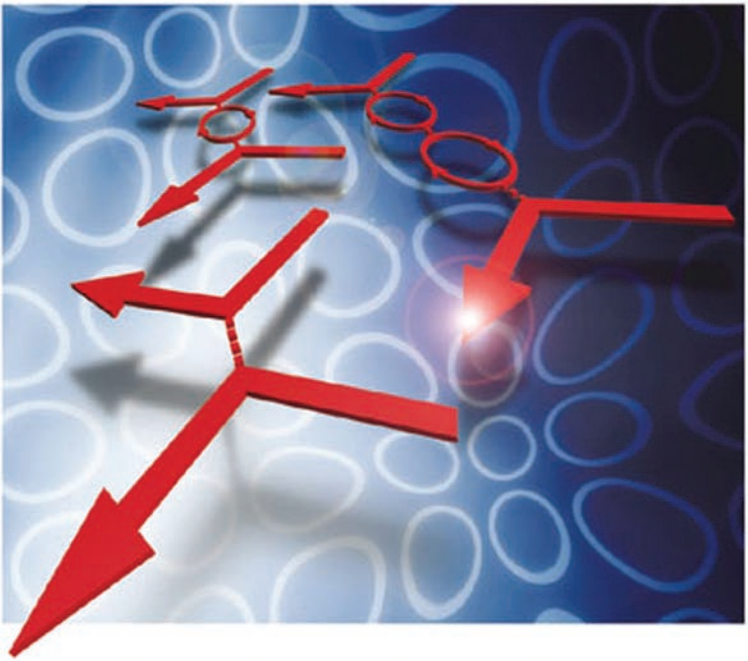
PHYSICS TEXTBOOK

Kerson Huang

WILEY-VCH

Quantum Field Theory

From Operators to Path Integrals



QUANTUM FIELD THEORY

From Operators to Path Integrals

Kerson Huang

Massachusetts Institute of Technology
Cambridge, Massachusetts

WILEY-VCH Verlag GmbH & Co. KGaA

This page is intentionally left blank

QUANTUM FIELD THEORY

This page is intentionally left blank

QUANTUM FIELD THEORY

From Operators to Path Integrals

Kerson Huang

Massachusetts Institute of Technology
Cambridge, Massachusetts

WILEY-VCH Verlag GmbH & Co. KGaA

Library of Congress Card No.: Applied for
British Library Cataloging-in-Publication Data:
A catalogue record for this book is available from the British Library

Bibliographic information published by
Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

Printed in the Federal Republic of Germany
Printed on acid-free paper

Cover Design Grafik-Design Schulz, Fußgönnheim
Printing Strauss GmbH, Mörlenbach
Bookbinding Großbuchbinderei J. Schäffer GmbH & Co. KG, KG Grünstadt

ISBN-13: 978- 0-471-14120-4
ISBN-10: 0-471-14120-8

To *Rosemary*

This page is intentionally left blank

Contents

This page is intentionally left blank

Preface

Quantum field theory, the quantum mechanics of continuous systems, arose at the beginning of the quantum era, in the problem of blackbody radiation. It became fully developed in quantum electrodynamics, the most successful theory in physics. Since that time, it has been united with statistical mechanics through Feynman's path integral, and its domain has been expanded to cover particle physics, condensed-matter physics, astrophysics, and wherever path integrals are spoken.

This book is a textbook on the subject, aimed at readers conversant with what is usually called "advanced quantum mechanics," the equivalent of a first-year graduate course. Previous exposure to the Dirac equation and "second quantization" would be very helpful, but not absolutely necessary. The mathematical level is not higher than what is required in advanced quantum mechanics; but a degree of maturity is assumed.

In physics, a continuous system is one that appears to be so at long wavelengths or low frequencies. To model it as mathematically continuous, one runs into difficulties, in that the high-frequency modes often give rise to infinities. The usual procedure is to start with a discrete version, by discarding the high-frequency modes beyond some cutoff, and then try to approach the continuum limit, through a process called *renormalization*.

Renormalization is a relatively new concept, but its workings were already evident in classical physics. At the beginning of the atomic era, Boltzmann noted that classical equipartition of energy presents conceptual difficulties, when one seriously considers the atomic structure of matter. Since atoms are expected to contain smaller subunits, which in turn should composed of even smaller subunits, and so *ad infinitum*, and each degree of freedom contributes equally to the thermal energy of a substance, the specific heat of matter would be infinite.. The origin of this divergence lies in the extrapolation of known physical laws into the high-frequency domain, a characteristic shared by the infinities in quantum field theory.

Boltzmann's "paradox," however, matters not a whit when it comes to practical calculations, as evidenced by the great success of classical physics. The reason is that most equations of macroscopic physics, such as those in thermodynamics and

hydrodynamics, make no explicit reference to atoms, but depend on coefficients like the specific heat, which can be obtained from experiments. From a modern perspective, we say that such theories are "renormalizable," in that the microstructure can be absorbed into measurable quantities.

One goal of this book is to explain what renormalization is, how it works, and what makes some systems appear "renormalizable" and others not. We follow the historical route, discovering it in quantum electrodynamics through necessity, and then realizing its physical meaning through Wilson's path-integral formulation.

This book, then, starts with a thorough introduction to the usual operator formalism, including Feynman graphs, from Chapters 1–10. This is followed by Chapters 11–14 on quantum electrodynamics, which illustrates how to do practical calculations, and includes a complete discussion of perturbative renormalization. The last part, Chapters 15–19, introduces the Feynman path integral, and discusses "modern" subjects, including the physical approach to renormalization, spontaneous symmetry breaking, and topological excitations. I have entirely omitted non-Abelian gauge fields and the standard model of particle physics, because these subjects are discussed in another book: K. Huang, *Quarks, Leptons, and Gauge Field,* 2nd ed. (World Scientific, Singapore, 1992).

I have chosen to introduce path integrals only after the canonical approach is fully developed and applied. Others might want them discussed earlier. To accommodate different tastes, I have tried to make each chapter self-contained in as much as possible, so that a knowledgeable reader can pick and skip.

There is definitely a change in flavor when quantum field theory is conveyed through the path integral. Apart from the union with statistical mechanics, which immeasurably enriches the subject, it liberates our imagination by making it possible to contemplate virtual but fantastic deformations, such as altering the structure of space–time. I am reminded of the classification of things as "gray" or "green" by Freeman Dyson, in his book *Disturbing the Universe* (Harper & Row, New York, 1979). He classified physics gray (and I suppose that included quantum field theory,) as opposed to things green, such as poems and horse manure. In a private letter dated August 3, 1983, Dyson wrote, "Everyone has to make his own choice of what to call gray and green. I took my choice from Goethe:

Grau, tenerer freund, ist alle Theorie,
Und grün des Lebens Goldner Baum.

Dear friend, all theory is grey,
And green is the golden tree of life.

I must admit that Hilbert space does seem a bit dreary at times; but, with Feynman's path integral, quantum field theory has surely turned green.

KERSON HUANG

December, 1997
Marblehead, Massachusetts

Acknowledgments

I wish to thank my colleagues and students for helping me to learn the stuff in this book over the years, in particular Herman Feshbach, from whom I took the first course on the subject. I thank Jeffrey Goldstone, Roman Jackiw, and Ken Johnson, who can usually be relied on to provide correct and simple answers to difficult questions; and Patrick Lee, who taught a course with me on the subject, thereby broadening my horizon. Last but not least, I thank my editor Greg Franklin for his understanding and support.

K. H.

QUANTUM FIELD THEORY

This page is intentionally left blank

This page is intentionally left blank

CHAPTER ONE

Introducing Quantum Fields

1.1 THE CLASSICAL STRING

We obtain a quantum field by quantizing a classical field, of which the simplest example is the classical string. To be on firm mathematical grounds, we define the latter as the long-wavelength limit of a discrete chain. Consider $N + 2$ masses described by the classical Lagrangian

$$L(q, \dot{q}) = \sum_{j=0}^{N+1} \left[\frac{m}{2} \dot{q}_j^2 - \frac{\kappa}{2}(q_j - q_{j+1})^2 \right] \tag{1.1}$$

where m is the mass and κ a force constant. The coordinate $q_j(t)$ represents the lateral displacement of the jth mass along a one-dimensional chain. We impose fixed-endpoint boundary conditions, by setting

$$q_0(t) = q_{N+1}(t) = 0 \tag{1.2}$$

The equations of motion for the N remaining movable masses are then

$$m\ddot{q}_j - \kappa(q_{j+1} - 2q_j + q_{j-1}) = 0 \qquad (j = 1, \ldots, N) \tag{1.3}$$

The normal modes have the form

$$q_j(t) = \cos(\omega t) \sin(jp) \tag{1.4}$$

To satisfy the boundary conditions, choose p to have one of the N possible values

$$p_n = \frac{\pi n}{N+1} \qquad (n = 1, \ldots, N) \tag{1.5}$$

Substituting this into the equations of motion, we obtain N independent normal frequencies ω_n:

$$\omega_n^2 = \omega_0^2 \sin^2\left(\frac{\pi}{2}\frac{n}{N+1}\right) \qquad (n = 1 \cdots N) \tag{1.6}$$

where

$$\omega_0 = 2\sqrt{\frac{\kappa}{m}} \tag{1.7}$$

This is a cutoff frequency, for the modes with $n > N$ merely repeat the lower ones. For $N = 4$, for example, the independent modes correspond to $n = 1,2,3,4$. The case $n = 5$ is trivial, since $p = \pi$, and hence $q_j(t) = 0$ by (1.4). The case $n = 6$ is the same as that for $n = 4$, since $\omega_6 = \omega_4$, and $\sin(jp_6) = -\sin(jp_4)$.

When N is large, and we are not interested in the behavior near the endpoints, it is convenient to use periodic boundary conditions:

$$q_{j+N}(t) = q_j(t) \tag{1.8}$$

In this case the normal modes are

$$q_j(t) = e^{i(jp-\omega t)} \tag{1.9}$$

For N even, the boundary conditions can be satisfied by putting

$$p_n = \frac{2\pi n}{N} \qquad \left(n = 0,\pm 1, \ldots, \pm\frac{N}{2}\right) \tag{1.10}$$

The corresponding normal frequencies are

$$\omega_n^2 = \omega_0^2 \sin^2\left(\frac{\pi n}{N}\right) \tag{1.11}$$

Compared to the fixed-end case, the spacing between normal frequencies is now doubled; but each frequency is twofold degenerate, and the number of normal modes remains the same. A comparison of the two cases for $N = 8$ is shown in Fig. 1.1.

The equilibrium distance a between masses does not explicitly appear in the Lagrangian; it merely supplies a length scale for physical distances. For example, it appears in the definition of the distance of a mass from an end of the chain:

$$x \equiv ja \qquad (j = 1,\ldots, N) \tag{1.12}$$

The total length of the chain is then defined as

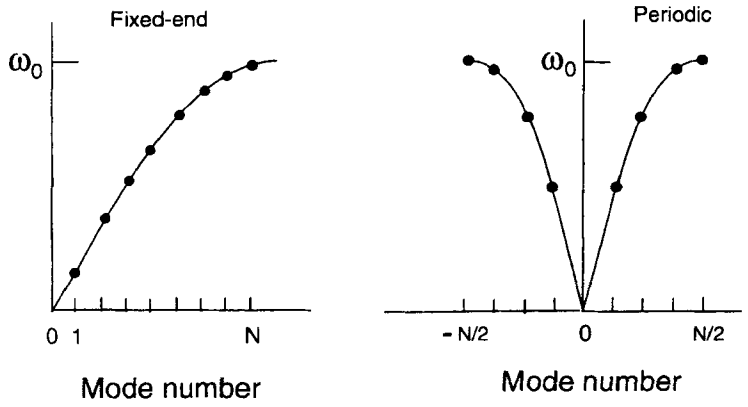


Figure 1.1 Normal modes of the classical chain for fixed-end and periodic boundary conditions.

$$R = Na \tag{1.13}$$

In the continuum limit

$$a \to 0 \qquad N \to \infty \qquad (R = Na \text{ fixed}) \tag{1.14}$$

the discrete chain approaches a continuous string, and the coordinate approaches a classical field defined by

$$q(x, t) \equiv q_j(t) \tag{1.15}$$

The Lagrangian in the continuum limit can be obtained by making the replacements

$$(q_{n+1} - q_j)^2 \to a^2 \left[\frac{\partial q(x, t)}{\partial x} \right]^2$$

$$\sum_j \to \frac{1}{a} \int_0^R dx \tag{1.16}$$

Assuming that the mass density ρ and string tension σ approach finite limits

$$\rho = \frac{m}{a} \tag{1.17}$$

$$\sigma = \kappa a \tag{1.18}$$

we obtain the limit Lagrangian

$$L_{\text{cont}} = \frac{1}{2}\int_0^R dx\left[\rho\left(\frac{\partial q(x,\,t)}{\partial t}\right)^2 - \sigma\left(\frac{\partial q(x,\,t)}{\partial x}\right)^2\right] \tag{1.19}$$

This leads to the equation of motion

$$\frac{\partial^2 q(x,\,t)}{\partial t^2} - \frac{1}{c^2}\,\frac{\partial^2 q(x,\,t)}{\partial x^2} = 0 \tag{1.20}$$

which is a wave equation, with propagation velocity

$$c = \sqrt{\frac{\sigma}{\rho}} \tag{1.21}$$

The general solutions are the real and imaginary parts of

$$q(x,\,t) = e^{i(kx-\omega t)} \tag{1.22}$$

with a linear dispersion law

$$\omega = ck \tag{1.23}$$

For fixed-end boundary conditions

$$q(0,\,t) = q(R,\,t) = 0 \tag{1.24}$$

the normal modes of the continuous string are

$$q_n(x,\,t) = \cos(\omega_n t)\,\sin(k_n x) \tag{1.25}$$

with $\omega_n = ck_n$, and

$$k_n = \frac{\pi n}{R} \qquad (n = 0,1,2,\ldots) \tag{1.26}$$

The normal frequencies ω_n are the same as those for the discrete chain for $n/N \ll 1$, as given in (1.6). However, the number of modes of the continuum string is infinite, and only the first N modes have correspondence with those of the discrete string. This is illustrated in Fig. 1.2 for $N = 4$. Thus, there is a cutoff frequency

$$\omega_c \equiv \omega_N = \frac{\pi c}{a} \tag{1.27}$$

This is of the same order, but not same as the maximum frequency defined earlier, $\omega_0 = 2c/a$, for ω_c is based on a linear dispersion law. The continuum model is an accurate representation of the discrete chain only for $\omega \ll \omega_c$.

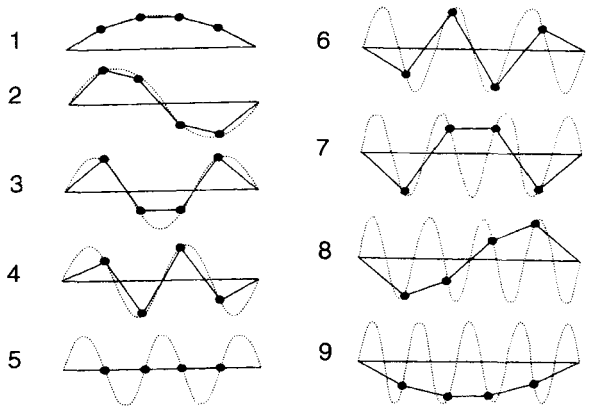


Figure 1.2 Normal modes of a discrete chain of four masses, compared with those of a continuous string. The former repeat themselves after the first four modes. (After J. C. Slater and N. H. Frank, Mechanics, McGraw-Hill, New York, 1947.)

For periodic boundary conditions

$$q(0, t) = q(R, t) \tag{1.28}$$

the allowed wave numbers are

$$k_n = \frac{2\pi n}{R} \qquad (n = 0, \pm 1, \pm 2, \ldots) \tag{1.29}$$

We obtain the cutoff frequency ω_c by setting $n = N/2$.

The high-frequency cutoff is a theoretical necessity. Without it, the specific heat of the string will diverge, since each normal mode contributes an amount kT. The value of the cutoff cannot be determined from the long-wavelength effective theory, because only the combination $c = a\omega_c/\pi$ occurs. Absorbing the cutoff into measurable parameters, as done in (1.17), is called *renormalization*. A theory for which this can be done is said to be *renormalizable*.

Nonrenormalizable systems exhibit behavior that is sensitive to details on an atomic scale. Such behavior would appear to be random on a macroscopic scale, as in the propagation of cracks in materials, and the nucleation of raindrops.

1.2 THE QUANTUM STRING

We now quantize the classical chain, to obtain a quantum field in the continuum limit. The Hamiltonian of the classical discrete chain is given by

$$H(p, q) = \sum_{j=1}^{N} \left[\frac{p_j^2}{2m} + \frac{\kappa}{2}(q_j - q_{j+1})^2 \right] \tag{1.30}$$

where $p_j = m\dot{q}_j$. The system can be quantized by replacing p_j and q_j by Hermitian operators satisfying the commutation relations

$$[p_j, q_k] = -i\delta_{jk} \tag{1.31}$$

We impose periodic boundary conditions, and expand these operators in Fourier series:

$$\begin{aligned} q_j &= \frac{1}{\sqrt{N}} \sum_{n=-N/2}^{N/2} Q_n e^{i2\pi nj/N} \\ p_j &= \frac{1}{\sqrt{N}} \sum_{n=-N/2}^{N/2} P_n e^{i2\pi nj/N} \end{aligned} \tag{1.32}$$

where P_n and Q_n are operators satisfying

$$\begin{aligned} [P_n^\dagger, Q_m] &= -i\delta_{nm} \\ P_n^\dagger &= P_{-n} \\ Q_n^\dagger &= Q_{-n} \end{aligned} \tag{1.33}$$

The system is reduced to a sum of independent harmonic oscillators:

$$\begin{aligned} H &= \sum_{n=-N/2}^{N/2} \left[\frac{1}{2m} P_n^\dagger P_n + \frac{1}{2} m\omega_n^2 Q_n^\dagger Q_n \right] \\ \omega_n^2 &= \frac{4\kappa}{m} \sin^2\left(\frac{\pi n}{N} \right) \end{aligned} \tag{1.34}$$

The eigenvalues are labeled by a set of occupation numbers $\{\alpha_n\}$:

$$E_\alpha = \sum_{n=-N/2}^{N/2} \omega_n(\alpha_n + \tfrac{1}{2}) \tag{1.35}$$

where $\alpha_n = 0,1,2, \ldots$. The frequency ω_n is taken to be the positive root of ω_n^2, since H is positive-definite.

In the continuum limit (1.14) the Hamiltonian becomes

$$H_{\text{cont}} = \int_0^R dx \left[\frac{1}{2\rho} p^2(x, t) + \frac{\sigma}{2} \left(\frac{\partial q(x, t)}{\partial x} \right)^2 \right] \tag{1.36}$$

where, with $x = ja$,

$$p(x, t) = \frac{p_j(t)}{a} = \rho \frac{\partial q(x, t)}{\partial t} \tag{1.37}$$

The quantum field $q(x, t)$ and its canonical conjugate $p(x, t)$ satisfy the equal-time commutation relation

$$[p(x, t), q(x', t)] = -i\delta(x - x') \tag{1.38}$$

Just as in the classical case, we have to introduce a cutoff frequency ω_c. General properties of the quantum field will be discussed more fully in Chapter 2.

1.3 SECOND QUANTIZATION

Another way to obtain a quantum field is to consider a collection of identical particles in quantum mechanics. In this case, the quantum field is an equivalent description of the system. Identical particles are defined by a Hamiltonian that is (1) invariant under a permutation of the particle coordinates and (2) has the same form for any number of particles. The quantized-field description is called "second quantization" for historical reasons, but quantization was actually done only once.

Let $\mathcal{H}_N$ be the Hilbert space of a system of N identical nonrelativistic particles. The union of all $\mathcal{H}_N$ is called the Fock space:

$$\mathcal{F} = \bigcup_{N=0}^{\infty} \mathcal{H}_N \tag{1.39}$$

The subspace with $N = 0$ contains the vacuum state as its only member. We assume that N is the eigenvalues of a "number operator" N_{op}, which commutes with the Hamiltonian. It is natural to introduce operators on Fock space that connect subspaces of different N. An elementary operator of this kind creates or annihilates one particle at a point in space. Such an operator is a quantum field operator, since it is a spatial function. This is why a quantum-mechanical many-particle system automatically gives rise to a quantum field.

