big]+\cdots,\end{aligned}$$

where all coefficient functions, V, A,...*etc.*, are now to be regarded as functions of $\cos^2\theta$.

4.3.3 Kinds of Explicit Symmetry Breaking (Quantitative)

Eq. (4.3.2) does not yet use any information concerning the nature or size of the explicit symmetry breaking (apart from the inclusion of μ). This we must now do if we are to quantify the predictions of approximate $SO(5)$ invariance. We do so by making an assumption as to how the symmetry-breaking terms transform under $SO(5)$.

In §2 we saw, for pions, that the quark masses were responsible for explicitly breaking the would-be chiral symmetry of the underlying microscopic theory (QCD). Although the same reasoning can be applied to $SO(5)$ breaking due to electromagnetic interactions and chemical potential dependence, incomplete understanding of the dynamics of the microscopic theory so far precludes a similar identification of the other symmetry-breaking parameters within some underlying condensed-matter system. For these, we instead are forced to make an assumption.

We therefore assume all $SO(5)$-breaking terms of the effective lagrangian to be proportional to one of two possible quantities:

1. *Chemical Potential:* Since we know how the chemical potential appears in the lagrangian, we know in detail how it breaks $SO(5)$. It does so by an amount which is

proportional to the electric charge. For the fields appearing in $\vec{N}$, this is represented by the five-by-five electric charge matrix: $Q = \mathrm{diag}(q, q, 0, 0, 0)$.

2. *Intrinsic Symmetry Breaking:* In the absence of more information, we make the simplest assumption for the form taken in the effective lagrangian by all other microscopic effects which explicitly break $SO(5)$. Since these break $SO(5)$ to $SO(3) \times SO(2)$ we take them to be proportional to a five-by-five matrix, M, where $M = \epsilon\,\mathrm{diag}\,(3, 3, -2, -2, -2)$. Here $\epsilon \ll 1$ is a measure of the quality of the approximation that $SO(5)$ is a symmetry.

With these choices the Lagrangian is then the most general function of the fields $\vec{N} = \binom{n_Q}{n_S}$, μQ and M, subject to the following $SO(5)$ transformation property

$$\mathcal{L}(\mathcal{O}_5\vec{N}, \mathcal{O}_5\mu Q\mathcal{O}_5^T, \mathcal{O}_5 M\mathcal{O}_5^T) = \mathcal{L}(\vec{N}, \mu Q, M), \tag{4.3.2}$$

where $\mathcal{O}_5$ is an $SO(5)$ transformation. The implications of the approximate $SO(5)$ invariance may then be extracted by expanding $\mathcal{L}$ in powers of the small quantities ϵ and μ. Since M and Q always appear premultiplied by these small numbers, this expansion restricts the kinds of symmetry breaking which can arise order by order, which in turn constrains the possible θ-dependence of the coefficient functions in $\mathcal{L}$.

For example, a term in the scalar potential involving $2n$ powers of $\vec{N}$ must have the following form:

$$V_{(n)} = \sum_{(k_1,l_1)\neq(0,0)} \cdots \sum_{(k_n,l_n)\neq(0,0)} C_{k_1 l_1,\ldots,k_n,l_n}\left[\vec{N}\cdot(\epsilon M)^{k_1}(\mu Q)^{2l_1}\vec{N}\right]\cdots\left[\vec{N}\cdot(\epsilon M)^{k_n}(\mu Q)^{2l_n}\vec{N}\right]. \tag{4.3.3}$$

Only even powers of Q enter here due to its antisymmetry, and the term $k_i = l_i = 0$ is excluded from the sums due to the constraint $\vec{N}^T\vec{N} = 1$. Clearly, expanding $\mathcal{L}$ to low

order in the $SO(5)$-breaking parameters ϵ and μ necessarily also implies keeping only the lowest powers of $n_Q \cdot n_Q = \cos^2\theta$ in V.

Similar conclusions may be obtained for the other coefficient functions in the Lagrangian of eq. (4.3.2). Working to $O(\epsilon^2, \epsilon\mu^2, \mu^4)$ in V, and to $O(\epsilon, \mu^2)$ in the two-derivative terms then gives:

$$V = V_0 + V_2 \cos^2\theta + \frac{1}{2} V_4 \cos^4\theta, \tag{4.3.4}$$

and

$$A = A_0 + A_2 \cos^2\theta, \qquad B = B_0 + A_2 \cos^2\theta, \qquad C = C_0,$$
$$D = D_0 + D_2 \cos^2\theta, \qquad E = E_0 + D_2 \cos^2\theta, \qquad F = F_0,$$

