INTRODUCTION TO EFFECTIVE FIELD THEORY

C.P. BURGESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9780521195478
DOI: 10.1017/9781139048040

First published 2021

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalogue record for this publication is available from the British Library.

ISBN 978-0-521-19547-8 Hardback

Contents

Tables

by aficianados in a hurry. The symbol ♠ indicates a section which may require a bit more digging for new students to digest, but which is reasonably self-contained and worth a bit of spadework. Finally, readers wishing to beat their heads against sections containing more challenging topics should seek out those marked with ♣.

The lion's share of the book is aimed at applications, since this most effectively brings out both the utility and the unity of the approach. The examples also provide a pedagogical framework for introducing some specific techniques. Since many of these applications are independent of one another, a course can be built by starting with Part I's introductory material and picking and choosing amongst the later sections that are of most interest.

Acknowledgements

This book draws heavily on the insight and goodwill of many people: in particular my teachers of quantum and classical field theory – Bryce De Witt, Willy Fischler, Joe Polchinski and especially Steven Weinberg – who shaped the way I think about this subject.

Special thanks go to Fernando Quevedo for a life-long collaboration on these subjects and his comments over the years on many of the topics discussed herein.

I owe a debt to Patrick Labelle, Sung-Sik Lee, Alexander Penin and Ira Rothstein for clarifying issues to do with nonrelativistic EFTs; to John Donoghue for many insights on gravitational physics; to Thomas Becher for catching errors in early versions of the text; to Jim Cline for a better understanding of the practical implications of Goldstone boson infrared effects; to Claudia de Rham, Luis Lehner, Adam Solomon, Andrew Tolley and Mark Trodden for helping better understand applications to time-dependent systems; to Subodh Patil and Michael Horbatsch for helping unravel multiple scales in scalar cosmology; to Michele Cicoli, Shanta de Alwis, Sven Krippendorf and Anshuman Maharana for shepherding me through the perils of string theory; to Mike Trott for help understanding the subtleties of power-counting and SMEFT; to Peter Adshead, Richard Holman, Greg Kaplanek, Louis Leblond, Jerome Martin, Sarah Shandera, Gianmassimo Tasinato, Vincent Vennin and Richard Woodard for understanding EFTs in de Sitter space and their relation to open systems, and to Ross Diener, Peter Hayman, Doug Hoover, Leo van Nierop, Duncan Odell, Ryan Plestid, Markus Rummel, Matt Williams and Laszlo Zalavari for helping clarify how EFTs work for massive first-quantized sources.

Collaborators and students too numerous to name have continued to help deepen my understanding in the course of many conversations about physics.

CERN, ICTP, KITP Santa Barbara and the Institute Henri Poincaré have at various times provided me with pleasant environs in which to focus undivided time on writing, and with stimulating discussions when taking a break from it. The book would not have been finished without them. The same is true of McMaster University and Perimeter Institute, whose flexible work environments allowed me to take on this project in the first place.

Heaven holds a special place for Simon Capelin and his fellow editors, both for encouraging the development of this book and for their enormous patience in awaiting it.

Most importantly, I am grateful to my late parents for their gift of an early interest in science, and to my immediate family (Caroline, Andrew, Ian, Matthew and Michael) for their continuing support and tolerance of time taken from them for physics.

Part I

Theoretical Framework


image
not
available

1 Decoupling and Hierarchies of Scale

The world around us contains a cornucopia of length scales, ranging (at the time of writing) down to quarks and leptons at the smallest and up to the universe as a whole at the largest, with qualitatively new kinds of structures – nuclei, atoms, molecules, cells, organisms, mountains, asteroids, planets, stars, galaxies, voids, and so on – seemingly arising at every few decades of scales in between. So it is remarkable that all of this diversity seems to be described in all of its complexity by a few simple laws.

How can this be possible? Even given that the simple laws exist, why should it be possible to winkle out an understanding of what goes on at one scale without having to understand everything all at once? The answer seems to be a very deep property of nature called *decoupling*, which states that most (but not all) of the details of very small-distance phenomena tend to be largely irrelevant for the description of much larger systems. For example, not much needs to be known about the detailed properties of nuclei (apart from their mass and electrical charge, and perhaps a few of their multipole moments) in order to understand in detail the properties of electronic energy levels in atoms.

Decoupling is a very good thing, since it means that the onion of knowledge can be peeled one layer at a time: our initial ignorance of nuclei need not impede the unravelling of atomic physics, just as ignorance about atoms does not stop working out the laws describing the motion of larger things, like the behaviour of fluids or motion of the moon.

It so happens that this property of decoupling is also shared by the mathematics used to describe the laws of nature [1]. Since nowadays this description involves quantum field theories, it is gratifying that these theories as a group tend to predict that short distances generically decouple from long distances, in much the same way as happens in nature.

This book describes the way this happens in detail, with two main purposes in mind. One purpose is to display decoupling for its own sake since this is satisfying in its own right, and leads to deep insights into what precisely is being accomplished when writing down physical laws. But the second purpose is very practical; the simplicity offered by a timely exploitation of decoupling can often be the difference between being able to solve a problem or not. When exploring the consequences of a particular theory for short distance physics it is obviously useful to be able to identify efficiently those observables that are most sensitive to the theory's details and those from which they decouple. As a consequence the mathematical tools – *effective field theories* – for exploiting decoupling have become ubiquitous in some areas of theoretical physics, and are likely to become more common in many more.

The purpose of the rest of Chapter 1 is twofold. One goal is to sketch the broad outlines of decoupling, effective lagrangians and the physical reason why they work,

all in one place. The second aim is to provide a toy model that can be used as a concrete example as the formalism built on decoupling is fleshed out in more detail in subsequent chapters.

1.1 An Illustrative Toy Model ◇

The first step is to set up a simple concrete model to illustrate the main ideas. To be of interest this model must possess two kinds of particles, one of which is much heavier than the other, and these particles must interact in a simple yet nontrivial way. Our focus is on the interactions of the two particles, with a view towards showing precisely how the heavy particle decouples from the interactions of the light particle at low energies.

To this end consider a complex scalar field, ϕ, with action[1]

$$S := -\int \mathrm{d}^4x \left[\partial_\mu \phi^* \partial^\mu \phi + V(\phi^*\phi)\right], \tag{1.1}$$

whose self-interactions are described by a simple quartic potential,

$$V(\phi^*\phi) = \frac{\lambda}{4}\left(\phi^*\phi - v^2\right)^2, \tag{1.2}$$

where λ and v^2 are positive real constants. The shape of this potential is shown in Fig. 1.1.

1.1.1 Semiclassical Spectrum

The simplest regime in which to explore the model's predictions is when $\lambda \ll 1$ and both v and $|\phi|$ are $O\left(\lambda^{-1/2}\right)$. This regime is simple because it is one for which the semiclassical approximation provides an accurate description. (The relevance of the semiclassical limit in this regime can be seen by writing $\phi := \varphi/\lambda^{1/2}$ and $v := \mu/\lambda^{1/2}$ with φ and μ held fixed as $\lambda \to 0$. In this case the action depends on λ only through an overall factor: $S[\phi, v, \lambda] = (1/\lambda)S[\varphi, \mu]$. This is significant because the action appears in observables only in the combination $S/\hbar$, and so the small-λ limit is equivalent to the small-$\hbar$ (classical) limit.)[2]

In the classical limit the ground state of this system is the field configuration that minimizes the classical energy,

$$E = \int \mathrm{d}^3x \left[\partial_t \phi^* \partial_t \phi + \nabla\phi^* \cdot \nabla\phi + V(\phi^*\phi)\right]. \tag{1.3}$$

Since this is the sum of positive terms it is minimized by setting each to zero; the classical ground state is any constant configuration (so $\partial_t\phi = \nabla\phi = 0$), with $|\phi| = v$ (so $V = 0$).

[1] Although this book presupposes some familiarity with quantum field theory, see Appendix C for a compressed summary of some of the relevant ideas and notation used throughout. Unless specifically stated otherwise, units are adopted for which $\hbar = c = 1$, so that time ~ length and energy ~ mass ~ 1/length, as described in more detail in Appendix A.

[2] The connection between small coupling and the semi-classical limit is explored more fully once power-counting techniques are discussed in §3.

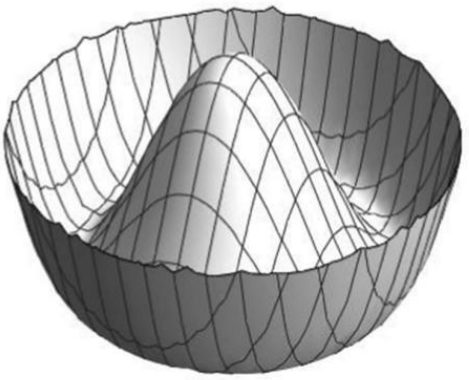

Fig. 1.1 The potential $V(\phi_R, \phi_I)$, showing its sombrero shape and the circular line of minima at $|\phi| = v$.

In the semi-classical regime, particle states are obtained by expanding the action about the classical vacuum, $\phi = v + \tilde{\phi}$,

$$S = -\int \mathrm{d}^4x \left\{ \partial_\mu \tilde{\phi}^* \partial^\mu \tilde{\phi} + \frac{\lambda}{4} \left[v(\tilde{\phi} + \tilde{\phi}^*) + \tilde{\phi}^* \tilde{\phi} \right]^2 \right\}, \tag{1.4}$$

and keeping the leading (quadratic) order in the quantum fluctuation $\tilde{\phi}$. In terms of the field's real and imaginary parts, $\tilde{\phi} = \frac{1}{\sqrt{2}}(\tilde{\phi}_R + \mathrm{i}\tilde{\phi}_I)$, the leading term in the expansion of S is

$$S_0 = -\frac{1}{2} \int \mathrm{d}^4x \left[\partial_\mu \tilde{\phi}_R \, \partial^\mu \tilde{\phi}_R + \partial_\mu \tilde{\phi}_I \, \partial^\mu \tilde{\phi}_I + \lambda v^2 \, \tilde{\phi}_R^2 \right]. \tag{1.5}$$

The standard form (see §C.3.1) for the action of a free, real scalar field of mass m is proportional to $\partial_\mu \psi \, \partial^\mu \psi + m^2 \psi^2$, and so comparing with Eq. (1.5) shows $\tilde{\phi}_R$ represents a particle with mass $m_R^2 = \lambda v^2$ while $\tilde{\phi}_I$ represents a particle with mass $m_I^2 = 0$. These are the heavy and light particles whose masses provide a hierarchy of scales.

1.1.2 Scattering

For small λ the interactions amongst these particles are well-described in perturbation theory, by writing $S = S_0 + S_{\text{int}}$ and perturbing in the interactions

$$S_{\text{int}} = -\int \mathrm{d}^4x \left[\frac{\lambda v}{2\sqrt{2}} \tilde{\phi}_R \left(\tilde{\phi}_R^2 + \tilde{\phi}_I^2 \right) + \frac{\lambda}{16} \left(\tilde{\phi}_R^2 + \tilde{\phi}_I^2 \right)^2 \right]. \tag{1.6}$$

Using this interaction, a straightforward calculation – for a summary of the steps involved see Appendix B – gives any desired scattering amplitude order-by-order in λ. Since small λ describes a semiclassical limit (because it appears systematically together with $\hbar$ in $S/\hbar$, as argued above), the leading contribution turns out to come from evaluating Feynman graphs with no loops[3] (*i.e.* tree graphs).

[3] A connected graph with no loops (or a 'tree' graph) is one which can be broken into two disconnected parts by cutting any internal line. Precisely how to count the number of loops and why this is related to powers of the small coupling λ is the topic of §3.

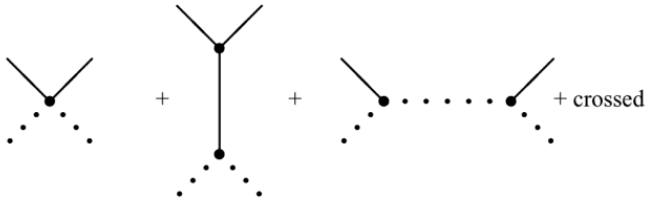

Fig. 1.2 The tree graphs that dominate $\tilde{\phi}_R\,\tilde{\phi}_I$ scattering. Solid (dotted) lines represent $\tilde{\phi}_R$ ($\tilde{\phi}_I$), and 'crossed' graphs are those with external lines interchanged relative to those displayed.

Consider the reaction $\tilde{\phi}_R(p) + \tilde{\phi}_I(q) \to \tilde{\phi}_R(p') + \tilde{\phi}_I(q')$, where $p^\mu = \{p^0, \mathbf{p}\}$ and $q^\mu = \{q^0, \mathbf{q}\}$ respectively denote the 4-momenta of the initial $\tilde{\phi}_R$ and $\tilde{\phi}_I$ particle, while p'^μ and q'^μ are 4-momenta of the final $\tilde{\phi}_R$ and $\tilde{\phi}_I$ states. The Feynman graphs of Fig. 1.2 give a scattering amplitude proportional to[4] $\mathcal{A}_{RI\to RI}\delta^4(p+q-p'-q')$, where the Dirac delta function, $\delta^4(p+q-p'-q')$, expresses energy–momentum conservation, and

$$\begin{aligned}\mathcal{A}_{RI\to RI} &= 4\mathrm{i}\left(-\frac{\lambda}{8}\right) + \left(\frac{\mathrm{i}^2}{2}\right)\left(-\frac{\lambda v}{2\sqrt{2}}\right)^2\left[\frac{24(-\mathrm{i})}{(p-p')^2+m_R^2} + \frac{8(-\mathrm{i})}{(p+q)^2} + \frac{8(-\mathrm{i})}{(p-q')^2}\right] \\ &= -\frac{\mathrm{i}\lambda}{2} + \frac{\mathrm{i}(\lambda v)^2}{2m_R^2}\left[\frac{3}{1-2q\cdot q'/m_R^2} - \frac{1}{1-2p\cdot q/m_R^2} - \frac{1}{1+2p\cdot q'/m_R^2}\right].\end{aligned} \tag{1.7}$$

Here the factors like 4, 24 and 8 in front of various terms count the combinatorics of how many ways each particular graph can contribute to the amplitude. The second line uses energy–momentum conservation, $(p-p')^\mu = (q'-q)^\mu$, as well as the kinematic conditions $p^2 = -(p^0)^2 + \mathbf{p}^2 = -m_R^2$ and $(q')^2 = q^2 = -(q^0)^2 + \mathbf{q}^2 = 0$, as appropriate for relativistic particles whose energy and momenta are related by $E = p^0 = \sqrt{\mathbf{p}^2 + m^2}$.

Notice that the terms involving the square bracket arise at the same order in λ as the first term, despite nominally involving two powers of S_{int} rather than one (provided that the square bracket itself is order unity). To see this, keep in mind $m_R^2 = \lambda v^2$ so that $(\lambda v/m_R)^2 = \lambda$.

For future purposes it is useful also to have the corresponding result for the reaction $\tilde{\phi}_I(p) + \tilde{\phi}_I(q) \to \tilde{\phi}_I(p') + \tilde{\phi}_I(q')$. A similar calculation, using instead the Feynman graphs of Fig. 1.3, gives the scattering amplitude

$$\begin{aligned}\mathcal{A}_{II\to II} &= 24\mathrm{i}\left(-\frac{\lambda}{16}\right) + 8\left(\frac{\mathrm{i}^2}{2}\right)\left(-\frac{\lambda v}{2\sqrt{2}}\right)^2 \\ &\quad\times\left[\frac{-\mathrm{i}}{(p+q)^2+m_R^2} + \frac{-\mathrm{i}}{(p-p')^2+m_R^2} + \frac{-\mathrm{i}}{(p-q')^2+m_R^2}\right] \\ &= -\frac{3\mathrm{i}\lambda}{2} + \frac{\mathrm{i}(\lambda v)^2}{2m_R^2}\left[\frac{1}{1+2p\cdot q/m_R^2} + \frac{1}{1-2q\cdot q'/m_R^2} + \frac{1}{1-2p\cdot q'/m_R^2}\right].\end{aligned} \tag{1.8}$$

[4] See Exercise 1.1 and Appendix B for the proportionality factors.

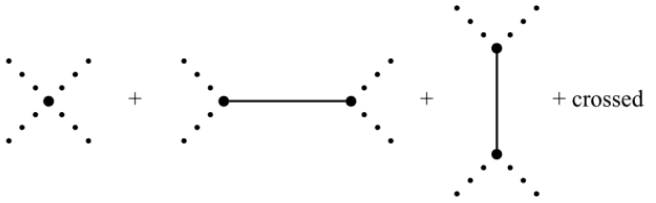

Fig. 1.3 The tree graphs that dominate the $\tilde{\phi}_I\,\tilde{\phi}_I$ scattering amplitude. Solid (dotted) lines represent $\tilde{\phi}_R$ and $\tilde{\phi}_I$ particles.

1.1.3 The Low-Energy Limit

For the present purposes it is the low-energy regime that is of most interest: when the centre-of-mass kinetic energy and momentum transfers during scattering are very small compared with the mass of the heavy particle. This limit is obtained from the above expressions by taking $|p{\cdot}q|$, $|p{\cdot}q'|$ and $|q{\cdot}q'|$ all to be small compared with m_R^2.

Taylor expanding the above expressions shows that both $\mathcal{A}_{RI\to RI}$ and $\mathcal{A}_{II\to II}$ are suppressed in this limit by powers of (q or q')/m_R, in addition to the generic small perturbative factor λ:

$$\mathcal{A}_{RI\to RI} \simeq 2\mathrm{i}\lambda\left(\frac{q\cdot q'}{m_R^2}\right) + O\left(m_R^{-4}\right), \tag{1.9}$$

while

$$\mathcal{A}_{II\to II} \simeq 2\mathrm{i}\lambda\left[\frac{(p\cdot q)^2 + (p\cdot q')^2 + (q\cdot q')^2}{m_R^4}\right] + O\left(m_R^{-6}\right). \tag{1.10}$$

Both of these expressions use 4-momentum conservation, and kinematic conditions like $q^2 = 0$ *etc.* to simplify the result, and both expressions end up being suppressed by powers of q/m_R and/or q'/m_R once this is done.

The basic simplicity of physics at low energies arises because physical quantities typically simplify when Taylor expanded in powers of any small energy ratios (like scattering energy/m_R in the example above). It is this simplicity that ultimately underlies the phenomenon of decoupling: in the toy model the low-energy implications of the very energetic $\tilde{\phi}_R$ states ultimately can be organized into a sequence in powers of m_R^{-2}, with only the first few terms relevant at very low energies.

1.2 The Simplicity of the Low-Energy Limit ◇

Now imagine that your task is to build an experiment to test the above theory by measuring the cross section for scattering $\tilde{\phi}_I$ particles from various targets, using only accelerators whose energies, E, do not reach anywhere near as high as the mass m_R. Since the experiment is more difficult if the scattering is rare, the suppression of the order-λ cross sections by powers of q/m_R and/or q'/m_R at low energies presents a potential problem. But maybe this suppression is an accident of the leading, $O(\lambda)$,

where $U(t,t') = \exp[-iH(t-t')]$ with a Hamiltonian[5] $H = H(\hat{\phi}_R, \hat{\phi}_l)$ depending on both the heavy and light fields. But if the initial state has an energy $E_i < m_R$ it cannot contain any $\hat{\phi}_R$ particles, and energy conservation then precludes $\hat{\phi}_R$ particles from ever being produced by subsequent time evolution.

This means that time evolution remains a linear and unitary transformation even when it is restricted to low-energy states. That is, suppose we define

$$U_{\rm eff}(t,t') := P_\Lambda\, U(t,t')\, P_\Lambda := \exp\left[-iH_{\rm eff}(t-t')\right], \tag{1.18}$$

with $P_\Lambda^2 = P_\Lambda$ being the projection operator onto states with low energy $E < \Lambda \ll m_R$. P_Λ commutes with H and so also with time evolution. Because $H_{\rm eff} = P_\Lambda H P_\Lambda$ if H is hermitian then so must be $H_{\rm eff}$ and so if $U(t,t')$ is unitary then so must be $U_{\rm eff}(t,t')$ when acting on low-energy states.

Furthermore, because the action of $H_{\rm eff}$ is well-defined for states having energy $E < \Lambda$, it can be written as a linear combination of products of creation and annihilation operators for the $\hat{\phi}_l$ field only (since these form a basis for operators that transform among only low-energy states).[6] As a consequence, it must be possible to write $H_{\rm eff} = H_{\rm eff}[\hat{\phi}_l]$, without making any reference to the heavy field $\hat{\phi}_R$ at all.

But there is no guarantee that the expression for $H_{\rm eff}[\hat{\phi}_l]$ obtained in this way is anywhere as simple as is $H[\hat{\phi}_R, \hat{\phi}_l]$. So the real puzzle is why the effective interaction found above is so simple. In particular, why is it local,

$$H_{\rm eff}[\hat{\phi}_l] = \int \mathrm{d}^3x\, \mathcal{H}_{\rm eff}(x), \tag{1.19}$$

with $\mathcal{H}_{\rm eff}(x)$ a simple polynomial in $\hat{\phi}_l(x)$ and its derivatives, all evaluated at the same spacetime point?

Ultimately, the simplicity of this local form can be traced to the uncertainty principle. Interactions, like Eq. (1.12), in $H_{\rm eff}$ not already present in H describe the influence on low-energy $\hat{\phi}_l$ particles of virtual processes involving heavy $\hat{\phi}_R$ particles. These virtual processes are not ruled out by energy conservation even though the production of real $\hat{\phi}_R$ particles is forbidden. One way to understand why they are possible is because the uncertainty principle effectively allows energy conservation to be violated,[7] $E_f = E_i + \Delta E$, but only over time intervals that are sufficiently short, $\Delta t \lesssim \hbar/\Delta E$. The effects of virtual $\hat{\phi}_R$ particles are necessarily localized in time over intervals that are of order $1/m_R$, which are unobservably short for observers restricted to energies $E \ll m_R$. Consequently, they are described at these energies by operators all evaluated at effectively the same time.

In relativistic theories, large momenta necessarily involve large energies and since the uncertainty principle relates large momenta to short spatial distances, a similar argument can be made that the effect of large virtual momentum transfers on the

[5] The convention here is to use $\tilde{\phi}$ to denote the fluctuation when this is a non-operator field (appearing within a path integral, say) and instead use $\hat{\phi}$ for the quantum operator fluctuation field.