For definiteness, consider N nonrelativistic particles in three spatial dimensions, with coordinates $\{\mathbf{r}_1, \ldots, \mathbf{r}_N\}$. The Hamiltonian is

$$H = -\frac{1}{2m} \sum_{i=1}^{N} \nabla_i^2 + V(\mathbf{r}_1, \ldots, \mathbf{r}_N) \tag{1.40}$$

where ∇_i^2 is the Laplacian with respect to $\mathbf{r}_i$, and where V is a symmetric function of its arguments. The eigenfunctions Ψ_n are defined by

$$H\Psi_n(\mathbf{r}_1, \ldots, \mathbf{r}_N) = E_n \Psi_n(\mathbf{r}_1, \ldots, \mathbf{r}_N) \tag{1.41}$$

For Bose or Fermi statistics, Ψ_n is respectively symmetric or antisymmetric under an interchange of any two coordinates $\mathbf{r}_i$ and $\mathbf{r}_j$. The particles are called *bosons* or *fermions*, respectively.

We now describe the equivalent quantum field theory, and justify it later. Let $\psi(\mathbf{r})$ be the Schrödinger-picture operator that annihilates one particle at $\mathbf{r}$. Its Hermitian conjugate $\psi^\dagger(\mathbf{r})$ will create one particle at $\mathbf{r}$. They are defined through the commutation relations

$$\begin{aligned} [\psi(\mathbf{r}), \psi^\dagger(\mathbf{r}')]_\pm &= \delta^3(\mathbf{r}-\mathbf{r}') \\ [\psi(\mathbf{r}), \psi(\mathbf{r}')]_\pm &= 0 \end{aligned} \tag{1.42}$$

where $[A,B]_\pm = AB \pm BA$, with the plus sign corresponding to bosons and the minus sign to fermions. The Fock-space Hamiltonian is defined in such a manner that it reduces to (1.40) in the N-particle subspace.

A general N-particle Hamiltonian has the structure

$$H = \sum_i f(\mathbf{r}_i) + \sum_{i<j} g(\mathbf{r}_i, \mathbf{r}_j) + \sum_{i<j<k} h(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \cdots \tag{1.43}$$

where the functions g, h, and so on are symmetric functions of their arguments. The first term is a "one-particle operator," a sum of operators of the form $f(\mathbf{r})$, which act on one particle only. The second term is a "two-particle operator," a sum of operators of the form $g(\mathbf{r}_1, \mathbf{r}_2)$, over all distinct pairs. Generally, an "n-particle operator" is a sum of operators that depend only on a set of n coordinates. To construct the Hamiltonian on Fock space, we associate an n-particle operator with an operator on Fock space, with the following correspondences:

$$\begin{aligned} \sum_i f(\mathbf{r}_i) &\to \int d^3r\, \psi^\dagger(\mathbf{r}) f(\mathbf{r}) \psi(\mathbf{r}) \\ \sum_{i<j} g(\mathbf{r}_i, \mathbf{r}_j) &\to \frac{1}{2}\int d^3r_1\, d^3r_2\, \psi_1^\dagger \psi_2^\dagger g_{12} \psi_2 \psi_1 \\ \sum_{i<j<k} h(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) &\to \frac{1}{3!}\int d^3r_1\, d^3r_2\, d^3r_3\, \psi_1^\dagger \psi_2^\dagger \psi_3^\dagger h_{123} \psi_3 \psi_2 \psi_1 \\ &\vdots \end{aligned} \tag{1.44}$$

where for brevity we have written $\psi_1 = \psi(\mathbf{r}_1)$, $g_{12} = g(\mathbf{r}_1, \mathbf{r}_2)$, and so on.

As an example, suppose the potential in (1.40) is a sum of two-body potentials:

$$V(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_{i<j} v(\mathbf{r}_i, \mathbf{r}_j) \tag{1.45}$$

Then the corresponding Fock-space Hamiltonian, also denoted H, takes the form

$$H = -\frac{1}{2m}\int d^3r\psi^\dagger(\mathbf{r})\nabla^2\psi(\mathbf{r})$$

$$+\frac{1}{2}\int d^3r_1\, d^3r_2\psi^\dagger(\mathbf{r}_1)\psi^\dagger(\mathbf{r}_2)v(\mathbf{r}_1,\mathbf{r}_2)\psi(\mathbf{r}_2)\psi(\mathbf{r}_1) \tag{1.46}$$

The particle number is the eigenvalue of the number operator, defined as

$$N_{\text{op}} = \int d^3r\ \psi^\dagger(\mathbf{r})\psi(\mathbf{r}) \tag{1.47}$$

By using (2.18), we can verify the relations

$$\begin{aligned}[N_{\text{op}},H] &= 0\\ [\psi(\mathbf{r}),N_{\text{op}}] &= \psi(\mathbf{r})\\ [\psi^\dagger(\mathbf{r}),N_{\text{op}}] &= -\psi^\dagger(\mathbf{r})\end{aligned} \tag{1.48}$$

These imply that the action of $\psi(\mathbf{r})$ on a eigenstate of N_{op} is to decrease its eigenvalue by 1, while that of $\psi^\dagger(\mathbf{r})$ is to increase it by 1. Thus $\psi(\mathbf{r})$ is an annihilation operator, while $\psi^\dagger(\mathbf{r})$ is a creation operator. The vacuum state $|0\rangle$ is defined as the eigenstate of N_{op} with eigenvalue zero. It is annihilated by all annihilation operators:

$$\psi(\mathbf{r})|0\rangle = 0 \tag{1.49}$$

By applying $\psi^\dagger(\mathbf{r})$ to the vacuum state repeatedly, it is easy to show that the eigenvalues of N_{op} are nonnegative integers.

To demonstrate that the quantum field is equivalent to the many-particle system, consider a complete set of states $|E,N\rangle$ of the quantum field, which are simultaneous eigenstate of H and N_{op}:

$$\begin{aligned}H|E,N\rangle &= E|E,N\rangle\\ N_{\text{op}}|E,N\rangle &= N|E,N\rangle\end{aligned}$$

We define the N-particle wave function $\Psi_E(\mathbf{r}_1,\ldots,\mathbf{r}_N)$ corresponding to $|E,N\rangle$ by

$$\Psi_E(\mathbf{r}_1,\ldots,\mathbf{r}_N) \equiv \frac{1}{\sqrt{N!}}\langle 0|\psi(\mathbf{r}_1)\cdots\psi(\mathbf{r}_N)|E,N\rangle \tag{1.50}$$

which has the correct symmetry with respect to particle permutation. It tells us that the probability amplitude for finding N particles at the positions $\mathbf{r}_1,\ldots,\mathbf{r}_N$ can be found by annihilating the particles at the respective locations from the state $|E,N\rangle$, and evaluating the overlap between the resulting state and the vacuum state. We leave it as an exercise to show that this wave function satisfies the N-particle Schrödinger equation (1.41). (See Problem 1.3.)

1.4 CREATION AND ANNIHILATION OPERATORS

The field operator $\psi(\mathbf{r})$ annihilates a particle at $\mathbf{r}$. That is, it annihilates a particle whose wave function is a δ function. Since the latter can be written as a linear superposition of a complete set of wave functions, we can express $\psi(\mathbf{r})$ as a linear superposition of operators that annihilate particles with specific types of wave functions. Suppose that $u_k(\mathbf{r})$ is a member of a complete orthonormal set of single-particle wave functions:

$$\int d^3r\, u_k^*(\mathbf{r}) u_{k'}(\mathbf{r}) = \delta_{kk'}$$

$$\sum_k u_k(\mathbf{r}) u_k^*(\mathbf{r}') = \delta^3(\mathbf{r} - \mathbf{r}')$$

An example of such a set is plane waves:

$$u_k(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{1.51}$$

We can expand the field operators with respect to such a basis:

$$\psi(\mathbf{r}) = \sum_k u_k(\mathbf{r}) a_k$$

$$\psi^\dagger(\mathbf{r}) = \sum_k u_k^*(\mathbf{r}) a_k^\dagger$$

The coefficient a_k and $a_k^\dagger$ are operators that satisfy the commutation relations

$$[a_k, a_{k'}^\dagger]_\pm = \delta_{kk'}$$
$$[a_k, a_{k'}]_\pm = 0 \tag{1.52}$$

where the + sign is for bosons and the – sign is for fermions. These relations follow from (1.42) and the orthonormality of the functions $u_k(\mathbf{r})$.

It follows from the commutation relations that, for each k, the eigenvalues of $a_k^\dagger a_k$ are integers n_k, called the "occupation number of the single-particle state k":

$$a^\dagger a|n\rangle = n|n\rangle$$
$$\langle n|m\rangle = \delta_{nm} \tag{1.53}$$

where we have omitted the label k for brevity. The allowed values of the occupation number are given by

$$n = \begin{cases} 0,1,2,\ldots,\infty & \text{(Bose statistics)} \\ 0,1 & \text{(Fermi statistics)} \end{cases}$$

The actions of a and $a^\dagger$ have the following results:

$$a|n\rangle = \sqrt{n}|n-1\rangle$$

$$a^\dagger|n\rangle = \sqrt{1 \pm n}|n+1\rangle \tag{1.54}$$

where the ±sign corresponds respectively to Bose (+) and Fermi (−) statistics, which show that a annihilates a particle in the state with wave function $u(\mathbf{r})$, and $a^\dagger$ creates such a particle. We leave it as an exercise to derive these basic results. (See Problem 1.2.)

The state $|0\rangle$ corresponding to $n = 0$ is the vacuum state, which satisfies

$$a|0\rangle = 0 \tag{1.55}$$

We assume that it is normalizable:

$$\langle 0|0\rangle = 1 \tag{1.56}$$

Obviously all other states can be obtained by creating particles from the vacuum:

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|n\rangle \tag{1.57}$$

We can simultaneously diagonalize $a_k^\dagger a_k$ for all k. The eigenstates are then labeled by a set of occupation numbers $\{n_0, n_1, \ldots\}$, and they constitute a basis for the Fock space. The total number of particles present is $N = \Sigma_k n_k$. We have

$$a_k^\dagger a_k|n_0, \ldots, n_k, \ldots\rangle = n_k|n_0, \ldots, n_k, \ldots\rangle$$
$$a_k|n_0, \ldots, n_k, \ldots\rangle = (-1)^s\sqrt{n_k}|n_0, \ldots, n_k - 1, \ldots\rangle$$
$$a_k^\dagger|n_0, \ldots, n_k, \ldots\rangle = (-1)^s\sqrt{1 \pm n_k}|n_0, \ldots, n_k + 1, \ldots\rangle$$

where

$$s = \begin{cases} 0 & \text{(Bose statistics)} \\ \Sigma_{p<k} n_p & \text{(Fermi statistics)} \end{cases} \tag{1.58}$$

That is, $s = \pm 1$ for fermions, depending on whether the number of fermions with quantum numbers less than k is even or odd, and the meaning of "less than" is set by an arbitrary but fixed ordering. This phase factor arises from the fact that fermion creation operators anticommute: $a_k^\dagger a_p^\dagger = -a_p^\dagger a_k^\dagger$.

A complete set of states can be constructed by creating particles from the vacuum:

$$|k\rangle = a_k^\dagger|0\rangle$$

$$|k,p\rangle = a_k^\dagger a_p^\dagger |0\rangle$$
$$\vdots \tag{1.59}$$

These states are not normalized to unity. When there are many particles present, it is more convenient to label the state with occupation numbers $\{n_k\}$, where n_k is the number of particles with single-particle quantum number k:

$$|n_0, \ldots, n_k, \ldots\rangle = C\prod_k [a_k^\dagger]^{n_k}|0\rangle \tag{1.60}$$

These states can be normalized to unity by choosing

$$C = \left[\prod_k n_k!\right]^{-1/2} \tag{1.61}$$

1.5 BOSE AND FERMI STATISTICS

The term "statistics" refers to the rule for counting the degeneracy of an energy level of a many-particle system. In three-dimensional (3D) space, it depends on the symmetry of the wave function under a permutation of the particle coordinates. Technically speaking, the different possible symmetries correspond to the different irreducible representations of the permutation group.

The completely symmetric and the completely antisymmetric representations correspond respectively to Bose and Fermi statistics. They are the only possible ones in a two-particle system; but for more than two particles other possibilities exist, in which the wave function is symmetric with respect to permutations among one subset S of coordinates, and antisymmetric for the complementary set. Called "parastatistics, " such representations correspond to the Young's tableaux with more than one row, or more than one column. Since the particle are identical, there is more than one way to choose the subset S. Consequently, such "para" representations must be multidimensional. That is, the carrier space for such a representation must be spanned by states having the same energy eigenvalue, and they mix under a permutation of the coordinates. Therefore, the energy levels of particles obeying parastatistics must have intrinsic degeneracies, which cannot be removed by any interaction that treat the particles as identical.

The Bose and Fermi statistics can be set apart from the parastatistics by virtue of the following properties:

- Under particle permutation, the symmetry character of wave functions is independent of the number of particles present.
- Energy eigenfunctions do not mix under particle permutation.

Parastatistics does occur in atomic physics, but only in the context of "incomplete" permutations, which interchange the positions of atomic electrons but not their

spins. With respect to permutations of both position and spin, electrons obey Fermi statistics, as we know. No known examples of parastatistics have been found in nature. Perhaps the simple properties itemized above are essential for consistency on some level.

Although we live in a 3D world, some interesting physical systems are effectively two-dimensional (2D). These include the electron sheets that exhibit the quantum Hall effect, the copper oxide planes in a high-temperature superconductor, and thin films of superfluid helium on various substrates. In a 2D system, the variety of statistics is far richer, because the exchange of two particles in a plane is not a unique process; we may rotate the particles about a center through angle $n\pi$, where n is any odd integer, and the paths corresponding to different n are not necessarily equivalent. Consequently, the symmetry group relevant to particle exchange is not the permutation group, but the much larger *braid group*. This circumstance allows for fractional spin and statistics; but we shall not discuss this, except for a brief discussion on fractional spin in Chapter 19.

PROBLEMS

1.1 Consider an actual string made of atoms spaced $a = 10^{-8}$ cm apart. Suppose the length of the string is 1 m, and it is kept at such a tension that the fundamental frequency is 100 cycles per second (Hz). Find the cutoff frequency, and show that it lies in the infrared region of the spectrum. (This gives the Debye temperature.)

1.2 **(a)** The basic commutation relation for boson annihilation and creation operator is

$$[a, a^\dagger] = 1 \qquad [a, a] = 0$$

where $[A, B] = AB - BA$. From this definition, show that eigenstates $|n\rangle$ of $a^\dagger a$ have the properties

$$a^\dagger a|n\rangle = n|n\rangle \qquad (n = 0,1,2, \ldots)$$
$$a|n\rangle = \sqrt{n}|n\rangle$$
$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$$

(b) Fermion annihilation and creation operators are defined by an anticommutation relation

$$\{a, a^\dagger\} = 1 \qquad \{a, a\} = 0$$

where $\{A, B\} = AB + BA$. Show

$$a^\dagger a|n\rangle = \sqrt{n}|n\rangle \qquad (n = 0, 1)$$
$$a|n\rangle = n|n\rangle$$
$$a^\dagger|n\rangle = \sqrt{1-n}|n+1\rangle$$

1.3 Consider the N-particle wave function defined in (1.50)

$$\Psi_E(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}}\langle 0|\psi(\mathbf{r}_1)\cdots\psi(\mathbf{r}_N)|E,N\rangle$$

where $|E,N\rangle$ is an N-particle energy eigenstate with respect to the Hamiltonian H given in (1.46).

(a) Show that it is normalized to unity:

$$\int d^3r_1 \cdots d^3r_N |\Psi_E(\mathbf{r}_1, \ldots, \mathbf{r}_N)|^2 = 1$$

(b) Show

$$E\Psi_E(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}}\langle 0|\psi(\mathbf{r}_1)\cdots\psi(\mathbf{r}_N)H|E,N\rangle$$

(c) Show that the wave function satisfies the N-particle Schrödinger equation

$$\left[-\sum_{i=1}^{N}\frac{1}{2m}\nabla_i^2 + \sum_{i<j} v(\mathbf{r}_i, \mathbf{r}_j)\right]\Psi_E(\mathbf{r}_1, \ldots, \mathbf{r}_N) = E\Psi_E(\mathbf{r}_1, \ldots, \mathbf{r}_N)$$

by going to the result in (b), and commute H all the way to the left, where it gives zero operating on the vacuum.

All the results stated hold for both Bose and Fermi statistics (see Huang [1]).

1.4 A nonrelativistic boson or fermion field $\psi(\mathbf{x})$ is governed by the Hamiltonian

$$H = -\frac{1}{2m}\int d\mathbf{x}\ \psi^\dagger(\mathbf{x})\nabla^2\psi(\mathbf{x}) + \frac{1}{2}\int d\mathbf{x}\, d\mathbf{y}\ \psi^\dagger(\mathbf{x})\psi^\dagger(\mathbf{y})v(\mathbf{x}-\mathbf{y})\psi(\mathbf{y})\psi(\mathbf{x})$$

The system is enclosed in a large cubical box of volume Ω ($\Omega \to \infty$), with periodic boundary conditions. Expand the field in terms of annihilation operators $a_\mathbf{k}$ for free-particle states of momentum $\mathbf{k}$, and show that

$$H = \sum_{\mathbf{k}}\frac{\mathbf{k}^2}{2m}a_\mathbf{k}^\dagger a_\mathbf{k} + \frac{1}{2\Omega}\sum_{\mathbf{p},\mathbf{q},\mathbf{k}}\tilde{v}(\mathbf{k})a_{\mathbf{p}+\mathbf{k}}^\dagger a_{\mathbf{q}-\mathbf{k}}^\dagger a_\mathbf{p} a_\mathbf{q}$$

where $\tilde{v}(\mathbf{k}) = \int d^3r\, e^{i\mathbf{k}\cdot\mathbf{r}}v(\mathbf{r})$.

1.5 Consider a system of N nonrelativistic electrons and N positive ions with Coulomb interactions, enclosed in a periodic box of volume Ω. The Hamiltonian is given by

$$H = \sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{2m} + \sum_{i<j}\frac{e^2}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{i,j=1}^{N}\frac{e^2}{|\mathbf{r}_i-\mathbf{R}_j|} + \sum_{i<j}\frac{e^2}{|\mathbf{R}_i-\mathbf{R}_j|} + \sum_{i=1}^{N}\frac{\mathbf{P}_i^2}{2M}$$

The ions are heavy. Hence consider $\mathbf{R}_i$ to be fixed numbers, neglect $\mathbf{P}_i$, and drop the last two terms.

(a) Label single-electron states by momentum $\mathbf{k}$ and spin s, designated collectively as $\alpha = \{\mathbf{k},s\}$. The corresponding wave function is

$$u_\alpha(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}\cdot\mathbf{r}}\zeta_s$$

$$\zeta_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \zeta_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

(b) To go to the quantized-field representation, replace one- and two-particle operators by the rules

$$\sum_{i=1}^{N} K(\mathbf{r}_i) \to \sum_{\alpha,\beta} \langle \alpha | K | \beta \rangle \, a_\alpha^\dagger a_\beta$$

$$\sum_{i<j} v(\mathbf{r}_i - \mathbf{r}_j) \to \frac{1}{2} \sum_{\alpha\beta\gamma\lambda} (a_\alpha a_\beta)^\dagger \langle \alpha\beta | v | \gamma\lambda \rangle (a_\gamma a_\lambda)$$

(c) Define Fourier transforms:

$$\left\langle \mathbf{k}s \left| \frac{e^2}{|\mathbf{r}-\mathbf{R}|} \right| \mathbf{k}'s' \right\rangle \equiv \frac{\delta_{ss'}}{\Omega} \int d^3r e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}} \frac{e^2}{|\mathbf{r}-\mathbf{R}|} = \frac{\delta_{ss'}}{\Omega} \frac{4\pi e^2}{|\mathbf{r}-\mathbf{R}|} e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{R}}$$

$$\langle \alpha\beta | v | \gamma\lambda \rangle \equiv \frac{\delta_{s_1 s_3}\delta_{s_2 s_4}}{\Omega} \int d^3x \, d^3y \frac{e^2}{|\mathbf{x}-\mathbf{y}|} e^{i[(\mathbf{k}_3-\mathbf{k}_1)\cdot\mathbf{x} + (\mathbf{k}_4-\mathbf{k}_2)\cdot\mathbf{y}]}$$

$$= \frac{\delta_{s_1 s_3}\delta_{s_2 s_4}}{\Omega} \delta_{\mathrm{K}}(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4) \frac{4\pi e^2}{|\mathbf{k}_3 - \mathbf{k}_1|}$$

where δ_{K} is the Kronecker delta.