for the coefficient functions in eq. (4.3.2). Notice that the terms proportional to $\cos^2\theta$ in A and B are identical, as are the corresponding terms in D and E. Expanding in powers of ϵ and μ, the constants in eqs. (4.3.4) and (4.3.5) start off linear in ϵ and μ^2: $A_i = A_i^{10}\epsilon + A_i^{01}\mu^2 + \cdots$ *etc.*. The only exceptions to this statement are: $B_0, E_0 \propto \epsilon$ (no μ^2 term), $C_0, F_0 \propto \mu^2$ (no ϵ term), and $V_4 = V_4^{20}\epsilon^2 + V_4^{11}\epsilon\mu^2 + V_4^{02}\mu^4$. Furthermore, since the $\mu^2\ n_Q \cdot n_Q$ term in V arises from substituting $\partial_t \to \partial_t - i\mu Q$ in the kinetic term for n_Q, we have: $V_2^{01} = -\frac{1}{2} f_t^2 q^2$ to leading order. Higher powers of μ originate from terms in $\mathcal{L}$ which involve more than two derivatives.

4.4 Pseudo-Goldstone Dispersion Relations

We now turn to the calculation of the pseudo-Goldstone boson dispersion relations. The scalar potential of eq. (4.3.2) has three types of extrema:

$$(1) \qquad \theta_0 = 0 \quad \text{or} \quad \pi;$$

$$
\begin{aligned}
&(2) \qquad \theta_0 = \frac{\pi}{2} \quad \text{or} \quad \frac{3\pi}{2}; \\
&(3) \qquad \theta_0 \quad \text{where } c = \cos\theta_0 \text{ satisfies } V'(c^2) = 0.
\end{aligned} \tag{4.4.1}
$$

This leads to the four possible phases: (*i*) SC phase: extremum (1) is a minimum, and (2) is a maximum; (*ii*) AF phase: (2) is a minimum, and (1) is a maximum; (*iii*) MX phase: both (1) and (2) are maxima, and (3) is a minimum; or (*iv*) metastable phase: both (1) and (2) are minima, and (3) is a maximum. We focus here purely on the AF and SC phases.

4.4.1 Superconducting Phase

An expansion about the superconducting mimimum, $\theta_0 = 0$, gives the dispersion relations in this phase for the four bosons. Three of these — θ, α and β — form a spin triplet of pseudo-Goldstone modes for which

$$
E(k) = \Big[c^2\, k^2 + \mathcal{E}^2\Big]^{\frac{1}{2}}, \tag{4.4.2}
$$

with the phase speed, $c_\alpha^2(SC)$, and gap, $\mathcal{E}_{SC}^2$, given to lowest order in $SO(5)$-breaking parameters by:

$$
\begin{aligned}
c_\alpha^2(SC) &= \frac{f_s^2}{f_t^2}\Big[1 + 2\Big(E(1) - B(1)\Big)\Big] \\
&= \frac{f_s^2}{f_t^2}\Big[1 + 2\Big(E_0 - B_0\Big) + 2\Big(D_2 - A_2\Big)\Big], \\
\mathcal{E}_{SC}^2 &= \frac{-2V'(1)}{f_t^2} \\
&= \frac{-2(V_2 + V_4)}{f_t^2}.
\end{aligned} \tag{4.4.3}
$$

In both of these results the first equation uses the general effective theory, eq. (4.3.2), while the second equality incorporates the additional information of eqs. (4.3.4) and (4.3.5).

An important part of the $SO(5)$ proposal is that these states have been seen in neutron-scattering experiments in the superconducting phase of the high-T_c cuprates, even quite far away from the antiferromagnetic regime.

The remaining field, ϕ, would have been a *bona fide* gapless Goldstone mode in the absence of electromagnetic interactions. Its dispersion relation, $E(k)$ is a more complicated function of c^2k^2 and $eq\mu$, whose form is not required here. The quantity c which appears with k throughout its dispersion relation is given explicitly by

$$\begin{aligned} c_\phi^2(SC) &= \frac{f_s^2}{f_t^2}\Big[1+2\Big(D(1)-A(1)\Big)\Big] \\ &= \frac{f_s^2}{f_t^2}\Big[1+2\Big(D_0-A_0\Big)+2\Big(D_2-A_2\Big)\Big]. \end{aligned} \tag{4.4.4}$$

4.4.2 Antiferromagnetic Phase

Expanding about the AF minimum gives the usual two magnons, as in §3, satisfying dispersion relation of eq. (4.4.2) with:

$$\begin{aligned} c_{GB}^2(AF) &= \frac{f_s^2}{f_t^2}\Big[1+2\Big(E(0)-B(0)\Big)\Big] \\ &= \frac{f_s^2}{f_t^2}\Big[1+2\Big(E_0-B_0\Big)\Big], \\ \mathcal{E}_{GB}^2(AF) &= 0. \end{aligned} \tag{4.4.5}$$