[6] See the discussion around Eq. (C.9) of Appendix C for details.

[7] More precisely, energy need not be conserved at each vertex when organized in old-fashioned Rayleigh–Schrödinger perturbation theory from undergraduate quantum mechanics classes. Once reorganized into manifestly relativistic Feynman–Schwinger–Dyson perturbation theory energy actually *is* preserved at each vertex, but internal particles are not on-shell: $E \neq \sqrt{\mathbf{p}^2 + m^2}$. Either way the locality consequences are the same.

low-energy theory can also be captured by effective interactions localized at a single spatial point. Together with the localization in time just described, this shows that the effects of very massive particles are local in both space and time, as found in the toy model above.

Locality arises explicitly in relativistic calculations when expanding the propagators of massive particles in inverse powers of m_R, after which they become local in spacetime since

$$G(x,y) := \langle 0|T\hat{\phi}_R(x)\hat{\phi}_R(y)|0\rangle = -\mathrm{i}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\,\frac{e^{ip(x-y)}}{p^2+m_R^2} \tag{1.20}$$
$$\simeq -\frac{\mathrm{i}}{m_R^2}\sum_{k=0}^{\infty}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\left(-\frac{p^2}{m_R^2}\right)^k e^{ip(x-y)} = -\frac{\mathrm{i}}{m_R^2}\sum_{k=0}^{\infty}\left(\frac{\Box}{m_R^2}\right)^k \delta^4(x-y),$$

where the 'T' denotes time ordering, $p(x-y) := p\cdot(x-y) = p_\mu(x-y)^\mu$ and $\Box = \partial_\mu\partial^\mu = -\partial_t^2 + \nabla^2$ is the covariant d'Alembertian operator.

The upshot is this: to any fixed order in $1/m_R$ the full theory usually can be described by a local effective lagrangian.[8] The next sections develop tools for its efficient calculation and use.

1.2.3 Symmetries: Linear vs Nonlinear Realization

Before turning to the nitty gritty of how the effective action is calculated and used, it is worth first pausing to extract one more useful lesson from the toy model considered above. The lesson is about symmetries and their low-energy realization, and starts by asking why it is that the self-interactions among the light $\hat{\phi}_I$ particles – such as the amplitudes of Eqs. (1.9) and (1.10) – are so strongly suppressed at low energies by powers of $1/m_R^2$.

That is, although it is natural to expect some generic suppression of low-energy interactions by powers of $1/m_R^2$, as argued above, why does nothing at all arise at zeroeth order in $1/m_R$ despite the appearance of terms like $\lambda\hat{\phi}_I^4$ in the full toy-model potential? And why are there so very many powers of $1/m_R$ in the case of $2\hat{\phi}_I \to N\hat{\phi}_I$ scattering in the toy model? (Specifically, why is the amplitude for two $\hat{\phi}_I$ particles scattering to N $\hat{\phi}_I$ particles suppressed by $(1/m_R)^{N+2}$?)

This suppression has a very general origin, and can be traced to a symmetry of the underlying theory [3–5]. The symmetry in question is invariance under the $U(1)$ phase rotation, $\phi \to e^{\mathrm{i}\omega}\phi$, of Eqs. (1.1) and (1.2). In terms of the real and imaginary parts this acts as

$$\begin{pmatrix}\phi_R\\ \phi_I\end{pmatrix} \to \begin{pmatrix}\cos\omega & -\sin\omega\\ \sin\omega & \cos\omega\end{pmatrix}\begin{pmatrix}\phi_R\\ \phi_I\end{pmatrix}. \tag{1.21}$$

A symmetry such as this that acts linearly on the fields is said to be *linearly realized*. As summarized in Appendix C.4, if the symmetry is also linearly realized on particle states then these states come in multiplets of the symmetry, all elements of which share the same couplings and masses. However (as is also argued in

[8] For nonrelativistic systems locality sometimes breaks down in space (*e.g.* when large momenta coexist with low energy). It can also happen that the very existence of a Hamiltonian (without expanding the number of degrees of freedom) breaks down for open systems – the topic of §16.

Appendix C.4) linear transformations of the fields – such as (1.21) – are insufficient to infer that the symmetry also acts linearly for particle states, $|\mathbf{p}\rangle = \mathfrak{a}^*_{\mathbf{p}}|0\rangle$, unless the ground-state, $|0\rangle$, is also invariant. If a symmetry of the action does not leave the ground state invariant it is said to be *spontaneously broken*.

For instance, in the toy model the ground state satisfies $\langle 0|\phi(x)|0\rangle = v$, and so the ground state is only invariant under $\phi \to e^{i\omega}\phi$ when $v = 0$. Indeed, for the toy model if $v = 0$ both particle masses are indeed equal: $m_R = m_I = 0$, as are all of their self-couplings. By contrast, when $v \neq 0$ the masses of the two types of particles differ, as does the strength of their cubic self-couplings. Although $\phi \to e^{i\omega}\phi$ always transforms linearly, the symmetry acts inhomogeneously on the deviation $\hat{\phi} = \phi - v = \frac{1}{\sqrt{2}}(\hat{\phi}_R + i\hat{\phi}_I)$ that creates and destroys the particle states. It is because the deviation does not transform linearly (and homogeneously) that the arguments in Appendix C.4 no longer imply that particle states need have the same couplings and masses when $v \neq 0$.

To see why this symmetry should suppress low-energy $\hat{\phi}_I$ interactions, consider how it acts within the low-energy theory. Even though ϕ transforms linearly in the full theory, because the low-energy theory involves only the single real field $\hat{\phi}_I$, the symmetry cannot act on it in a linear and homogeneous way. To see what the action of the symmetry becomes purely within the low-energy theory, it is useful to change variables to a more convenient set of fields than $\hat{\phi}_R$ and $\hat{\phi}_I$.

To this end, define the two real fields χ and ξ by[9]

$$\phi = \left(v + \frac{\chi}{\sqrt{2}}\right) e^{i\xi/\sqrt{2}v}. \tag{1.22}$$

These have the advantage that the action of the $U(1)$ symmetry, $\phi \to e^{i\omega}\phi$ takes a particularly simple form,

$$\xi \to \xi + \sqrt{2}\, v\, \omega, \tag{1.23}$$

with χ unchanged, so ξ carries the complete burden of symmetry transformation.

In terms of these fields the action, Eq. (1.1), becomes

$$S = -\int \mathrm{d}^4x \left[\frac{1}{2}\partial_\mu\chi\partial^\mu\chi + \frac{1}{2}\left(1 + \frac{\chi}{\sqrt{2}\,v}\right)^2 \partial_\mu\xi\partial^\mu\xi + V(\chi)\right], \tag{1.24}$$

with

$$V(\chi) = \frac{\lambda}{4}\left(\sqrt{2}\,v\,\chi + \frac{\chi^2}{2}\right)^2. \tag{1.25}$$

Expanding this action in powers of χ and ξ gives the perturbative action $S = S_0 + S_{\text{int}}$, with unperturbed contribution

$$S_0 = -\frac{1}{2}\int \mathrm{d}^4x \left[\partial_\mu\chi\partial^\mu\chi + \partial_\mu\xi\partial^\mu\xi + \lambda v^2\,\chi^2\right]. \tag{1.26}$$

This shows that χ is an alternative field representation for the heavy particle, with $m_\chi^2 = m_R^2 = \lambda v^2$. ξ similarly represents the massless field.

It also shows the symmetry is purely realized on the massless state, as an inhomogeneous shift (1.23) rather than a linear, homogeneous transformation.

[9] Numerical factors are chosen here to ensure fields are canonically normalized.

Such a transformation – often called a *nonlinear realization* of the symmetry (both to distinguish it from the linear realization discussed above, and because the transformations turn out in general to be nonlinear when applied to non-abelian symmetries) – is a characteristic symmetry realization in the low-energy limit of a system which spontaneously breaks a symmetry.

The interactions in this representation are given by

$$S_{\rm int} = -\int \mathrm{d}^4x \left[\left(\frac{\chi}{\sqrt{2}\,v} + \frac{\chi^2}{4\,v^2}\right)\partial_\mu\xi\partial^\mu\xi + \frac{\lambda v}{2\sqrt{2}}\,\chi^3 + \frac{\lambda}{16}\,\chi^4\right]. \tag{1.27}$$

For the present purposes, what is important about these expressions is that ξ always appears differentiated. This is a direct consequence of the symmetry transformation, Eq. (1.23), which requires invariance under constant shifts: $\xi \to \xi + \text{constant}$. Since this symmetry forbids a ξ mass term, which would be $\propto m_I^2\,\xi^2$, it ensures ξ remains exactly massless to all orders in the small expansion parameters. ξ is what is called a *Goldstone boson* for the spontaneously broken $U(1)$ symmetry: it is the massless scalar that is guaranteed to exist for spontaneously broken (global) symmetries. Because ξ appears always differentiated it is immediately obvious that an amplitude describing N_i ξ particles scattering into N_f ξ particles must be proportional to at least $N_i + N_f$ powers of their energy, explaining the low-energy suppression of light-particle scattering amplitudes in this toy model.

For instance, explicitly re-evaluating the Feynman graphs of Fig. 1.3, using the interactions of Eq. (1.27) instead of (1.6), gives the case $N_i = N_f = 2$ as

$$\begin{aligned}
&\mathcal{A}_{\xi\xi\to\xi\xi} \\
&= 0 + 8\left(\frac{\mathrm{i}^2}{2}\right)\left(-\frac{1}{\sqrt{2}\,v}\right)^2\left[\frac{-\mathrm{i}(p\cdot q)(p'\cdot q')}{(p+q)^2+m_R^2} + \frac{-\mathrm{i}(p\cdot p')(q\cdot q')}{(p-p')^2+m_R^2} + \frac{-\mathrm{i}(p\cdot q')(q\cdot p')}{(p-q')^2+m_R^2}\right] \\
&= \frac{2\mathrm{i}\lambda}{m_R^4}\left[\frac{(p\cdot q)^2}{1+2p\cdot q/m_R^2} + \frac{(q\cdot q')^2}{1-2q\cdot q'/m_R^2} + \frac{(p\cdot q')^2}{1-2p\cdot q'/m_R^2}\right],
\end{aligned} \tag{1.28}$$

in precise agreement with Eq. (1.8) – as may be seen explicitly using the identity $(1+x)^{-1} = 1 - x + x^2/(1+x)$ – but with the leading low-energy limit much more explicit.

This representation of the toy model teaches several things. First, it shows that scattering amplitudes (and, more generally, arbitrary physical observables) do not depend on which choice of field variables are used to describe a calculation [8–10]. Some kinds of calculations (like loops and renormalization) are more convenient using the variables $\hat{\phi}_R$ and $\hat{\phi}_I$, while others (like extracting consequences of symmetries) are easier using χ and ξ.

Second, this example shows that it is worthwhile to use the freedom to perform field redefinitions to choose those fields that make life as simple as possible. In particular, it is often very useful to make symmetries of the high-energy theory as explicit as possible in the low-energy theory as well.

Third, this example shows that once restricted to the low-energy theory it need not be true that a symmetry remains linearly realized by the fields [11–13], even if this were true for the full underlying theory including the heavy particles. The necessity of realizing symmetries nonlinearly arises once the scales defining the

low-energy theory (*e.g.* $E \ll m_R$) are smaller than the mass difference (*e.g.* m_R) between particles that are related by the symmetry in the full theory, since in this case some of the states required to fill out a linear multiplet are removed as part of the high-energy theory.

1.3 Summary

This first chapter defines a toy model, in which a complex scalar field, ϕ, self-interacts *via* a potential $V = \frac{\lambda}{4}(\phi^*\phi - v^2)^2$ that preserves a $U(1)$ symmetry: $\phi \to e^{i\omega}\phi$. Predictions for particle masses and scattering amplitudes are made as a function of the model's two parameters, λ and v, in the semiclassical regime $\lambda \ll 1$. This model is used throughout the remaining chapters of Part I as a vehicle for illustrating how the formalism of effective field theories works in a concrete particular case.

The semiclassical spectrum of the model has two phases. If $v = 0$ the $U(1)$ symmetry is preserved by the semiclassical ground state and there are two particles whose couplings and masses are the same because of the symmetry. When $v \neq 0$ the symmetry is spontaneously broken, and one particle is massless while the other gets a nonzero mass $m = \sqrt{\lambda}\, v$.

The model's symmetry-breaking phase has a low-energy regime, $E \ll m$, that provides a useful illustration of low-energy methods. In particular, the massive particle decouples at low energies in the precise sense that its virtual effects only play a limited role for the low-energy interactions of the massless particles. In particular, explicit calculation shows the scattering of massless particles at low energies in the full theory to be well-described to leading order in λ and E/m in terms of a simple local 'effective' interaction with lagrangian density $\mathfrak{L}_{\text{eff}} = a_{\text{eff}}(\partial_\mu \xi\, \partial^\mu \xi)^2$, with effective coupling: $a_{\text{eff}} = \lambda/(4m^4)$. The $U(1)$ symmetry of the full theory appears in the low-energy theory as a shift symmetry $\xi \to \xi +$ constant.

Exercises

Exercise 1.1 Use the Feynman rules coming from the action $S = S_0 + S_{\text{int}}$ given in Eqs. (1.5) and (1.6) to evaluate the graphs of Fig. 1.2. Show from your result that the corresponding S-matrix element is given by

$$\langle \hat{\phi}_R(p'), \hat{\phi}_I(q') | \mathcal{S} \,|\, \hat{\phi}_R(p), \hat{\phi}_I(q) \rangle = -i(2\pi)^4 \mathcal{A}_{RI\to RI}\, \delta^4(p + q - p' - q'),$$

with $\mathcal{A}_{RI\to RI}$ given by Eq. (1.7). Taylor expand your result for small q, q' to verify the low-energy limit given in Eq. (1.9). [Besides showing the low-energy decoupling of Goldstone particles, getting right the cancellation that provides this suppression in these variables is a good test of – and a way to develop faith in – your understanding of Feynman rules.]

Exercise 1.2 Using the Feynman rules coming from the action $S = S_0 + S_{\text{int}}$ given in Eqs. (1.5) and (1.6) evaluate the graphs of Fig. 1.3 to show

$$\langle \hat{\phi}_I(p'), \hat{\phi}_I(q') | \mathcal{S} \,|\, \hat{\phi}_I(p), \hat{\phi}_I(q) \rangle = -i(2\pi)^4 \mathcal{A}_{II\to II}\, \delta^4(p + q - p' - q'),$$

The basic connection between operator correlation functions and path integrals is the expression

$$G^{a_1\cdots a_n}(x_1,\cdots,x_n) = \int \mathcal{D}\phi\,\left[\phi^{a_1}(x_1)\cdots\phi^{a_n}(x_n)\right]\exp\left\{\mathrm{i}S[\phi\,]\right\}, \tag{2.4}$$

where $\mathcal{D}\phi = \mathcal{D}\phi^{a_1}\cdots\mathcal{D}\phi^{a_n}$ denotes the functional measure for the sum over all field configurations, $\phi^a(x)$, with initial and final times weighted by the wave functional, $\Psi_i[\phi]$ and $\Psi^*_o[\phi]$, appropriate for the initial and final states, ${}_o\langle\Omega|$ and $|\Omega\rangle_i$. The special case $n = 0$ is the example most frequently encountered in elementary treatments, for which

$${}_o\langle\Omega|\Omega\rangle_i = \int \mathcal{D}\phi\ \exp\left\{\mathrm{i}S[\phi]\right\}. \tag{2.5}$$

Direct use of the definitions then leads to the following expression for $Z[J]$:

$$Z[J] = \int \mathcal{D}\phi\ \exp\left\{\mathrm{i}S[\phi\,] + \mathrm{i}\int \mathrm{d}^4x\,\phi^a(x)J_a(x)\right\}, \tag{2.6}$$

and so $Z[J=0] = {}_o\langle\Omega|\Omega\rangle_i$.

Semiclassical Evaluation

Semiclassical perturbation theory can be formulated by expanding the action within the path integral about a classical background:[3] $\phi^a(x) = \varphi^a_{\rm cl}(x) + \tilde\phi^a(x)$, where $\varphi^a_{\rm cl}$ satisfies

$$\left(\frac{\delta S}{\delta\phi^a}\right)_{\phi=\varphi_{\rm cl}} + J_a = 0. \tag{2.7}$$

The idea is to write the action, $S_J[\phi\,] := S[\phi\,] + \int \mathrm{d}^4x\,(\phi^a J_a)$, as

$$S_J[\varphi_{\rm cl} + \tilde\phi\,] = S_J[\varphi_{\rm cl}] + S_2[\varphi_{\rm cl},\tilde\phi\,] + S_{\rm int}[\varphi_{\rm cl},\tilde\phi\,], \tag{2.8}$$

with

$$S_2 = -\int \mathrm{d}^4x\ \tilde\phi^a\,\Delta_{ab}(\varphi_{\rm cl})\tilde\phi^b, \tag{2.9}$$

being the quadratic part in the expansion (for some differential operator Δ_{ab}). The 'interaction' term, $S_{\rm int}$, contains all terms cubic and higher order in $\tilde\phi^a$; no linear terms appear because the background field satisfies Eq. (2.7).

The relevant path integrals can then be evaluated by expanding

$$\exp\left\{\mathrm{i}S[\phi\,] + \mathrm{i}\int \mathrm{d}^4x\,\phi^a J_a\right\} = \exp\left\{\mathrm{i}S_J[\varphi_{\rm cl}] + \mathrm{i}S_2[\varphi_{\rm cl},\tilde\phi\,]\right\}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\mathrm{i}S_{\rm int}[\varphi_{\rm cl},\tilde\phi\,]\right)^r, \tag{2.10}$$

in the path integral (2.6) and explicitly computing the resulting gaussian functional integrals.

[3] It is sometimes useful to make a more complicated, nonlinear, split $\phi = \phi(\varphi_{\rm cl},\tilde\phi)$ in order to make explicit convenient properties (such as symmetries) of the action.

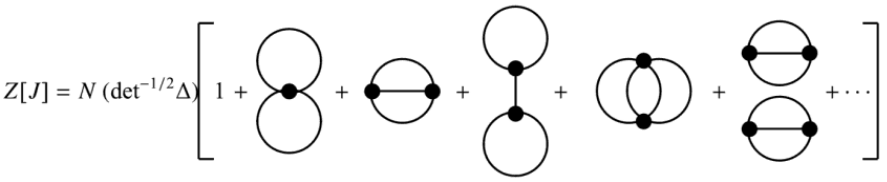

Fig. 2.1 A sampling of some leading perturbative contributions to the generating functional $Z[J]$ expressed using Eq. (2.11) as Feynman graphs. Solid lines are propagators (Δ^{-1}) and solid circles represent interactions that appear in $S_{\rm int}$. 1-Particle reducible and 1PI graphs are both shown as examples at two loops and a disconnected graph is shown at four loops. The graphs shown use only quartic and cubic interactions in $S_{\rm int}$.

This process leads in the usual way to the graphical representation of any correlation function. Gaussian integrals ultimately involve integrands that are powers of fields, leading to integrals of the schematic form[4]

$$\begin{aligned}\int \mathcal{D}\tilde{\phi}\, e^{i\tilde{\phi}^a \Delta_{ab} \tilde{\phi}^b}\, \tilde{\phi}^{c_1}(x_1)\cdots\tilde{\phi}^{c_n}(x_n) &\propto \left(\det{}^{-1/2}\Delta\right) \\ &\times \left[(\Delta^{-1})^{c_1 c_2}\cdots(\Delta^{-1})^{c_{n-1}c_n} + (\text{perms})\right],\end{aligned} \tag{2.11}$$

if n is even, while the integral vanishes if n is odd. Here, the evaluation ignores a proportionality constant that is background-field independent (and so isn't important in what follows). The interpretation in terms of Feynman graphs comes because the combinatorics of such an integral correspond to the combinatorics of all possible ways of drawing graphs whose internal lines represent factors of Δ^{-1} and whose vertices correspond to interactions within $S_{\rm int}$.

Within this type of graphical expression $Z[J]$ is given as the sum over all vacuum graphs, with no external lines. All of the dependence on J appears through the dependence of the result on $\varphi_{\rm cl}$, which depends on J because of (2.7). The graphs involving the fewest interactions (vertices) first arise with two loops, a sampling of which are shown in Fig. 2.1 that can be built using cubic and quartic interactions within $S_{\rm int}$.

As mentioned earlier, this includes *all* graphs, including those that are disconnected, like the right-most four-loop graph involving four cubic vertices in Fig. 2.1. Graphs like this are disconnected in the sense that it is not possible to get between any pair of vertices along some sequence of contiguous internal lines.

Although simple to state, the perturbation expansion outlined above in terms of vacuum graphs is not yet completely practical for explicit calculations. The problem is the appearance of the background field φ in the propagator $(\Delta^{-1})_{ab}$. Although $\Delta_{ab}(x,y) = -\delta^2 S/\delta\phi^a(x)\delta\phi^b(y)$ itself is easy to compute, it is often difficult to invert explicitly for generic background fields. For instance, for a single scalar field interacting through a scalar potential $U(\phi)$ one has $\Delta(x,y) = [-\Box + U''(\varphi)]\,\delta^4(x-y)$ and although this is easily inverted in momentum space when φ is constant, it is more difficult to invert for arbitrary $\varphi(x)$.

This problem is usually addressed by expanding in powers of $J_a(x)$, so that the path integral is evaluated as a semiclassical expansion about a simple background

[4] This expression assumes a bosonic field, but a similar expression holds for fermions.

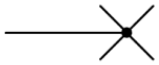

Fig. 2.2 The Feynman rule for the vertex coming from the linear term, $S_{\rm lin}$, in the expansion of the action. The cross represents the sum $\delta S/\delta\varphi^a + J_a$.

configuration, $\varphi^a_{\rm cl}$, that satisfies $(\delta S/\delta\phi^a)(\phi = \varphi_{\rm cl}) = 0$ instead of Eq. (2.7). The Feynman graphs for this modified expansion differ in two ways from the expansion described above: (*i*) the propagators Δ^{-1} now are evaluated at a J-independent configuration, $\varphi^a_{\rm cl}$, which can be explicitly evaluated if this configuration is simple enough (such as, for instance, if $\varphi^a_{\rm cl} = 0$); and (*ii*) the term $\phi^a J_a$ in the exponent of the integrand in (2.6) is now treated as an interaction. Since this interaction is linear in ϕ^a it corresponds graphically to a 'tadpole' contribution (as in Fig. 2.2), with the line ending in a cross whose Feynman rule is $J_a(x)$.