(d) Obtain the Hamiltonian in quantized-field form:

$$H = \sum_{\mathbf{k}s} \frac{\hbar^2 \mathbf{k}^2}{2m} a_{\mathbf{k}s}^\dagger a_{\mathbf{k}s} + \frac{2\pi e^2}{\Omega} \sum_{\mathbf{p},\mathbf{q},\mathbf{k}} \sum_{s,s'} \frac{1}{k^2} (a_{\mathbf{p}+\mathbf{k},s} a_{\mathbf{q}-\mathbf{k},s'})^\dagger a_{\mathbf{p},s} a_{\mathbf{q},s'}$$

$$- \frac{4\pi e^2}{\Omega} \sum_{\mathbf{p},\mathbf{k}} \sum_{s} \frac{1}{k^2} a_{\mathbf{p},s}^\dagger a_{\mathbf{k}+\mathbf{p},s} \sum_{i=1}^{N} e^{i\mathbf{k}\cdot\mathbf{R}_i}$$

(e) Show that the second term gives, for small $\mathbf{k}$,

$$\frac{2\pi e^2}{\Omega k^2} \sum_{\mathbf{p},\mathbf{q}} \sum_{s,s'} (a_{\mathbf{p},s} a_{\mathbf{q},s'})^\dagger a_{\mathbf{p},s} a_{\mathbf{q},s'} = \frac{2\pi e^2 N(N-1)}{\Omega k^2}$$

which is divergent at $\mathbf{k} = 0$. Show that the divergent term proportional to N^2 is canceled by the $\mathbf{k} = 0$ limit of the third term. The $O(N)$ term above remains divergent. The source of this divergence is the periodic boundary conditions, by which the set of coordinates $\{\mathbf{r}_1, \ldots, \mathbf{r}_N\}$ is being repeated an infinite number of times in space. Consequently, the Coulomb energy of an electron diverges, due to long-range interactions with an infinite number of distant copies.

(f) It is clear that this is a mathematical artifact, and to avoid it we should treat the ions dynamically; but we do not wish to add that complication. The expedient way out is to simply leave out the $\mathbf{k} = 0$ contribution in both the second and third terms. Hav-

ing done this, we can ignore the third term altogether, because it sums to zero under the assumption that the ions are uniformly distributed in space. Thus we take as effective Hamiltonian

$$H = \sum_{\mathbf{k}s} \frac{k^2}{2m} a^\dagger_{\mathbf{k}s} a_{\mathbf{k}s} + \frac{2\pi e^2}{\Omega} \sum_{\substack{\mathbf{p},\mathbf{q} \\ \mathbf{k}\neq 0}} \sum_{s,s'} \frac{1}{k^2} (a_{\mathbf{p}+\mathbf{k},s} a_{\mathbf{q}-\mathbf{k},s'})^\dagger a_{\mathbf{p},s} a_{\mathbf{q},s'}$$

This describes electrons immersed in a uniform positively charged background that makes the whole system electrically neutral.

1.6 Imposing periodic boundary conditions means filling infinite space with identical cells that contain copies of our system. This problem illustrates the effect of long-range interactions among the cells. Consider a unit point charge at the center $\mathbf{r} = 0$ of a cubic cell of volume L^3, which contains a uniform negative charge density, so that the total charge in the cell is zero. Impose periodic boundary conditions, and calculate the potential in the neighborhood of the unit point charge.

(a) Show that the potential is given by

$$V(\mathbf{r}) = \frac{4\pi}{L^3} \sum_{\mathbf{k}\neq 0} \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{k^2} \qquad k_i = \frac{2\pi n_i}{L} \qquad (n_i = 0, \pm 1, \pm 2, \ldots)$$

by showing

$$\nabla^2 V(\mathbf{r}) = 4\pi \left[\sum_{\mathbf{n}} \delta^3(\mathbf{r} - \mathbf{n}L) - \frac{1}{L^3} \right]$$

(b) For $r/L \ll 1$, show[2]

$$V(\mathbf{r}) = \frac{1}{r} - \frac{c}{L} + O(L^{-2})$$

$$c = 2.837297 \cdots$$

REFERENCES

1. K. Huang, *Statistical Mechanics,* 2nd ed., Wiley, New York, 1987, Appendix.
2. K. Huang and C. N. Yang, *Phys. Rev.* **105,** 767 (1957).
3. Lüscher, *Commun. Math. Phys.* **105,** 153 (1986).
4. C. N. Yang, *Chin. J. Phys.* **25,** 80 (1987).

[2]Huang and Yang [2] first calculated c, but gave an incorrect value, 2.37. The correct value was obtained by Lüscher [3]. An elementary derivation was given by Yang [4].

CHAPTER TWO

Scalar Fields

2.1 KLEIN–GORDON EQUATION

A fast way to go from classical mechanics to quantum mechanics is to replace the energy and momentum of a particle by operators, according to the prescription

$$E \to i\frac{\partial}{\partial t}$$
$$\mathbf{p} \to -i\nabla \tag{2.1}$$

Making the replacements in the nonrelativistic relation $E = \mathbf{p}^2/2m$, and applying the result to the wave function, we obtain the Schrödinger equation:

$$-\frac{1}{2m}\nabla^2\psi(\mathbf{r}, t) = i\frac{\partial}{\partial t}\psi(\mathbf{r}, t) \tag{2.2}$$

Of course, this is not covariant under Lorentz transformations. For a covariant equation, we use the same trick on the relativistic relation $E^2 = \mathbf{p}^2 + m^2$. The result is the Klein–Gordon equation

$$(\Box^2 + m^2)\psi(x) = 0 \tag{2.3}$$

where x stands for $x^\mu = (t, \mathbf{x})$, and

$$\Box^2 \equiv \partial^\mu\partial_\mu = \frac{\partial^2}{\partial t^2} - \nabla^2 \tag{2.4}$$

in units with $c = 1$. Assuming that ψ is invariant under Lorentz transformations, we have a covariant equation—one that has the same form in all Lorentz frames. What is not clear is how to interpret $\psi(x)$.

By analogy with the Schrödinger equation, we might interpret $\psi(x)$ to be the

wave function of a relativistic particle. That would require the existence of a 4-vector probability current density j^μ, which should be conserved: $\partial_\mu j^\mu = 0$. However, the obvious choice $j_0 = \psi^*\psi$ is untenable, for $\psi^*\psi$ is Lorentz-invariant by assumption, and cannot be part of a 4-vector.

As in the case of the Schrödinger equation, we can construct a conserved current as follows. Multiply (2.3) from the left by ψ^* to obtain

$$\partial^\mu(\psi^*\partial_\mu\psi) - (\partial^\mu\psi^*)(\partial_\mu\psi) + m^2\psi^*\psi = 0 \tag{2.5}$$

Subtracting this from its complex conjugate leads to

$$\partial_\mu j^\mu = 0 \tag{2.6}$$

where

$$j^\mu = \psi\partial^\mu\psi^* - \psi^*\partial^\mu\psi \tag{2.7}$$

However, the time component

$$j^0 = \psi^*\frac{\partial\psi}{\partial t} - \frac{\partial\psi^*}{\partial t}\psi \tag{2.8}$$

is not positive-definite, and therefore cannot be a probability density.

The root of the difficulty lies in the second time derivative in the Klein–Gordon equation. As we shall see, this leads to negative frequencies corresponding to antiparticles. The relativistic kinematics makes it impossible to have a one-particle theory. We shall regard $\psi(x)$ not as a wave function but as a classical wave field, and as such should be quantized.

2.2 REAL SCALAR FIELD

Consider a real scalar field $\phi(\mathbf{r}, t)$, which is invariant under Lorentz transformations. The current j^μ vanishes identically in this case. We enclose the system in a large periodic box of volume Ω, and expand the field in a Fourier series:

$$\phi(\mathbf{r}, t) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{k}} q_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{r}} \tag{2.9}$$

where

$$q_{\mathbf{k}}^*(t) = q_{-\mathbf{k}}(t) \tag{2.10}$$

because the field is real. Assuming that $\phi(\mathbf{r}, t)$ satisfies the Klein–Gordon equation, we have

$$\ddot{q}_{\mathbf{k}} + \omega_k^2 q_{\mathbf{k}} = 0 \tag{2.11}$$

where

$$\omega_k^2 = \mathbf{k}^2 + m^2 \tag{2.12}$$

The system is equivalent to a collection of harmonic oscillators, and may be quantized by imposing the commutation relations

$$\begin{aligned} i[\dot{q}_{\mathbf{k}}^\dagger(0), q_{\mathbf{k}'}(0)] &= \delta_{\mathbf{k}\mathbf{k}'} \\ [q_{\mathbf{k}}(0), q_{\mathbf{k}'}(0)] &= 0 \end{aligned} \tag{2.13}$$

where $q_{\mathbf{k}}^\dagger(0)$ is the hermitian conjugate of $q_{\mathbf{k}}(0)$. This fixes the normalization of the field, left arbitrary by the Klein–Gordon equation.

Since the Klein–Gordon equation is invariant under time translation, the origin of time is arbitrary. The commutations relations in fact hold at any time t:

$$\begin{aligned} i[\dot{q}_{\mathbf{k}}^\dagger(t), q_{\mathbf{k}'}(t)] &= \delta_{\mathbf{k}\mathbf{k}'} \\ [q_{\mathbf{k}}(t), q_{\mathbf{k}'}(t)] &= 0 \end{aligned} \tag{2.14}$$

which are equivalent to

$$\begin{aligned} i[\dot{\phi}(\mathbf{r}, t), \phi(\mathbf{r}', t)] &= \delta^3(\mathbf{r} - \mathbf{r}') \\ [\phi(\mathbf{r}, t), \phi(\mathbf{r}', t)] &= 0 \end{aligned} \tag{2.15}$$

These are called *equal-time commutators,* which serve as initial conditions for the equation of motion. The unequal-time commutators must be calculated from the solutions, and contain dynamical information.

In the present free-particle case, the equation of motion (2.11) is trivial to solve. For given wave number k there are two frequencies $\pm\omega_k$, with

$$\omega_k \equiv +\sqrt{\mathbf{k}^2 + m^2} \tag{2.16}$$

Taking into account the reality property (2.10), we write the solution in the form

$$q_{\mathbf{k}}(t) = \frac{1}{\sqrt{2\omega_k}}[a_{\mathbf{k}} e^{-i\omega_k t} + a_{-\mathbf{k}}^\dagger e^{i\omega_k t}] \tag{2.17}$$

where $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^\dagger$ are operators, with commutation relations determined by (2.14):

$$\begin{aligned} [a_{\mathbf{k}}, a'^\dagger_{\mathbf{k}}] &= \delta_{\mathbf{k}\mathbf{k}'} \\ [a_{\mathbf{k}}, a_{\mathbf{k}'}] &= 0 \end{aligned} \tag{2.18}$$

The normalization factor $(2\omega)^{-1/2}$ in (2.17) is chosen to make the commutators simple. We recognize that $a_{\mathbf{k}}$ is an annihilation operator, and $a_{\mathbf{k}}^{\dagger}$ a creation operator for a boson, as introduced in Section 1.4. The time-dependent quantized-field operator can be represented in the form

$$\phi(\mathbf{r}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k}} [a_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} + a_{\mathbf{k}}^{\dagger} e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}] \tag{2.19}$$

The positive-frequency part (the first term) annihilates a particle, and the negative-frequency part creates a particle. The negative-frequency part is absent in a nonrelativistic field, because the kinematic relation $E = p^2/2m$ allows only one sign of the frequency.

2.3 ENERGY AND MOMENTUM

Analogy with the harmonic oscillator suggests that the Hamiltonian of the free scalar field should be

$$H = \frac{1}{2} \sum_{\mathbf{k}} [|\dot{q}_{\mathbf{k}}|^2 + \omega_k^2 |q_{\mathbf{k}}|^2] \tag{2.20}$$

In terms of the field $\phi(x) = \phi(\mathbf{r}, t)$, it has the form

$$H = \int d^3r\, \mathcal{H}(x)$$

$$\mathcal{H}(x) = \frac{1}{2}\left[\left(\frac{\partial \phi}{\partial t}\right)^2 + |\nabla \phi|^2 + m^2\phi^2\right] \tag{2.21}$$

where $\mathcal{H}(x)$ is called the *Hamiltonian density.* The Lagrangian of the system can be obtained through the general relation $L(q, \dot{q}) = p\dot{q} - L(p, q)$:

$$L = \int d^3r\, \mathcal{L} \tag{2.22}$$

where $\mathcal{L}(x)$ is the Lagrangian density given by

$$\mathcal{L}(x) = \frac{1}{2}\left[\left(\frac{\partial \phi}{\partial t}\right)^2 - |\nabla \phi|^2 - m^2\phi^2\right] = \frac{1}{2}\left[\partial^{\mu}\phi\partial_{\mu}\phi - m^2\phi^2\right] \tag{2.23}$$

The last form shows that the Lagrangian density is Lorentz-invariant. In contrast, the Hamiltonian density, which is an energy density, cannot be invariant. For this reason, relativistic theories are usually specified via the Lagrangian density.

In terms of creation and annihilation operators, the Hamiltonian takes the diagonalized form

$$H = \sum_{\mathbf{k}} \omega_k (a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \tfrac{1}{2}) \tag{2.24}$$

The zero-point energy diverges unless there is a cutoff; but the cutoff has no physical relevance, since the energy of any state relative to that of the vacuum is independent of it. The energy of a particle is given by

$$\omega_k = \sqrt{\mathbf{k}^2 + m^2} \tag{2.25}$$

where $\mathbf{k}$ is its momentum and m is the rest mass. Accordingly, the total momentum operator is

$$\mathbf{P} = \sum_{\mathbf{k}} \mathbf{k} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \tag{2.26}$$

According to the general principles of quantum mechanics, the Hamiltonian is the generator of time evolution, through the Heisenberg equation of motion:

$$i[H, \phi(\mathbf{r}, t)] = \frac{\partial \phi(\mathbf{r}, t)}{\partial t} \tag{2.27}$$

The formal solution yields

$$\phi(\mathbf{r}, t) = e^{iHt} \phi(\mathbf{r}, 0) e^{-iHt} \tag{2.28}$$

For consistency, we must show that this is consistent with the Klein–Gordon equation, which we used to arrive at the solution (2.19). Substituting $\phi(\mathbf{r}, 0)$ from (2.19) into (2.28), we obtain

$$\phi(\mathbf{r}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k}} e^{iHt} [a_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} + a_{\mathbf{k}}^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}}] e^{-iHt} \tag{2.29}$$

For the free-field Hamiltonian, we have (see Problem 2.1)

$$e^{iHt} a_{\mathbf{k}} e^{-iHt} = a_{\mathbf{k}} e^{-i\omega_k t} \tag{2.30}$$

This demonstrates that (2.28) is the same as (2.19).

Again, according to general principles, the momentum operator $\mathbf{P}$ should generate spatial translations:

$$-i[\mathbf{P}, \phi(\mathbf{r}, t)] = \nabla \phi(\mathbf{r}, t) \tag{2.31}$$

with formal solution

$$\phi(\mathbf{r}, t) = e^{-i\mathbf{P}\cdot\mathbf{r}} \phi(0, t) e^{i\mathbf{P}\cdot\mathbf{r}} \tag{2.32}$$

As a straightforward calculation shows, this is the same as (2.19).

2.4 PARTICLE SPECTRUM

The vacuum state $|0\rangle$ is the state of lowest energy, defined by

$$a_{\mathbf{k}}|0\rangle = 0 \qquad (\text{all } \mathbf{k}) \tag{2.33}$$

and normalized to

$$\langle 0|0\rangle = 1 \tag{2.34}$$

A one-particle state is defined by

$$|\mathbf{p}\rangle = a_{\mathbf{p}}^{\dagger}|0\rangle \tag{2.35}$$

Using the commutation relations, we find

$$a_{\mathbf{k}}|\mathbf{p}\rangle = a_{\mathbf{k}}a_{\mathbf{p}}^{\dagger}|0\rangle = (a_{\mathbf{p}}^{\dagger}a_{\mathbf{k}} + \delta_{\mathbf{kp}})\,|0\rangle = \delta_{\mathbf{kp}}|0\rangle \tag{2.36}$$

The field operator has nonvanishing matrix elements only between a one-particle state and the vacuum:

$$\langle 0|\phi(x)|\mathbf{p}\rangle = \frac{e^{i(\mathbf{p}\cdot\mathbf{r}-\omega_p t)}}{\sqrt{2\omega_p\Omega}} \tag{2.37}$$

This is the wave function of a particle of momentum $\mathbf{p}$, normalized to a density $(2\omega_p\Omega)^{-1}$. By successively creating particles from the vacuum, we can build a complete set of states:

Vacuum:	$\|0\rangle$
1-particle states:	$\|\mathbf{p}\rangle = a_{\mathbf{p}}{}^{\dagger}\|0\rangle$
2-particle states:	$\|\mathbf{p}_1\mathbf{p}_2\rangle = a_{\mathbf{p}_1}{}^{\dagger}a_{\mathbf{p}_2}{}^{\dagger}\|0\rangle$
	$\vdots$

2.5 CONTINUUM NORMALIZATION

In the limit $\Omega \to \infty$, the allowed values of $\mathbf{k}$ approach a continuum, and we can make the replacements

$$\frac{1}{\Omega}\sum_{\mathbf{k}} \to \int \frac{d^3k}{(2\pi)^3}$$

$$\Omega\delta_{\mathbf{k}\mathbf{k}'} \to (2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}') \tag{2.38}$$

We define continuum versions of the annihilation and creation operators by putting

$$\begin{aligned} a(\mathbf{k}) &\equiv \Omega a_{\mathbf{k}} \\ a^\dagger(\mathbf{k}) &\equiv \Omega a_{\mathbf{k}}^\dagger \end{aligned} \tag{2.39}$$

The commutators then have limiting forms:

$$\begin{aligned} [a(\mathbf{k}), a^\dagger(\mathbf{k}')] &= (2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}') \\ [a(\mathbf{k}), a(\mathbf{k}')] &= 0 \end{aligned} \tag{2.40}$$

The field operator can be represented as a Fourier integral:

$$\phi(\mathbf{r}, t) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2\omega_k}}[a(\mathbf{k})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} + a^\dagger(\mathbf{k})e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}] \tag{2.41}$$

As before, the vacuum state $|0\rangle$ is defined by $a(\mathbf{k})|0\rangle = 0$ with $\langle 0|0\rangle = 1$, and a one-particle state is defined by

$$|\mathbf{p}\rangle = a^\dagger(\mathbf{p})|0\rangle \tag{2.42}$$

Using the commutation relations, we find

$$a(\mathbf{k})|\mathbf{p}\rangle = (2\pi)^3\delta^3(\mathbf{p}-\mathbf{k})|0\rangle \tag{2.43}$$

The single-particle wave function is

$$\langle 0|\phi(x)|\mathbf{p}\rangle = \frac{e^{i(\mathbf{p}\cdot\mathbf{r}-\omega_{\mathbf{p}}t)}}{\sqrt{2\omega_{\mathbf{p}}}} \tag{2.44}$$

with a particle density $(2\omega_{\mathbf{p}})^{-1}$. The normalization is such that the number of particles in volume element d^3r is the Lorentz-invariant combination $d^3r/(2\omega_{\mathbf{p}})$. The Hamiltonian and total momentum now take the forms

$$\begin{aligned} H &= \int \frac{d^3k}{(2\pi)^3}\omega_k a^\dagger(\mathbf{k})a(\mathbf{k}) \\ \mathbf{P} &= \int \frac{d^3k}{(2\pi)^3}\mathbf{k}a^\dagger(\mathbf{k})a(\mathbf{k}) \end{aligned} \tag{2.45}$$

The choice between discrete or continuum normalization is a matter of notation, since we always regard Ω as large but finite in intermediate steps of calcula-

tions. The limit $\Omega \to \infty$ is taken only in final answers. This is done to avoid irrelevant concerns about mathematical rigor, such as how to define the Hilbert space when the dimensionality is noncountably infinite. The continuum normalization merely anticipates this limit.

2.6 COMPLEX SCALAR FIELD

A complex scalar field is just two real scalar fields constituting the real and imaginary parts. What is new is a symmetry between the two fields, and this leads to a conserved current. In physical terms, a complex field can have electric charge, whereas a real field must be neutral.