The remaining two states group into an electrically-charged pseudo-Goldstone state satisfying:

$$E_\pm(k) = \Big[c^2\,k^2+\mathcal{E}^2\Big]^{\frac{1}{2}} \pm q\mu, \tag{4.4.6}$$

with:

$$c_{pGB}^2(AF) = \frac{f_s^2}{f_t^2}\Big[1+2\Big(D(0)-A(0)\Big)\Big]$$

$$
\begin{aligned}
&= \frac{f_s^2}{f_t^2}\Big[1+2\Big(D_0-A_0\Big)\Big],\\
\mathcal{E}^2_{AF}=\mathcal{E}^2_{pGB}(AF) &= \frac{2V'(0)}{f_t^2}\\
&= \frac{2V_2}{f_t^2}.
\end{aligned}
\tag{4.4.7}
$$

These expressions imply a simple dependence of the gap on the chemical potential:

$$
\begin{aligned}
\mathcal{E}^2_{AF} &= m^2-\kappa\mu^2,\\
\mathcal{E}^2_{SC} &= -m^2+\kappa\mu^2-\xi\mu^4,
\end{aligned}
\tag{4.4.8}
$$

where $m^2 := 2V_2^{10}\epsilon/f_t^2+O(\epsilon^2)$, $\kappa := -2V_2^{01}/f_t^2+O(\epsilon) = q^2+O(\epsilon)$ and $\xi := 2V_4^{02}/f_t^2+O(\epsilon)$. Within the AF phase the pseudo-Goldstone boson gap is predicted to fall linearly with μ^2:

$$
\mathcal{E}^2_{AF} \approx \mathcal{E}^2_{AF}(0)[\mu^2_{AF}-\mu^2],
\tag{4.4.9}
$$

where μ_{AF} represents the doping for which one leaves the AF regime. Similarly $\mathcal{E}^2_{SC}$ varies quadratically with μ^2.

Robust consequences of $SO(5)$ invariance are obtained from expressions such as these by eliminating the free parameters to obtain relations amongst observables. For example, if one eliminates parameters in favour of properties of the gap as a function of μ we find:

$$
\begin{aligned}
\varepsilon^2_{AF}(\mu) &= \frac{\varepsilon^2_{AF}(0)}{\mu^2_{AF}}\Big[\mu^2_{AF}-\mu^2\Big],\\
\varepsilon^2_{SC}(\mu) &= \frac{\varepsilon^2_{SC}(\mathrm{opt})}{\mu^4_{\mathrm{opt}}}(\mu^2-\mu^2_{SC-})(2\mu^2_{\mathrm{opt}}-\mu^2),\\
\frac{\varepsilon^2_{AF}(0)}{\mu^2_{AF}} &= 2\,\frac{\varepsilon^2_{SC}(\mathrm{opt})}{\mu^2_{\mathrm{opt}}},\\
\mu^2_{AF} &= \mu^2_{SC-}+O(\epsilon^2).
\end{aligned}
\tag{4.4.10}
$$

Here μ_{opt} denotes the chemical potential corresponding to the maximum gap, $\mathcal{E}_{SC}$.

Similarly, the phase velocities for all modes in both SC and AF phases are equal to one another, and to f_t^2/f_s^2, in the limit of strict $SO(5)$ invariance. (The parameters f_t and f_s may be related to other observables, such as the electric and magnetic screening lengths.) It turns out that the $O(\epsilon)$ corrections to this limit are not arbitrary, but also satisfy some model-independent relations, which follow by eliminating parameters from the above expressions:

$$c_\phi^2(SC) - c_\phi^2(AF) = c_\alpha^2(SC) - c_\alpha^2(AF) = O(\epsilon). \tag{4.4.11}$$

4.5 Summary

Approximate $SO(5)$ invariance clearly carries real implications for the low-energy excitations of the system. It predicts, in particular, the existence of a spin-triplet pseudo-Goldstone state in the SC phase, and an electrically-charged state in the AF phase. Furthermore, $SO(5)$ invariance unambiguously relates the properties of these states, like their gap and phase velocity, to one another. Better yet, these properties are claimed to have been measured in the SC phase, since the spin-triplet state is believed to have been observed, with a gap (at optimal doping) of 41 meV. If true, this permits the inference of the size of the rough order of the $SO(5)$-breaking parameter ϵ, and hence to predictions for the properties of the hitherto undetected boson in the AF phase.

It is extremely unlikely that an electrically-charged state having a gap of only ~ 40 meV can exist deep within the AF phase. Among other things it would make its presence felt through the electromagnetic response of these systems in the AF phase. At the very least, one can therefore conclude that $SO(5)$ invariance cannot penetrate very far into the AF part of the phase diagram, despite its appearance fairly deep in the SC phase. Being based purely on pseudo-Goldstone boson properties, this conclusion comes independent

of the details of how the underlying electrons are interacting on more microscopic scales.

There is not yet a consensus as to how uncomfortable this conclusion should make one feel about the remarkable $SO(5)$ hypothesis. Either way, robust predictions based on the low-energy consequences of symmetries are likely to play a key role in forming any such final consensus.