Within this framework, the Feynman graphs giving $Z[J]$ are obtained from those given in Fig. 2.1 by inserting external lines in all possible ways (both to propagators and vertices), with the understanding that the end of each external line represents a factor of $J_a(x)$. This kind of modified expansion gives $Z[J]$ explicitly as a Taylor expansion in powers of J.

2.1.1 Connected Correlations

As Fig. 2.1 shows, the graphical expansion for $Z[J]$ in perturbation theory includes both connected and disconnected Feynman graphs. It is often useful to work instead with a generating functional, $W[J]$, whose graphical expansion contains only connected graphs. As shown in Exercise 2.4, this is accomplished simply by defining $Z[J] := \exp\{iW[J]\}$ [5, 15], since taking the logarithm has the effect of subtracting out the disconnected graphs. This implies the path integral representation

$$\exp\{iW[J]\} = \int \mathcal{D}\phi \ \exp\left\{iS[\phi\,] + i\int d^4x\,\phi^a J_a\right\}. \tag{2.12}$$

The connected, time-ordered correlation functions are then given by functional differentiation:

$$\langle T[\phi^{a_1}(x_1)\cdots\phi^{a_n}(x_n)]\rangle_c := (-i)^{n-1}\left(\frac{\delta^n W[J]}{\delta J_{a_1}(x_1)\cdots\delta J_{a_n}(x_n)}\right)_{J=0}. \tag{2.13}$$

For example,

$$\langle\phi^a(x)\rangle_c = \left(\frac{\delta W[J]}{\delta J_a(x)}\right)_{J=0} = -i\left(\frac{1}{Z[J]}\frac{\delta Z[J]}{\delta J_b(y)}\right)_{J=0} = \frac{{}_o\langle\Omega|\phi^a(x)|\Omega\rangle_i}{{}_o\langle\Omega|\Omega\rangle_i}, \tag{2.14}$$

while,

$$\begin{aligned}\langle T[\phi^a(x)\,\phi^b(y)]\rangle_c &= -i\left(\frac{\delta^2 W[J]}{\delta J_a(x)\delta J_b(y)}\right)_{J=0}\\ &= \frac{{}_o\langle\Omega|T[\phi^a(x)\,\phi^b(y)]|\Omega\rangle_i}{{}_o\langle\Omega|\Omega\rangle_i}\\ &\quad - \left(\frac{{}_o\langle\Omega|\phi^a(x)|\Omega\rangle_i}{{}_o\langle\Omega|\Omega\rangle_i}\right)\left(\frac{{}_o\langle\Omega|\phi^b(y)|\Omega\rangle_i}{{}_o\langle\Omega|\Omega\rangle_i}\right),\end{aligned} \tag{2.15}$$

and so on. As is easily verified, the graphical expansion of the factor ${}_o\langle\Omega|T[\phi^a(x)\,\phi^b(y)]|\Omega\rangle_i$ in this last expression corresponds to the sum over all Feynman graphs with precisely two external lines, corresponding to the fields $\phi^a(x)$ and $\phi^b(y)$. The graphical representation of a term like ${}_o\langle\Omega|\phi^a(x)|\Omega\rangle_i$ is similarly given by the sum over all Feynman graphs (called tadpole graphs) with precisely one external line, corresponding to $\phi^a(x)$.

Dividing all terms by the factors of ${}_o\langle\Omega|\Omega\rangle_i$ in the denominator is precisely what is needed to cancel disconnected vacuum graphs (*i.e.* those disconnected subgraphs having no external lines). But this does not remove graphs in ${}_o\langle\Omega|T[\phi^a(x)\,\phi^b(y)]|\Omega\rangle_i$ corresponding to a pair of disconnected 'tadpole' graphs, each of which has a single external line. These disconnected graphs precisely correspond to the product ${}_o\langle\Omega|\phi^a(x)|\Omega\rangle_i\,{}_o\langle\Omega|\phi^b(y)|\Omega\rangle_i$ in (2.15), whose subtraction therefore cancels the remaining disconnected component from $\langle T[\phi^a(x)\,\phi^b(y)]\rangle_c$.

A similar story goes through for the higher functional derivatives, and shows how correlations obtained by differentiating W have their disconnected parts systematically subtracted off. Indeed Eqs. (2.13) and (2.12) can be used as non-perturbative definitions of what is meant by connected correlations functions and their generators [15].

2.1.2 The 1PI (or Quantum) Action ♠

As Eqs. (2.12) and (2.4) show, the functional $Z[J] = \exp\{iW[J]\}$ can be physically interpreted as the 'in-out' vacuum amplitude in the presence of an applied current $J_a(x)$. Furthermore, the applied current can be regarded as being responsible for changing the expectation value of the field, since not evaluating Eq. (2.14) at $J_a = 0$ gives

$$\varphi^a(x) := \langle\phi^a(x)\rangle_J = \frac{\delta W}{\delta J_a(x)}, \tag{2.16}$$

as a functional of the current $J_a(x)$. However, it is often more useful to have the vacuum-to-vacuum amplitude expressed directly as a functional of the expectation value, $\varphi^a(x)$, itself, rather than $J_a(x)$. This is accomplished by performing a Legendre transform, as follows.

Legendre Transform

To perform a Legendre transform, define [15]

$$\Gamma[\varphi\,] := W[J] - \int \mathrm{d}^4x\; \varphi^a J_a, \tag{2.17}$$

with $J_a(x)$ regarded as a functional of $\varphi^a(x)$, implicitly obtained by solving Eq. (2.16). Once $\Gamma[\varphi\,]$ is known, the current required to obtain the given $\varphi^a(x)$ is found by directly differentiating the definition, Eq. (2.17), using the chain rule together with Eq. (2.16) to evaluate the functional derivative of $W[J]$:

$$\frac{\delta\Gamma}{\delta\varphi^a(x)} = \int \mathrm{d}^4y\; \frac{\delta J_b(y)}{\delta\varphi^a(x)}\,\frac{\delta W}{\delta J_b(y)} - J_a(x) - \int \mathrm{d}^4y\; \varphi^b(y)\,\frac{\delta J_a(y)}{\delta\varphi^a(x)} = -J_a(x). \tag{2.18}$$

In particular, this last equation shows that the expectation value for the 'real' system with $J_a = 0$ is a stationary point of $\Gamma[\varphi\,]$. In this sense $\Gamma[\varphi\,]$ is related to $\langle\phi^a\rangle$ in the same way that the classical action, $S[\phi\,]$, is related to a classical background configuration, $\varphi^a_{\rm cl}$. For this reason $\Gamma[\varphi\,]$ is sometimes thought of as the quantum generalization of the classical action, and known as the theory's *quantum action*.

This similarity between $\Gamma[\varphi\,]$ and the classical action is also reinforced by other considerations. For instance, because the classical action is usually the difference, $S = K - V$, between kinetic and potential energies, for time-independent configurations (for which the kinetic energy is $K = 0$) the classical ground state actually minimizes $V = -S$. It can be shown that for time-independent systems – *i.e.* those where the ground state $|\Omega\rangle$ is well-described in the adiabatic approximation – the configuration $\varphi^a = \langle\phi^a\rangle$ similarly minimizes the quantity $-\Gamma$. In particular, for configurations φ^a independent of spacetime position the configuration minimizes the quantum 'effective potential' $V_{\rm q}(\varphi) = -\Gamma[\varphi]/({\rm Vol})$, where 'Vol' is the overall volume of spacetime.

One way to prove this [16, 17] is to show that, for any static configuration, φ^a, the quantity $-\Gamma[\varphi\,]$ can be interpreted as the minimum value of the energy, $\langle\Psi|H|\Psi\rangle$, extremized over all normalized states, $|\Psi\rangle$, that satisfy the condition $\langle\Psi|\phi^a(x)|\Psi\rangle = \varphi^a(x)$. The global minimum to $-\Gamma[\varphi\,]$ then comes once φ^a is itself varied over all possible values.

Semiclassical Expansion

How is $\Gamma[\varphi\,]$ computed within perturbation theory? To find out, multiply the path integral representation for $W[J]$, Eq. (2.12), on both sides by $\exp\left\{-{\rm i}\int {\rm d}^4x\,(\varphi^a J_a)\right\}$. Since neither φ^a nor J_a are integration variables, this factor may be taken inside the path integral, giving

$$\begin{aligned}\exp\left\{{\rm i}\Gamma[\varphi\,]\right\} &= \exp\left\{{\rm i}W[J] - {\rm i}\int {\rm d}^4x\,\varphi^a J_a\right\}\\ &= \int \mathcal{D}\phi\ \exp\left\{{\rm i}S[\phi\,] + {\rm i}\int {\rm d}^4x\,(\phi^a - \varphi^a)J_a\right\}\\ &= \int \mathcal{D}\tilde{\phi}\ \exp\left\{{\rm i}S[\varphi + \tilde{\phi}\,] + {\rm i}\int {\rm d}^4x\,\tilde{\phi}^a J_a\right\}.\end{aligned} \tag{2.19}$$

The last line uses the change of integration variable $\phi^a \to \tilde{\phi}^a := \phi^a - \varphi^a$.

At face value, Eq. (2.19) doesn't seem so useful in practice, since the dependence on J_a inside the integral is to be regarded as a functional of φ^a, using Eq. (2.18). This means that $\Gamma[\varphi\,]$ is only given implicitly, since it appears on both sides. But on closer inspection, the situation is much better than this, because the implicit appearance of Γ through J_a on the right-hand side is actually very easy to implement in perturbation theory.

To see how this works, imagine evaluating Eq. (2.19) perturbatively by expanding the action inside the path integral about the configuration $\phi^a = \varphi^a$, using

$$S[\varphi + \tilde{\phi}\,] = S[\varphi\,] + S_2[\varphi, \tilde{\phi}\,] + S_{\rm lin}[\varphi, \tilde{\phi}\,] + S_{\rm int}[\varphi, \tilde{\phi}\,]. \tag{2.20}$$

This is very similar to the expansion in Eq. (2.10), apart from the term linear in $\tilde{\phi}^a$,

displacement of φ^a away from a sufficiently simple background for which Δ^{-1}_{ab} can be evaluated).

In this case, using

$$\left(\frac{\delta^2 S}{\delta\phi^a(x)\,\delta\phi^b(y)}\right)_\varphi = \left(\frac{\delta^2 S}{\delta\phi^a(x)\,\delta\phi^b(y)}\right)_0 + \int \mathrm{d}^4 z \left(\frac{\delta^3 S}{\delta\phi^a(x)\,\delta\phi^b(y)\,\delta\phi^c(z)}\right)_0 \varphi^c(z) + \cdots, \tag{2.26}$$

and $(\Delta_0 - \delta\Delta)^{-1} = \Delta_0^{-1}\sum_{n=0}^{\infty}(\delta\Delta\,\Delta_0^{-1})^n$ shows that the required Feynman graphs are obtained by inserting external lines in all possible ways (both to the internal lines and the vertices) in the 1PI vacuum graphs of Fig. 2.1, with the external lines representing the Feynman rule $\varphi^a(x)$.

2.2 The High-Energy/Low-Energy Split ◇

So far, so good, but how can the above formalism be used to compute and use low-energy effective actions? The rest of this chapter specializes to theories having two very different intrinsic mass scales – like $m_l \ll m_h$ of the toy model in Chapter 1 – in order to address this question. After formalizing the split into low- and high-energy theory in this section, the following two sections identify two useful ways of defining a low-energy effective action.

2.2.1 Projecting onto Low-Energy States

The starting point, in this section, is to define more explicitly the split between low- and high-energy degrees of freedom. There are a variety of ways to achieve this split. Most directly, imagine dividing the quantum field ϕ^a into a low-energy and high-energy part: $\phi^a(x) = l^a(x) + \mathfrak{h}^a(x)$, with

$$l^a(x) := P_\Lambda \phi^a(x) P_\Lambda, \tag{2.27}$$

where $P_\Lambda = P_\Lambda^2$ denotes the projector onto states having energy $E < \Lambda$. To be of practical use, the scale Λ should lie somewhere between the two scales (such as m_l and m_h) that define the underlying hierarchy ($m_l \ll m_h$) in terms of which the low-energy limit is defined for the theory of interest.

This can be made more explicit in semiclassical perturbation theory, where $\phi^a = \varphi^a_{\rm cl} + \hat\phi^a$. Since in the interaction representation the quantum field satisfies the linearized field equation, $\Delta_{ab}\hat\phi^b = 0$, one can decompose $\hat\phi^a(x)$ in terms of a basis of eigenmodes, $u_p(x)$,

$$\hat\phi^a(x) = \sum_p \left[\mathfrak{a}_p u_p^a(x) + \mathfrak{a}_p^* u_p^{a*}(x)\right]. \tag{2.28}$$

For time-independent backgrounds these eigenmodes can be chosen to simultaneously diagonalize the energy, $i\partial_t u_p = \varepsilon_p\, u_p$, and so the low-energy part of the field

is that part of the sum for which the mode energies are smaller than the reference scale Λ. That is,

$$\hat{l}^a(x) := \sum_{\varepsilon_p<\Lambda} \left[\mathfrak{a}_p u_p^a(x) + \mathfrak{a}_p^* u_p^{a*}(x)\right], \tag{2.29}$$

and so

$$\hat{\mathfrak{h}}^a := \hat{\phi}^a - \hat{l}^a = \sum_{\varepsilon_p>\Lambda} \left[\mathfrak{a}_p u_p^a(x) + \mathfrak{a}_p^* u_p^{a*}(x)\right]. \tag{2.30}$$

Of course, one might also implement a cutoff more smoothly, by weighting high-energy states in amplitudes by some suitably decreasing function of energy rather than completely cutting them off above Λ.

It is natural at this point to worry that a division into high- and low-energy modes introduces an explicit frame-dependence into the problem. After all, a collision that appears to involve only low energies to one observer would appear to involve very high energies to another observer who moves very rapidly relative to the first one. Although this is true in principle, in practice frame-independent physical quantities (like the scattering amplitudes examined for the toy model in Chapter 1) only depend on invariant quantities like centre-of-mass energies, and all observers agree when these are large or small. For scattering calculations the natural split between low- and high-energies is therefore made in the centre-of-mass frame. The point is that in order to profit from the simplification of physics at low energies, it suffices that there exist *some* observers who see a process to be at low energies; it is not required that *all* observers do so.

Notice that correlation functions of low-energy states necessarily do not vary very quickly with time. This may be seen by inserting a complete set of intermediate energy eigenstates between any two pairs of low-energy fields, such as

$$\begin{aligned} {}_o\langle\Omega|\, \hat{l}^a(x)\, \hat{l}^b(y)\, |\Omega\rangle_i &= {}_o\langle\Omega|P_\Lambda \hat{\phi}^a(x) P_\Lambda^2 \hat{\phi}^b(y) P_\Lambda|\Omega\rangle_i \\ &= \sum_{\varepsilon_p<\Lambda} {}_o\langle\Omega|\hat{\phi}^a(x)|p\rangle\langle p\,|\hat{\phi}^b(y)|\Omega\rangle_i, \end{aligned} \tag{2.31}$$

which uses $P_\Lambda|\Omega\rangle = |\Omega\rangle$ and $P_\Lambda|p\rangle = |p\rangle$ for low-energy states, while $P_\Lambda|p\rangle = 0$ for high-energy states. This result clearly has support only for frequencies $\omega = \varepsilon_p < \Lambda$. In relativistic and translation-invariant theories, for which low energy also means low momentum, the same argument shows that correlations also have slow spatial variation.

Example: The Toy Model

To make this concrete, consider the toy model of Chapter 1. In this case, there are two quantum fields, $\hat{\phi}_l$ and $\hat{\phi}_R$ (or equivalently, $\hat{\xi}$ and $\hat{\chi}$), and the two intrinsic mass scales are $m_l = 0 \ll m_R$. The energy eigenmodes for these are labeled by 4-momentum, $u_p(x) \propto e^{ipx}$, and the linearized field equation $(-\Box + m^2)\,\hat{\phi} = 0$, relates the energy to the momentum by $q^0 = \epsilon_q = q$ for $\hat{\phi}_l$ and $p^0 = \varepsilon_p = \sqrt{p^2 + m_R^2}$ for $\hat{\phi}_R$.

In this case, the useful choice is $m_l \ll \Lambda \ll m_R$, which is possible because of the hierarchy $m_l \ll m_R$. With this choice, the light fields consist only of the long-wavelength modes of $\hat{\phi}_l$,

$$\hat{l}(x) = \sum_{\epsilon_q < \Lambda} \left[\mathfrak{c}_q \, e^{iqx} + \mathfrak{c}_q^* \, e^{-iqx} \right], \tag{2.32}$$

while the heavy fields contain all of the $\hat{\phi}_R$ modes together with the short-wavelength modes of $\hat{\phi}_I$:

$$\hat{\mathfrak{h}}_I(x) = \sum_{\epsilon_q > \Lambda} \left[\mathfrak{c}_q \, e^{iqx} + \mathfrak{c}_q^* \, e^{-iqx} \right] \quad \text{and} \quad \hat{\mathfrak{h}}_R(x) = \sum_p \left[\mathfrak{b}_p \, e^{ipx} + \mathfrak{b}_p^* \, e^{-ipx} \right]. \tag{2.33}$$

2.2.2 Generators of Low-Energy Correlations ♠

The next step is to seek generating functionals specifically for low-energy correlation functions, and to investigate their properties. The key tool for this purpose is the observation made above that the correlation functions themselves can vary only over time and length scales larger than Λ^{-1}.

Imagine now defining the generating functional, $Z_{\rm le}[J]$, for the time-ordered in-out correlations of only the low-energy fields, $\hat{l}^a(x)$. This can be done simply by restricting the definition, Eq. (2.2), of $Z[J]$, to include only correlation functions that vary slowly in space and time (*i.e.* only over scales larger than Λ^{-1}), leading to the result

$$Z_{\rm le}[J] := \sum_{n=0}^{\infty} \frac{i^n}{n!} \int d^4x_1 \cdots d^4x_n G_{\rm le}^{a_1 \cdots a_n}(x_1, \cdots, x_n) J_{a_1}(x_1) \cdots J_{a_n}(x_n). \tag{2.34}$$

Because the low-energy correlation functions only vary slowly in space and time, the same is true of any currents, $J_a(x)$, appearing in $Z_{\rm le}[J]$. That is, if the current is split into long- and short-wavelength Fourier modes, $J_a(x) = j_a(x) + \mathcal{J}_a(x)$, with

$$j_a(x) = \sum_{\text{slowly varying}} j_a(p) \, e^{ipx}, \tag{2.35}$$

then the generating functional for low-energy correlations, $Z_{\rm le}[J]$, is simply the restriction of the full generating functional to slowly varying currents:

$$Z_{\rm le}[j] = Z[j, \mathcal{J} = 0]. \tag{2.36}$$

Here, the precise definition of 'slowly varying' in Eq. (2.35) depends on the details of the particle masses and the way the cutoff Λ is implemented – *c.f.* Eq. (2.29) for example – for the quantum states.

It might seem bothersome that the generating functionals for low-energy correlation functions depend explicitly on the value of Λ, as well as on all of the details of precisely how the high-energy modes are cut off. One of the tasks of later chapters is to show how this dependence can be absorbed into appropriate renormalizations of effective couplings, so that predictions for physical processes (like scattering amplitudes) only depend on physical mass scales like m_R (rather than Λ or other definitional details).

For relativistic, translationally invariant theories a slightly more convenient way to break Fourier modes into slowly and quickly varying parts is to Wick rotate [18] to euclidean signature, $\{x^0, \mathbf{x}\} \to \{ix^4, \mathbf{x}\}$. In this case, the time-components of any 4-vector must be similarly rotated, so the invariant inner product of two 4-vectors becomes

$$p \cdot q = p_\mu q^\mu = -p^0 q^0 + \mathbf{p} \cdot \mathbf{q} \to +p^4 q^4 + \mathbf{p} \cdot \mathbf{q} = (p \cdot q)_E. \tag{2.37}$$

This ensures that the invariant condition $p_\mu p^\mu = (p^4)^2 + \mathbf{p}^2 < \Lambda^2$ excludes large values of both $|\mathbf{p}|$ and p^4 (unlike for Minkowski signature, where $p_\mu p^\mu = -(p^0)^2 + \mathbf{p}^2 < \Lambda^2$ allows both $|\mathbf{p}|$ and p^0 to be arbitrarily large but close to the light cone).

The generator, $W_{\rm le}[j]$, of low-energy connected correlations can be defined as before, by taking the logarithm of $Z_{\rm le}[j]$, leading to the path integral representation

$$\exp\Big\{\mathrm{i}W_{\rm le}[j]\Big\} = \int \mathcal{D}\phi \ \exp\left\{\mathrm{i}S[\phi] + \mathrm{i}\int \mathrm{d}^4x\, \phi^a j_a\right\}. \tag{2.38}$$

The main difference between this and the expression for $W[J]$ is the absence of currents coupled to the high-frequency components of ϕ^a. That is, if $\phi^a = l^a + \mathfrak{h}^a$ is split into slowly varying ('light', l^a) and rapidly varying ('heavy', $\mathfrak{h}^a$) parts, along the same lines as Eq. (2.35) for J_a, then Eq. (2.38) becomes

$$\exp\Big\{\mathrm{i}W_{\rm le}[j]\Big\} = \int \mathcal{D}l\mathcal{D}\mathfrak{h} \ \exp\left\{\mathrm{i}S[l+\mathfrak{h}] + \mathrm{i}\int \mathrm{d}^4x\, l^a j_a\right\}. \tag{2.39}$$

Physically, this states that a restriction to low-energy correlations can be obtained simply by restricting oneself only to probing the system with slowly varying currents.