We denote the complex scalar field by $\psi(\mathbf{r}, t)$, and decompose it into real and imaginary parts in the form

$$\psi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \tag{2.46}$$

The Lagrangian density is taken to be

$$\begin{aligned}\mathcal{L}(x) &= \partial^\mu \psi^\dagger(x)\partial_\mu \psi(x) - m^2\psi^\dagger(x)\psi(x) \\ &= \tfrac{1}{2}\sum_{j=1}^{2}[\partial^\mu \phi_j(x)\partial_\mu \phi_j(x) - m^2\phi_j(x)\phi_j(x)]\end{aligned} \tag{2.47}$$

The normalization factor $1/\sqrt{2}$ in (2.46) is chosen in order that ϕ_j has the same normalization as the real scalar field discussed previously.

To quantize the system, we impose the commutation relations

$$\begin{aligned}i[\dot{\phi}_i(\mathbf{r}, t), \phi_j(\mathbf{r}', t)] &= \delta_{ij}\delta^3(\mathbf{r} - \mathbf{r}') \\ [\phi_i(\mathbf{r}, t), \phi_j(\mathbf{r}', t)] &= 0\end{aligned} \tag{2.48}$$

The complex field $\psi(\mathbf{r}, t)$ becomes a non-Hermitian operator satisfying

$$\begin{aligned}i[\dot{\psi}^\dagger(\mathbf{r}, t), \psi(\mathbf{r}', t)] &= \delta^3(\mathbf{r} - \mathbf{r}') \\ [\dot{\psi}(\mathbf{r}, t), \psi(\mathbf{r}', t)] = [\psi^\dagger(\mathbf{r}, t), \psi(\mathbf{r}', t)] &= [\psi(\mathbf{r}, t), \psi(\mathbf{r}', t)] = 0\end{aligned} \tag{2.49}$$

Thus, the canonical conjugate to ψ is $\dot{\psi}^\dagger$.

In accordance with (2.19), we can expand ϕ_j in terms of annihilation and creation operators:

$$\phi_j(\mathbf{x}, t) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_{\mathbf{k}}}}(a_{j\mathbf{k}}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} + a_{j\mathbf{k}}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}) \quad (j = 1, 2) \tag{2.50}$$

with the commutation relations

$$[a_{i\mathbf{k}}, a^{\dagger}_{j\mathbf{k}'}] = \delta_{ij}\delta_{\mathbf{k}\mathbf{k}'}$$
$$[a_{i\mathbf{k}}, a_{j\mathbf{k}}{}'] = 0 \tag{2.51}$$

In the complex representation, we have

$$\psi(\mathbf{r}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_{\mathbf{k}}}} (b_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} + c^{\dagger}_{\mathbf{k}} e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)})$$
$$\psi^{\dagger}(\mathbf{r}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_{\mathbf{k}}}} (b^{\dagger}_{\mathbf{k}} e^{i-(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} + c_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}) \tag{2.52}$$

where

$$b_{\mathbf{k}} = \frac{1}{\sqrt{2}}(a_{1\mathbf{k}} + ia_{2\mathbf{k}})$$
$$c_{\mathbf{k}} = \frac{1}{\sqrt{2}}(a_{1\mathbf{k}} - ia_{2\mathbf{k}}) \tag{2.53}$$

with commutation relations

$$[b_{\mathbf{k}}, b^{\dagger}_{\mathbf{p}}] = \delta_{\mathbf{k}\mathbf{p}}$$
$$[c_{\mathbf{k}}, c^{\dagger}_{\mathbf{p}}] = \delta_{\mathbf{k}\mathbf{p}}$$
$$[b_{\mathbf{k}}, b_{\mathbf{p}}] = [c_{\mathbf{k}}, c_{\mathbf{p}}] = [b_{\mathbf{k}}, c_{\mathbf{p}}] = 0 \tag{2.54}$$

The total energy and momentum can be expressed as follows:

$$H = \sum_{\mathbf{k}} \omega_{\mathbf{k}}(a^{\dagger}_{1\mathbf{k}}a_{1\mathbf{k}} + a^{\dagger}_{2\mathbf{k}}a_{2\mathbf{k}}) = \sum_{\mathbf{k}} \omega_{\mathbf{k}}(b^{\dagger}_{\mathbf{k}}b_{\mathbf{k}} + c^{\dagger}_{\mathbf{k}}c_{\mathbf{k}})$$
$$\mathbf{P} = \sum_{\mathbf{k}} \omega_{\mathbf{k}}\mathbf{k}(a^{\dagger}_{1\mathbf{k}}a_{1\mathbf{k}} + a^{\dagger}_{2\mathbf{k}}a_{2\mathbf{k}}) = \sum_{\mathbf{k}} \omega_{\mathbf{k}}\mathbf{k}(b^{\dagger}_{\mathbf{k}}b_{\mathbf{k}} + c^{\dagger}_{\mathbf{k}}c_{\mathbf{k}}) \tag{2.55}$$

There are two type of quanta, which can be designated either as a_1 and a_2 quanta, or as b and c quanta. The energy and the momentum do not distinguish between these descriptions. We shall see, however, that only the b and c quanta have definite charge.

2.7 CHARGE AND ANTIPARTICLE

The current density for the complex scalar field is given by

$$j^{\mu} = \psi\partial^{\mu}\psi^{*} - \psi^{*}\partial^{\mu}\psi = \tfrac{1}{2}(\phi_2\partial^{\mu}\phi_1 - \phi_1\partial^{\mu}\phi_2) \tag{2.56}$$

which satisfies the conservation law $\partial_\mu j^\mu = 0$, or

$$\frac{\partial}{\partial t} j^0 + \nabla \cdot \mathbf{j} = 0 \tag{2.57}$$

Integrating both sides over the spatial volume, we obtain

$$\int d^3x \frac{\partial}{\partial t} j^0(x) = \frac{d}{dt} \int d^3x j^0(x) = 0 \tag{2.58}$$

or

$$\frac{dQ}{dt} = 0 \tag{2.59}$$

where Q is the total charge operator

$$\begin{aligned} Q &= \int d^3x j^0(x) \\ &= \sum_{\mathbf{k}} (a_{1\mathbf{k}}^\dagger a_{2\mathbf{k}}^\dagger - a_{2\mathbf{k}}^\dagger a_{1\mathbf{k}}) \\ &= \sum_{\mathbf{k}} (b_{\mathbf{k}}^\dagger b_{\mathbf{k}} - c_{\mathbf{k}}^\dagger c_{\mathbf{k}}) \end{aligned} \tag{2.60}$$

As we can see, a b quantum carries one unit of positive charge, and a c quantum carries one unit of negative charge. The a_1, a_2 quanta, which are linear combinations of those of b and c, do not have definite charge. By convention, we refer to a c quantum as an "antiparticle." Thus, the positive-frequency part of ψ annihilates a particle, and its negative-frequency part creates an antiparticle. Similarly, $\psi^\dagger$ either creates a particle or annihilates an antiparticle. In light of this, we can say that for the real field, the particle is its own antiparticle.

The term "charge" is used in a generic sense, and does not necessarily mean the electric charge, since we have not turned on the electromagnetic coupling. The unit of charge is arbitrary, because j^μ is defined only up to a multiplicative constant.

It is straightforward to verify that charge is conserved:

$$[Q, H] = 0 \tag{2.61}$$

This implies that the number of particles N_+ minus the number of antiparticles N_- is a constant of the motion. In the free-field case, this conservation law is trivial, since $N_\pm$ are separately conserved. It becomes significant when, in the presence of interactions, $N_\pm$ are no longer conserved. In that case, $N_+ - N_-$ is still conserved as long as (2.61) holds.

2.8 MICROCAUSALITY

A classical signal propagating according to the Klein–Gordon equation has a group velocity

$$v_{\text{group}} = \frac{\partial \omega_k}{\partial |\mathbf{k}|} = \frac{|\mathbf{k}|}{\sqrt{\mathbf{k}^2 + m^2}} \tag{2.62}$$

which never exceeds 1. This means that events at two space–time points lying outside of each other's light cone (or separated by a spacelike interval) cannot influence each other. In the quantum theory, this means that two field operators at points separated by a spacelike interval must commute with each other:

$$[\phi(x), \phi(x')] = 0 \qquad \text{if} \quad (x - x')^2 < 0 \tag{2.63}$$

This condition is called *microcausality*. We must verify that our quantized field theory satisfies this condition.

To compute the commutator in (2.63), note that it is a c-number,[1] and at fixed x' it satisfies the Klein–Gordon equation, because $\phi(x)$ does. The initial condition at $x_0 = x'_0$ is the equal-time commutator (2.15), which is a c-number. Therefore the commutator remains a c-number at all times, and we can equate it with its vacuum expectation value:

$$[\phi(x), \phi(y)] = \langle 0|[\phi(x), \phi(y)]|0\rangle \equiv i\Delta(x - y) \tag{2.64}$$

This defines a Lorentz-invariant correlation function $\Delta(x - y)$, which depends on $x - y$, and not on x and y separately, because of the translational invariance of the vacuum state. (See Problem 2.1.) We use the expansion (2.41) to obtain

$$\begin{aligned} \langle 0|\phi(\mathbf{r}, t)\phi(0)|0\rangle &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} \\ \langle 0|\phi(0)\phi(\mathbf{r}, t)|0\rangle &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} \end{aligned} \tag{2.65}$$

Subtracting one from the other, we have

$$\Delta(x) = -i\langle 0|[\phi(\mathbf{r}, t), \phi(0)]|0\rangle = -\int \frac{d^3k}{(2\pi)^3} \frac{\sin(\omega_k t)}{\omega_k} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{2.66}$$

[1] A c-number is a "classical" number, a multiple of the identity operator.

Since $\Delta(x)$ is Lorentz-invariant, it can depend only on the invariant

$$x^2 = t^2 - \mathbf{x}^2 \tag{2.67}$$

If x is spacelike, for which $x^2 < 0$, we can put $t = 0$. By (2.66) this gives $\Delta(x) = 0$. ■

The proof of microcausality depends on the initial condition from the commutator in (2.15), which quantizes the system according to Bose statistics. Had we used Fermi statistics by replacing commutator with anticommutator, microcausality would have been violated. The particles here have spin 0, since there does not exist discrete degrees of freedom corresponding to spin. Our result is partial demonstration of the *spin-statistics connection,* which states that particles with integer spin are bosons, while those with half-integer spin are fermions. The second half of the statement will be shown in Chapter 6 on the Dirac field.

2.9 THE FEYNMAN PROPAGATOR

The propagation of a free particle in the vacuum can be described by the correlation function

$$\Delta^{(+)}(x_2 - x_1) = -i\langle 0|\psi(x_2)\psi^\dagger(x_1)|0\rangle \tag{2.68}$$

in which $\psi^\dagger(x_1)$ creates a particle from the vacuum at x_1, which is annihilated by $\psi(x_2)$ at x_2. This makes sense physically when $t_2 > t_1$. Similarly, the correlation function

$$\Delta^{(-)}(x_2 - x_1) = -i\langle 0|\psi^\dagger(x_1)\psi(x_2)|0\rangle \tag{2.69}$$

describes the propagation of a test antiparticle from x_2 to x_1, and is physically meaningful when $t_1 > t_2$. To obtain a correlation function that has physical meaning, we use either $\Delta^{(+)}$ or $\Delta^{(-)}$ depending on the sign of the relative time. The result is the *Feynman propagator,* or *causal propagator:*

$$\Delta_F(x_2 - x_1) = -i\langle 0|T\psi(x_2)\psi^\dagger(x_1)|0\rangle \tag{2.70}$$

where the time-ordering operator T rearranges the operators, if necessary, such that the operators stand in such order that time increases from right to left:

$$TA(t_2)B(t_1) \equiv \begin{cases} A(t_2)B(t_1) & \text{if} \quad t_2 > t_1 \\ B(t_1)A(t_2) & \text{if} \quad t_2 < t_1 \end{cases} \tag{2.71}$$

If $t_2 > t_1$, the Feynman propagator describes the propagation of a particle from x_1 to x_2; if $t_2 < t_1$, it describes the propagation of an antiparticle from x_2 to x_1. This is the

basis of Feynman's famous remark: "An antiparticle is a particle traveling backwards in time."

To calculate the propagator, we start with the expression

$$\Delta_F(x) = -i\begin{cases} \langle 0|\psi(x)\psi^\dagger(0)|0\rangle & \text{if} \quad x^0 > 0 \\ \langle 0|\psi^\dagger(0)\psi(x)|0\rangle & \text{if} \quad x^0 < 0 \end{cases} \tag{2.72}$$

and insert a complete set of states between the operators. Since the field operator connects the vacuum to one-particle states only, we have

$$\Delta_F(x) = -i\int \frac{d^3k}{(2\pi)^3} \begin{cases} \langle 0|\psi(x)|\mathbf{k}\rangle\,\langle \mathbf{k}|\psi^\dagger(0)|0\rangle & \text{if} \quad x^0 > 0 \\ \langle 0|\psi^\dagger(0)|\mathbf{k}\rangle\,\langle \mathbf{k}|\psi(x)|0\rangle & \text{if} \quad x^0 > 0 \end{cases} \tag{2.73}$$

Using $\psi(x) = e^{iP\cdot x}\psi(0)e^{-iP\cdot x}$, and changing the integration variable from $\mathbf{k}$ to $-\mathbf{k}$ in the lower formula, we obtain

$$\Delta_F(x) = -i\int \frac{d^3k}{(2\pi)^3}|\langle 0|\psi(0)|0\rangle|^2 e^{i\mathbf{k}\cdot\mathbf{r}}e^{-i\omega_k|t|} \tag{2.74}$$

Using the following integral representation

$$e^{-i\omega_k|t|} = \frac{i\omega}{\pi}\int_{-\infty}^{\infty} dk_0 \frac{e^{ik_0 t}}{k_0^2 - \omega_k^2 + i\eta} \qquad (\eta \to 0^+) \tag{2.75}$$

we obtain

$$\Delta_F(x) = \int \frac{d^4k}{(2\pi)^4} 2\omega_k |\langle 0|\psi(0)|0\rangle|^2 \frac{e^{ik\cdot x}}{k^2 - m^2 + i\eta} \qquad (\eta \to 0^+) \tag{2.76}$$

From (2.44) we have

$$|\langle 0|\psi(0)|0\rangle|^2 = \frac{1}{2\omega_k} \tag{2.77}$$

Therefore

$$\Delta_F(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik\cdot x}}{k^2 - m^2 + i\eta} \qquad (\eta \to 0^+) \tag{2.78}$$

Operating on both sides of this equation by $\Box^2 + m^2$ gives

$$(\Box^2 + m^2)\Delta_F(x) = -\delta^4(x) \tag{2.79}$$

This shows that the Feynman propagator is a Green's function of the Klein–Gordon equation.

The Fourier transform of the Feynman propagator is

$$\tilde{\Delta}_F(k) = \frac{1}{k^2 - m^2 + i\eta} \tag{2.80}$$

which has poles at

$$k_0 = \pm\sqrt{\mathbf{k}^2 + m^2} \tag{2.81}$$

corresponding to a particle or antiparticle of mass m. The residue $2\omega_k|\langle 0|\psi(0)|0\rangle|^2 = 1$ reflects the normalization of the wave function. We may view the propagator as the propagation amplitude of a virtual particle of 4-momentum k^μ. The virtual particle, whose squared mass k^2 ranges between $-\infty$ and ∞, becomes a real particle when it "goes on mass shell," at $k^2 = m^2$.

To obtain $\Delta_F(x)$ explicitly, we integrate over the angles of $\mathbf{k}$ in (2.74) to obtain

$$\Delta_F(x) = \frac{i}{4\pi^2}\int_0^\infty dk \frac{k^2}{\omega_k}\frac{\sin kr}{kr} e^{i\omega_k|t|} \tag{2.82}$$

By Lorentz invariance $\Delta_F(x)$ can depend only on

$$s \equiv x^2 = t^2 - \mathbf{r}^2 \tag{2.83}$$

For $s > 0$, we can put $\mathbf{r} = 0$ to obtain the representation

$$\Delta_F(x) = \frac{i}{4\pi^2}\int_0^\infty dk \frac{k^2}{\omega_k} e^{i\omega_k\sqrt{s}} = \frac{m}{8\pi\sqrt{s}} H_1^{(1)}(m\sqrt{s}) \qquad (s > 0) \tag{2.84}$$

For $s < 0$, we put $t = 0$ to obtain

$$\Delta_F(x) = \frac{i}{4\pi^2}\int_0^\infty dk \frac{k^2}{\omega_k}\frac{\sin k\sqrt{-s}}{k\sqrt{-s}} = -\frac{im}{4\pi^2\sqrt{-s}} K_1(m\sqrt{-s}) \qquad (s < 0) \tag{2.85}$$

where $H_1^{(1)}$ and K_1 are Bessel functions. In the timelike region $s > 0$ the function describes an outgoing wave for large s. This corresponds to the $i\eta$ prescription in (2.80). The $-i\eta$ prescription would have yielded an incoming wave. In the spacelike region $s < 0$ it damps exponentially for large $|s|$. On the light cone $s = 0$ there is a delta-function singularity not covered by the preceding formulas:

$$\lim_{x^2 \to 0} \Delta_F(x) = -\frac{1}{4\pi}\delta(x^2) \tag{2.86}$$

2.10 THE WAVE FUNCTIONAL

In quantum mechanics, the coordinate representation is defined by basis states $|\mathbf{r}\rangle$ satisfying

$$\mathbf{r}_{\text{op}}|\mathbf{r}\rangle = \mathbf{r}|\mathbf{r}\rangle \tag{2.87}$$

with the basic commutator realized through the replacements

$$\begin{aligned} \mathbf{r}_{\text{op}} &\rightarrow \mathbf{r} \qquad \text{(c-number)} \\ \mathbf{p}_{\text{op}} &\rightarrow -i\nabla \end{aligned} \tag{2.88}$$

A state $|A\rangle$ is represented by the wave function

$$\psi_A(\mathbf{r}) = \langle \mathbf{r}|A\rangle \tag{2.89}$$

and inner products are defined by

$$\langle A|B\rangle = \int d^3r \psi_A^*(\mathbf{r})\, \psi_B(\mathbf{r}) \tag{2.90}$$

In the analogous *field representation* in quantum field theory, we diagonalize the field operator, thus representing it by its eigenvalue, which is a c-number function. For a real scalar field $\phi_{\text{op}}(\mathbf{r})$ at a fixed time $t = 0$, we denote its eigenstates by $|\phi\rangle$:

$$\phi_{\text{op}}(\mathbf{r})|\phi\rangle = \phi(\mathbf{r})|\phi\rangle \tag{2.91}$$

where the eigenvalue $\phi(\mathbf{r})$ is a real-valued function of $\mathbf{r}$. The commutation relations (2.15) are realized through the replacements

$$\begin{aligned} \phi_{\text{op}}(\mathbf{r}) &\rightarrow \phi(\mathbf{r}) \qquad \text{(c-number function)} \\ \dot{\phi}_{\text{op}}(\mathbf{r}) &\rightarrow -i\frac{\delta}{\delta\phi(\mathbf{r})} \end{aligned} \tag{2.92}$$

where $\delta/\delta\phi(\mathbf{r})$ is denotes the functional derivative with respect to the value of the function ϕ at $\mathbf{r}$.

A state $|A\rangle$ is represented in the field representation by the wave functional

$$\Psi_A[\phi] = \langle \phi|A\rangle \tag{2.93}$$

which is a complex-valued function whose argument is a function; that is, its value depends on the form of the function. Inner products between wave functional are functional integrals:

$$(\Psi_A, \Psi_B) = \int D\phi\, \Psi_A^*[\phi]\Psi_B[\phi] \tag{2.94}$$

where $D\phi$ denotes the measure on the space of functions. Writing the Hamiltonian in the field representation, we have the Schrödinger equation for the wave functional:

$$\frac{1}{2}\int d^3r\left[-\frac{\delta^2}{\delta\phi(\mathbf{r})^2} + |\nabla\phi(\mathbf{r})|^2 + m^2\phi^2(\mathbf{r})\right]\Psi[\phi, t] = i\frac{\partial\Psi[\phi, t]}{\partial t} \tag{2.95}$$

2.11 FUNCTIONAL OPERATIONS

We digress on functional operations on a functional $F[\phi]$. First, the functional derivative $\delta F[\phi]/\delta\phi(x)$ is defined as follows. We make a small change $\phi \to \phi + \delta\phi$, where the function $\delta\phi(x)$ is different from zero only in the neighborhood of x. Then the functional derivative is given by

$$\frac{\delta F[\phi]}{\delta\phi(x)} = \lim_{\delta\phi\to 0}\left[\frac{F[\phi + \delta\phi] - F[\phi]}{\delta\phi(x)}\right] \tag{2.96}$$

To calculate any functional derivative, we need the elementary functional derivative $\delta\phi(x)/\delta\phi(y)$, which is obviously proportional to $\delta(x - y)$. To determine the proportionality constant, we replace the continuous space of x by a lattice of spacing a, and denote by ϕ_j the value of the function on site j. Clearly,

$$\frac{\partial\phi_j}{\partial\phi_k} = \delta_{jk} \tag{2.97}$$

In the continuum limit

$$\sum_j \to a^{-1}\int dx \tag{2.98}$$

$$\delta_{jk} \to a\delta(x - y)$$

we have

$$\frac{\delta\phi(x)}{\delta\phi(y)} = \delta(x - y) \tag{2.99}$$

With this formula, we can calculate a general functional derivative. As illustration, take $F[\phi] = \int d^3y|\nabla\phi(\mathbf{y})|^2$. Then

$$\frac{\delta F[\phi]}{\delta \phi(x)} = -2\int d^3y\, \delta^3(\mathbf{x}-\mathbf{y})\nabla^2\phi(\mathbf{y})$$
$$= -2\nabla^2\phi(\mathbf{x}) \tag{2.100}$$

where we have assumed that boundary conditions are chosen such that surface integrals from partial integrations vanish.

A functional integral is defined as the limit of an ordinary integral. Again, let us replace continuous space by a lattice of discrete points labeled by j. A function $\phi(x)$ becomes a discrete set of values $\{\phi_j\}$, and a functional $F[\phi]$ becomes a function of this set $F(\phi_1,\phi_2,\ldots)$. The functional integral of F is defined as

$$\int D\phi\, F[\phi] = C\prod_j \int_{-\infty}^{\infty} d\phi_j\, F(\phi_1,\phi_2,\ldots) \tag{2.101}$$

where C is a normalization constant. The continuum limit is to be taken eventually.

Alternatively, we put the system in a large cube of volume Ω, with periodic boundary conditions. The Fourier components $\tilde{\phi}(k)$ of $\phi(x)$ are then discrete. The functional integral over ϕ can be defined as the multiple integral over all Fourier components independently:

$$\int D\phi\, F[\phi] = C\prod_k \int_{-\infty}^{\infty} d\tilde{\phi}(k) F[\phi] \tag{2.102}$$

Eventually we take the limit $\Omega \to \infty$.

In either of the preceding methods, the integral by itself may not have a continuum limit; but matrix elements of the form

$$O_{AB} = \frac{(\Psi_A, O\Psi_B)}{\sqrt{(\Psi_A,\Psi_A)(\Psi_B,\Psi_B)}} \tag{2.103}$$

usually has a definite continuum limit.

2.12 VACUUM WAVE FUNCTIONAL

We now calculate the wave functional for the free vacuum state. First let us express the annihilation operator $a(\mathbf{k})$ in the field representation. From (2.41) we have, at $t = 0$,

$$\int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}}\phi(\mathbf{r}) = \frac{1}{\sqrt{2\omega_k}}[a(\mathbf{k}) + a^\dagger(\mathbf{k})]$$

$$\int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}}\dot{\phi}(\mathbf{r}) = -i\sqrt{\frac{\omega_k}{2}}[a(\mathbf{k}) - a^\dagger(\mathbf{k})] \tag{2.104}$$

Solving for $a(\mathbf{k})$, we obtain

$$a(\mathbf{k}) = \frac{1}{\sqrt{2\omega_k}} \int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}} [\omega_k \phi(\mathbf{r}) + i\dot{\phi}(\mathbf{r})] \tag{2.105}$$

Now we write

$$e^{-i\mathbf{k}\cdot\mathbf{r}} \omega_k = \sqrt{\mathbf{k}^2 + m^2}\, e^{-i\mathbf{k}\cdot\mathbf{r}} = \sqrt{-\nabla^2 + m}\, e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{2.106}$$

so that

$$\begin{aligned} a(\mathbf{k}) &= \frac{1}{\sqrt{2\omega_k}} \int d^3r \left[\phi(\mathbf{r}) \sqrt{-\nabla^2 + m} + i\dot{\phi}(\mathbf{r}) \right] e^{-i\mathbf{k}\cdot\mathbf{r}} \\ &= \frac{1}{\sqrt{2\omega_k}} \int d^3r \left\{ \left[\sqrt{-\nabla^2 + m}\, \phi(\mathbf{r}) \right] + i\dot{\phi}(\mathbf{r}) \right\} e^{-i\mathbf{k}\cdot\mathbf{r}} \end{aligned} \tag{2.107}$$

The last step is obtained by expanding $\sqrt{-\nabla^2 + m}$ in a power series in ∇^2, performing a partial integration in every term, and summing the series again. The surface integrals generated in the partial integrations vanish as a result of periodic boundary conditions. Replacing $i\dot{\phi}(\mathbf{r})$ by $\delta/\delta\phi(\mathbf{r})$, we obtain

$$a(\mathbf{k}) = \frac{1}{\sqrt{2\omega_k}} \int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}} \left[\sqrt{-\nabla^2 + m^2}\, \phi(\mathbf{r}) + \frac{\delta}{\delta\phi(\mathbf{r})} \right] \tag{2.108}$$

The virtue of this representation is that the Fourier coefficient in the integrand is independent of $\mathbf{k}$.

The wave functional of the free vacuum satisfies the equation

$$a(\mathbf{k})\Psi_0[\phi] = 0 \tag{2.109}$$

Thus, it must be annihilated by the Fourier coefficient in (2.108):

$$\left[\sqrt{-\nabla^2 + m^2}\, \phi(\mathbf{r}) + \frac{\delta}{\delta\phi(\mathbf{r})} \right] \Psi_0[\phi] = 0 \tag{2.110}$$

The solution to this equation is

$$\Psi_0[\psi] = C \exp\left[-\tfrac{1}{2} \int d^3r\, \phi(\mathbf{r})\, \sqrt{-\nabla^2 + m^2}\, \phi(\mathbf{r}) \right] \tag{2.111}$$

where C is a normalization constant. This gives the probability amplitude that the field has the functional form $\phi(\mathbf{r})$ in the vacuum state. The relative probability for the field to have a functional form lying in the neighborhood of ϕ in the volume element $D\phi$ of function space is

$$|\Psi[\phi]|^2 D\phi$$

The most probable form is $\phi \equiv 0$, and deviations from it occur with a Gaussianlike distribution.

The exponent in (2.111) can be rewritten in different forms. Introducing the Fourier transform

$$\tilde{\phi}(\mathbf{k}) = \int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}} \phi(\mathbf{r}) \tag{2.112}$$

we can write

$$\int d^3r\, \phi(\mathbf{r}) \sqrt{-\nabla^2 + m^2}\, \phi(\mathbf{r}) = \int \frac{d^3k}{(2\pi)^3} \sqrt{\mathbf{k}^2 + m^2}\, |\tilde{\phi}(\mathbf{k})|^2$$

$$= \int d^3r\, d^3r'\, \phi(\mathbf{r}) K(\mathbf{r} - \mathbf{r}')\, \phi(\mathbf{r}') \tag{2.113}$$

where

$$K(\mathbf{r}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} \sqrt{\mathbf{k}^2 + m^2} \tag{2.114}$$

For a complex scalar field, there are now two coordinates, which can be taken as either $\{\phi_1, \phi_2\}$ or $\{\psi, \psi^\dagger\}$. Inner products of wave functionals now take the form

$$(\Psi_A, \Psi_B) = \int D\phi_1\, D\phi_2 \Psi_A^*[\phi_2, \phi_2] \Psi_B[\phi_2, \phi_2] \tag{2.115}$$

or equivalently

$$(\Psi_A, \Psi_B) = \int D\psi D\psi^* \Psi_A^*[\psi, \psi^*] \Psi_B[\psi, \psi^*] \tag{2.116}$$

The complex measure is defined in terms of the real and imaginary parts:

$$D\psi\, D\psi^* \equiv D\phi_1\, D\phi_2 \tag{2.117}$$

The vacuum wave functional for the free complex field is just the product of those for the two independent real fields. Reexpressing the result in terms of the complex field, we have

$$\Psi_0[\psi] = C \exp\left[-\int d^3r\, \psi^*(\mathbf{r}) \sqrt{-\nabla^2 + m^2}\, \psi(\mathbf{r})\right] \tag{2.118}$$

2.13 THE ϕ^4 THEORY

As the simplest example of an interacting field theory, consider the Lagrangian density of the so-called ϕ^4 theory:

$$\mathcal{L}(x) = \tfrac{1}{2}\partial^\mu\phi\partial_\mu\phi - \tfrac{1}{2}m^2\phi^2 - g\phi^4 \tag{2.119}$$

where $g > 0$. The quartic term makes the equation of motion nonlinear:

$$(\Box^2 + m^2)\phi + g\phi^3 = 0 \tag{2.120}$$

The Hamiltonian density takes the form

$$\mathcal{H}(x) = \frac{1}{2}\left[\left(\frac{\partial\phi}{\partial t}\right)^2 + |\nabla\phi|^2\right] + V(\phi(x))$$

$$V(\phi(x)) = \frac{1}{2}m^2\phi^2(x) + g\phi^4(x) \tag{2.121}$$

which suggests that $V(\phi(x))$ is a potential.

To quantize the theory, we impose the equal-time commutators (2.15), which can be satisfied by taking as initial condition

$$\phi(\mathbf{r}, 0) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{k}}\frac{1}{\sqrt{2\omega_k}}[a_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{r}} + a_{\mathbf{k}}^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}}]$$

$$= \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{k}}\frac{1}{\sqrt{2\omega_k}}e^{i\mathbf{k}\cdot\mathbf{r}}(a_{\mathbf{k}} + a_{-\mathbf{k}}^\dagger) \tag{2.122}$$

where the creation and annihilation operators $a_{\mathbf{k}}^\dagger$, $a_{\mathbf{k}}$ are defined by the commutation relations

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}^\dagger] = \delta_{\mathbf{k}\mathbf{k}'}$$

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}] = 0 \tag{2.123}$$

The equation of motion is not soluble unless $g = 0$. We can always write, as a formal solution,

$$\phi(\mathbf{r}, t) = e^{iHt}\phi(\mathbf{r}, 0)e^{-iHt} \tag{2.124}$$

but this is not simple unless $g = 0$.

To see the effects of the interactions, separate the Hamiltonian into a "free" term and an "interaction" term, at some arbitrary time $t = 0$:

$$H = H_0 + H_{\text{int}} \tag{2.125}$$

where

$$H_0 = \tfrac{1}{2}\int d^3r[(\dot{\phi}^2(\mathbf{r}, 0) + |\nabla\phi(\mathbf{r}, 0)|^2]$$

$$H_{\text{int}} = g\int d^3r\phi^4(\mathbf{r}, 0) \tag{2.126}$$

In terms of the creation and annihilation operators we have

$$H_0 = C_0 + \sum_{\mathbf{k}} \omega_k a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$$

$$H_{\text{int}} = \frac{g}{\Omega} + \sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} \frac{\delta_K(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_4)}{2\sqrt{\omega_1\omega_2\omega_3\omega_4}}(a_1^\dagger + a_1)(a_2^\dagger + a_2)(a_3^\dagger + a_3)(a_4^\dagger + a_4) \tag{2.127}$$

where C_0 is an irrelevant zero-point energy and δ_K denotes the Kronecker δ. We use the shorthand $\omega_1 = \omega_{\mathbf{k}_1}$, $a_1 = a_{\mathbf{k}_1}$ and so on. The interaction Hamiltonian H_{int} describes four-particle processes that conserve momentum. Substitution of this expansion into (2.124) generates a complicated series for the time-dependent field operator. We shall learn how to organize such terms in a systematic manner in Chapter 9.

PROBLEMS

2.1 Space–Time Translation Consider a free scalar field $\phi(x)$, which can be expanded in terms of the annihilation operators $a_{\mathbf{k}}$. This problem illustrates the fact that the 4-momentum $P^\mu = \Sigma_{\mathbf{k}} k^\mu a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$ is the generator of space-time translations.

(a) As a useful tool show that, for two operators A and B,

$$e^A B e^{-A} = B + [A, B] + \tfrac{1}{2}[A, [A, B]] + \tfrac{1}{3!}[A, [A, [A, B]]] + \cdots$$

(b) Use this formula to show

$$e^{iP\cdot x} a_{\mathbf{k}} e^{-iP\cdot x} = a_{\mathbf{k}} e^{-ik\cdot x}$$

and the infinitesimal form

$$[P^\mu, a_{\mathbf{k}}] = -k^\mu a_{\mathbf{k}}$$

(c) Establish that P^μ is the generator of space–time translations by showing

$$[P^\mu, \phi(x)] = i\partial^\mu\phi(x)$$

(d) Let $|K\rangle$ be an eigenstate of P^μ, satisfying $P^\mu|K\rangle = K^\mu|K\rangle$. Show that this state is translationally invariant:

$$\langle K|\phi(x)\,\phi(y)|K\rangle = \langle K|\phi(x - y)\,\phi(0)|K\rangle$$

2.2 Charge Conjugation The designation of particle and antiparticle is a matter of convention, and we can freely reverse the labels. More specifically, for a complex scalar

field $\psi(x)$, construct an operator C that takes $b_{\mathbf{k}}$ to $c_{\mathbf{k}}$ and vice versa, and commutes with Hamiltonian H:

$$C\psi(x)C^{-1} = \psi^{\dagger}(x)$$
$$[H, C] = 0$$
$$C^{\dagger}C = 1$$
$$C^2 = 1$$

The operation C is called *charge conjugation,* or *particle–antiparticle* conjugation.

2.3 Lorentz Invariance We have calculated the function $\Delta(x) = \langle 0|[\phi(x), \phi(0)]|0\rangle$ in (2.66), but not in a manifestly Lorentz-invariant form. Show that it can be put into the desired form

$$\Delta(x) = 2\pi i \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot x}\delta(k^2 - m^2)\epsilon(k_0)$$

2.4 Spin and Statistics Quantize the real scalar field according to Fermi statistics; in other words replace the commutators in (2.15) by anticommutators. Show that this will violate microcausality.

2.5 External Source Consider a real scalar field $\phi(x)$\coupled to an external source function $J(x)$, with Lagrangian density

$$\mathcal{L}(x) = \tfrac{1}{2}\partial^{\mu}\phi\partial_{\mu}\phi + \tfrac{1}{2}m^2\phi^2 + J\phi$$

(a) Obtain the Hamiltonian in terms of the creation and annihilation operators $a_{\mathbf{k}}$, $a_{\mathbf{k}}^{\dagger}$ for plane-wave states.

(b) Suppose that the source is static, that is, that $J(\mathbf{x},t)$ is independent of t. Using perturbation theory, show that there is no scattering from the fixed source to second order.

(c) Show that there is no scattering at all, to any order. (*Hint:* Show that a linear canonical transformation of $a_{\mathbf{k}}$ reduces the Hamiltonian to the source-free case.)

2.6 Level Shift Suppose that the external source in the previous problem consists of a single static point source: $J(x) = g\delta^3(\mathbf{r})$.

(a) Calculate the change in the energy of the vacuum state to order g^2. The result will be a divergent integral. Cut it off at a large momentum Λ. This illustrates a prototype of divergence in quantum field theory.

(b) Show that all levels of the system shift by the same amount and therefore that the divergence in this case has no physical relevance.

2.7 Yukawa Potential Continuing with the last two problems, suppose the source function $J(x)$ consists of two static point sources located at $\mathbf{r}_1$,$\mathbf{r}_2$:

$$J(\mathbf{r}, t) = g[\delta^3(\mathbf{r} - \mathbf{r}_1) + \delta^3(\mathbf{r} - \mathbf{r}_2)]$$

Treating g as a perturbation, calculate the change in the vacuum energy to second order in g, and show that there is an attractive potential between the two point sources:

$$V(R) = -\frac{g^2}{4\pi}\frac{e^{-mR}}{R}$$

where $R = |\mathbf{r}_1 - \mathbf{r}_2|$. This is the Yukawa potential, originally proposed as the potential between two nucleons due to interactions with scalar mesons.

2.8 Vacuum Fluctuations Consider a free scalar field in a large periodic box of volume Ω. Let the Fourier transform be denoted

$$\tilde{\phi}(\mathbf{k}) = \frac{1}{\sqrt{\Omega}}\int d^3r e^{-i\mathbf{k}\cdot\mathbf{r}}\phi(\mathbf{r})$$

(a) Show that the vacuum expectation value of $\tilde{\phi}(\mathbf{k})$ is zero.

(b) By expanding the field in terms of creation and annihilation operators, show that the mean-square fluctuation of the Fourier transform is given by

$$\langle 0|\tilde{\phi}^2(\mathbf{k})|0\rangle = \frac{1}{2\sqrt{\mathbf{k}^2 + m^2}}$$

(c) The mean-square average can be expressed in the field representation as

$$\langle 0|\tilde{\phi}^2(\mathbf{k})|0\rangle = \frac{\int D\phi\phi^2(\mathbf{k})\Psi_0^2[\phi]}{\int D\phi\Psi_0^2[\phi]}$$

where $\Psi_0[\phi]$ is the wave functional ofproduce the last result from this formula.

(d) Calculate the mean-square fluctuation $\langle 0|\phi^2(x)|0\rangle$ in coordinate space. The result is divergent because of the high-momentum modes. Exhibit its dependence on the cutoff momentum Λ.

CHAPTER THREE

Relativistic Fields

3.1 LORENTZ TRANSFORMATIONS

Relativistic quantum fields can be classified according to the way they transform under Lorentz transformations. More specifically, they transform according to irreducible representations of the Lorentz group. The different representations give rise to particles with different values of the spin angular momentum.

According to the principle of special relativity, the laws of physics should be covariant with respect to Lorentz transformations; that is, they should have the same forms in all reference frames connected by Lorentz transformations. The simplest Lorentz transformation is a "boost" of the reference frame with velocity v along some axis, say, the x axis:

$$t' = \frac{t - vx}{\sqrt{1 - v^2}}$$

$$x' = \frac{x - vt}{\sqrt{1 - v^2}} \tag{3.1}$$

This may be supplemented by a rotation of the coordinate system, say, about the z axis through an angle θ:

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta \tag{3.2}$$

Defining a boost "angle" ϕ by

$$\cosh\phi = 1/\sqrt{1 - v^2}$$

$$\sinh\phi = v/\sqrt{1 - v^2} \tag{3.3}$$

we can write the matrices of these transformations as follows:

$$\text{Lorentz boost:} \quad \begin{pmatrix} \cosh\phi & -\sinh & 0 & 0 \\ -\sinh & \phi & 0 & 0 \\ \phi & \cosh\phi & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\text{Rotation:} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.4}$$

The inverses of these matrices can be obtained by reversing the signs of ϕ and θ. The rotation matrices are orthogonal matrices, while the Lorentz boosts are not, because the invariant form $t^2 - x^2$ for the Lorentz boost is not positive-definite.

The angles of rotation are not additive, unless the rotations are all made about the same axis. Similarly, the velocities of successive Lorentz boosts are not additive, unless the boosts are all made along the same direction.

We use a relativistic notation in which the coordinate 4-vector is denoted by $x^\mu = (t, \mathbf{r})$ and the metric tensor is diagonal:

$$g^{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{3.5}$$

A general Lorentz transformation is a linear transformation Λ on x that leaves $x^2 = t^2 - \mathbf{r}^2$ invariant:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{3.6}$$

with the requirement

$$g_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = g_{\alpha\beta} \tag{3.7}$$

which ensures the invariance of x^2. In shorthand, we write the transformation in the form

$$x' = \Lambda x \tag{3.8}$$

The transformations above form the continuous Lorentz group, which is characterized by six parameters: three velocity components and three angles of rotation. As we can see from (3.4), they are represented by matrices with determinant +1. In contrast, the discrete transformations

$$\begin{aligned} \text{Spatial reflection:} \quad & t' = t \qquad \mathbf{x}' = -\mathbf{x} \\ \text{Time reversal:} \quad & t' = -t \qquad \mathbf{x}' = \mathbf{x} \end{aligned} \tag{3.9}$$

have determinant −1. These discrete elements together with the continuous Lorentz transformations form the general Lorentz group. We shall reserve the name "Lorentz transformation" for the continuous Lorentz transformations.

Any element of the Lorentz group can be built up from infinitesimal ones, with the general form

$$\Lambda^{\mu}{}_{\nu} = g^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} \tag{3.10}$$

We write in shorthand

$$\Lambda = 1 + \omega \tag{3.11}$$

Lorentz transformations generally do not commute with one other; but the infinitesimal transformations do, because their commutators are of second-order smallness:

$$(1 + \omega_1)(1 + \omega_2) = 1 + \omega_1 + \omega_2 + O(\omega^2) \tag{3.12}$$

Thus, group multiplication is equivalent to addition of the ω's.