2.2.3 The 1LPI Action

At this point, it is hard to stop from performing a Legendre transformation to obtain the generating functional, $\Gamma_{\rm le}[\ell]$, directly in terms of the low-energy field configurations, ℓ^a, rather than j_a. To this end, define

$$\Gamma_{\rm le}[\ell] := W_{\rm le}[j] - \int \mathrm{d}^4x\, \ell^a j_a, \tag{2.40}$$

with $j_a = j_a[\ell]$ regarded as a functional of ℓ^a found by inverting the relation $\ell^a = \ell^a[j]$ obtained from

$$\ell^a := \frac{\delta W_{\rm le}}{\delta j_a}, \tag{2.41}$$

with the result — *c.f.* Eq. (2.18)

$$j_a = -\frac{\delta \Gamma_{\rm le}}{\delta \ell^a}. \tag{2.42}$$

It is important to realize that although $\Gamma_{\rm le}[\ell]$ obtained in this way only has support on slowly varying field configurations, $\ell^a(x)$, it is *not* simply the restriction of $\Gamma[\varphi\,] = \Gamma[\ell, h]$ to long-wavelength configurations: $h^a := \delta W/\delta \mathcal{J}_a = 0$. To see why not, consider its path integral representation:

$$\begin{aligned}\exp\Big\{\mathrm{i}\Gamma_{\rm le}[\ell]\Big\} &= \int \mathcal{D}l\, \mathcal{D}\mathfrak{h} \ \exp\left\{\mathrm{i}S[l,\mathfrak{h}] + \mathrm{i}\int \mathrm{d}^4x\, (l^a - \ell^a) j_a\right\} \\ &= \int \mathcal{D}\tilde{l}\, \mathcal{D}\mathfrak{h} \ \exp\left\{\mathrm{i}S[\ell + \tilde{l}, \mathfrak{h}] + \mathrm{i}\int \mathrm{d}^4x\, \tilde{l}^{\,a} j_a\right\}.\end{aligned} \tag{2.43}$$

For comparison, the earlier result, Eq. (2.4), for $\Gamma[\varphi\,] = \Gamma[\ell, h]$ is

$$\exp\Big\{\mathrm{i}\Gamma_{\rm le}[\ell, h]\Big\} = \int \mathcal{D}\tilde{l}\, \mathcal{D}\tilde{\mathfrak{h}} \ \exp\left\{\mathrm{i}S[\ell + \tilde{l}, h + \tilde{\mathfrak{h}}] + \mathrm{i}\int \mathrm{d}^4x\, (\tilde{l}^{\,a} j_a + \tilde{\mathfrak{h}}^a \mathcal{J}_a)\right\}. \tag{2.44}$$

The key point is that the condition $h = 0$ is *not* generically equivalent to the condition $\mathcal{J}_a(x) = 0$ that relates $\Gamma_{\rm le}[\ell]$ to $\Gamma[\ell, h]$. Instead, the condition $\mathcal{J}_a = 0$ states that the short-wavelength part of the field should be chosen as that configuration, $h^a = h^a_{\rm le}(\ell)$, that satisfies

$$\left(\frac{\delta\Gamma}{\delta h^a}\right)_{h=h_{\rm le}(\ell)} = 0. \tag{2.45}$$

In particular, the vanishing of $\mathcal{J}_a$ means that the short-wavelength components of the current are not available to take the values $\mathcal{J}_a = -\delta\Gamma/\delta h^a$ they would have taken in the Legendre transform of $W[J] = W[j, \mathcal{J}]$. They are therefore not able to cancel the one-particle-reducible graphs that can be broken in two by cutting a single $\hat{\mathfrak{h}}^a$ line. The quantity $\Gamma_{\rm le}[\varphi]$ is therefore given as the sum of *one-light-particle irreducible* (or 1LPI) graphs, which are only irreducible in the sense that they cannot be broken into two disconnected pieces by cutting a *light*-particle, $\hat{l}^a$, line.

Example: The Toy Model

How does all this look in the toy model of Chapter 1? In this case, with Λ chosen to satisfy $m_l \ll \Lambda \ll m_R$, the 'light' fields consist only of the low-energy modes of the massless field, ξ (or $\hat{\phi}_l$), and the 'heavy' fields consist of both the high-energy modes of ξ together with all of the modes of the massive field χ (or $\hat{\phi}_R$). The 1LPI generator of low-energy connected correlation functions then is

$$\exp\left\{\mathrm{i}\Gamma_{\rm le}[\xi]\right\} = \int \mathcal{D}\tilde{\xi}\mathcal{D}\tilde{\chi}\ \exp\left\{\mathrm{i}S[\xi + \tilde{\xi}, \tilde{\chi}] + \mathrm{i}\int \mathrm{d}^4x\, \tilde{\xi}^a j_a\right\}, \tag{2.46}$$

with ξ^a and $j_a = -\delta\Gamma_{\rm le}/\delta\xi^a$ only varying over times and distances longer than Λ^{-1}.

Recall that small λ controls a semiclassical expansion, and imagine computing $\Gamma_{\rm le}[\xi]$ in the leading, classical approximation. As argued earlier (and elaborated in §3 below), in this limit the full 1PI generator reduces to the classical action: $\Gamma[\xi, \chi] \simeq S[\xi, \chi]$, explicitly given in Eq. (1.24),

$$S[\xi, \chi] = -\int \mathrm{d}^4x\ \left[\frac{1}{2}\,\partial_\mu\chi\partial^\mu\chi + \frac{1}{2}\left(1 + \frac{\chi}{\sqrt{2}\,v}\right)^2 \partial_\mu\xi\partial^\mu\xi + V(\chi)\right], \tag{2.47}$$

with

$$V(\chi) = \frac{m_R^2}{2}\,\chi^2 + \frac{\lambda\, v}{2\sqrt{2}}\,\chi^3 + \frac{\lambda}{16}\,\chi^4. \tag{2.48}$$

In general, the above arguments say that $\Gamma_{\rm le}[\xi] = \Gamma_{\rm le}[\xi, \chi_{\rm le}(\xi)]$, where $\chi_{\rm le}(\xi)$ is obtained by solving $\delta\Gamma[\xi, \chi]/\delta\chi = 0$ (*c.f.* Eq. (2.45)). So in the classical approximation $\Gamma_{\rm le}[\xi] \simeq S[\xi, \chi_{\rm le}(\xi)]$, where Eq. (2.45) in the classical approximation says $\chi_{\rm le}(\xi)$ is found by solving the classical field equation

$$\left(-\Box + m_R^2 + \frac{1}{2v^2}\,\partial_\mu\xi\partial^\mu\xi\right)\chi_{\rm le} = -\frac{1}{\sqrt{2}\,v}\,\partial_\mu\xi\partial^\mu\xi - \frac{3\lambda v}{2\sqrt{2}}\,\chi_{\rm le}^2 - \frac{\lambda}{4}\,\chi_{\rm le}^3. \tag{2.49}$$

Using this in the classical action leads (after an integration by parts) to

$$S[\xi, \chi_{\rm le}(\xi)] = \int \mathrm{d}^4x\ \left[-\frac{1}{2}\left(1 + \frac{\chi_{\rm le}}{\sqrt{2}\,v}\right)\partial_\mu\xi\partial^\mu\xi + \frac{\lambda v}{4\sqrt{2}}\,\chi_{\rm le}^3 + \frac{\lambda}{16}\,\chi_{\rm le}^4\right], \tag{2.50}$$

where $\chi_{\rm le}(\xi)$ is to be interpreted as the function of ξ obtained by solving (2.49).

2.3.1 Definitions

A good starting point for describing the Wilson action is the path integral expression for the 1LPI generator, Eq. (2.43):

$$\exp\Big\{\mathrm{i}\Gamma_{\rm le}[\ell]\Big\} = \int \mathcal{D}\tilde{l}\,\mathcal{D}\mathfrak{h}\;\exp\left\{\mathrm{i}S[\ell+\tilde{l}+\mathfrak{h}] + \mathrm{i}\int \mathrm{d}^4x\,\tilde{l}^{\,a}j_a\right\}. \tag{2.57}$$

What is noteworthy about this expression is that – because the currents are chosen to explore only low-energy quantities – the heavy field, $\mathfrak{h}^a$, appears only in the classical action and not in the current term. As a consequence, all possible low-energy influences of the heavy field must be captured in the quantity

$$\exp\Big\{\mathrm{i}S_{\!W}[l]\Big\} := \int \mathcal{D}\mathfrak{h}\;\exp\Big\{\mathrm{i}S[l+\mathfrak{h}]\Big\}, \tag{2.58}$$

in terms of which the full 1LPI action is given by

$$\exp\Big\{\mathrm{i}\Gamma_{\rm le}[\ell]\Big\} = \int \mathcal{D}\tilde{l}\;\exp\left\{\mathrm{i}S_{\!W}[\ell+\tilde{l}] + \mathrm{i}\int \mathrm{d}^4x\,\tilde{l}^{\,a}j_a\right\}. \tag{2.59}$$

Eq. (2.58) defines the Wilson action, obtained by *integrating out* all heavy degrees of freedom having energies above the scale Λ. It has several noteworthy features, which are explored in detail throughout the rest of the book.

- As the definition shows, the Wilson action provides the earliest place in a calculation to systematically identify, once and for all, the low-energy influence of the heavy degrees of freedom $\mathfrak{h}$. Best of all, this can be done in one fell swoop, before choosing precisely which observable or correlation function is of interest in a particular application.
- For practical applications, most real interest is in obtaining the Wilson action as a series expansion in inverse powers of the heavy mass scales in the problem of interest. As shall be seen in some detail, at any fixed order in this expansion the Wilson action is a local functional, $S_{\!W} = \int \mathrm{d}^4x\;\mathfrak{L}_{\!W}(x)$, with $\mathfrak{L}_{\!W}(x)$ being a function of the light fields and their derivatives all evaluated at the same spacetime point.
- What is striking about Eq. (2.59) is that the Wilson action, $S_{\!W}$, appears in the expression for the generator, $\Gamma_{\rm le}$, of low-energy correlators, in precisely the way that the classical action, S, appears in the expression, Eq. (2.19), for the generator, Γ, of generic correlators. This suggests that the classical action of the full theory might itself be better regarded as the Wilson action from some even higher-energy theory.
- Eq. (2.58) shows that $S_{\!W}$ depends in detail on things like Λ and precisely how the split is made between the high- and low-energy sectors, since these are buried in the definitions of the split between $\mathfrak{h}^a$ and l^a. So it is misleading to speak about 'the' Wilson action, rather than 'a' Wilson action. Yet we know that Λ cannot appear in any physical observables, because it is just an arbitrary artificial scale that is introduced for calculational convenience. Part of the story to follow must therefore be why all these calculational details in $S_{\!W}$ drop out of observables. The outlines of this argument are already clear in Eq. (2.59), which shows that the Λ dependence introduced by performing the integration over $\mathcal{D}\mathfrak{h}$ is later canceled when integrating over the rest of the fields, $\mathcal{D}l$, since the total measure $\mathcal{D}\phi = \mathcal{D}l\,\mathcal{D}\mathfrak{h}$ is Λ-independent.

In semiclassical perturbation theory, the arguments of earlier sections show that Eq. (2.58) gives S_W as the sum over all connected vacuum graphs – not just 1PI graphs, say – using Feynman rules computed for the 'high-energy' fields with the 'low-energy' fields regarded as fixed background values. (Recall in this split that high-energy fields can include the high-energy modes of particles with small masses.) Eq. (2.59) then says to construct Feynman graphs using propagators and vertices for the light fields defined from S_W, with $\Gamma_{\rm le}$ then obtained by computing all 1PI graphs in this low-energy theory. This combination reproduces the set of 1LPI graphs using the Feynman rules of the full theory.

In particular, since any tree graph with an internal line is one-particle reducible, this means that $\Gamma_{\rm le}[\ell] \simeq S_W[\ell]$ within the classical approximation (no loops). Furthermore, in the same approximation both are related to the classical action of the full theory by

$$\Gamma_{\rm le}[\ell] \simeq S_W[\ell] \simeq S[\ell, h_{\rm le}(\ell)] \qquad \text{(classical approximation)}, \tag{2.60}$$

where $h^a = h^a_{\rm le}(\ell)$ is obtained (in the classical approximation) by solving $(\delta S/\delta h^a)_{h=h_{\rm le}(\ell)} = 0$. But – as is seen more explicitly below – $\Gamma_{\rm le}[\ell]$ and $S_W[\ell]$ need no longer agree once loops are included.

It is the Wilson action that is the main tool used in the rest of this book. But why bother with S_W, given that $\Gamma_{\rm le}$ also captures all of the information relevant for low-energy predictions? As later examples show in more detail, in real applications it is often the Wilson action that is the easier to use, since it exploits the simplicity of the low-energy limit as early as possible. It plays such a central role because it has two very attractive properties.

First, it contains enough information to be useful. That is, any low-energy observable can be constructed from low-energy correlation functions (and so also from $\Gamma_{\rm le}$), and because $\Gamma_{\rm le}$ can be computed from S_W using only low-energy degrees of freedom, it follows that S_W carries all of the information necessary to extract the predictions for any low-energy observable.

But it is the second property that makes it such a practical tool: it doesn't contain too much information. That is, the Wilson action is the bare-bones quantity that contains all of the information about the system's high-energy degrees of freedom without polluting it with any low-energy details. Unlike $\Gamma_{\rm le}$, the Wilson action is constructed by integrating only over the high-energy degrees of freedom. This means that there is a maximal labour saving in exploiting any simplicity in S_W, since this simplicity is present *before* performing the rest of the low-energy part of the calculation.

Example: The Toy Model

To better understand how the Wilson action is defined, and how it is related to the low-energy 1LPI generator, it is useful to have a concrete example to examine in detail. Once again the toy model of Chapter 1 provides a useful place to start.

Since $\Gamma_{\rm le}$ and S_W only begin to differ beyond the classical approximation, imagine computing both $\Gamma_{\rm le}$ and S_W at one loop. According to its definition, the Feynman graphs contributing to S_W can involve only the high-energy degrees of freedom in the internal lines, while those contributing to $\Gamma_{\rm le}$ also involve virtual low-energy states.

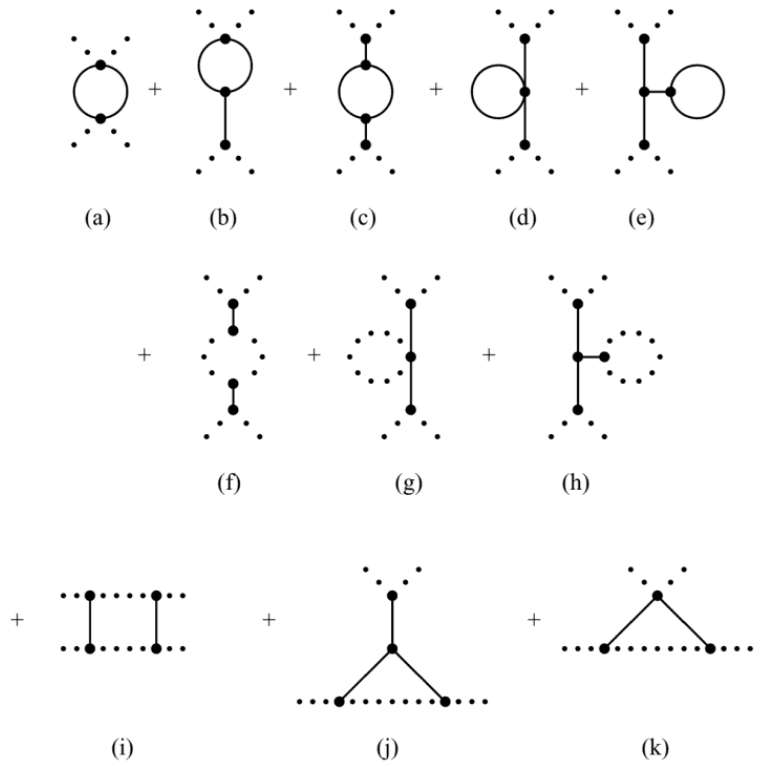

Fig. 2.4 One-loop graphs that contribute to the $(\partial_\mu \xi \partial^\mu \xi)^2$ interaction in the Wilson and 1LPI actions using the interactions of Eqs. (1.24) and (1.25). Solid (dotted) lines represent χ (and ξ) fields. Graphs involving wave-function renormalizations of ξ are not included in this list.

For both S_W and $\Gamma_{\rm le}$ the graphs can be one-particle reducible when cutting heavy-particle lines, but for $\Gamma_{\rm le}$ the graphs must be one-particle irreducible when light-particle lines are cut.

For concreteness' sake, for the toy model consider the one-loop contributions to the effective interaction

$$a \int \mathrm{d}^4 x \; (\partial_\mu \xi \partial^\mu \xi)^2, \tag{2.61}$$

in both $\Gamma_{\rm le}$ and S_W. The relevant Feynman graphs are shown in Fig. 2.4, using Feynman rules appropriate for the χ and ξ fields using the interactions given in Eqs. (1.24) and (1.25). (An equivalent set of graphs could also be written for the variables $\hat{\phi}_R$ and $\hat{\phi}_I$. Although both ultimately give the same physical results, they can differ in intermediate steps, and which is more useful depends on the application one has in mind.)

Since all of the internal lines for Feynman graphs (a) through (e) involve only the field χ, and since all modes of this field are classified as 'high-energy' – *c.f.* the discussion in §2.2.1 – all five of these graphs contribute to both S_W and $\Gamma_{\rm le}$. In order of magnitude, each contributes to the effective interaction an amount, as follows

$$a_{\text{graph(a)}} \propto \left(-\frac{1}{4v^2}\right)^2 L_1 = \frac{L_1}{16\,v^4}, \tag{2.62}$$

$$a_{\text{graph(b)}} \propto \left(-\frac{1}{4v^2}\right)\left(-\frac{\lambda v}{2\sqrt{2}}\right)\left(-\frac{1}{\sqrt{2}\,v}\right)\frac{L_1}{m_R^2} = -\frac{L_1}{16\,v^4}, \tag{2.63}$$

$$a_{\text{graph(c)}} \propto \left(-\frac{\lambda v}{2\sqrt{2}}\right)^2\left(-\frac{1}{\sqrt{2}\,v}\right)^2\frac{L_1}{m_R^4} = \frac{L_1}{16\,v^4}, \tag{2.64}$$

$$a_{\text{graph(d)}} \propto \left(-\frac{\lambda}{16}\right)\left(-\frac{1}{\sqrt{2}\,v}\right)^2\frac{L_2}{m_R^4} = -\frac{L_2}{32\lambda v^6}, \tag{2.65}$$

and

$$a_{\text{graph(e)}} \propto \left(-\frac{\lambda v}{2\sqrt{2}}\right)^2\left(-\frac{1}{\sqrt{2}\,v}\right)^2\frac{L_2}{m_R^6} = \frac{L_2}{16\lambda v^6}, \tag{2.66}$$

where the powers of $1/m_R^2$ come from the internal lines that do not appear within a loop, since these are evaluated at momenta much smaller than m_R. The contribution of the loop integrals themselves are of order

$$L_1 = \int^{\Omega}\frac{\mathrm{d}^4p}{(2\pi)^4}\left(\frac{1}{p^2+m_R^2}\right)^2 \propto \frac{1}{16\pi^2}\,\ln\left(\frac{\Omega}{m_R}\right)$$

$$\text{and}\quad L_2 = \int^{\Omega}\frac{\mathrm{d}^4p}{(2\pi)^4}\left(\frac{1}{p^2+m_R^2}\right) \propto \frac{\Omega^2}{16\pi^2}. \tag{2.67}$$

Here $\Omega \gg m_R$ is a cutoff that is introduced because the loop in the full theory is UV divergent. This divergence is ultimately dealt with by renormalizing the couplings of the microscopic theory; a point to be returned to in more detail shortly.

For the present purposes – keeping in mind that $m_R^2 = \lambda\, v^2$ – what is important is that graphs (a) through (c) clearly contribute to the coupling a (in both S_W and Γ_{le}) an amount of order $a_{1-\text{loop}} \propto L_1/v^4 \propto (1/4\pi v^2)^2 \ln\left(\Omega^2/m_R^2\right)$. Graphs (d) and (e) instead contribute an amount of order $a_{1-\text{loop}} \propto (1/4\pi v^2)^2\left(\Omega^2/m_R^2\right)$. Once the UV divergent function of Ω/m_R is renormalized into an appropriate coupling, the remaining coefficient for each of these loop contributions is suppressed by a factor of $\lambda/16\pi^2$ relative to the tree-level result, $a_{\text{tree}} = 1/(4\lambda\, v^4)$, in agreement with the discussion surrounding Eq. (2.24). As such, they all contribute to the coefficient a_2 of Eq. (1.16).

The difference between S_W and Γ_{le} arises in graphs (f) through (k), with S_W only receiving contributions where the momentum in the internal ξ propagators is larger than Λ, whereas there is no such a restriction for Γ_{le}. The contribution from graphs (i) through (k) when all loop momenta are large is of order

$$\begin{aligned}
a_{\text{graph(i)}} &\propto \left(-\frac{1}{\sqrt{2}\,v}\right)^4 L_1 = \frac{L_1}{4v^4}\\
a_{\text{graph(j)}} &\propto \left(-\frac{1}{\sqrt{2}\,v}\right)^3\left(-\frac{\lambda v}{2\sqrt{2}}\right)\frac{L_1}{m_R^2} = \frac{L_1}{8v^4}\\
a_{\text{graph(k)}} &\propto \left(-\frac{1}{\sqrt{2}\,v}\right)^2\left(-\frac{1}{4v^2}\right)L_1 = -\frac{L_1}{8v^4},
\end{aligned} \tag{2.68}$$

where the new loop integrals are also logarithmically divergent in the UV, and so up to numerical factors are again or order L_1 in size. These contribute to a an amount comparable to the size of graphs (a) through (c). By contrast, graphs (f) through (h) give the results,

$$\begin{aligned} a_{\text{graph(f)}} &\propto \left(-\frac{1}{\sqrt{2}\,v}\right)^4 \frac{L_3}{m_R^4} = \frac{L_3}{4v^4 m_R^4} \\ a_{\text{graph(g)}} &\propto \left(-\frac{1}{\sqrt{2}\,v}\right)^2 \left(-\frac{1}{4v^2}\right) \frac{L_3}{m_R^4} = -\frac{L_3}{8v^4 m_R^4} \\ a_{\text{graph(h)}} &\propto \left(-\frac{1}{\sqrt{2}\,v}\right)^3 \left(-\frac{\lambda v}{2\sqrt{2}}\right) \frac{L_3}{m_R^6} = \frac{L_3}{8v^4 m_R^4}, \end{aligned} \tag{2.69}$$

and so are of order $(1/v^4)(L_3/m_R^4)$, where the ξ loop has the ultraviolet behaviour

$$L_3 = \int^{\Omega} \frac{\mathrm{d}^4 p}{(2\pi)^4} \left(\frac{p^2}{p^2}\right)^k \propto \frac{\Omega^4}{16\pi^2}, \tag{2.70}$$

where $k = 2$ for graph (f) and $k = 1$ for graphs (g) and (h).