An infinitesimal transformation of the coordinate system, characterized by boosts with velocities v^j along the x^j axes, and rotations of angles θ^k about the x^k axes, is described by the tensor

$$\omega^{\mu}{}_{\nu} = \begin{pmatrix} 0 & -v^1 & -v^2 & -v^3 \\ -v^1 & 0 & \theta^3 & -\theta^2 \\ -v^2 & -\theta^3 & 0 & \theta^1 \\ -v^3 & \theta^2 & -\theta^1 & 0 \end{pmatrix} \tag{3.13}$$

By raising the lower index, we obtain an antisymmetric tensor

$$\omega^{\mu\nu} = g^{\nu\lambda}\omega^{\mu}{}_{\nu} = \begin{pmatrix} 0 & v^1 & v^2 & v^3 \\ -v^1 & 0 & -\theta^3 & \theta^2 \\ -v^2 & \theta^3 & 0 & -\theta^1 \\ -v^3 & -\theta^2 & \theta^1 & 0 \end{pmatrix} \tag{3.14}$$

whose elements can be summarized as follows:

$$\begin{aligned} \omega^{0k} &= -\omega^{k0} = v^k \\ \omega^{ij} &= -\omega^{ji} = -\epsilon^{ijk}\theta^k \end{aligned} \tag{3.15}$$

3.2 MINIMAL REPRESENTATION: $SL(2C)$

It is well known that the smallest faithful representation of the rotation group is $SU(2)$, the group of 2×2 unitary matrices of unit determinant. For the Lorentz group, the minimal representation is $SL(2C)$, the linear group of 2×2 complex matrices of unit determinant. To see this, let us organize the coordinates into a 2×2 complex matrix:

$$X \equiv t + (\boldsymbol{\sigma} \cdot \mathbf{x}) = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} \tag{3.16}$$

where σ^k are the Pauli matrices, with the following properties:

$$\begin{aligned} &\{\sigma^j, \sigma^k\} = \delta_{jk} \\ &\sigma^1\sigma^2 = i\sigma^3 \qquad \text{(and cyclic permutations)} \\ &[\sigma^1, \sigma^2] = 2i\sigma^3 \qquad \text{(and cyclic permutations)} \end{aligned} \tag{3.17}$$

We see that

$$\det X = x^2 \tag{3.18}$$

A Lorentz transformation that takes X into X' can be represented by the operation

$$X' = L(\Lambda) X L^\dagger(\Lambda) \tag{3.19}$$

where $L(\Lambda)$ is a 2×2 complex matrix and $L^\dagger(\Lambda)$ its Hermitian conjugate. Taking the determinant of both sides, we have

$$\det X' = \det X |\det L(\Lambda)|^2 \tag{3.20}$$

To preserve x^2, we must have $\det X' = \det X$, and hence

$$\det L(\Lambda) = \pm 1 \tag{3.21}$$

Consequently $\det \Lambda = \pm 1$. This is a more formal proof of a result stated earlier. The matrices $L(\Lambda)$ with $\det L = 1$ constitute the group $SL(2C)$.

Any 2×2 can be represented in the form

$$A + (\mathbf{B} \cdot \boldsymbol{\sigma}) = \begin{pmatrix} A + B_3 & B_1 - iB_2 \\ B_1 + iB_2 & A - B_3 \end{pmatrix} \tag{3.22}$$

where A and B_k are complex numbers. The determinant of the preceding equation is

$A^2 - \Sigma_k B_k^2$. Hence $L(\Lambda)$ is a matrix of this form, with $A^2 - \Sigma_k B_k^2 = 1$. We leave it as exercises to show that a pure boost and pure rotation are represented by the following:

$$\text{Boost along } \hat{\mathbf{n}}: \quad L(\Lambda) = e^{-\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\phi/2} = \cosh\frac{\phi}{2} - (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sinh\frac{\phi}{2}$$

$$\text{Rotation about } \hat{\mathbf{n}}: \quad L(\Lambda) = e^{i\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\theta/2} = \cos\frac{\theta}{2} + i(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2} \tag{3.23}$$

where $\hat{\mathbf{n}}$ is a unit vector, ϕ is the boost angle defined in (3.3), and θ is the rotation angle.

3.3 THE POINCARÉ GROUP

The laws of physics should be covariant with respect to space–time translations as well as Lorentz transformations. These transformations combined constitute the inhomogeneous Lorentz group, or the Poincaré group. The transformation law is as follows:

$$x'^{\mu} = a^{\mu} + \Lambda^{\mu}{}_{\nu} x^{\nu} \tag{3.24}$$

where a^{μ} is a 4-vector. The infinitesimal version has the form

$$x'^{\mu} = x^{\mu} + a^{\mu} + \omega^{\mu}{}_{\nu} x^{\nu} \tag{3.25}$$

which contains 10 independent parameters: a^{μ} and $\omega^{\mu\nu} = -\omega^{\nu\mu}$.

We can realize the Poincaré group on the space of functions $f(x)$, through the transformation

$$\begin{aligned} f(x') &= f(x + a + \omega x) \\ &= f(x) + a_{\mu}\partial_{\mu} f(x) + \omega^{\mu}{}_{\nu} x^{\nu}\partial_{\mu} f(x) \\ &= [1 + a_{\mu}\partial^{\mu} - \tfrac{1}{2}\omega_{\mu\nu}(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu})] f(x) \end{aligned} \tag{3.26}$$

where we have use the fact that $\omega_{\mu\nu}$ is antisymmetric. We can rewrite

$$f(x') = \left(1 - ia^{\mu}P^{\mu} + \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu}\right) f(x) \tag{3.27}$$

which defines the generators

$$P^{\mu} = i\partial^{\mu}$$

$$M^{\mu\nu} = x^\mu P^\nu - x^\nu P^\mu \tag{3.28}$$

Of these, 10 are independent operators, constituting the Lie algebra of the Poincaré group. An arbitrary element of the Poincaré group can be written in the form

$$\exp\left(ia_\mu P^\mu - i\omega_{\mu\nu} M^{\mu\nu}\right) \tag{3.29}$$

where a^μ and $\omega^{\mu\nu}$ represent 10 real independent parameters.

From (3.28) we obtain the commutator

$$[x^\mu, P^\nu] = -ig^{\mu\nu} \tag{3.30}$$

Although derived from an explicit representation, we consider the preceding equations as abstract algebraic relations. Such a procedure is analogous to obtaining the Lie algebra $[J^j, J^k] = i\epsilon^{jkl}J^l$ for angular momentum from the special representation $\mathbf{J} = -i\mathbf{r} \times \nabla$. As abstract relations, the Lie algebra admits half-integer representations.

The Lie algebra of the Poincaré group consists of the following commutators:

$$\begin{aligned}
[M^{\mu\nu}, M^{\alpha\beta}] &= -i(g^{\mu\alpha}M^{\nu\beta} - g^{\nu\alpha}M^{\mu\beta} + g^{\nu\beta}M^{\mu\alpha} - g^{\mu\beta}M^{\nu\alpha}) \\
[M^{\lambda\mu}, P^n] &= i(g^{\mu\nu}P^\lambda - g^{\lambda\nu}P^\mu) \\
[P^\mu, P^n] &= 0
\end{aligned} \tag{3.31}$$

which can be obtained through a straightforward calculation. In physical terms, the four generators

$$P^\mu = (H, P^1, P^2, P^3) \tag{3.32}$$

make up the total 4-momentum operator, and $P^0 = H$ is the Hamiltonian. The six independent components of $M^{\mu\nu}$ are generalized angular momentum operators made up of the angular momentum $\mathbf{J}$ and the Lorentz boost $\mathbf{K}$:

$$\begin{aligned}
M^{jk} &= \epsilon^{jkl}J^l \\
M^{0j} &= K^j
\end{aligned} \tag{3.33}$$

We can recast the Poincaré algebra as follows. The last two lines in (3.31) are equivalent to

$$\begin{aligned}
&[P^j, P^k] = [P^j, H] = [J^j, H] = 0 \\
&[J^j, P^k] = -i\epsilon^{jkl}P^l \\
&[K^j, H] = -iP^j \\
&[K^j, P^k] = -i\delta_{jk}H
\end{aligned} \tag{3.34}$$

These relations all involve the inhomogeneous part of the group. The first equation above expresses the independence of the spatial translations among themselves, of spatial and time translations, and of rotations and time translations. The second equation is what one can deduce from $\mathbf{J} = -i\mathbf{r} \times \nabla$ and $\mathbf{P} = -i\nabla$. The other equations above describe how energy and momentum change under a Lorentz boost. In addition to these, we obtain from (3.31) a closed set of commutation relations among angular momentum and boost operators:

$$\begin{aligned} [J^j, J^k] &= i\epsilon^{jkl}J^l \\ [K^j, K^k] &= -i\epsilon^{jkl}J^l \\ [J^j, K^k] &= i\epsilon^{jkl}K^l \end{aligned} \tag{3.35}$$

These form the Lie algebra of the Lorentz group.

3.4 SCALAR, VECTOR, AND SPINOR FIELDS

In quantum mechanics, the wave functions in a central potential can be classified according to orbital angular momenta, which correspond to irreducible representations of the rotation group, with possible dimensions $2l + 1$, ($l = 0,1,2, \ldots$). In a similar way, relativistic fields transform according to irreducible representations of the Lorentz group, which have definite dimensions. Accordingly, a relativistic field has a definite number of components, related to the spin angular momentum of the field.

The simplest relativistic field is a scalar field, which may have more than one component, but each component $\phi(x)$ must be invariant under Lorentz transformations:

$$\phi'(x') = \phi(x) \tag{3.36}$$

This says that the transformed field called ϕ', at the transformed coordinate x', is the same as the original field called ϕ, at the old coordinate x. It expresses the fact that x' and x are different labels that we use for the same physical point, and the scalar field is unaffected by this; but for us the functional form of the field must change:

$$\phi'(x) = \phi(\Lambda^{-1}x) \tag{3.37}$$

As we shall see, the spin of a scalar field is zero.

A vector field, such as the electromagnetic field $A^\mu(x)$, *is* affected by a change in the coordinate system, since by definition its four components transform among themselves like x^μ. The transformation law is

$$A'^{\mu}(x') = \Lambda^{\mu}{}_{\nu} A^{\nu}(x) \qquad (\mu = 0,1,2,3) \tag{3.38}$$

The spin of a vector field is 1. This will be demonstrated in Section 5.5.

In general, a tensor field of rank n transforms like a product of n x^{μ} terms, and corresponds to spin n. For example, the gravitational field is a symmetric tensor of rank 2.

There are "half-integer" representations, analogous to those for the rotation group. The latter are representations of $SU(2)$, which generalizes to $SL(2C)$ in the present case. To accommodate space–time reflections, we have to include two copies of $SL(2C)$, so that they transform into each other under a reflection. Accordingly, the minimal representation space is spanned by a four-component complex field, called the *Dirac spinor field* $\psi_i(x)$, which transforms according to

$$\psi'_r(x') = S_{rs}(\Lambda)\psi_s(x) \qquad (r = 1,2,3,4) \tag{3.39}$$

where $S(\Lambda)$ is a 4×4 complex matrix, discussed in more detail in Chapter 6. The spin of a spinor field is $\frac{1}{2}$.

In general, a field forming a K-dimensional irreducible representation of the Lorentz group has K components:

$$\phi_a(x) \qquad (a = 1,2, \ldots, K) \tag{3.40}$$

which transform under a Lorentz transformation Λ according to

$$\phi'_a(x') = S_{ab}(\Lambda)\phi_b(x) \tag{3.41}$$

For an infinitesimal transformation $\Lambda = 1 + \omega$, we can put $S(\Lambda)$ in the form

$$S_{ab} = \delta_{ab} + \tfrac{1}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}_{ab} \tag{3.42}$$

this defines the coefficients $\Sigma^{\mu\nu}_{ab}$, which, as we will show, constitute the spin matrix.

Under an infinitesimal Lorentz transformation, then, a general field transforms according to

$$\phi'_a(x') = \phi_a(x) + \tfrac{1}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}_{ab}\phi_b(x) \tag{3.43}$$

The change in the functional form of the field can be found by writing

$$\begin{aligned}\phi'_a(x') = \phi'_a(x + \omega x) &= \phi'_a(x) + \omega_{\mu\nu}x^{\nu}\partial_{\mu}\phi'_a(x) \\ &= \phi'_a(x) - \tfrac{1}{2}\omega_{\mu\nu}(x^{\mu}\partial^{\nu} - x^{\nu}\partial_{\mu})\phi'_a(x)\end{aligned} \tag{3.44}$$

Thus

$$\phi'_a(x) = \phi'_a(x') + \tfrac{1}{2}\omega_{\mu\nu}(x^\mu\partial^\nu - x^\nu\partial^\mu)\phi_a(x) \tag{3.45}$$

Substituting $\phi'_a(x')$ from (3.42), we obtain

$$\phi'_a(x) = \phi_a(x) + \tfrac{1}{2}\omega_{\mu\nu}[(x^\mu\partial^\nu - x^\nu\partial^\mu)\delta_{ab} + \Sigma^{\mu\nu}_{ab}]\phi_b(x) \tag{3.46}$$

This identifies the $K \times K$ matrices $\Sigma^{\mu\nu} = -\Sigma^{\nu\mu}$ as *spin matrices*, since they are added to the generalized orbital angular momentum.

The spin matrix for a scalar field is obviously zero. For the vector field, we can find it from its transformation law under an infinitesimal Lorentz transformation

$$A'_\alpha(x') = A_\alpha(x) + \omega_{\alpha\beta}A^\beta(x) \tag{3.47}$$

Putting $\omega_{\alpha\beta} = \frac{1}{2}\omega_{\mu\nu}\Sigma^{\mu\nu}_{\alpha\beta}$, we obtain for the vector field

$$\Sigma^{\mu\nu}_{\alpha\beta} = g^\mu_\alpha g_\beta{}^\nu - g^\mu_\beta g_\alpha{}^\nu \tag{3.48}$$

As we shall show in Section 5.5, this gives spin 1. The case of the spinor field will be discussed in Chapter 6, and is included in the following summary for reference:

$$\begin{aligned} &\text{Scalar field:} && \Sigma^{\mu\nu} = 0 \\ &\text{Vector field:} && \Sigma^{\mu\nu}_{\alpha\beta} = g^\mu_\alpha g^n_\beta - g^\mu_\beta g^\nu_\alpha \\ &\text{Spinor field:} && \Sigma^{\mu\nu}_{rs} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)_{rs} \end{aligned} \tag{3.49}$$

where γ^μ are the 4×4 Dirac matrices defined in Chapter 6.

3.5 RELATIVISTIC QUANTUM FIELDS

Since quantum fields are operators that act on a Hilbert space, we can represent Lorentz transformations by transformations on the Hilbert space. Recall that a Lorentz transformation changes the functional form of a classical field:

$$\phi_a(x) \rightarrow \phi'_a(x) \tag{3.50}$$

In the quantized version, this means that the operator ϕ_a attached to point x is replaced by ϕ'_a. Since ϕ_a and ϕ'_a act on the same Hilbert space, the transformation is a mapping of the Hilbert space into itself. Since ϕ_a and ϕ'_a are physically equivalent, the transformation must be unitary. Thus, there should exist a unitary operator $U(\Lambda)$ on the Hilbert space, corresponding to the Lorentz transformation Λ, such that

$$\phi'_a(x) = U(\Lambda)\phi_a(x)U^{-1}(\Lambda) \tag{3.51}$$

The fact that the transformation is unitary means

$$U^{\dagger}(\Lambda) = U^{-1}(\Lambda) \tag{3.52}$$

From the definition of the primed fields $\phi_a'(x) = S_{ab}\phi_b(x)$, we obtain the condition

$$U\phi_a(x)U^{-1} = S_{ab}\phi_b(\Lambda^{-1}x) \tag{3.53}$$

The set of operators $U(\Lambda)$ forms an infinite-dimensional unitary representation of the Lorentz group. In contrast to this, the finite-dimensional representations of the Lorentz group are nonunitary. As examples, we have

$$\begin{aligned} &\text{Scalar field:} && U\phi(x)U^{-1} = \phi(\Lambda^{-1}x) \\ &\text{Vector field:} && UA^{\mu}(x)U^{-1} = \Lambda^{\mu}_{\nu}A^{\nu}(\Lambda^{-1}x) \\ &\text{Spinor field:} && U\psi_r(x)U^{-1} = S_{rs}\psi_s(\Lambda^{-1}x) \end{aligned} \tag{3.54}$$

We can immediately extend this consideration to Poincaré transformations:

$$U\phi_a'(x)U^{-1} = S_{ab}\phi_b(\Lambda^{-1}x - a) \tag{3.55}$$

For infinitesimal Poincaré transformations, U must be in the neighborhood of the identity operator, and linear in the parameters of the Poincaré group:

$$U = 1 - ia^{\mu}P^{\mu} + \frac{i}{2}\omega_{\mu\nu}M^{\mu\nu} \tag{3.56}$$

This defines the Hermitian operators P^{μ} and $M^{\mu\nu}$, which represent the generators of the Poincaré group on the Hilbert space. In contrast, the generators denoted by the same symbols in (3.28) are finite matrices, generally non-Hermitian.

Substituting (3.56) into (3.55), we obtain

$$\begin{aligned} &\left(1 - ia\cdot P + \frac{i}{2}\omega\cdot M\right)\phi_a(x)\left(1 + ia\cdot P - \frac{i}{2}\omega\cdot M\right) \\ &\quad = \left(\delta_{ab} + \frac{1}{2}\omega\cdot\Sigma_{ab}\right)\phi_b(x - \omega x - a) \end{aligned} \tag{3.57}$$

which is written in an obvious abbreviated notation. Expanding both sides to first order in a^{μ} and $\omega^{\mu\nu}$, and equating their coefficients, we obtain

$$\begin{aligned} i[P^{\mu}, \phi_a(x)] &= \partial^{\mu}\phi_a(x) \\ i[M^{\mu\nu}, \phi_a(x)] &= (x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu})\,\phi_a(x) + \Sigma^{\mu\nu}_{ab}\phi_b(x) \end{aligned} \tag{3.58}$$

This shows that P^{μ} is the 4-momentum operator, since it generates space–time

translations, and $M^{\mu\nu}$ is a generalized angular momentum operator, since it generates space–time rotations. The spin matrix Σ_{ab} induces a mixing of the field components undergoing space–time rotation.

The generators P^μ and $M^{\mu\nu}$ can be constructed explicitly from the field operators ϕ_a. Rather than doing this on a case-by-case basis, we shall do it via a unified approach in the next chapter.

3.6 ONE-PARTICLE STATES

A one-particle state is an eigenstate of P^μ, with energy eigenvalue $E > 0$, and momentum eigenvalue $\mathbf{p}$, such that the invariant mass squared

$$P^2 \equiv E^2 - \mathbf{p}^2 = m^2 \tag{3.59}$$

is a fixed number. Such a state corresponds to a particle of mass m. The P^2 of any state generally lies in a continuum, but those of one-particle states form a discrete set. If there are no massless particles, the invariant-mass spectrum of a field theory consists of the vacuum value 0 as a lower bound, a discrete set of particle masses, and a continuum separated from the vacuum value by a finite gap. The continuum corresponds to states containing two or more particles, whose masses occur within the gap. There can be particles whose mass occurs in the continuum, but only if these particles are stable against decay, due to selection rules. The gap vanishes when there are massless particles, such as the photon. In this case, there is a discrete mass in the continuum corresponding to the electron, which cannot decay into photons because of conservation laws.

The one-particle states of a free field can be generated by applying creation operators $a_k^\dagger$ to the vacuum state. If we do this for a nonfree field, we will not get one-particle states, because we will not get eigenstates of P^μ. Instead, we will have a mixture of states involving interacting particles. Nevertheless, we can discuss properties of a one-particle state through general considerations, without constructing it explicitly.

We confine our attention to massive particles, with $m > 0$. There exists a Lorentz frame in which $\mathbf{p} = 0$, called the rest frame. The spin operator $\mathbf{S}$ of the one-particle state is defined as the total angular momentum in the *rest frame*. The eigenvalue of $\mathbf{S}^2$ has the form $S(S + 1)$, where S is called the spin of the particle. The projection s of $\mathbf{S}$ along the momentum of the particle is called the *helicity*, which for $m > 0$ has $2S + 1$ possible values $S, S - 1, \ldots, -S$. We can label a one-particle state by momentum $\mathbf{p}$ and helicity s:

$$|1\text{-particle state}\rangle = |\mathbf{p}, s\rangle \tag{3.60}$$

The parameters m and S are suppressed, because they are constants.