All of these graphs are dominated by large momenta (small wavelengths), which is why they diverge for large Ω. Although a more systematic treatment of these UV divergences (in particular how to treat them using dimensional regularization) is given in Chapter 3, there is a conceptual point to be made concerning their *lower* limit of integration. The point is that for graphs (f) through (k) this lower limit differs when computing Γ_{le} and S_W. For Γ_{le} the contributions to the effective interaction

$$a_{\text{le}} \int \mathrm{d}^4 x\, (\partial_\mu \xi \partial^\mu \xi)^2 \subset \Gamma_{\text{le}} \tag{2.71}$$

integrate over all momenta. But for S_W, in the contribution to

$$a_W \int \mathrm{d}^4 x\, (\partial_\mu \xi \partial^\mu \xi)^2 \subset S_W, \tag{2.72}$$

the integrations exclude momenta smaller than Λ (for which the internal ξ propagators are then 'light' degrees of freedom) that are not integrated out in the path integral representation of S_W.

Take, for instance, graphs (i) through (k) of Fig. 2.4. Since $\Lambda \ll m_R$, the predicted coefficient for Γ_{le} differs from the coefficient in S_W by an amount of order

$$a_{\text{le}} - a_W(\Lambda) \simeq \frac{1}{v^4} \int_0^{\Lambda} \frac{\mathrm{d}^4 p}{(2\pi)^4} \left(\frac{1}{p^2 + m_R^2}\right)^2 \left(\frac{p^2}{p^2}\right)^k \propto \frac{1}{16\pi^2 v^4} \left(\frac{\Lambda^4}{m_R^4}\right). \tag{2.73}$$

The suppression by powers of Λ/m_R ensures this difference is numerically small, although that turns out to be an artefact of this particular example. It is nonetheless conceptually important. In particular, the Λ-dependence of the right-hand side is associated with the Wilsonian coupling $a_W(\Lambda)$ because the scale Λ does not appear at all in the definition of a_{le} (which, after all, is defined in terms of integrations over modes at *all* scales).

How can these different values, $a_{\text{le}} \neq a_W$, lead to the same physical predictions for observables? The answer lies in Eq. (2.59), which states that Γ_{le} is obtained from S_W by integrating over the light degrees of freedom, using S_W rather than S as the action. It is this that fills in those parts of Γ_{le} not produced through loops

with

$$\mathfrak{L}_W = \sum_n c_n \, O_n(\phi, \partial\phi, \cdots), \tag{2.78}$$

a sum of powers of ϕ and its derivatives all evaluated at the same point, and (for relativistic systems) built to transform like a Lorentz scalar. If the low-energy theory is unitary then $\mathfrak{L}_W$ should also be real. The goal is to use dimensional analysis to identify the power of M appearing in each effective coupling, c_n.

This book uses 'fundamental' units for which[8] $\hbar = c = 1$, and so the (engineering) dimension of any quantity can be regarded as a power of energy or mass (see Appendix A for conversions between these and more conventional units). In these units the action, S_W, itself is dimensionless – *i.e.* has dimension (energy)^0 – or, more precisely, $S_W/\hbar$ is dimensionless. Similarly, time and space coordinates, t and $\mathbf{x}$, have dimension (energy)$^{-1}$, while derivatives like ∂_μ have dimension of energy. It is common to use the notation $[A] = p$ as a short form for the statement 'quantity A has dimension (energy)^p in units where $\hbar = c = 1$', and in this notation $[S_W] = 0$, $[x^\mu] = -1$ and $[\partial_\mu] = 1$.

Because the action is related to the lagrangian density by Eq. (2.77), in four spacetime dimensions it follows that $\mathfrak{L}_W$ has dimension (energy)^4 – *i.e.* $[\mathfrak{L}_W] = 4$ – because the measure, d^4x, satisfies $\left[\mathrm{d}^4x\right] = -4$. If M is the only relevant mass scale in the problem and if a particular interaction, $\mathcal{O}_n$, has dimension $[\mathcal{O}_n] = \Delta_n$, then because $[c_n\mathcal{O}_n] = 4$ it follows that $[c_n] = 4 - \Delta_n$, and so one expects

$$[\mathcal{O}_n] = \Delta_n \quad \Rightarrow \quad c_n = \frac{a_n}{M^{p_n}} \quad \text{with} \quad p_n = \Delta_n - 4, \tag{2.79}$$

where a_n is dimensionless. To the extent that it is M that sets the scale of c_n in this way (and much of the next chapter is devoted to showing that the low-energy theory often can be set up so that it is), higher-dimensional interactions in S_W should be expected to be more suppressed at low energies by higher powers of M.

Further progress requires a way to compute the dimension, Δ_n, of a given operator, $\mathcal{O}_n$. For weakly interacting systems dimensions can be computed in perturbation theory. To see how this works, consider a real scalar field, ϕ, and suppose the regime of interest is one where it is relativistic and very weakly interacting. This means the action $S_W = S_0 + S_{\rm int}$ is dominated by its kinetic term

$$S_0 = -\frac{1}{2}\int \mathrm{d}^4x \; \partial_\mu\phi \, \partial^\mu\phi, \tag{2.80}$$

while all remaining terms, lumped together into $S_{\rm int} = \int \mathrm{d}^4x \; \mathfrak{L}_{\rm int}$ with

$$\mathfrak{L}_{\rm int} = -\frac{m^2}{2}\,\phi^2 + c_{4,0}\,\phi^4 + \frac{c_{6,0}}{\Lambda^2}\,\phi^6 + \frac{c_{4,2}}{\Lambda^2}\,\phi^2\,\partial_\mu\phi\,\partial^\mu\phi + \frac{c_{4,4}}{\Lambda^4}(\partial_\mu\phi\,\partial^\mu\phi)^2 + \cdots, \tag{2.81}$$

are assumed to be small. In this expression a symmetry of the form $\phi \to -\phi$ is imposed (for simplicity) so that only terms involving an even power of ϕ need to be considered. Furthermore, appropriate powers of the cutoff, Λ, for the Wilsonian EFT

[8] When temperatures are considered, units are also chosen with Boltzmann's constant satisfying $k_B = 1$, so temperature can also be measured in units of energy.

are included explicitly in the coefficients for each effective interaction for reasons now to be explained.

Any effective coupling premultiplying S_0 is imagined to be removed by appropriately rescaling ϕ, with the choice of $\frac{1}{2}$ in Eq. (2.80) called canonical normalization. (The reasons for using this normalization are elaborated below, and in Appendix C.3.) Given this normalization, the dimension, $[\phi\,]$, of the (scalar) field ϕ is then determined by demanding that $[\mathfrak{L}_0] = 4$, and so

$$4 = [\partial_\mu\phi\,\partial^\mu\phi\,] = 2 + 2[\phi\,], \tag{2.82}$$

which implies that $[\phi\,] = 1$ (or ϕ has dimensions of energy). With this choice, an identical argument shows $\left[m^2\right] = 2$, and so m also has dimensions of energy (as expected, since m is interpreted as the ϕ-particle's mass).

A similar story applies to all of the other terms in $S_{\rm int}$, and shows that the factors of Λ in (2.81) are extracted so that the remaining effective couplings are dimensionless: $[c_{n,d}] = 0$. For later purposes notice that a term in $S_{\rm int}$ involving n powers of ϕ and d derivatives comes premultiplied by Λ^p with $p = 4 - n - d$, and so the infinite number of local interactions that are not written explicitly in (2.81) have effective couplings with only negative powers of Λ.

The goal is to identify the domain of validity of the assumed perturbative hierarchy between S_0 and $S_{\rm int}$. The next few paragraphs argue that perturbative arguments are appropriate when the dimensionless couplings are assumed to be small: $|c_{n,d}| \ll 1$, following arguments made in [19]. To this end, consider evaluating $S_w[\phi_k]$ at a wave-packet configuration $\phi_k(x) = f_k(x)\,e^{ikx}$, where $f_k(x)$ is a smooth envelope that is order unity for a spacetime region of linear size $2\pi/k$ in all four rectangular spacetime directions. For such a configuration spacetime derivatives are of order $\partial_\mu\phi_k \sim k_\mu\phi_k$ and the spacetime volume integral is of order $\int \mathrm{d}^4x \sim (2\pi/k)^4$, and so

$$\begin{aligned} S_w[\phi_k] &\sim \left(\frac{2\pi}{k}\right)^4 \left[k^2\,\phi_k^2 + m^2\phi_k^2 + c_{4,0}\,\phi_k^4 + \frac{c_{6,0}}{\Lambda^2}\,\phi_k^6 \right. \\ &\qquad \left. + c_{4,2}\left(\frac{k^2}{\Lambda^2}\right)\phi_k^4 + c_{4,4}\left(\frac{k^4}{\Lambda^4}\right)\phi_k^4 + \cdots\right] \\ &\sim (2\pi)^4\left[\varphi_k^2 + \frac{m^2}{k^2}\,\varphi_k^2 + c_{4,0}\,\varphi_k^4 + c_{6,0}\left(\frac{k^2}{\Lambda^2}\right)\varphi_k^6 \right. \\ &\qquad \left. + c_{4,2}\left(\frac{k^2}{\Lambda^2}\right)\varphi_k^4 + c_{4,4}\left(\frac{k^4}{\Lambda^4}\right)\varphi_k^4 + \cdots\right], \end{aligned} \tag{2.83}$$

where $\varphi_k := \phi_k/k$ is a new dimensionless variable. What is key about φ_k is that the path integral over φ_k would be dominated by $\varphi_k \lesssim O(1)$ in the absence of interactions[9] (*i.e.* when $S_w = S_0$). This conclusion that dominant configurations for φ_k are order unity is contingent on the coefficient of $(\partial\phi)^2$ in S_0 being order unity due to the choice of canonical normalization.

Perturbation theory in $S_{\rm int}$ requires $|S_{\rm int}| \ll |S_0|$ throughout the regime from which the path integral receives significant contributions, which from the above

[9] This is clearest if the problem is Wick-rotated to euclidean signature by going to imaginary time, so that $e^{iS} \to e^{-S}$. This estimate also ignores factors of 2π, though their inclusion somewhat broadens the domain of validity of perturbative methods.

considerations is the regime $\varphi_k \lesssim O(1)$. Consider first choosing k as large as possible: near the UV cutoff $|k^2| \sim \Lambda^2$. Since $\Lambda \gg m$, it follows that $|k^2| \gg m^2$, and so perturbation theory in this regime requires $|c_{n,d}| \ll 1$.

How does this conclusion change if k is now dialled down to smaller values? Since all interactions except the ϕ^2 and ϕ^4 interactions come pre-multiplied by positive powers of k^2/Λ^2, they become less and less important for smaller k. Interactions like this, which are less important at lower energies, are called *irrelevant*. Irrelevant interactions are also often called 'non-renormalizable'.[10]

By contrast, the ϕ^4 interaction is k-independent and so has strength controlled by $c_{4,0}$ for all k. Interactions like this, whose strength does not vary with k, are called *marginal*.

Finally, the mass (or ϕ^2) term is the only interaction that grows in importance for smaller k, the defining property of a *relevant* interaction. Once $|k^2| \lesssim m^2$, the mass term competes with S_0 and so changes the nature of the dominant path integral configurations. This nonrelativistic regime is, of course, important to many applications, and so is returned to in some detail as the topic of Part III. Relevant interactions are sometimes also called 'super-renormalizable', while marginal and relevant interactions taken together are called 'renormalizable'.

A similar story goes through for fields representing other spins at weak coupling. For instance, a field, ψ, describing a free relativistic spin-half particle with lagrangian density

$$S_{1/2} = -\int \mathrm{d}^4x\, \overline{\psi}\, \partial\!\!\!/\, \psi, \tag{2.84}$$

with $\partial\!\!\!/ = \gamma^\mu \partial_\mu$ for dimensionless Dirac matrices, γ^μ (see Appendices A.2.3 and C.3.2), must have dimension $[\psi] = 3/2$. The kinetic term for an electromagnetic potential, A_μ, is

$$S_1 = -\frac{1}{4}\int \mathrm{d}^4x\, F_{\mu\nu}F^{\mu\nu} = -\frac{1}{2}\int \mathrm{d}^4x\, \partial_\mu A_\nu(\partial^\mu A^\nu - \partial^\nu A^\mu), \tag{2.85}$$

and so the potential has dimension $[A_\mu] = 1$, while the field-strength satisfies $[F_{\mu\nu}] = 2$.

It is an important fact that all of the weakly interacting fields most commonly dealt with – such as ϕ, ψ and A_μ, as well as the derivative ∂_μ – have positive dimension, so more complicated interactions involving more powers of fields and/or derivatives always have higher and higher dimension. The corresponding effective couplings must therefore be proportional to more and more powers of $1/\Lambda$ (and so be less and less important at low energies). This is what ensures that all but a handful of effective interactions are irrelevant at low energies, in the sense defined above. Precisely how irrelevant they are for any given k depends on the power of k^2/Λ^2 involved, so it makes sense to organize any list of potential interactions in order of increasing operator dimension, since the leading terms on the list are likely to be the most important at low energies.

[10] The recognition that it is useful to classify interactions according to their dimension came early [6], as did the connection to renormalizable and non-renormalizable interactions [7].

From this point of view it is clear that the limited number of renormalizable (marginal and relevant) interactions with dimension $[\mathcal{O}_n] \leq 4$ are special, since their importance is not diminished (and can be enhanced) at low energies.

For concreteness' sake the introductory discussion given above is phrased in perturbation theory, so it is worth mentioning in passing that this is not, in principle, necessary. That is, although EFT methods always exploit expansions in ratios of energy scales (like E/m_R for the toy model), it is *not* a requirement of principle that the dimensionless couplings of the underlying UV theory (like λ for the toy model) are perturbatively small.[11] Although it goes beyond the scope of this chapter to show in detail, strong underlying couplings can change some of the detailed statements used above, such as by changing the dimension of the field to be $[\phi\,] = 1 + \delta$, with $|\delta| \to 0$ as these couplings are taken to zero. (Examples along these lines where δ is perturbatively small are considered in later sections.) Differences like δ are called 'anomalous dimensions' for the quantities involved. What counts in the dimensional arguments to follow is that the full scaling dimensions (including these anomalous contributions) are used, rather than the lowest-order 'naive' scaling dimension.

Example: The Toy Model

As applied to the toy model, because the kinetic terms for the two fields ξ and χ have the form of Eq. (2.80), the dimensions of both are $[\chi] = [\xi] = 1$. Using this with the full classical action, Eqs. (1.24) and (1.25), shows that $[\lambda\,] = 0$ and $[v] = [m_R] = 1$, as expected.

Applied to the Wilson action, the effective coupling a appearing in the interaction $\mathfrak{L}_W \supset a\,(\partial_\mu \xi \partial^\mu \xi)^2$ must have dimension (energy)$^{-4}$, consistent with its computed tree-level value $a_{\rm tree} = \lambda/4m_R^4$. This shows that at leading order it is explicitly the heavy scale m_R that plays the role of the dimensional parameter of the general discussion. Powers of $\Lambda \ll m_R$ also arise once loops are included, and subsequent sections are devoted to identifying which scale is important in any particular application.

2.4.2 Scaling

It is worth rephrasing the above discussion more formally in terms of a scaling transformation. This is useful for several reasons: because it sets up the use of renormalization-group methods; and because it provides a framework that is more easy to generalize in more complicated settings, such as in the nonrelativistic limit considered in Part III.

To this end, consider again the scalar-field Wilson action of Eqs. (2.80) and (2.81):

$$S_W[\phi(x); m, c_{4,0}, c_{6,0}, \cdots] = S_0[\phi(x)] + S_{\rm int}[\phi(x); m, c_{4,0}, \cdots], \tag{2.86}$$

[11] In an unfortunate use of language the breakdown of the low-energy (*e.g.* E/m_R) expansion has in some quarters come to be called 'strong coupling'. This is misleading because it can happen that the physics appropriate to these energies is weakly coupled, in the sense that it involves dimensionless couplings that are small. For this reason in this book 'strong coupling' never means 'breakdown of the low-energy limit', and it is reserved for situations where underlying dimensionless couplings (like λ in the toy model) are not small.

where the notation explicitly highlights the dependence on the effective couplings as well as on the field ϕ. Now perform the scale transformation, $x^\mu \to sx^\mu$ and $\phi(x) \to \phi(sx)$, where s is a real parameter. For a configuration like $\phi_k(x) \propto e^{ikx}$ this becomes $\phi_k(sx) \propto e^{iskx} \propto \phi_{sk}(x)$ and so taking $s \to 0$ corresponds to taking the infrared limit where $k \to 0$.

Inserting these definitions into S_0 gives

$$\begin{aligned} S_0[\phi(sx)] &= -\frac{1}{2}\int \mathrm{d}^4x\ \partial_\mu\phi(sx)\,\partial^\mu\phi(sx) \\ &= -\frac{1}{2}\int \frac{\mathrm{d}^4x'}{s^4}\ \Big[s^2\partial_{\mu'}\phi(x')\partial^{\mu'}\phi(x')\Big], \end{aligned} \tag{2.87}$$

in which the spacetime integration variable is changed from x^μ to $x'^\mu = sx^\mu$. This shows that S_0 remains unchanged if the scalar field variable is also rescaled according to $\phi(x) \to \phi_s(x) := \phi(x)/s$. Requiring $S_0[\phi(sx)] = S_0[\phi_s(x)]$ is natural in the weak-coupling limit, since this keeps fixed the configurations that dominate in the path integral over ϕ and ϕ_s.

With these choices, the effects of rescaling on the interaction terms can be read off, giving

$$\begin{aligned} S_{\rm int}[\phi(sx); m, c_{4,0}, c_{6,0}, c_{4,2}, \cdots] &= \int \mathrm{d}^4x' \left[-\frac{m^2}{2s^2}\,\phi_s^2 + c_{4,0}\phi_s^4 + \frac{s^2c_{6,0}}{\Lambda^2}\phi_s^6 \right. \\ &\qquad \left. + \frac{s^2c_{4,2}}{\Lambda^2}\,\phi_s^2(\partial_{\mu'}\phi_s\partial^{\mu'}\phi_s) + \cdots\right] \\ &= S_{\rm int}\left[\phi_s(x); \frac{m}{s}, c_{4,0}, s^2c_{6,0}, s^2c_{4,2}\ \cdots\right]. \end{aligned} \tag{2.88}$$

This shows that changes of scale can be compensated by appropriately rescaling all effective couplings. For instance, for an interaction $S_n[\phi; a_n] = \int \mathrm{d}^4x\ a_n\mathcal{O}_n[\phi] \in S_{\rm int}$, where $\mathcal{O}_n$ has engineering dimension $[\mathcal{O}_n] = \Delta_n$, the required scaling is

$$S_n[\phi(sx); a_n] = S_n[\phi_s(x); s^{p_n}a_n], \tag{2.89}$$

where $p_n = -[a_n] = \Delta_n - 4 =$ (so that $a_n = c_n/\Lambda^{p_n}$ for dimensionless c_n).

Rescalings can be regarded as motions within the space of coupling constants. This provides an alternative way to define relevant, marginal and irrelevant interactions. Since low energies correspond to $s \to 0$, an effective interaction is irrelevant if $p_n > 0$, it is marginal if $p_n = 0$ and it is relevant if $p_n < 0$. This definition clearly agrees with the one presented earlier.

2.5 Redundant Interactions ◊

It is generally useful to have in mind what are the most general possible kinds of effective interactions that can arise in a Wilson action at any given dimension, and it is tempting to think that this means simply listing all possible combinations of local interactions involving the given fields and their derivatives. In practice, such

$$S_w[\xi] = -\int \mathrm{d}^4x \left[\frac{1}{2}\,\partial_\mu\xi\partial^\mu\xi - a(\partial_\mu\xi\partial^\mu\xi)^2 - a'(\partial_\mu\xi\partial^\mu\xi)\Box(\partial_\mu\xi\partial^\mu\xi) - b\,(\partial_\mu\xi\partial^\mu\xi)^3 - b'(\partial_\mu\xi\partial^\mu\xi)\Box(\partial_\mu\xi\partial^\mu\xi)^3 + \cdots\right], \tag{2.97}$$

where

$$a = \frac{1}{m_R^4}\left[\frac{\lambda}{4} + O(\lambda^2)\right], \quad a' = \frac{1}{m_R^6}\left[\frac{\lambda}{4} + O(\lambda^2)\right], \quad b = \frac{1}{m_R^8}\left[0 + O(\lambda^3)\right] \tag{2.98}$$

and so on. The question is: are these the most general kinds of interactions possible? In particular, are there terms suppressed by only two powers of $1/m_R$? If not, why not?

Of course, symmetries restrict the form of S_w, and for the toy model symmetry under the shift $\xi \to \xi + \sqrt{2}\,v\,\omega$ requires ξ always to appear in S_w differentiated, so all interactions must involve at least as many derivatives as powers of ξ. Furthermore, to be a Lorentz scalar it must involve an even number of derivatives, so that all Lorentz indices can be contracted. But these conditions allow interactions that do not appear in Eq. (2.97). For instance, they allow the following effective interactions with dimension (energy)^6,

$$\begin{aligned}\mathfrak{L}_6 &= \frac{a_1}{m_R^2}(\partial_\mu\xi\Box\partial^\mu\xi) + \frac{a_2}{m_R^2}(\partial_\mu\partial_\nu\xi\,\partial^\mu\partial^\nu\xi)\\ &= \frac{(a_1 - a_2)}{m_R^2}(\partial_\mu\xi\Box\partial^\mu\xi) + \frac{a_2}{m_R^2}\partial_\nu(\partial_\mu\xi\,\partial^\mu\partial^\nu\xi),\end{aligned} \tag{2.99}$$

where (given the explicit dimensional factor of m_R^{-2}) a_1 and a_2 are dimensionless effective couplings.

The point is that both of these interactions are redundant, in the sense outlined above. The second line shows that one combination can be regarded as a total derivative, and so it is redundant to the extent that boundaries (or topology) do not play an important role in the physics of interest. The remaining term, involving the combination $\partial_\mu\xi\Box\partial^\mu\xi$, vanishes once evaluated using the equations of motion, $\Box\,\xi = 0$, for the lowest-order action. It can therefore be removed using the field redefinition

$$\xi \to \xi + \frac{a_2 - a_1}{m_R^2}\,\Box\,\xi, \tag{2.100}$$

since in this case

$$-\frac{1}{2}\,\partial_\mu\xi\partial^\mu\xi \to -\frac{1}{2}\,\partial_\mu\xi\partial^\mu\xi + \frac{a_2 - a_1}{m_R^2}\,(\partial_\mu\xi\Box\partial^\mu\xi), \tag{2.101}$$

up to terms of order $1/m_R^4$. This shows that (in the absence of boundaries) the first low-energy effects of virtual heavy particles arise at order $1/m_R^4$ rather than $1/m_R^2$.

What about interactions with dimension (energy)^8: is $(\partial_\mu\xi\,\partial^\mu\xi)^2$ the only allowed dimension-8 interaction? Since total derivatives are dropped, integration by parts can be used freely to simplify any candidate interactions. The most general possible Lorentz-scalar interactions invariant under constant shifts of ξ then are

$$\mathfrak{L}_8 = -a(\partial_\mu\xi\partial^\mu\xi)^2 - \frac{a_3}{m_R^4}\,(\partial_\mu\xi\Box^2\partial^\mu\xi), \tag{2.102}$$

where a_3 is a new dimensionless coupling and the freedom to integrate by parts is used (for the terms quadratic in ξ) to ensure all of the derivatives but one act on only one of the fields. The second term in (2.102) can be removed using the field redefinition

$$\xi \to \xi + \frac{a_3}{m_R^4}\,\Box^2\,\xi, \tag{2.103}$$

without changing the coefficient a (or coefficients of any lower-dimension interactions), showing that a captures all of the effects that can arise at order $1/m_R^4$.

2.6 Summary

This chapter lays one of the cornerstones for the rest of the book; laying out how effective lagrangians fit into the broader context of generating functionals and the quantum (1PI) action.[15] By doing so, it provides a constructive framework for defining and explicitly building effective actions for a broad class of physical systems.

The 1PI action is a useful starting point for this purpose because it already plays a central role in quantum field theory. It does so partly because it is related to the full correlation functions and the energetics of field expectation values in the same way that the classical action is related to the classical correlation functions and the energetics that fixes the values of classical background fields. The low-energy 1LPI action is the natural generalization of the 1PI action because it is constructed in precisely the same way, but with the proviso that it only samples slowly varying field configurations. As such, it contains all the information needed to construct any observable that involves only low-energy degrees of freedom.

In this chapter, the Wilson action, S_w, emerges as the minimal object for capturing the implications of high-energy degrees of freedom for the low-energy theory. The Wilson action is related to the 1LPI action in the low-energy theory in precisely the same way that the classical action is related to the 1PI generator in the full theory. Because S_w is obtained by integrating out only high-energy states, its interactions efficiently encode their low-energy implications. And because knowledge of S_w allows the calculation of the 1LPI action it contains all of the information required to compute any low-energy observable.

The chapter concludes with a few tools that will prove useful in later chapters when computing and using the Wilson action. The first tool is simply dimensional analysis, which classifies effective interactions based on their operator dimension (in powers of energy). More and more interactions exist with larger operator dimension, but it is the relatively few lower-dimension interactions in this classification that are more important at low energies. This chapter also describes the related renormalization-group scaling satisfied by the effective couplings. These express how the effective couplings differentially adjust as more and more modes are integrated out, lowering the energy scale Λ that differentiates low energies from high. (The next chapter also has more to say about this dimensional scaling and its utility for identifying which interactions are important at low energies.)

The second tool described in this chapter identifies classes of effective interactions that are redundant in the sense that they do not contribute at all to physical processes. They do not contribute for one of two

[15] In some of the earlier literature the quantum action is also called the effective action, unlike the modern usage, where effective action usually means the Wilson action.

reasons: either they are total derivatives and so are only sensitive to physics that depends in detail on the information at the system's boundaries; or because a change of variables exists that allows them to be completely removed.

Exercises

Exercise 2.1 Prove Eq. (2.6) starting from Eqs. (2.2) and (2.4).

Exercise 2.2 Draw all possible two-loop vacuum diagrams that contribute to $Z[J]$ in a theory involving both cubic and quartic interactions (such as a scalar potential $V(\phi) = g\phi^3 + \lambda\phi^4$ for scalar-field self-interactions). Which of these diagrams contribute to $W[J]$ and to $\Gamma[\varphi]$?

Exercise 2.3 For a scalar field self-interacting through the potential $V = g\phi^3 + \lambda\phi^4$ express Eq. (2.15) as a sum of Feynman graphs with two external lines. Draw all graphs that contribute out to two-loop order. Show how the disconnected graphs cancel in the result.

Exercise 2.4 Prove that the graphical expansion of $W[J]$, defined by $Z[J] = \exp\{iW[J]\}$, is obtained by simply omitting any disconnected graphs that contribute to $Z[J]$. Do so by showing that the exponential of the sum of all connected graphs reproduces all of the combinatorial factors in the sum over all (connected and disconnected) graphs. For this argument it is not necessary to assume that only cubic or quartic interactions arise in $S_{\rm int}$.

Exercise 2.5 Consider a single scalar field, φ, self-interacting through a scalar potential $U(\varphi)$. Evaluate Eq. (2.23) in one-loop approximation for φ specialized to a constant spacetime-independent configuration. To do so, use the identity $\ln\det\Delta = \mathrm{Tr}\ln\Delta$ and work in momentum space, for which $\Delta(p,p') = (p^2 + m^2 - i\epsilon)\,\delta^4(p-p')$, where $m^2 := U'' := \partial^2 U/\partial\varphi^2$. Evaluate the trace explicitly and Wick rotate to Euclidean signature ($p^0 = ip_E^4$) to derive the following expression

$$V_{\rm q}(\varphi) = U(\varphi) + \frac{1}{2}\int \frac{{\rm d}^4 p_E}{(2\pi)^4}\,\ln(p_E^2 + m^2) = U_\infty(\varphi) + \frac{1}{64\pi^2}\,m^4\ln\left(\frac{m^2}{\mu^2}\right),$$

for the quantum effective potential. Regulate the UV divergences using dimensional regularization (for which μ is the arbitrary scale: see Appendix A.2.4), and show that $U_\infty(\varphi) = U(\varphi) + A + B\,m^2(\varphi) + Cm^4(\varphi)$, where A, B and C are divergent constants in four spacetime dimensions. Show that if $U(\varphi) = U_0 + U_2\varphi^2 + U_4\varphi^4$ is quartic (and so renormalizable) then all divergences can be absorbed into the constants U_0, U_2 and U_4.

Exercise 2.6 Prove that the quantum effective potential is always convex [16, 21] when constructed about a stable vacuum. That is, show that for $0 \le s \le 1$

$$V_{\rm q}[s\varphi_1 + (1-s)\varphi_2] \le sV_{\rm q}(\varphi_1) + (1-s)V_{\rm q}(\varphi_2).$$

Exercise 2.7 Suppose the action $S[\phi]$ for a field theory is invariant under a symmetry transformation of the form $\delta\phi^a = \omega\,\zeta^a(\phi)$, where $\zeta^a(\phi)$ is a possibly nonlinear function and ω is an infinitesimal symmetry parameter. Show that the 1PI generator, $\Gamma[\varphi]$, is invariant under the symmetry $\delta\varphi^a = \omega\,\langle\zeta^a\rangle_J$,

where the matrix element is taken in the adiabatic vacuum in the presence of the current $J_a(\varphi)$ defined by Eq. (2.16). In the special case of a linear transformation, with $\zeta^a(\phi) = M^a{}_b\,\phi^b$ the condition $\langle\phi^a\rangle_J = \varphi^a$ implies that both $\Gamma[\varphi]$ and $S[\phi]$ share a symmetry with the same functional form.

The invariance condition $\delta\Gamma = 0$ for this transformation can be expressed as

$$\int \mathrm{d}^4x\,\langle\zeta^a(x)\rangle_J\,\frac{\delta\Gamma}{\delta\varphi^a(x)} = 0.$$

Since this is true for all $\varphi^a(x)$ repeated functional differentiation leads to a sequence of relations – called Taylor–Slavnov identities [22, 23] – amongst the 1PI correlation functions obtained by differentiating $\Gamma[\varphi]$.

Exercise 2.8 For the toy model of §1 draw all tree-level (no loops) Feynman graphs that can contribute to the effective interaction $\mathfrak{L}_W \supset c\,(\partial_\mu\xi\,\partial^\mu\xi)^4$ within the Wilson action. Evaluate these graphs and compute the effective coupling c at tree level.

Exercise 2.9 Construct the most general possible renormalizable relativistic interactions for a single real scalar field ϕ in $D = 4$ spacetime dimensions. Repeat this exercise for $D = 6$ spacetime dimensions. For $D = 4$ find the most general possible renormalizable relativistic interactions for a real scalar field coupled to a spin-half Dirac field ψ.

Exercise 2.10 For a real scalar field, ξ, subject to a shift symmetry, $\xi \to \xi + \text{constant}$, every appearance of ξ in the Wilson action must be differentiated at least once. Show that the most general effective interactions possible for such a field involving six or fewer derivatives is

$$\begin{aligned}\mathfrak{L}_W = &-\frac{1}{2}\,(\partial_\mu\xi\,\partial^\mu\xi) + a\,(\partial_\mu\xi\,\partial^\mu\xi)^2 + b\,(\partial_\mu\xi\,\partial^\mu\xi)^3\\ &+ c\,(\partial_\mu\xi\,\partial^\mu\xi)(\partial_\lambda\partial_\rho\xi)\,(\partial^\lambda\partial^\rho\xi),\end{aligned}$$

up to redundant interactions, for effective couplings a, b and c.

3 Power Counting and Matching

The previous chapter argues that the Wilson action captures the influence of virtual high-energy states on all low-energy observables, but a good number of questions remain to be addressed before it becomes a tool of practical utility. In particular, the Wilson action in principle contains an infinite number of interactions of various types, and although these are local (once expanded in inverse powers of the heavy scale) they ultimately involve arbitrarily many powers of the low-energy fields and their derivatives. What is missing is a simple way to identify systematically which interactions are required to calculate any given observable to a given order in the low-energy (and any other) expansions.

In principle, as argued in §2.4, what makes the Wilson action useful is dimensional analysis, which shows that more complicated interactions (with more derivatives or powers of fields) have coupling constants more suppressed by inverse powers of the physical heavy mass scale, M (like m_R in the toy model). This suggests that a dimension-Δ interaction can be ignored at low energies, E, provided effects of order $(E/M)^p$, with $p = \Delta - 4$, are negligible.

Sounds simple. Unfortunately, there is a confounding factor that complicates the simple dimensional argument. Although each use of an effective interaction within a Feynman graph costs inverse powers of a heavy scale like M, it is also true that the 4-momenta of virtual particles circulating within loops can include energies that are not small. This means that heavy scales can appear in numerators of calculations as well as in denominators, making it trickier to quantify the size of higher-order effects. Power counting – the main subject of this chapter – makes this argument more precise, and is the tool with which to identify which effective interactions are relevant to any particular order in the low-energy expansion.

To see how scales appear in calculations, for some purposes it is useful to explicitly track cutoffs, like Λ, that label the highest energies allowed to circulate within loops. Depending on the relative size of scales like M and Λ, it can happen that loop effects can cause effective couplings to acquire coefficients $c_n \propto \Lambda^{-p}$ rather than $c_n \propto M^{-p}$. For $p > 0$ these naively dominate because $\Lambda \ll M$. Estimates of the size of such corrections are discussed in this chapter in the section devoted to the 'exact renormalization group' (or 'exact RG').

But it is also true that Λ ultimately drops out of physical quantities, making its presence an unnecessary complication when formulating dimensional arguments. (Λ drops out of physical quantities because the precise split between low- and high-energy quantities is ultimately a book-keeping device for making calculations convenient, so Λ is not a physical scale. As this chapter shows, the disappearance of Λ in physical predictions happens in detail because the explicit Λ-dependence of the effective couplings in S_W cancels the Λ-dependence implicit in the definition of the low-energy path integral in which S_W is used.)

A second useful identity defines the number of loops, $\mathcal{L}$, for each (connected) graph:

$$\mathcal{L} = 1 + \mathcal{I} - \sum_n \mathcal{V}_n, \qquad \text{(definition of } \mathcal{L}\text{)}. \tag{3.5}$$

As mentioned around Eq. (2.24), this definition is motivated by the topological identity that applies to any graph that can be drawn on a plane, that states that $\mathcal{L} - \mathcal{I} + \sum_n \mathcal{V}_n = 1$ (which is the Euler number of a disc). In what follows, Eqs. (3.4) and (3.5) are used to eliminate $\mathcal{I}$ and $\sum_n f_n \mathcal{V}_n$.

Feynman Rules

The next step is to use the action of Eqs. (3.2) and (3.55) to construct the Feynman rules for the graph of interest. This is done here in momentum space, but since the argument to be made is in essence a dimensional one, it could equally well be made in position space.

Schematically, in momentum space the product of all of the vertices contributes the following factor to the amplitude:

$$\text{(Vertices)} = \prod_n \left[\mathrm{i}(2\pi)^4 \delta^4(p)\, \mathfrak{f}^4 \left(\frac{p}{M}\right)^{d_n} \left(\frac{1}{v}\right)^{f_n} \right]^{\mathcal{V}_n}, \tag{3.6}$$

where p generically denotes the various momenta running through the vertex. The product of all of the internal line factors gives the additional contribution:

$$\text{(Internal Lines)} = \left[-\mathrm{i} \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \left(\frac{M^2 v^2}{\mathfrak{f}^4}\right) \frac{1}{p^2 + m^2} \right]^{\mathcal{I}}, \tag{3.7}$$

where p again denotes the generic momentum flowing through the lines. m is the mass of the light particle (or their generic order of magnitude – for simplicity taken to be similar in size – should there be more than one light field) coming from the unperturbed term, Eq. (3.2). For the 'amputated' Feynman graphs relevant to scattering amplitudes and contributions to effective couplings in Γ_{le} the external lines are removed, and so no similar factors are included for external lines.

The momentum-conserving delta functions appearing in (3.6) can be used to perform many of the integrals appearing in (3.7) in the usual way. Once this is done, one delta function remains that depends only on external momenta, $\delta^4(q)$, and so cannot be used to perform additional integrals. This is the delta function that enforces the overall conservation of energy and momentum for the amplitude. It is useful to extract this factor once and for all, by defining the reduced amplitude, $\mathcal{A}_{\mathcal{E}}$, by

$$A_{\mathcal{E}}(q) = \mathrm{i}(2\pi)^4 \delta^4(q)\, \mathcal{A}_{\mathcal{E}}(q). \tag{3.8}$$

The total number of integrations that survive after having used all of the momentum-conserving delta functions is then $\mathcal{I} - \sum_n \mathcal{V}_n + 1 = \mathcal{L}$. This last equality uses the definition, Eq. (3.5), of the number of loops, $\mathcal{L}$.

3.1.2 Power Counting Using Cutoffs

The hard part in computing $\mathcal{A}_{\mathcal{E}}(q)$ is to evaluate the remaining multi-dimensional integrals. Things are not so bad if the only goal is to track how the result depends on

the scales ħ, M and v, however, since then it suffices to use dimensional arguments to estimate the size of the result. Since the integrals typically diverge in the ultraviolet, they are most sensitive to the largest momenta in the loop, and according to Eq. (2.59) this is set by the cutoff Λ. (The contributions of loops having momenta higher than Λ are the ones used when computing S_W itself from the underlying theory.)

This leads to the following dimensional estimate for the result of the integration

$$\int \cdots \int \left[\frac{\mathrm{d}^4 p}{(2\pi)^4}\right]^A \frac{p^B}{(p^2+m^2)^C} \sim \left(\frac{1}{4\pi}\right)^{2A} \Lambda^{4A+B-2C}. \tag{3.9}$$

For the purposes of counting 2π's, a factor of π^2 is included for each $\mathrm{d}^4 p$ integration corresponding to the result of performing the three angular integrations.[3]

The idea is to Taylor expand the amplitude $\mathcal{A}_{\mathcal{E}}(q)$ in powers of external momentum, q, using Eq. (3.9) to estimate the size of the coefficients. Schematically,

$$\mathcal{A}_{\mathcal{E}}(q) \simeq \sum_{\mathcal{D}} \mathcal{A}_{\mathcal{E}\mathcal{D}}\, q^{\mathcal{D}}, \tag{3.10}$$

where the coefficients require an estimate for the following integral

$$\begin{aligned}\mathcal{A}_{\mathcal{E}\mathcal{D}}\, q^{\mathcal{D}} &\propto \int \cdots \int \left[\frac{\mathrm{d}^4 p}{(2\pi)^4}\right]^{\mathcal{L}} \frac{1}{(p^2+m^2)^{\mathcal{I}}} \left(\frac{q}{p}\right)^{\mathcal{D}} \prod_n p^{d_n \mathcal{V}_n} \\ &\sim \left(\frac{1}{4\pi}\right)^{2\mathcal{L}} \left(\frac{q}{\Lambda}\right)^{\mathcal{D}} \Lambda^{4\mathcal{L}-2\mathcal{I}+\sum_n d_n \mathcal{V}_n}.\end{aligned} \tag{3.11}$$

Combining this with the powers of ħ, M and v given by the Feynman rules then gives, after using identities (3.4) and (3.5),

$$\mathcal{A}_{\mathcal{E}\mathcal{D}}\, q^{\mathcal{D}} \sim \hbar^4 \left(\frac{1}{v}\right)^{\mathcal{E}} \left(\frac{q}{\Lambda}\right)^{\mathcal{D}} \left(\frac{M\Lambda}{4\pi\hbar^2}\right)^{2\mathcal{L}} \left(\frac{\Lambda}{M}\right)^{2+\sum_n (d_n-2)\mathcal{V}_n}. \tag{3.12}$$

This is the main result of this section, whose properties are now explored.

A reality check for this formula comes if it is applied to the simplest graph of all (see Fig. 3.1): one including no internal lines (so $\mathcal{L} = 0$) and only a single vertex, $n = n_0$, with $f_{n_0} = \mathcal{E}$ external lines and $d_{n_0} = \mathcal{D}$ derivatives (so $\sum_n \mathcal{V}_n = 1$ and $\sum_n d_n \mathcal{V}_n = \mathcal{D}$). In this case, (3.12) implies that the amplitude depends on the scales M, Λ and ħ in precisely the same way as does the starting lagrangian (3.2): $\mathcal{A}_{\mathcal{E}\mathcal{D}}\, q^{\mathcal{D}} \sim \hbar^4 (1/v)^{\mathcal{E}} (q/M)^{\mathcal{D}}$.

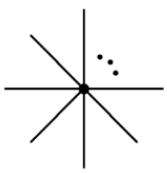

Fig. 3.1 The graph describing the insertion of a single effective vertex with $\mathcal{E}$ external lines and no internal lines.

[3] The factor of π^2 is clearest to see if momenta are Wick rotated to euclidean signature, since it there represents the volume of the unit 3-sphere corresponding to the integration over the three directions taken by a 4-vector.

A second reality check applies (3.12) to the special case where all scales are set by Λ (*i.e.* $\dagger = M = v = \Lambda$) since this corresponds to the choices made in the dimensional arguments of §2.4.1. In this limit (3.12) becomes

$$\mathcal{A}_{\mathcal{E}\mathcal{D}}\, q^{\mathcal{D}} \sim \Lambda^4 \left(\frac{1}{\Lambda}\right)^{\mathcal{E}} \left(\frac{q}{\Lambda}\right)^{\mathcal{D}} \left(\frac{1}{4\pi}\right)^{2\mathcal{L}}. \tag{3.13}$$

Since $\mathcal{E}$ and $\mathcal{D}$ are fixed by external characteristics (the total number of external legs and power of q in the final answer), the last factor is the only part of (3.13) that changes for more and more complicated diagrams that share these external properties. This factor simply says that $1/(4\pi)^2$ is the price for each additional loop; that is to say, all graphs with a fixed number of loops are similar in size assuming the (unwritten) dimensionless couplings – *i.e.* the $\hat{c}_n$ of (3.2) – are also similar in size. Furthermore, perturbation theory in this regime is ultimately controlled by the ratio of the dimensionless $\hat{c}_n$ compared with $16\pi^2$. The statement that perturbation theory applies for small enough $\hat{c}_n$ agrees with (and refines) the simple estimate of §2.4.1.

Validity of the Perturbative Expansion

More broadly, Eq. (3.12) outlines the domain of validity of the perturbative expansion itself. If the contribution estimated in Eq. (3.12) is small for all choices of $\mathcal{D}$, $\mathcal{E}$, $\mathcal{L}$ and $\mathcal{V}_n$, then this ensures that the perturbative expansion used in its derivation is a good approximation, particularly if more complicated graphs (higher $\mathcal{L}$ and $\mathcal{V}_n$) are more suppressed than less complicated ones. Conversely, if there is a choice for $\mathcal{D}$, $\mathcal{E}$, $\mathcal{L}$ and $\mathcal{V}_n$ for which Eq. (3.12) is not small, then the perturbative expansion fails unless some other small parameter – such as the dimensionless couplings $\hat{c}_n$ of (3.2) – can be found that can systematically suppress more complicated graphs. Furthermore, since the semiclassical expansion is an expansion in loops,[4] the perturbative expansion becomes a semiclassical expansion when it is the $\mathcal{L}$-dependent factor that controls perturbation theory.

Eq. (3.12) shows that there are three small quantities whose size can help control perturbative corrections: q/Λ, Λ/M and $\Lambda M/4\pi\dagger^2$. Some remarks are in order for each of these.

Derivative Expansion

Consider first the factor q/Λ, which controls the suppression of higher powers of external momenta. There is no question that $q/\Lambda \ll 1$ since the entire construction of the low-energy theory presupposes Λ can be chosen much smaller than the scale of the heavy physics that is being integrated out, but much larger than the low energies, $q \simeq E$ of applications. But when $\Lambda \ll M$ the ratio q/Λ is much bigger than q/M, even if both are separately very small. Eq. (3.12) therefore shows that it could happen that the derivative expansion in physical quantities (like scattering amplitudes) and in quantities like S_w ends up being controlled by powers of q/Λ rather than the powers of q/M assumed for the original action, (3.2). This point is returned to in §3.1, but it means that (all other things being equal) derivatives like to be suppressed by the

[4] The connection between loops and the semiclassical expansion is established in the discussion surrounding (2.24). In essence, the semiclassical expansion counts loops because it is an expansion in powers of $\hbar$, which appears as an overall factor in the quantity $S/\hbar$ within the path integral.

lowest possible UV scale available: in this case Λ. The case $M \simeq \Lambda$ is one that should be taken seriously in what follows.