For $m > 0$ a one-particle state in the rest frame is denoted by $|s\rangle$, and we can obtain $|\mathbf{p}, s\rangle$ from $|s\rangle$ through a Lorentz boost $L(\mathbf{p})$:

$$|\mathbf{p}, s\rangle = U(L(\mathbf{p}))|s\rangle \tag{3.61}$$

Applying a Lorentz transformation Λ to both sides, we obtain

$$U(\Lambda)|\mathbf{p}, s\rangle = U(\Lambda)U(L(\mathbf{p}))|s\rangle \tag{3.62}$$

Now insert in front of the right side the identity operator in the form

$$1 \equiv U(L(\mathbf{p}))U^{-1}(L(\mathbf{p})) \tag{3.63}$$

and regroup the factors in the following manner:

$$U(\Lambda)|\mathbf{p}, s\rangle = U(L(\mathbf{p})[U^{-1}(L(\mathbf{p}))U(\Lambda)U(L(\mathbf{p}))]|s\rangle \tag{3.64}$$

By the group property, the operator within the square brackets can be rewritten as

$$U^{-1}(L(\mathbf{p}))U(\Lambda)U(L(\mathbf{p})) = U(L^{-1}(\Lambda\,\mathbf{p})\Lambda L(\mathbf{p})) \tag{3.65}$$

which represents a pure rotation called the *Wigner rotation*:

$$R(\Lambda, \mathbf{p}) \equiv L^{-1}(\Lambda\mathbf{p})\Lambda L(\mathbf{p}) \tag{3.66}$$

It boosts a particle from rest to momentum $\mathbf{p}$, makes a Lorentz transformation Λ, and then brings the particle back to rest. The operation has no effect on the state vector except possibly multiplying it by a phase factor, which represents a rotation. Thus, the general effect of a Lorentz transformation on a one-particle state is to boost the momentum, and rotate the spin by a Wigner rotation:

$$U(\Lambda)|\mathbf{p}, s\rangle = U(L(\Lambda\mathbf{p}))U(R)|s\rangle \tag{3.67}$$

For a more explicit representation of R, we insert a complete set of helicity states to obtain

$$\begin{aligned} U(\Lambda)|\mathbf{p}, s\rangle &= \sum_{s'} U(L(\Lambda\mathbf{p}))|s'\rangle\,\langle s'|U(R)|s\rangle \\ &= \sum_{s'} \mathcal{D}^{J}_{ss'}(R)|\Lambda\mathbf{p}, s'\rangle \end{aligned} \tag{3.68}$$

where $\mathcal{D}^{J}_{ss'}(R)$ are the rotation coefficients. An example of the Wigner rotation is the Thomas precession discussed in Section 6.8.

Massless particles are special, in that there is no rest frame. A massless particle of spin S can have the values $\pm S$ only. We shall explicitly demonstrate this for photons in Section 5.5. A general proof may be found in books on representations of the Lorentz group (see, e.g., Tung [1]).

PROBLEMS

3.1 Verify the Poincaré algebra (3.31).

3.2 Verify the spin matrix (3.48) for the vector field.

3.3 Verify that the $SL(2C)$ matrices $L(\Lambda)$ given in (3.23) correctly represent Lorentz transformations. It is necessary to verify them only for infinitesimal transformations.

3.4 Show the following identity, which is useful when working with $SL(2C)$ matrices:

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot A\times\mathbf{B}$$

where the components of $\boldsymbol{\sigma}$ are Pauli matrices and the components of $\mathbf{A}$ and $\mathbf{B}$ are numbers.

3.5 Consider two infinitesimal successive infinitesimal Lorentz boosts with angles $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$. Show that the result is equivalent to a boost $\boldsymbol{\phi}_1 + \boldsymbol{\phi}_2$ plus a rotation $\frac{1}{2}\boldsymbol{\phi}_1 \times \boldsymbol{\phi}_2$. Here, $\boldsymbol{\phi} = \hat{\mathbf{v}}\tanh^{-1}v$, where $\mathbf{v}$ is the velocity of the boost. Lorentz boosts.

3.6 **(a)** Under the action of a Lorentz boost with velocity $\mathbf{v}$, a 4-momentum p^μ is transformed to p'^μ. show

$$\mathbf{p}' = \mathbf{p} + \mathbf{v}\left(\frac{\gamma-1}{v^2}\mathbf{v}\cdot\mathbf{p} - \gamma p_0\right)$$

$$p_0' = \gamma(p_0 - \mathbf{v}\cdot\mathbf{p})$$

where $\gamma = (1-v^2)^{-1/2}$.

(b) Writing $p'^\mu = \Lambda^\mu{}_\nu p^\nu$, obtain the transformation matrix

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & \gamma v_k \\ -\gamma v^k & \delta_k^j - \dfrac{(\gamma-1)v^j v_k}{v^2} \end{pmatrix}$$

(c) Let $L(\mathbf{p})$ be the transformation matrix corresponding to a Lorentz boost that transforms the rest frame of a particle of mass m into a frame in which the particle has momentum $\mathbf{p}$ and energy E. Show

$$[L(\mathbf{p})]^\mu_\nu = \begin{pmatrix} \dfrac{E}{m} & -\dfrac{p_k}{m} \\ \dfrac{p^k}{m} & \delta_k^j - \dfrac{p^j p_k}{m(E+m)} \end{pmatrix}$$

3.7 **(a)** Consider a particle of mass m and helicity s, moving with momentum $\mathbf{p}$ along the z axis. Make a Lorentz boost of velocity v along the x axis, so that $p \to p'$. Find the rotation matrix $R^i{}_j$ for the Wigner rotation.

(b) Show that for an ultrarelativistic particle

$$R^i_j\frac{p^j}{E} \approx R^i_j\frac{p'^j}{E'}$$

That is, the Wigner rotation is the same as that taking the initial velocity $\mathbf{p}/E$ to the final velocity $\mathbf{p}'/E'$. This shows that the helicity of a massless particle such as the photon is Lorentz-invariant.

REFERENCE

1. W. K. Tung, *Group Theory in Physics,* World Scientific, Singapore, 1985, Section 10.4.4.

CHAPTER FOUR

Canonical Formalism

4.1 PRINCIPLE OF STATIONARY ACTION

The equations of motion for a classical field can be derived from a Lagrangian through the principle of stationary action. This approach gives a unified treatment of topics discussed previously through special examples. It also makes clear that symmetries of the system give rise to conservation laws. Consider a set of classical fields collectively denoted by $\phi(x)$:

$$\phi(x) = \{\phi_1(x), \ldots, \phi_K(x)\} \tag{4.1}$$

We denote their space–time derivatives by

$$\phi_\mu(x) \equiv \partial_\mu \phi(x) \tag{4.2}$$

The Lagrangian density is assumed to depend on the fields and their first derivatives:

$$\mathcal{L}(x) \equiv \mathcal{L}(\phi(x), \phi_\mu(x)) \tag{4.3}$$

This will ensure that the equations of motion are second-order differential equations in the time, as in classical mechanics. We assume that, unless external fields are explicitly introduced, space–time is homogeneous. Thus, $\mathcal{L}(x)$ depends on x not explicitly, but only implicitly through $\phi(x)$ and $\phi_\mu(x)$. We consider only local field theories, for which the Lagrangian density at x depends only on properties of the field at x. Nonlocal terms of the form

$$\int d^4y \phi_a(y) K_{ab}(x-y) \phi_b(y) \tag{4.4}$$

are ruled out, unless $K_{ab}(x-y) \propto \delta(x-y)$.

The classical action of the system is

$$S[\phi] = \int_{\Omega} d^4x \mathcal{L}(x) \tag{4.5}$$

where Ω is the space–time volume, which eventually goes to infinity. We impose definite boundary conditions on the surface of Ω, say, $\phi = 0$. The principle of stationary action, which is a generalization of that in classical mechanics, states the following:

> Suppose that $\phi(x)$ is a solution to the equations of motion. If we vary its functional form by adding an arbitrary infinitesimal function $\delta\phi(x)$ that preserves the boundary condition:
>
> $$\phi(x) \rightarrow \phi(x) + \delta\phi(x) \qquad \delta\phi(x) = 0 \text{ on boundary of } \Omega$$
>
> then the variation of the action will be of second-order smallness:
>
> $$\delta S \equiv S[\phi + \delta\phi] - S[\phi] = 0$$

This means that $S[\phi]$ is at an extremum when $\phi(x)$ is a solution to the equation of motion.

To find the equation of motion according to this principle, let us calculate the variation of the Lagrangian density:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\phi_\mu}\delta\phi_\mu \tag{4.6}$$

Using the fact that

$$\delta\phi_\mu = \delta(\partial_\mu\phi) = \partial_\mu(\delta\phi) \tag{4.7}$$

we get

$$\delta\mathcal{L} = \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\phi_\mu}\right)\right]\delta\phi + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\phi_\mu}\delta\phi\right) \tag{4.8}$$

The last term is a total 4-divergence. It vanishes when integrated over the space–time volume, since it then becomes a surface integral by Gauss' theorem, and $\delta\phi = 0$ on the surface. Thus

$$\delta S = \int_{\Omega} d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\phi_\mu}\right)\right]\delta\phi = 0 \tag{4.9}$$

Since $\delta\phi(x)$ is arbitrary, its coefficient must vanish. We thus obtain the equation of motion

$$\partial_\mu \pi^\mu - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{4.10}$$

where

$$\pi^\mu \equiv \frac{\partial \mathcal{L}}{\partial \phi_\mu} \tag{4.11}$$

The canonical conjugate to $\phi(x)$ is defined by analogy with classical mechanics:

$$\pi(x) \equiv \pi^0(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \tag{4.12}$$

where a dot denotes partial time derivative. The Hamiltonian density is defined by

$$\mathcal{H}(x) \equiv \mathcal{H}(\pi(x), \phi(x), \nabla \phi(x)) \equiv \pi\dot{\phi} - \mathcal{L}(x) \tag{4.13}$$

where $\dot{\phi}(x)$ should be re-expressed in terms of $\pi(x)$ according to (4.12). The Hamiltonian is given by

$$H = \int d^3x \mathcal{H}(x) \tag{4.14}$$

To quantize the system, we replace the field and its canonical conjugate by operators, which are defined by the equal-time commutators

$$\begin{aligned} i[\pi(\mathbf{x}, t), \phi(\mathbf{y}, t)]_\pm &= \delta^3(\mathbf{x} - \mathbf{y}) \\ i[\pi(\mathbf{x}, t), \pi(\mathbf{y}, t)]_\pm &= [\phi(\mathbf{x}, t), \phi(\mathbf{y}, t)]_\pm = 0 \end{aligned} \tag{4.15}$$

where for bosons we use the commutator

$$[A, B]_- \equiv [A, B] \equiv AB - BA \tag{4.16}$$

and for fermions we use the anticommutator

$$[A, B]_+ \equiv \{A, B\} \equiv AB + BA \tag{4.17}$$

If $\pi(\mathbf{x}, t) \equiv 0$, as is the case for the electromagnetic field, then the field conjugate to $\pi(\mathbf{x}, t)$ is not an independent dynamical variable, and should not be independently quantized.

4.2 NOETHER'S THEOREM

A transformation $\phi(x) \rightarrow \phi(x) + \delta\phi(x)$ is called a *symmetry transformation* of the system if it changes the Lagrangian density only by the addition of a 4-divergence. As we have seen, this does not change the equations of motion. More specifically, the change must be of the form $\delta\mathcal{L}(x) = \partial_\mu W^\mu(x)$ for arbitrary $\phi(x)$, regardless of whether it obeys the equation of motion. If the symmetry transformation is continuous, then there is an associated conserved current density. The formal statement is as follows.

■ **NOETHER'S THEOREM [1]** *If, under a continuous infinitesimal transformation*

$$\phi(x) \rightarrow \phi(x) + \delta\phi(x)$$

the change in the Lagrangian density is found to be of the form

$$\delta\mathcal{L}(x) = \partial^\mu W_\mu(x)$$

without using the equations of motion, then there exists a current density

$$j^\mu(x) = \pi^\mu(x)\delta\phi(x) - W^\mu(x) \tag{4.18}$$

which, for fields obeying the equations of motion, satisfies

$$\partial_\mu j^\mu(x) = 0 \tag{4.19}$$

Proof. We calculate the $\delta\mathcal{L}(x)$ when the field changes by $\delta\phi$, using the equation of motion, but without assuming that $\delta\phi$ comes from a symmetry transformation:

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\phi_\mu}\delta\phi_\mu = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \pi^\mu\partial_\mu(\delta\phi) \\ &= \partial_\mu(\pi^\mu\delta\phi) + \left(\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\pi^\mu\right)\delta\phi \\ &= \partial_\mu(\pi^\mu\delta\phi)\end{aligned}$$

where the equation of motion was used in the last step. Specializing the preceding to symmetry transformations, we equate it with $\partial_\mu W^\mu$ to obtain

$$\partial_\mu(\pi^\mu\delta\phi - W^\mu) = 0$$

■

The conserved current j^μ is called a Noether current, and is determined only up to an arbitrary normalization factor.

Noether's theorem was proved for classical fields, and one usually extends it to quantum theory by replacing the fields in j^μ by the corresponding quantized fields. This does not always give a conserved current in the quantum theory. When the quantum current so obtained fails to be conserved, the nonzero divergence $\partial_\mu j^\mu$ is called a "current anomaly." Some examples of this are discussed in Section 19.8.

4.3 TRANSLATIONAL INVARIANCE

An important symmetry for any system is Poincaré invariance, which is called a "space–time symmetry," because it is associated with the transformation of the field under a change in the coordinate system. We discuss this symmetry by breaking up the Poincaré group into the translation and Lorentz subgroups.

Invariance under the translation group should give rise to four independent Noether currents corresponding to the four possible space–time translations. Consider an infinitesimal translation

$$x^\mu \longrightarrow x^\mu + a^\mu \tag{4.20}$$

under which each component of $\phi(x)$ independently transforms according to $\phi'(x + a) = \phi(x)$. The functional form of the field changes according to

$$\phi'(x) = \phi(x - a) = \phi(x) - a_\mu \partial^\mu \phi(x) \tag{4.21}$$

We shall choose a_μ to have only one nonvanishing component, say, a_α, and take $\alpha = 0,1,2,3$ in turn. Thus, the change in functional form of the field is

$$\delta\phi(x) = \partial^\alpha \phi(x) \tag{4.22}$$

where we have dropped the proportionality constant $-a_\alpha$, because it enters all subsequent formulas only through the overall normalization of the Noether current. Differentiating the above with respect to x^μ gives

$$\delta\phi_\mu(x) = \partial^\alpha \phi_\mu(x) \tag{4.23}$$

The statement of translational invariance is that $\mathcal{L}(x)$ does not depend on x explicitly. Hence

$$\begin{aligned} \delta\mathcal{L}(x) &= \frac{\partial \mathcal{L}}{\partial \phi} \partial^\alpha \phi + \frac{\partial \mathcal{L}}{\partial \phi_\mu} \partial^\alpha \phi_\mu = \partial^\alpha \mathcal{L}(x) \\ &= \partial_\mu (g^{\mu\alpha} \mathcal{L}(x)) \end{aligned} \tag{4.24}$$

from which we read off

$$W^{\mu\alpha}(x) = g^{\mu\alpha}\mathcal{L}(x) \tag{4.25}$$

The four corresponding Noether currents are denoted by $T_c{}^{\mu\alpha}$, where α labels the direction of translation:

$$T_c^{\mu\alpha}(x) = \pi^\mu \partial^\alpha \phi(x) - g^{\mu\alpha}\mathcal{L}(x) \tag{4.26}$$

They satisfy the conservation law

$$\partial_\mu T_c^{\mu\alpha}(x) = 0 \tag{4.27}$$

This is called the *canonical energy–momentum tensor*. It is generally not a symmetric tensor. The subscript "c," which stands for "canonical," distinguishes it from a symmetrized version to be discussed later.

The conservation law can be rewritten in the form

$$\frac{\partial}{\partial t} T_c^{0\alpha} + \frac{\partial}{\partial x^k} T_c^{k\alpha} = 0 \tag{4.28}$$

Integrating both sides over all space, and assuming that surface contributions vanish, we obtain

$$\frac{d}{dt} P^\alpha = 0 \tag{4.29}$$

where

$$P^\alpha = \int d^3x T_c^{0\alpha} \tag{4.30}$$

is the total 4-momentum. Thus, T_c^{00} is the energy density and T_c^{0k} the kth component of the momentum density. Their conservation laws are given respectively by the time and spatial components of (4.28):

$$\frac{\partial}{\partial t} T_c^{00} + \frac{\partial}{\partial x^k} T_c^{k0} = 0$$

$$\frac{\partial}{\partial t} T_c^{0j} + \frac{\partial}{\partial x^k} T_c^{kj} = 0 \tag{4.31}$$

It is clear that T_c^{k0} is the kth component of the energy current and T_c^{kj} is the kth component of the current of the "jth component of momentum." The latter is called the *stress-energy tensor*. These components are displayed in the following matrix:

$$T_c^{\mu\alpha} = \begin{pmatrix} T_c^{00} & T_c^{01} & T_c^{02} & T_c^{03} \\ T_c^{10} & & & \\ T_c^{20} & & \boxed{T_c^{ij}} & \\ T_c^{30} & & & \end{pmatrix} \tag{4.32}$$

The identification of T^{00} as the energy density is consistent with the definition of the Hamiltonian density in (4.13).

The explicit expressions for the total energy and momentum are

$$P^0 = \int d^3x[\pi(x)^k\dot{\phi}(x) - \mathcal{L}(x)] = \int d^3x\mathcal{H}(x)$$

$$P^k = \int d^3x\pi(x)\partial^k\phi(x) \tag{4.33}$$

where a dot denotes time derivative and $\mathcal{H}(x)$ is the Hamiltonian density. We go over to the quantum theory by replacing $\pi(x)$ and $\phi(x)$ by the appropriate operators.

To show that P^k generates spatial translations, we calculate the commutator of $\phi(x)$ with the total momentum operator

$$P^k = \int d^3y\pi(y)\partial^k\phi(y) \tag{4.34}$$

This does not depend on y_0. We are therefore free to choose $y_0 = x_0 = t$, to take advantage of the simplicity of the equal-time commutators. Thus, we have

$$[P^k, \phi(\mathbf{x}, t)] = \int d^3y\left[\pi(\mathbf{y}, t)\frac{\partial\phi(y, t)}{\partial y_k}, \phi(\mathbf{x}, t)\right]$$

$$= -i\int d^3y\ \delta^3(\mathbf{x} - \mathbf{y})\frac{\partial\phi(\mathbf{y}, t)}{\partial y_k} = -i\frac{\partial\phi(\mathbf{x}, t)}{\partial x_k} \tag{4.35}$$

To show that P^0 generates time translations, we calculate

$$[P^0, \phi(\mathbf{x}, t)] = \int d^3y\{[\pi(\mathbf{y}, t)\dot{\phi}(\mathbf{y}, t), \phi(\mathbf{x}, t)] - [\mathcal{L}(\mathbf{y}, t), \phi(\mathbf{x}, t)]\}$$

$$= -i\dot{\phi}(\mathbf{x}, t) + \int d^3y\{\pi(\mathbf{y}, t)[\dot{\phi}(\mathbf{y}, t), \phi(\mathbf{x}, t)] - [\mathcal{L}(\mathbf{y}, t), \phi(\mathbf{x}, t)]\} \tag{4.36}$$

The integral in the second term identically vanishes.

Proof. Use the representation $\phi = i\delta/\delta\pi$ to write

$$[O(y), \phi(x)] = -i\frac{\delta O(y)}{\delta\pi(x)} = -i\delta^3(\mathbf{x} - \mathbf{y})\frac{\partial O(y)}{\partial\pi(x)}$$

Using $\pi = \partial\mathcal{L}/\partial\dot{\phi}$, we can calculate the integral as follows:

$$-i\int d^3y\left[\frac{\partial\mathcal{L}(y)}{\partial\dot{\phi}(y)}\frac{\delta\dot{\phi}(y)}{\delta\pi(x)} - \frac{\delta\mathcal{L}(y)}{\delta\pi(x)}\right] = -i\left[\frac{\partial\mathcal{L}(x)}{\partial\dot{\phi}(x)}\frac{\partial\dot{\phi}(x)}{\partial\pi(x)} - \frac{\partial\mathcal{L}(x)}{\partial\pi(x)}\right] \equiv 0 \qquad \blacksquare$$

Combining the preceding results, we have

$$i[P^\alpha, \phi(x)] = \partial^\alpha \phi(x) \tag{4.37}$$

This shows that the total 4-momentum generates space–time translations. If F is a function of ϕ and its derivatives, then

$$i[P^\alpha, F] = \partial^\alpha F \tag{4.38}$$

Choosing, in particular, $F = P^\mu$, we have $i[P^\alpha, P^\mu] = \partial^\alpha P^\mu$. The right side vanishes because P^μ is independent of space by construction, and independent of time by 4-momentum conservation. Thus

$$[P^\alpha, P^\mu] = 0 \tag{4.39}$$

which is part of the Poincaré algebra discussed in Chapter 3. It is realized here in terms of the field operators.