Field Expansion

Notice that the same thing does *not* happen for the expansion in powers of ϕ/v, since the factor $(1/v)^{\mathcal{E}}$ assumed to appear in S_w does not get converted into a power like $(1/\Lambda)^{\mathcal{E}}$ or $(1/M)^{\mathcal{E}}$ in $\mathcal{A}_{\mathcal{E}}$. This means that it is consistent to have the scale v that controls the field expansion be systematically different from the scales Λ or M that control the derivative expansion. These scales are logically independent because, in general, large fields need not imply large energies, so small-field expansions are not necessarily required in a low-energy limit.

Loop Expansion

Next consider the suppressions coming from loops and vertices. Notice first that if $M \simeq \Lambda$ then the only systematic perturbative suppression in (3.12) comes from loops, due to the factor $(\Lambda^2/4\pi\mathfrak{f}^2)^{2\mathcal{L}}$. If all dimensionless couplings are order unity then perturbation theory in this limit is revealed to be a semiclassical expansion (*i.e.* controlled purely by the number of loops) whose validity rests on the assumption $4\pi\mathfrak{f}^2 \gg \Lambda^2$.

This condition is automatically satisfied in the regime $\Lambda \ll M$ provided that $\mathfrak{f} \gtrsim M$ is also true. It is a much stronger condition on Λ, however, if $\mathfrak{f}$ should be much smaller than M. In the particularly interesting case where $\mathfrak{f}^2 \simeq Mv$ (corresponding to canonical normalization in (3.2)) the loop-suppression factor becomes $(M\Lambda/4\pi\mathfrak{f}^2)^{2\mathcal{L}} \simeq (\Lambda/4\pi v)^{2\mathcal{L}}$.

Dangerous Non-Derivative Interactions

Finally, consider the final factor in (3.12). If $\Lambda \lesssim M$ the power of $(\Lambda/M)^{\mathcal{P}}$ appearing in Eq. (3.12) represents a suppression rather than an enhancement, provided the power

$$\mathcal{P} := 2 + \sum_n (d_n - 2)\mathcal{V}_n, \tag{3.14}$$

is non-negative. It is this factor that expresses the suppression of the effects of interactions involving three or more derivatives.

Lorentz invariance often requires d_n to be even (*e.g.* for scalar fields), and in this case it is only interactions with no derivatives at all ($d_n = 0$) for which $\mathcal{P}$ can be negative. These interactions are potentially dangerous in that they can in principle allow an *enhancement* in $\mathcal{A}_{\mathcal{E}}$ when $\Lambda \ll M$. When such interactions exist a more detailed estimate is required to see whether higher-order effects really are suppressed.

As an example of non-derivative interactions, imagine the low-energy field, ϕ, self-interacts through a scalar potential,

$$S_w \supset -\int \mathrm{d}^4x\, V(\phi), \tag{3.15}$$

where

$$V(\phi) := \mathfrak{f}_v^4 \sum_r g_r \left(\frac{\phi}{v}\right)^r. \tag{3.16}$$

Here, g_r are dimensionless couplings and $\mathfrak{f}_V^4$ is the typical potential energy density associated with fields of order $\phi \simeq v$. If $\mathfrak{f}_V \neq \mathfrak{f}$ then repeating the above power counting argument shows that each appearance of a vertex drawn from $V(\phi)$ contributes an additional factor $g_r(\mathfrak{f}_V/\mathfrak{f})^4$, modifying Eq. (3.12) to

$$\mathcal{A}_{\mathcal{E}\mathcal{D}}\, q^{\mathcal{D}} \sim \frac{\Lambda^2 \mathfrak{f}^4}{M^2} \left(\frac{1}{v}\right)^{\mathcal{E}} \left(\frac{q}{\Lambda}\right)^{\mathcal{D}} \left(\frac{M\Lambda}{4\pi\,\mathfrak{f}^2}\right)^{2\mathcal{L}} \times \left\{\prod_r \left[g_r \left(\frac{\mathfrak{f}_V^4 M^2}{\mathfrak{f}^4 \Lambda^2}\right)\right]^{\mathcal{V}_{0,r}}\right\} \left\{\prod_{d\geq 2} \prod_{i_d} \left(\frac{\Lambda}{M}\right)^{(d-2)\mathcal{V}_{d,i_d}}\right\}, \tag{3.17}$$

where the product over vertex labels is now subdivided into groups involving precisely d derivatives: $\{n\} = \{d, i_d\}$, with $i_0 = r$.

This last expression shows that the potentially hazardous enhancement factor, $(M/\Lambda)^{2\mathcal{V}_{0,r}}$, need not be dangerous if the potential energy density in the low-energy theory is sufficiently small relative to the generic energy density, $\mathfrak{f}_V^4/\mathfrak{f}^4 \lesssim \Lambda^2/M^2$. But if this is not so, generic non-derivative interactions can be legitimate obstructions to having a well-behaved low-energy limit, a point that must be checked on a case-by-case basis.

Example: The Toy Model

The Wilson action for the toy model, Eq. (2.97), is a special case of the general form assumed in Eqs. (3.2), with $M = m_R$ and $\mathfrak{f}^2 = m_R v$ and no zero-derivative interactions for the low-energy field ξ. For this special case the estimate Eq. (3.12) becomes

$$\mathcal{A}_{\mathcal{E}\mathcal{D}}\, q^{\mathcal{D}} \sim v^2\Lambda^2 \left(\frac{1}{v}\right)^{\mathcal{E}} \left(\frac{q}{\Lambda}\right)^{\mathcal{D}} \left(\frac{\Lambda}{4\pi\, v}\right)^{2\mathcal{L}} \left(\frac{\Lambda}{m_R}\right)^{\sum_n (d_n-2)\mathcal{V}_n}, \tag{3.18}$$

which neglects dimensionless factors that come as a series in powers of the coupling λ.

Notice that Eq. (3.18) agrees with the calculations of the previous chapter for the size of tree and loop contributions to the effective vertex, $a_{\rm le}(\partial_\mu \xi\, \partial^\mu \xi)^2$ appearing in $\Gamma_{\rm le}$, for which $\mathcal{E} = \mathcal{D} = 4$. For instance, consider the three contributions of Fig. 2.5. Figure (a) has $\mathcal{L} = 0$ and $\mathcal{V}_{4,4} = 1$, and so Eq. (3.18) gives

$$\begin{aligned} \delta\, a_{\rm le} &\sim v^2\Lambda^2 (1/v)^4 (1/\Lambda)^4 (\Lambda/m_R)^2 \\ &\sim 1/(v^2 m_R^2) = \lambda/m_R^4\,; \end{aligned} \tag{3.19}$$

Figure (b) has $\mathcal{L} = 1$ and $\mathcal{V}_{4,4} = 2$, and so Eq. (3.18) gives

$$\begin{aligned} \delta\, a_{\rm le} &\sim v^2\Lambda^2 (1/v)^4 (1/\Lambda)^4 (\Lambda/4\pi v)^2 (\Lambda/m_R)^4 \\ &\sim [1/(16\pi^2 v^4)](\Lambda/m_R)^4\,; \end{aligned} \tag{3.20}$$

Figure (c) has $\mathcal{L} = 1$ and $\mathcal{V}_{6,6} = 1$, and so Eq. (3.18) gives

$$\begin{aligned} \delta\, a_{\rm le} &\sim v^2\Lambda^2 (1/v)^4 (1/\Lambda)^4 (\Lambda/4\pi v)^2 (\Lambda/m_R)^4 \\ &\sim [1/(16\pi^2 v^4)](\Lambda/m_R)^4. \end{aligned} \tag{3.21}$$

These all agree with the estimates performed in §2.3.

But Eq. (3.18) contains much more information than just this. Most importantly, since the symmetry $\xi \to \xi$ + constant implies that there are no interactions with

where the $O_n^{(d_n,f_n)}$ describe all possible local interactions involving f_n powers of the fields and d_n derivatives, and the last line specializes to a single scalar field for concreteness' sake.

The power counting result, (3.12), provides an estimate of the size of amputated Feynman graphs built using these interactions, involving fields defined below the cutoff Λ. But this also determines the Λ-dependence of perturbative corrections to the couplings in $S_{w\text{int}}$ because the direct contribution of interactions represented by graphs like Fig. 3.1 must precisely cancel the Λ-dependence coming from loop graphs, as estimated by Eq. (3.12).

For the action of (3.26) the estimate (3.12) gives the contribution to the effective coupling of a term in S_w involving $\mathcal{E}$ powers of ϕ and $\mathcal{D}$ derivatives to be

$$\delta\left[\hat{c}_n v^2\Lambda^2\left(\frac{1}{v}\right)^{\mathcal{E}}\left(\frac{1}{\Lambda}\right)^{\mathcal{D}}\right] \sim v^2\Lambda^2\left(\frac{1}{v}\right)^{\mathcal{E}}\left(\frac{1}{\Lambda}\right)^{\mathcal{D}}\left(\frac{\Lambda}{4\pi\, v}\right)^{2\mathcal{L}}, \tag{3.27}$$

and so $\delta\hat{c}_n$ acquires corrections from $\mathcal{L}$-loop graphs that are of order

$$\delta\,\hat{c}_n \sim \left(\frac{\Lambda}{4\pi\, v}\right)^{2\mathcal{L}} \times \quad \text{(combinations of other } \hat{c}_n\text{'s)}. \tag{3.28}$$

If $v \gtrsim \Lambda$, this shows that it is consistent to have the $\hat{c}_n$'s all be generically $\lesssim 1$ for all Λ. Some couplings can be much smaller than this if $v \gg \Lambda$ (or other hierarchies like powers of Λ/M or small dimensionless couplings are buried in the $\hat{c}_n$'s), provided these additional suppressions preserve any initially small values.

Log Running vs Power-Law Running

The exact cutoff-dependent renormalization group is not pursued further in this book, since the focus here is instead on more practical methods of approximate calculation. Before leaving the subject, though, it is useful to address a conceptual question and by so doing contrast the implications of logarithmic and power-law running of effective couplings, $\hat{c}_n(\Lambda)$, as Λ is varied.

The conceptual question is this: why does one care how couplings run with Λ if Λ itself ultimately does not appear in any physical results? It is emphasized many times in this book that Λ enters calculations purely as a convenient book-keeping device: it is useful to organize calculations by scales and integrate out physics one scale at a time. But in the end, physical quantities are obtained only after *all* scales are integrated out, after which the arbitrary separations between these scales disappear.

This section follows [28] to argue that understanding the running of couplings in S_w is useful to the extent that it helps track the dependence of physical quantities on large physical ratios of scale, M/m. In particular, there is often a precise connection between logarithmic dependence of low-energy quantities on the cutoff Λ and a logarithmic dependence on physical scales. The analogous connection is only qualitative for power-law dependence, however, and so is usually less useful.

To see why this is so, imagine a system characterized by two very different scales, $m \ll M$, such as the masses of two different particles. Further imagine that there is a physical quantity, A, whose dependence on M/m happens to be logarithmic, so

$$A = a_0 \ln\left(\frac{M}{m}\right) + a_1, \tag{3.29}$$

for some calculable constants a_0 and a_1. If both of these constants are similar in size, then the value of a_0 can be important in practice since the large logarithm can make it dominate numerically in the total result.

Next, suppose a Wilsonian calculation is performed that divides the contributions to A coming from physics above and below the scale Λ, with $m \ll \Lambda \ll M$. How does the large logarithm get into the low-energy part of the theory, given that it depends on scales that lie on opposite sides of Λ? Typically, this happens as follows:

$$\begin{aligned} A_{\rm le} &= a_{0\,\rm le} \ln\left(\frac{\Lambda}{m}\right) + a_{1\,\rm le} \\ A_{\rm he} &= a_{0\,\rm he} \ln\left(\frac{M}{\Lambda}\right) + a_{1\,\rm he} \\ \text{so that} \quad A &= A_{\rm le} + A_{\rm he} = a_0 \ln\left(\frac{M}{m}\right) + a_1. \end{aligned} \tag{3.30}$$

The requirement that Λ cancels implies that $a_{0\,\rm le} = a_{0\,\rm he}$ and then having the results agree with the full theory implies that $a_{0\,\rm le} = a_{0\,\rm he} = a_0$ and $a_1 = a_{1\,\rm le} + a_{1\,\rm he}$. What is significant is that the coefficient, a_0, of the logarithm in the full answer is calculable purely within the low-energy theory because Λ-cancellation dictates that $a_{0\,\rm le} = a_0$.

The same is not so for power-law dependence. Suppose, for example, that another observable, B, is computed that depends quadratically on masses, so

$$B = b_0 M^2 + b_1 m^2. \tag{3.31}$$

Again the coefficient b_0 is of practical interest since the large size of M can make this term dominate numerically. In this case, the low- and high-energy parts of the calculations instead are

$$\begin{aligned} B_{\rm le} &= b_{0\,\rm le} \Lambda^2 + b_{1\,\rm le} m^2 + \cdots \\ B_{\rm he} &= b_{0\,\rm he} M^2 + b_{1\,\rm he} \Lambda^2 + \cdots \\ \text{so that} \quad B &= B_{\rm le} + B_{\rm he} = b_0 M^2 + b_1 m^2, \end{aligned} \tag{3.32}$$

with $b_0 = b_{0\,\rm he}$ and $b_1 = b_{1\,\rm le}$, while Λ-cancellation requires $b_{0\,\rm le} + b_{1\,\rm he} = 0$.

Evidently, the b_0 term cannot be computed purely within the low-energy part of a Wilsonian calculation simply by tracking the dependence on Λ^2, unlike the way in which the $\ln\Lambda$ terms reproduce the value for a_0. This is a fairly generic result: quantitative predictions for quantities like b_0 really require detailed knowledge of the UV theory and cannot be computed using the low-energy Wilsonian theory alone. But logarithms can often be inferred purely from within the low-energy Wilsonian perspective. For this reason considerable attention is given to renormalization-group methods that allow efficient extraction of large logarithms using Wilsonian EFTs.

Method of Regions I

The cancellations of powers of cutoff and the utility of logarithms can be promoted to a useful tool – sometimes called the *method of regions* [29] – for estimating

Feynman integrals in situations where different integration regions compete in their contributions to the final result.

To illustrate the method, follow [30] and consider the integral

$$\int_0^\infty \frac{k\,\mathrm{d}k}{(k^2+m^2)(k^2+M^2)} = \frac{\ln(M/m)}{M^2-m^2} \simeq \frac{\ln(M/m)}{M^2}\left[1+\frac{m^2}{M^2}+\cdots\right], \tag{3.33}$$

and imagine trying to extract the dominant small-m/M expansion without first evaluating the full integral. Naively, one simply Taylor expands the integrand in powers of m and integrates term-by-term, but this has the problem that each term involves an integral that diverges in the infrared

$$\int_0^\infty \frac{k\,\mathrm{d}k}{k^2(k^2+M^2)}\left[1-\frac{m^2}{k^2}+\cdots\right], \tag{3.34}$$

as might be expected given that the full result (3.33) is not analytic at $m = 0$.

A better procedure instead separates the integral into two regions, an IR region $0 < k < \Lambda$ and a UV region $k > \Lambda$, and expands the integrand differently in each. For the IR the integrand is expanded with $k^2 \sim m^2 \ll M^2$, while in the UV one takes instead $m \ll k \sim M$. This leads to the result $\mathcal{I}(m,M) = \mathcal{I}^{IR}(m,M,\Lambda) + \mathcal{I}^{UV}(m,M,\Lambda)$, with

$$\mathcal{I}^{IR} = \int_0^\Lambda \frac{k\,\mathrm{d}k}{(k^2+m^2)M^2}\left[1-\frac{k^2}{M^2}+\cdots\right] \simeq \frac{\ln(1+\Lambda^2/m^2)}{2M^2} - \frac{\Lambda^2}{2M^4}+\cdots, \tag{3.35}$$

and

$$\mathcal{I}^{UV} = \int_\Lambda^\infty \frac{k\,\mathrm{d}k}{k^2(k^2+M^2)}\left[1-\frac{m^2}{k^2}+\cdots\right] \simeq \frac{\ln(1+M^2/\Lambda^2)}{2M^2} + O(m^2). \tag{3.36}$$

Once these last two formulae are summed, all Λ-dependence cancels (as it must), leaving residual logarithms of M/m as outlined above that reproduce the expansion of (3.33). Large logarithms like $\ln(M/m)$ are ultimately leftovers from the cancellation between the IR divergence of $\mathcal{I}^{UV}$ and the UV divergence of $\mathcal{I}^{IR}$.

3.1.4 Rationale behind Renormalization ◇

The above discussion about integrating out high-energies also provides physical insight into the entire framework of renormalization. This is because a central message is that the scale Λ is ultimately a calculational convenience that drops out of all physical quantities. In detail, Λ drops out because of a cancellation between: (*i*) the explicit Λ-dependence of the cutoff on the limits of integration for virtual low-energy states in loops, and (*ii*) the cutoff-dependence that is implicitly contained within the effective couplings of Ω_W.

But this cancellation is eerily reminiscent of how UV divergences are traditionally handled within *any* renormalizable theory, and in particular for the underlying UV theory from which S_W is calculated. The entire renormalization program relies on any UV-divergent cutoff-dependence arising from regulated loop integrals being cancelled by the regularization dependence of the counterterms of the renormalized lagrangian. There are, however, the following important differences.

1. The cancellations in the effective theory occur even though Λ is not sent to infinity, and even though $\mathfrak{L}_W$ contains arbitrarily many terms that are not renormalizable in the traditional sense.
2. The cancellation of regularization dependence in the traditional picture of renormalization appears completely ad-hoc and implausible, while the cancellation of Λ from observables within the effective theory is essentially obvious. It is obvious due to the fact that Λ was only introduced as an intermediate step in a calculation, and so cannot survive uncancelled in the answer.

This resemblance is likely not accidental. It suggests that rather than considering a model's classical lagrangian as something pristine or fundamental, it is better regarded as an effective lagrangian obtained by integrating out still-more-microscopic degrees of freedom. The cancellation of ultraviolet divergences within the renormalization program is, within this interpretation, simply the usual removal of an intermediate step in a calculation to whose microscopic part we are not privy.

This is the modern picture of what renormalization really means. When discovering successful theories, what is found is not a 'classical' action, to be quantized and compared with experiment. What is found is really a Wilsonian action describing the low-energy limit obtained by integrating out high-energy degrees of freedom in some more fundamental theory that describes what is really going on at much, much higher energies.

It is this Wilsonian theory, itself potentially already containing many high-energy quantum effects, whose low-energy states are quantized and compared with observations. Physics progresses by successively peeling back layer after layer of structure in nature, and our mathematics describes this through a succession of Wilsonian descriptions with ever-increasing accuracy.

This is how real progress often happens in science. Efforts to solve concrete practical questions – in this case, a desire to exploit hierarchies of scale as efficiently as possible – can ultimately provide deep insights about foundational issues – in this case, about what it is that is really achieved when new fundamental theories (be it Maxwell's equations, General Relativity or the Standard Model) are discovered.

3.2 Power Counting and Dimensional Regularization ♢

As previous sections make clear, there is a lot of freedom of definition when setting up a Wilson action: besides the freedom to make field redefinitions, there are also all the details of precisely how to differentiate between scales above and below Λ. Physical results do not depend on any of these choices at all since observables are independent of field redefinitions and are blind to the details of a regularization scheme. This freedom should be exploited to make the Wilson action as useful as possible for practical calculations. In particular, it should be used to optimize the efficiency with which effective interactions and Feynman graphs can be identified that completely capture the contributions to low-energy processes at any fixed order in low-energy expansion parameters like q/M.

Though instructive, the power counting analysis of the previous section does not yet do this, due to the appearance in all estimates of the cutoff Λ. Since Λ ultimately cancels in all physical quantities, it is inconvenient to have to rely on it when estimating the size of contributions from different interactions in the low-energy Wilson action. For this reason most practical applications (and most of the rest of this book) define the Wilson action using dimensional regularization rather than cutoffs [31, 32]. Dimensional regularization is useful because it is both simple to use and preserves more symmetries than do other regularization schemes. This section explores how this can be done.

3.2.1 EFTs in Dimensional Regularization

At first sight, it seems impossible to define a Wilson action in terms of dimensional regularization at all. After all, the entire purpose of the Wilson action is to efficiently encapsulate the high-energy part of a calculation, for later use in a variety of low-energy applications. This seems to *require* something like Λ to distinguish high energies from low energies. By contrast, although dimensional regularization is designed to regulate UV-divergent integrals, it does not do so by cutting them off at large momenta and energies. The regularization is instead provided by defining the integral (including contributions from arbitrarily large momenta) for complex dimension, D, taking advantage of the fact that the integral converges in the ultraviolet if D is sufficiently small or negative. The result still diverges in the limit $D \to 4$, but usually as a pole or other type of isolated singularity when D is a positive integer. The limit $D \to 4$ is taken at the end of a calculation, after any singularities are absorbed into the renormalization of the appropriate couplings.

This section describes how dimensional regularization can nonetheless be used to define a Wilson action, despite it not seeming to explicitly separate high from low energies. This is done first by briefly describing dimensional regularization itself, followed by a presentation of the logic of constructing an effective theory using it. (See also Appendix §A.2.4 for more details about dimensional regularization.)

What is Dimensional Regularization?

Consider the following integral over D-dimensional Euclidean momentum p^μ, where $p^2 = \delta_{\mu\nu} p^\mu p^\nu$ (and similarly for q^2),[6] [33, 34]

$$\begin{aligned} I_D^{(A,B)}(q) &:= \int \frac{\mathrm{d}^D p_E}{(2\pi)^D} \left[\frac{p^{2A}}{(p^2+q^2)^B}\right] \\ &= \frac{1}{(4\pi)^{D/2}} \left[\frac{\Gamma\left(A+\frac{D}{2}\right)\Gamma\left(B-A-\frac{D}{2}\right)}{\Gamma(B)\,\Gamma\left(\frac{D}{2}\right)}\right] \left(q^2\right)^{A-B+D/2}, \end{aligned} \tag{3.37}$$

where $\Gamma(z)$ is Euler's Gamma function, defined to satisfy $z\,\Gamma(z) = \Gamma(z+1)$ with $\Gamma(n+1) = n!$ when restricted to positive integers, n. This integral converges in the

[6] Such Euclidean expressions are given in Euclidean signature, obtained by Wick rotating with $\mathrm{d}^4 p = \mathrm{i}\mathrm{d}^4 p_E$, meaning that the Minkowski-signature result has an additional factor of i. Notice that there are no additional explicit signs in continuing positive q^2 from Euclidean to Minkowski signature because of the wisdom of using conventions with a $(-+++)$ metric.

and

$$\mathcal{I}_\epsilon^{UV} := \int_0^\infty \left(\frac{\mu}{k}\right)^\epsilon \frac{k\,\mathrm{d}k}{k^2(k^2+M^2)}\left[1-\frac{m^2}{k^2}+\cdots\right] \simeq \frac{(M/\mu)^{-\epsilon}}{M^2}\left[-\frac{1}{\epsilon}+O(\epsilon)\right] = \frac{1}{M^2}\left[-\frac{1}{\epsilon}+\ln\left(\frac{M}{\mu}\right)+O(\epsilon)\right]. \tag{3.42}$$

The approximate equality starting the second line for both of these formulae drops all subdominant terms in powers of m/M. Notice that the pole $1/\epsilon$ arises due to a UV divergence in $\mathcal{I}^{IR}$ as $\epsilon \to 0$, while it corresponds to the IR divergence in $\mathcal{I}^{UV}$ as $\epsilon \to 0$. Notice also that these poles cancel in the sum $\mathcal{I} = \mathcal{I}^{UV} + \mathcal{I}^{IR}$, after which the limit $\epsilon \to 0$ can be taken, revealing agreement with (3.40).

The surprise in this exercise was that it was not important to explicitly separate out the region with $k < \Lambda$ and $k > \Lambda$ when defining $\mathcal{I}^{UV}$ and $\mathcal{I}^{IR}$, which nevertheless reproduce the correct dependence on m/M once summed to give the full integral. This works because any cutoff-dependence in the definition of these integrals is guaranteed to cancel in any case, and so although including the cutoff-dependence could be done, it is wasted effort.

Several concrete examples of the use of dimensional regularization when matching between underlying and effective theories are examined in more detail (for relativistic theories) in Chapter 7, which also explores a modification to minimal subtraction called 'decoupling subtraction' that proves useful when matching is done at or above one-loop accuracy. Nonrelativistic examples of beyond-leading-order matching are similarly studied in §12 and §15.

3.2.2 Matching vs Integrating Out

Matching – the fixing of low-energy couplings by comparing the predictions of the full theory with the predictions of its low-energy Wilsonian approximation – is often much easier to carry out than is the process of explicitly integrating out a heavy state using a cut-off path integral. This is partly because the comparison can be made for *any* physical quantity, and, in particular, this quantity can be chosen to make the comparison as simple as possible. Furthermore, because the comparison is made at the level of renormalized interactions, for both the full and low-energy theories, there are no UV divergences to worry about.

Example: The Toy Model

As usual, the toy model helps make the above statements more concrete. For the toy model the heavy mass scale is m_R and the full theory describing the physics of the two fields χ and ξ above this scale is given by the action $S[\chi,\xi]$ of Eq. (1.24) and (1.25) (repeated for convenience here):

$$S = -\int \mathrm{d}^4x \left[\frac{1}{2}\,\partial_\mu\chi\partial^\mu\chi + \frac{1}{2}\left(1+\frac{\chi}{\sqrt{2}\,v}\right)^2 \partial_\mu\xi\partial^\mu\xi + V(\chi)\right], \tag{3.43}$$

with

$$V(\chi) = \frac{\lambda\,v^2}{2}\,\chi^2 + \frac{\lambda\,v}{2\sqrt{2}}\,\chi^3 + \frac{\lambda}{16}\,\chi^4. \tag{3.44}$$

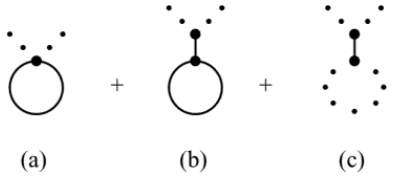

Fig. 3.3 One-loop graphs that contribute to the $\partial_\mu \xi \partial^\mu \xi$ kinetic term in the Wilson and 1LPI actions using the interactions of Eqs. (3.43) and (3.44). Solid (dotted) lines represent χ (and ξ) fields.

One could equivalently use the fields $\hat{\phi}_R$ and $\hat{\phi}_I$ with the action $S[\hat{\phi}_R, \hat{\phi}_I]$ of Eqs. (1.1) and (1.2), and renormalization is actually easier using these variables. This chapter sticks to χ and ξ to keep the symmetries of the problem more manifest. UV divergences in this theory are handled using dimensional regularization, and where necessary divergences are renormalized using modified minimal subtraction (see Appendix A.2.4 for details).

The low-energy Wilson action in this case is $S_W[\xi]$ (or $S_W[\hat{\phi}_I]$), depending only on the single light field, with UV divergent integrals again defined using dimensional regularization. Renormalization is again based on minimal subtraction, though with the difference that the finite part of the coupling is fixed by matching to predictions of the full theory (rather than again using modified minimal subtraction).

For present purposes the first step when matching is to write down all possible interactions in S_W up to some order in $1/m_R$, since this identifies the effective couplings whose value matching is meant to determine. For the toy model we know from §2.5 that all $1/m_R^2$ interactions are redundant, and the most general interactions (consistent with Lorentz invariance and the symmetry under constant shifts in ξ) are given to order $1/m_R^4$ by:

$$S_W = -\int \mathrm{d}^4x \, \left[\frac{z_W}{2} \, \partial_\mu \xi \, \partial^\mu \xi - a_W \, (\partial_\mu \xi \, \partial^\mu \xi)^2 + \cdots \right], \tag{3.45}$$

where on dimensional grounds z_W is dimensionless while $a_W \propto 1/m_R^4$ and terms not explicitly written are suppressed by at least $1/m_R^6$ (see §2.5).

Whereas earlier sections use the freedom to rescale ξ to set $z_W = 1$ (*i.e.* to canonically normalize ξ), writing (3.45) recognizes this has only been done at the classical level and not exactly, so at one loop $z_W = 1 + z_W^{(1)}$ with $z_W^{(1)} \simeq O(\lambda/16\pi^2)$ due to graphs like those of Fig. 3.3.

The contributions to $z_W^{(1)}$ found by Wick rotating these graphs, evaluating them in dimensional regularization and matching them to the contribution of (3.45) are

$$\mathrm{i} z_{W\,3.3(\mathrm{a})}^{(1)} = \mathrm{i}\left(-\frac{1}{4v^2}\right) \int \frac{\mathrm{i}\,\mathrm{d}^4 p_E}{(2\pi)^4} \, \frac{-\mathrm{i}}{p^2 + m_R^2} = -\frac{\mathrm{i}}{4v^2} \, I_D^{(0,1)}(m_R) \tag{3.46}$$

and

$$\begin{aligned} \mathrm{i} z_{W\,3.3(\mathrm{b})}^{(1)} &\simeq 3\left[2 \times \frac{\mathrm{i}^2}{2!}\right] \left(-\frac{1}{\sqrt{2}\,v}\right) \left(-\frac{\lambda v}{2\sqrt{2}}\right) \left(-\frac{\mathrm{i}}{m_R^2}\right) \\ &\quad \times \int \frac{\mathrm{i}\,\mathrm{d}^4 p_E}{(2\pi)^4} \, \frac{-\mathrm{i}}{p^2 + m_R^2} = \frac{3\mathrm{i}\lambda}{4m_R^2} \, I_D^{(0,1)}(m_R) \end{aligned} \tag{3.47}$$

while $z^{(1)}_{W\,3.3(c)} = 0$ because its dimensionally regularized loop evaluates to $I^{(1,1)}_D(0)$ and so vanishes. Here the integrals $I^{(0,1)}_D(m)$ and $I^{(1,1)}_D(m)$ are defined in (3.37), and the nonzero result evaluates to

$$\begin{aligned} I^{(0,1)}_D(m) &= \frac{m^2}{(4\pi)^{D/2}}\Gamma\left(1-\frac{D}{2}\right)\left(\frac{m^2}{\mu^2}\right)^{(D-4)/2} \\ &= \frac{m^2}{16\pi^2}\left[\frac{1}{(D/2)-2} + \gamma - 1 + \ln\left(\frac{m^2}{4\pi\mu^2}\right) + O(D-4)\right]. \end{aligned} \tag{3.48}$$

Summing these contributions (using $m_R^2 = \lambda v^2$) gives the one-loop prediction

$$z^{(1)}_W = \frac{I^{(0,1)}_D(m_R)}{2v^2} = \frac{\lambda}{16\pi^2}\left[\frac{1}{D-4} + \frac{1}{2}(\gamma-1) + \frac{1}{2}\ln\left(\frac{m_R^2}{4\pi\mu^2}\right) + O(D-4)\right]. \tag{3.49}$$

These corrections to z_W can again be absorbed into a rescaling of ξ – *i.e.* ξ is 're-normalized' by defining $\xi \to z_W^{-1/2}\xi$ – leading to the following rescaled version of (3.45):

$$S_W = -\int \mathrm{d}^4x\, \left[\frac{1}{2}\,\partial_\mu\xi\,\partial^\mu\xi - \frac{a_W}{z_W^2}\,(\partial_\mu\xi\,\partial^\mu\xi)^2 + \cdots\right]. \tag{3.50}$$

The remainder of the matching calculation computes another observable using (3.50) and (3.49) and compares the result to the calculation of the same quantity in the full theory to read off the coefficient a_W. It is relatively simple to do this with the same quantity as used at lowest order in §1.2.1: the amplitude for $\xi(p) + \xi(q) \to \xi(p') + \xi(q')$ scattering, keeping only terms up to order $1/m_R^4$. Since this calculation is already performed in chapter 1 at leading order in λ, it suffices here to sketch how things change once subdominant contributions are included.

To this end, write the coefficients in S_W as a series in λ,

$$a_W = a_W^{(0)} + a_W^{(1)} + \cdots \qquad \text{and} \qquad z_W = 1 + z_W^{(1)} + \cdots, \tag{3.51}$$

for which $z_W^{(1)}$ is given in (3.49). Starting on the Wilson side of the calculation, for $\xi - \xi$ scattering the required graphs up to one loop order are given by (a), (b) and (c) of Fig. 2.5. Of these, graphs (b) and (c) and their crossed counterparts both evaluate to give a contribution to $\xi\xi \to \xi\xi$ scattering that is suppressed by more than just four powers of $1/m_R$. This is easy to see in dimensional regularization because the coefficients of the interactions are suppressed by more than $1/m_R^4$ and the loop integrals only involve massless states and so cannot introduce compensating factors of m_R into the numerator. If one wishes to work only to lowest order in $1/m_R$ but to higher order in λ, it suffices to work with the tree contribution, graph (a), within the Wilsonian theory, but with λ-corrected effective coefficients $a_W^{(1)}$ and $z_W^{(1)}$.

Evaluating graph (a) using the Wilsonian coupling a_W/z_W^2 expanded out to subdominant order in λ, $a_W^{(1)} - 2a_W^{(0)}z_W^{(1)}$, then gives:

$$\begin{aligned} \mathcal{A}^{W(a)}_{\xi\xi\to\xi\xi} &= 8\mathrm{i}\left(a_W^{(0)} + a_W^{(1)} - 2a_W^{(0)}z_W^{(1)}\right) \\ &\quad\times\Big[(p\cdot q)(p'\cdot q') + (q\cdot q')(p\cdot p') + (p\cdot q')(p'\cdot q)\Big] + \cdots. \end{aligned} \tag{3.52}$$

Field-theory aficionados will recognize the $z_W^{(1)}$ term as the wave-function renormalization counter-term that cancels UV divergences due to the loops of Fig. 3.3 inserted into the external lines of the tree-level scattering graphs. (These graphs are not drawn explicitly in Fig. 2.4.)

This is to be compared with the one-loop contributions computed within the UV theory, working out to subdominant order in λ. The leading contribution comes from the tree graphs of Fig. 1.3, which evaluate to the result given in Eq. (1.28):

$$\mathcal{A}^{\text{full,tree}}_{\xi\xi\to\xi\xi} = \frac{2i\lambda}{m_R^4}\Big[(p\cdot q)(p'\cdot q') + (q\cdot q')(p\cdot p') + (p\cdot q')(p'\cdot q)\Big] + \cdots, \quad (3.53)$$

where the ellipses contain terms of higher order than $1/m_R^4$. Equating this to the lowest-order part of (3.52) then gives the previously obtained tree-level result: $a_W^{(0)} = \lambda/(4m_R^4) = 1/(4\lambda v^4)$.

Repeating this procedure including one-loop $O(\lambda/16\pi^2)$ corrections in the UV theory is less trivial, but in principle proceeds in precisely the same manner. This involves evaluating the graphs of Fig. 2.4, plus the 'wave-function renormalization' graphs obtained by inserting Fig. 3.3 into the external lines of the tree-level scattering graphs of Fig. 1.3. The $z_W^{(1)}$ contributions of (3.52) are important for reproducing the effects of these latter graphs in the full theory. The final result is a prediction for $a_W^{(1)}$ that is of order $\lambda^2/(16\pi^2 m_R^4) = 1/(4\pi v^2)^2$ in size.

This example shows how loops in the Wilsonian theory are not counted in the same way as are loops in the UV theory. Loops in the Wilsonian theory necessarily involve higher powers of E/m_R (more about this below), while loops in the UV theory are suppressed by factors of $\lambda/(16\pi^2)$ only, with all powers of E/m_R appearing at each loop order.

3.2.3 Power Counting Using Dimensional Regularization

The previous section made assertions about the size of the contributions of loop graphs – like graphs (b) and (c) of Fig. 2.5 in the toy model – which this section explores more systematically. More generally, this section's goal is to track how a generic Feynman graph computed using the Wilsonian action depends on a heavy scale like $1/m_R$, given that this scale does not appear in the same way for all interactions within $\mathfrak{L}_W$.

The logic here is much as used in §3.1.1, where dimensional analysis was employed to track how the cutoff Λ appears in amplitudes. The only difference now is to regulate the UV divergences in these graphs with dimensional regularization, since the size of a dimensionally regulated integral is set by the physical scales (light masses or external momenta) that appear in the integrand (rather than Λ). The power counting rules obtained in this way are much more useful since they directly track how amplitudes depend on physical variables, rather than unphysical quantities like Λ that in any case cancel from physical quantities.

The basic observation is that dimensional analysis applied to a dimensionally regulated integral estimates its size as

$$\int \cdots \int \left(\frac{d^D p}{(2\pi)^D}\right)^A \frac{p^B}{(p^2+q^2)^C} \sim \left(\frac{1}{4\pi}\right)^{DA} q^{DA+B-2C}, \tag{3.54}$$

with a dimensionless prefactor that depends on the dimension, D, of spacetime, and which may well be singular in the limit that $D \to 4$. Here, q represents the dominant scale appearing in the integrand of the momentum integrations. If the light particles appearing as external states in $A_{\mathcal{E}}(q)$ should be massless, or highly relativistic, then the typical external momenta are much larger than their masses and q in the above expression represents these momenta.[9] If all masses and momenta are comparable, then q is their common value. The important assumption is that there is only one low-energy scale (the more complicated case of multiple hierarchies is examined in later chapters, in particular in the nonrelativistic applications of Part III for which small speed, $v \sim E_{\rm kin}/p$, can be regarded as a ratio of two separate low-energy scales).

With this in mind, the idea is to repeat the steps of §3.1.1 and use the effective action, $S_W = S_{W,0} + S_{W,\rm int}$, of (3.2) – repeated here for ease of reference:

$$\begin{aligned} S_{W,0} &= -\frac{\mathfrak{f}^4}{M^2 v^2} \int \mathrm{d}^4x \left[\partial_\mu\phi\,\partial^\mu\phi + m^2\phi^2\right] \\ S_{W,\rm int} &= -\mathfrak{f}^4 \sum_n \frac{\hat{c}_n}{M^{d_n} v^{f_n}} \int \mathrm{d}^4x\, \mathcal{O}_n(\phi), \end{aligned} \tag{3.55}$$

to compute amputated Feynman amplitudes, $\mathcal{A}_{\mathcal{E}}(q)$, having $\mathcal{E}$ external lines, $\mathcal{I}$ internal lines, $\mathcal{L}$ loops and $\mathcal{V}_n$ vertices coming from the effective interaction with label 'n'. Respectively denoting (as before) the number of derivatives and fields appearing in this interaction as d_n and f_n, the amplitude becomes proportional to the following multiple integral:

$$\int \cdots \int \left(\frac{d^D p}{(2\pi)^D}\right)^{\mathcal{L}} \frac{p^{\mathcal{R}}}{(p^2+q^2)^{\mathcal{I}}} \sim \left(\frac{1}{4\pi}\right)^{2\mathcal{L}} q^{4\mathcal{L}-2\mathcal{I}+\mathcal{R}}, \tag{3.56}$$

where $\mathcal{R} = \sum_n d_n \mathcal{V}_n$ and the final estimate takes $D \to 4$. Liberally using the identities (3.4) and (3.5) then gives the following order of magnitude for $\mathcal{A}_{\mathcal{E}}(q)$:

$$\mathcal{A}_{\mathcal{E}}(q) \sim \mathfrak{f}^4 \left(\frac{1}{v}\right)^{\mathcal{E}} \left(\frac{Mq}{4\pi\mathfrak{f}^2}\right)^{2\mathcal{L}} \left(\frac{q}{M}\right)^{2+\sum_n (d_n-2)\mathcal{V}_n}. \tag{3.57}$$

This last formula is the main result, used extensively in many applications considered later. Its utility lies in the fact that it links the contributions of the various effective interactions in the effective lagrangian, (3.55), with the dependence of observables on small ratios of physical scales such as q/M. Notice in particular that more and more complicated graphs – for which $\mathcal{L}$ and $\mathcal{V}_n$ become larger and larger – are generically suppressed in their contributions to the graphical expansion if q is much smaller than the other scales M and $\mathfrak{f}^2/M$. This suppression assumes only that the powers appearing in (3.57) are all non-negative, and this is true so long as $d_n \geq 2$. The special cases where $d_n = 0, 1$ are potentially dangerous in this context, and require examination on a case-by-case basis.

[9] Any logarithmic dependence on q and infrared mass singularities that might arise in this limit are ignored here (for now), since the main interest is in following *powers* of ratios of the light and heavy mass scales.

propagator is order $1/q$, and so is suppressed by q relative to the bosonic propagator $1/q^2$. The second equality trades the dependence on $\mathcal{I}_F$ for $\mathcal{E}_F$ using (3.63).

The factor $(q/M)^{-\mathcal{E}_F/2}$ in (3.64) might also seem problematic at low energies, indicating as it does that more external lines necessarily imply more factors of the large ratio M/q. However, because such factors are fixed for all graphs contributing to any explicit process with a given number of external legs they usually do not in themselves undermine the validity of a perturbative expansion.

Furthermore, for the specific case where $\mathcal{A}_{\mathcal{E}_B\mathcal{E}_F}(q)$ represents a scattering amplitude each external fermion line corresponds to an initial-state or final-state spinor – $u_{q\,\sigma}$ or $\overline{u}_{q\,\sigma}$ – or the corresponding antiparticle spinor – $v_{q\,\sigma}$ or $\overline{v}_{q\,\sigma}$ – labelled by the corresponding state's momentum and spin. But each of these is itself proportional to an external particle – and so low-energy, $O(q)$ – scale, as can be seen from their appearance in spin-averaged expressions like $\sum_\sigma u_{q\,\sigma}\,\overline{u}_{q\,\sigma} = m - \mathrm{i}\,\not{q}$ and $\sum_\sigma v_{q\,\sigma}\,\overline{v}_{q\,\sigma} = -m - \mathrm{i}\,\not{q}$. This $q^{1/2}$ scaling of each external fermion line systematically cancels the factor $q^{-\mathcal{E}_F/2}$ in the amplitude $\mathcal{A}_{\mathcal{E}_B\mathcal{E}_F}(q)$, leading to non-singular predictions for scattering process at low energies. The same is true for effective couplings in S_W if these are obtained by matching to scattering processes.

Dangerous Interactions

As usual, interactions with the same number of fields and derivatives as the kinetic terms – either $f_n = 0$ and $d_n = 2$ (for bosons) or $f_n = 2$ and $d_n = 1$ (for fermions) – are unsuppressed by powers of q/M, beyond the usual loop factor. Interactions with more fields or derivatives than the kinetic terms additionally suppress a graph each time they are used. But interactions with no derivatives and two or fewer fermions can be potentially dangerous at low energies, introducing as they do negative powers of the small ratio q/M.

The kinds of interactions that are dangerous in this way are terms in a scalar potential ($d_n = f_n = 0$) and Yukawa couplings ($d_n = 0$ and $f_n = 2$). In principle, these kinds of interactions can be genuine threats to the consistency of the low-energy expansion, and whether such interactions are consistent with low-energy physics depends on the details.

What can make these interactions benign at low energies is if they do not carry too much energy for generic field configurations, $\phi \sim v_B$ and $\psi \sim v_F^{3/2}$. For instance, suppose, following the discussion around Eq. (3.17), that the scalar potential only carries energy density $\mathfrak{f}_V^4 \ll \mathfrak{f}^4$ when fields are order $\phi \sim v_B$ in size, such as if

$$V(\phi) \sim \mathfrak{f}_V^4 \sum_r g_r \left(\frac{\phi}{v_B}\right)^r . \tag{3.66}$$

In particular, the $r = 2$ term represents a mass for the field ϕ of order $m_B^2 \sim \mathfrak{f}_V^4/v_B^2$, so a natural criterion for ϕ to survive into the low-energy theory might be that $\mathfrak{f}_V^4 \sim m_B^2 v_B^2$ with $m_B \lesssim q$ for q a typical (possibly relativistic) momentum in the low-energy sector.

If this is the case then – assuming the couplings g_r are order unity – all the dimensionless couplings c_n of Eq. (3.61) for these particular interactions are secretly suppressed, with $c_n(d_n = 0) \sim g_r\,(\mathfrak{f}_V^4/\mathfrak{f}^4) \sim g_r\,(m_B^2 v_B^2/\mathfrak{f}^4)$. The contributions of these particular $d_n = f_n = 0$ interactions to (3.64) then become