4.4 LORENTZ INVARIANCE

Consider an infinitesimal Lorentz transformation in the direction labeled by $\{\alpha,\beta\}$. For example, a rotation about the x^3 axis corresponds to $\{1,2\}$, or a boost along the x^1 axis corresponds to $\{0,1\}$. Under the transformation, the functional form of the field changes by $\delta\phi(x)$, as given by (3.46). We have, up to a multiplicative constant,

$$\delta\phi(x) = (x^\alpha \partial^\beta - x^\beta \partial^\alpha + \Sigma^{\alpha\beta})\phi(x) \tag{4.40}$$

where $\Sigma^{\alpha\beta}$ is the spin matrix, antisymmetric in $\{\alpha, \beta\}$. Differentiating $\delta\phi$ with respect to x^μ, we find

$$\delta\phi_\mu(x) = [x^\alpha \partial^\beta - x^\beta \partial^\alpha + \Sigma^{\alpha\beta}]\phi_\mu(x) + (g^\alpha_\mu \phi^\beta - g^\beta_\mu \phi^\alpha) \tag{4.41}$$

Using the fact that $\mathcal{L}(x)$ has no explicit x dependence, we have

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial \mathcal{L}}{\partial \phi}\partial^\alpha \phi + \frac{\partial \mathcal{L}}{\partial \phi_\mu}\partial^\alpha \phi_\mu \\ &= (x^\alpha \partial^\beta - x^\beta \partial^\alpha)\mathcal{L} + \frac{\partial \mathcal{L}}{\partial \phi}\Sigma^{\alpha\beta}\phi + \pi^\mu \Sigma^{\alpha\beta}\phi_\mu + \pi^\alpha \phi^\beta - \pi^\beta \phi^\alpha\end{aligned} \tag{4.42}$$

The statement of Lorentz invariance is that the preceding is a 4-divergence. The first term is of the desired form, for it can be rewritten $\partial^\beta(x^\alpha \mathcal{L}) - \partial^\alpha(x^\beta \mathcal{L}) = \partial_\mu(g^{\mu\beta}x^\alpha \mathcal{L} - g^{\mu\alpha}x^\beta \mathcal{L})$. Therefore the rest must vanish:

$$\frac{\partial \mathcal{L}}{\partial \phi}\Sigma^{\alpha\beta}\phi + \pi^\mu \Sigma^{\alpha\beta}\phi_\mu + \pi^\alpha \phi^\beta - \pi^\beta \phi^\alpha = 0 \tag{4.43}$$

With this, we have

$$\delta\mathcal{L} = \partial_\mu(g^{\mu\beta}x^\alpha\mathcal{L} - g^{\mu\alpha}x^\beta\mathcal{L}) \tag{4.44}$$

which gives

$$W^\mu(x) = (g^{\mu\beta}x^\alpha - g^{\mu\alpha}x^\beta)\mathcal{L}(x) \tag{4.45}$$

This leads to six independent Noether currents labeled by $\{\alpha, \beta\}$:

$$M_c^{\mu\alpha\beta} = \pi^\mu(x^\alpha\partial^\beta - x^\beta\partial^\alpha + \Sigma^{\alpha\beta})\phi(x) - (g^{\mu\beta}x^\alpha - g^{\mu\alpha}x^\beta)\mathcal{L}(x) \tag{4.46}$$

which can be written in the form

$$M_c^{\mu\alpha\beta}(x) = x^\alpha T_c^{\mu\beta}(x) - x^\beta T_c^{\mu\alpha}(x) + \pi^\mu\Sigma^{\alpha\beta}\phi(x) \tag{4.47}$$

where $T_c^{\mu\nu}$ is the canonical energy–momentum tensor. It satisfies the conservation law

$$\partial_\mu M_c^{\mu\alpha\beta}(x) = 0 \tag{4.48}$$

and is called the *canonical angular momentum tensor.*

The conservation law gives rise to six constants of the motion, the Lorentz boost

$$K^j = \int d^3x\, M_c^{00j}(x) = \int d^3x[x^0 T_c^{0j}(x) - x^j T_c^{00}(x)] \tag{4.49}$$

and the angular momentum

$$J^k = \tfrac{1}{2}\epsilon^{ijk}\int d^3x\, M_c^{0ij}(x) = \tfrac{1}{2}\epsilon^{ijk}\int d^3x[x^i T_c^{0j}(x) - x^j T_c^{0i}(x) + \pi(x)\Sigma^{ij}\phi(x)] \tag{4.50}$$

where the first two terms represent orbital angular momentum and the last term is the intrinsic spin.

It is straightforward to verify that K^j and J^k generate Lorentz transformations by showing

$$i\left[\int d^3y\, M_c^{0\alpha\beta}(\mathbf{y}, t), \phi(\mathbf{x}, t)\right] = \delta\phi(\mathbf{x}, t) \tag{4.51}$$

where $\delta\phi$ is as given by (4.40). It can also be verified that the commutation relations among the operators P^μ and $M_c^{0\alpha\beta}$ realize the Poincaré algebra (3.31).

4.5 SYMMETRIZED ENERGY–MOMENTUM TENSOR

The canonical energy–momentum tensor $T_c^{\mu\alpha}$ is not unique, because the Lagrangian density is defined only up to a 4-divergence. We can replace it by any tensor of the form

$$T^{\mu\alpha} = T_c^{\mu\alpha} + \tfrac{1}{2}\partial_\lambda X_{\lambda\mu\alpha} \tag{4.52}$$

where $X_{\lambda\mu\alpha}$ is antisymmetric in $\lambda\mu$:

$$X_{\lambda\mu\alpha} = -X^{\mu\lambda\alpha} \tag{4.53}$$

The antisymmetry is a sufficient condition that the conservation law be unchanged:

$$\partial_\mu T^{\mu\alpha} = \partial_\mu T_c^{\mu\alpha} + \tfrac{1}{2}\partial_\mu\partial_\lambda X_{\lambda\mu\alpha} = 0 \tag{4.54}$$

A possible change in the total 4-momentum is

$$\int d^3x T^{0\alpha} - \int d^3x T_c^{0\alpha} = \tfrac{1}{2}\int d^3x \partial_\lambda X^{\lambda 0\alpha} \tag{4.55}$$

which vanishes for the following reasons: (1) the term with $\lambda = 0$ vanishes because $X^{00\alpha} = 0$ by antisymmetry and (2) the terms from $\lambda = k$ vanish because they give a surface integral. From a physical point of view, therefore, $T^{\mu\alpha}$ is equivalent to $T_c^{\mu\alpha}$, because they give the same total 4-momentum.

The fact that $T_c^{\mu\alpha}$ is not guaranteed to be a symmetric tensor poses a problem, if it is to be used as a source of the gravitational field. We can, however, replace it with a equivalent symmetric tensor $T^{\mu\alpha}$. The condition for symmetry is

$$T^{\mu\alpha} - T^{\alpha\mu} = [T_c^{\mu\alpha} - T_c^{\alpha\mu}] + \tfrac{1}{2}\partial_\lambda(X^{\lambda\mu\alpha} - X^{\lambda\alpha\mu}) = 0 \tag{4.56}$$

The term in brackets can be rewritten using the condition (4.43) for Lorentz invariance:

$$\begin{aligned} T_c^{\mu\alpha} - T_c^{\alpha\mu} &= \pi^\mu\phi^\alpha - \pi^\alpha\phi^\mu = -\frac{\partial\mathcal{L}}{\partial\phi}\Sigma^{\mu\alpha}\phi - \pi^\lambda\Sigma^{\mu\alpha}\partial_\lambda\phi \\ &= -\partial(\pi^\lambda\Sigma^{\mu\alpha}\phi) - \left(\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\lambda\pi^\lambda\right)\Sigma^{\mu\alpha}\phi \end{aligned} \tag{4.57}$$

where the last term vanishes by the equation of motion. Substituting this result into (4.56), we obtain the condition

$$X^{\lambda\mu\alpha} - X^{\lambda\alpha\mu} = 2\pi^\lambda\Sigma^{\mu\alpha}\phi \tag{4.58}$$

a solution to which is

$$X^{\lambda\mu\alpha} = (\pi^{\lambda}\Sigma^{\mu\alpha} - \pi^{\alpha}\Sigma^{\lambda\mu} - \pi^{\mu}\Sigma^{\lambda\alpha})\phi \tag{4.59}$$

This term is needed only when the spin is nonzero.

Corresponding to the symmetrized energy–momentum tensor, we can define a new angular momentum tensor:

$$M^{\mu\alpha\beta} \equiv x^{\alpha}T^{\mu\beta} - x^{\beta}T^{\mu\alpha} \tag{4.60}$$

This is related to the canonical angular momentum tensor through

$$M^{\mu\alpha\beta} = M_{c}^{\mu\alpha\beta} + \tfrac{1}{2}\partial_{\lambda}(x^{\alpha}X^{\lambda\mu\beta} - x^{\beta}X^{\lambda\mu\alpha}) \tag{4.61}$$

It is easy to show that $M^{\mu\alpha\beta}$ is conserved, and that it preserves the definition of the boost and the angular momentum:

$$\begin{aligned} \partial_{\mu}M^{\mu\alpha\beta} &= 0 \\ \int d^3x M^{0\alpha\beta} &= \int d^3x M_{c}^{0\alpha\beta} \end{aligned} \tag{4.62}$$

From a physical point of view, therefore, $M^{\mu\alpha\beta}$ and $M_{c}^{\mu\alpha\beta}$ are equivalent.

4.6 GAUGE INVARIANCE

In contrast to space–time symmetries, there are internal symmetries, which are x-independent transformations of the field that leaves the Lagrangian density invariant.

A simple example is a change of phase in a complex scalar field:

$$\begin{aligned} \psi(x) &\to e^{-i\alpha}\psi(x) \\ \psi^{\dagger}(x) &\to e^{i\alpha}\psi^{\dagger}(x) \end{aligned} \tag{4.63}$$

This is called a *global gauge transformation,* where the label "global" refers to the fact that α is independent of x. Invariance with respect to it means that the Lagrangian density is independent of the phase. This is true for the free field, in which the fields appear in the combination $\psi^{\dagger}(x)\psi(x)$ or $\partial^{\mu}\psi^{\dagger}(x)\partial_{\mu}\psi(x)$. The infinitesimal form of the transformation is

$$\begin{aligned} \delta\psi(x) &= \psi(x) \\ \delta\psi^{\dagger}(x) &= -\psi(x) \end{aligned} \tag{4.64}$$

where we left out a proportionality constant $i\alpha$.

In terms of the real and imaginary parts defined by

$$\psi(x) = \frac{1}{\sqrt{2}}[\phi_1(x) + i\phi_2(x)]$$

$$\psi^\dagger(x) = \frac{1}{\sqrt{2}}[\phi_1(x) - i\phi_2(x)] \tag{4.65}$$

the transformation is a rotation in internal space:

$$\phi_1'(x) = \phi_1(x)\cos\alpha + \phi_2(x)\sin\alpha$$
$$\phi_2'(x) = -\phi_1(x)\sin\alpha + \phi_2(x)\cos\alpha \tag{4.66}$$

The Noether current is just the conserved current mentioned in Chapter 2:

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu\psi)}\psi$$
$$= \psi\partial^\mu\psi^* - \psi^*\partial^\mu\psi$$
$$= \tfrac{1}{2}(\phi_2\partial^\mu\phi_1 - \phi_1\partial^\mu\phi_2) \tag{4.67}$$

More generally, internal symmetries are linear transformations for a multicomponent field $\phi_a(x)$, $(a = 1, \ldots, K)$, of the form

$$\phi_a'(x) = C_{ab}\phi_b(x) \tag{4.68}$$

where C_{ab} are elements of a $K \times K$ constant matrix. If the matrix belongs to a K-dimensional representation of some group G, we call G an *internal symmetry group.* The group can be continuous or discrete. In the previous example $G = U(1)$, the unitary group of dimension 1.

Physical examples of conserved charges are

- Electric charge = positive minus negative charge
- Baryon number = number of protons minus number of antiprotons
- Electron number = number of electrons minus number of positrons

A important case is isospin, which is discussed in Section 7.5.

PROBLEMS

4.1 Consider a field $\phi(x)$ with Lagrangian density $\mathcal{L}_0(x) + \mathcal{L}_1(x)$, where the first term has a certain symmetry, while the second term does not. That is, under a transformation $\phi(x) \to \phi(x) + \delta\phi(x)$,

$$\mathcal{L}_0(x) = \partial^\mu W_\mu(x),$$

whereas $\mathcal{L}_1(x)$ cannot be put into this form. If $\mathcal{L}_1(x)$ were absent, the system would have a conserved Noether current. Show that, in the presence of $\mathcal{L}_1(x)$, the divergence of the would-be Noether current is $\delta\mathcal{L}_1(x)$.

4.2 A condition for Lorentz invariance is (4.43). For scalar fields, for which $\Sigma^{\alpha\beta} = 0$, what restriction does this place on the Lagrangian density?

4.3 Nonrelativistic System A nonrelativistic many-particle system has a second-quantized Hamiltonian

$$H = \int d^3x\, \psi^\dagger(\mathbf{x})\left(-\frac{1}{2m}\nabla^2 + \mu\right)\psi(\mathbf{x})$$

where μ is the chemical potential. Usually one assumes the commutation relation $[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \delta^3(\mathbf{x} - y)$. We want to see whether this is consistent with the canonical formalism

(a) Find the equation of motion using $i\dot{\psi} = [\psi, H]$.

(b) Regard H as a classical Hamiltonian. Show that the corresponding Lagrangian density is

$$\mathcal{L} = i\psi^*\partial\phi\backslash\partial t + \psi^*\left(\frac{1}{2m}\right)\nabla^2 - \mu)\,\psi$$

Work out the equation of motion using the canonical formalism.

(c) Show that the usual commutation relation $[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \delta^3(\mathbf{x} - \mathbf{y})$ is canonical.

(d) Work out the Noether current associated with space–time translational invariance, and global gauge invariance.

4.4 Field Representation Since $\psi(\mathbf{r})$ and $i\psi^\dagger(\mathbf{r})$ are canonical conjugates in the nonrelativistic system, it would be awkward to introduce the field representation by diagonalizing $\psi(\mathbf{r})$. Show that in this case we can put

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{2}}\left[\varphi(\mathbf{r}) + \frac{\delta}{\delta\varphi(\mathbf{r})}\right]$$

$$\psi^\dagger(\mathbf{r}) = \frac{1}{\sqrt{2}}\left[\varphi(\mathbf{r}) - \frac{\delta}{\delta\varphi(\mathbf{r})}\right]$$

where $\varphi(\mathbf{r})$ is a c-number function.

4.5 First-Order Lagrangian The nonrelativistic Lagrangian in the last problem differs from a relativistic one, in that it is first-order instead of second-order in the time derivative. The Dirac field discussed in Chapter 7 also has a first-order Lagrangian. To fully explore the consistency of the canonical formalism, let us strip the problem down to bare essentials, and consider a classical system with two coordinates a and b, which are like $i\psi^*$ and ψ. Take the Lagrangian to be

$$L(a, b, \dot{b}) = a\dot{b} - V(a, b)$$

The canonical rule says that a has no canonical conjugate. It is the conjugate to b. Is this completely consistent with the Lagrangian and Hamiltonian equations of motion?

(a) Find the Lagrangian equations for motion for $\dot{a}$ and $\dot{b}$.

(b) Find canonical momenta and the Hamilton equations of motion. Check that they are the same as the Lagrangian ones.

(c) It is thus completely consistent to regard a and b as canonically conjugate. To quantize the system, impose $[a, b]_{\pm} = -i$.

REFERENCE

1. E. Noether, *Nachr. kgl. Wiss. Göttingen* 235 (1918); E. L. Hill, *Rev. Mod. Phys.* **23,** 253 (1951).

CHAPTER FIVE

Electromagnetic Field

5.1 MAXWELL'S EQUATIONS

The classical electromagnetic field is described by two 3-vector fields, the electric field $\mathbf{E}(\mathbf{r}, \mathbf{t})$ and the magnetic field $\mathbf{B}(\mathbf{r}, \mathbf{t})$, which obey Maxwell's equations. In rationalized units with $c = 1$, they read

$$\nabla \cdot \mathbf{E} = \rho$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \mathbf{B} = \mathbf{j} + \frac{\partial \mathbf{E}}{\partial t} \tag{5.1}$$

where $\rho(\mathbf{r}, t)$ and $\mathbf{j}(\mathbf{r}, t)$ are respectively the external charge density and external current density, which must satisfy the continuity equation

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \tag{5.2}$$

The second and third equations are solved by introducing the vector potential $\mathbf{A}(\mathbf{r}, t)$ and scalar potential $\phi(\mathbf{r}, t)$:

$$\mathbf{B} = \nabla \times \mathbf{A}$$

$$\mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \tag{5.3}$$

whereupon the remaining two equations become

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} = \mathbf{j} - \nabla\left(\nabla\cdot\mathbf{A} + \frac{\partial}{\partial t}\phi\right)$$

$$\nabla^2\phi + \frac{\partial}{\partial t}\nabla\cdot\mathbf{A} = -\rho \tag{5.4}$$

The potentials are determined only up to a local gauge transformation, which involves an arbitrary function $\chi(\mathbf{r}, t)$

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla\chi$$
$$\phi \rightarrow \phi - \frac{\partial\chi}{\partial t} \tag{5.5}$$

The Lorentz gauge corresponds to the condition

$$\nabla\cdot\mathbf{A} + \frac{\partial\phi}{\partial t} = 0 \qquad \text{(Lorentz gauge)} \tag{5.6}$$

In this gauge, both potentials satisfy the wave equation:

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} = \mathbf{j}$$

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi = \rho \tag{5.7}$$

The symmetric appearance of these equations is sometimes convenient, but it actually obscures the physics. These equations seem to indicate that there are four independent propagating modes, but actually there only two—the transverse components of **A**. This can be shown by going to the Coulomb gauge.

In Coulomb gauge (or radiation gauge), **A** is purely transverse:

$$\nabla\cdot\mathbf{A} = 0 \qquad \text{(Coulomb gauge)} \tag{5.8}$$

The equations for the potentials become

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} = \mathbf{j}_T$$

$$\nabla^2\phi = -\rho \tag{5.9}$$

where $\mathbf{j}_T$ is the transverse current density

$$\mathbf{j}_T = \mathbf{j} - \frac{\partial}{\partial t}\nabla\phi \tag{5.10}$$

which satisfies $\nabla\cdot\mathbf{j}_T = 0$. In this gauge, $\mathbf{A}$ describes transverse electromagnetic radiation, whose source is the transverse current density, while ϕ describes the instantaneous Coulomb interaction between charges. The potential between two unit charges located at $\mathbf{r}_1$ and $\mathbf{r}_2$ is given by

$$\phi(\mathbf{r}) = \frac{1}{4\pi|\mathbf{r}_1 - \mathbf{r}_2|} \tag{5.11}$$

To show that we can always impose the Coulomb gauge, suppose $\nabla\cdot\mathbf{A} = f$. To go to Coulomb gauge, we make the gauge transformation $\mathbf{A} \rightarrow \mathbf{A} + \nabla\chi$, with χ satisfying $\nabla^2\chi = -f$. The solution corresponds to the statement that χ is the electrostatic potential due to the charge distribution $\nabla\cdot\mathbf{A}$.

We are using rationalized instead of unrationalized units. The difference between these systems arises from the normalization convention for the free fields, and is tabulated as follows:

	Rationalized	*Unrationalized*
4-Current	j^μ	$4\pi j^\mu$
Coulomb's law	q^2/r	$q^2/4\pi r$
Energy density	$(\mathbf{E}^2 + \mathbf{B}^2)/2$	$(\mathbf{E}^2 + \mathbf{B}^2)/8\pi$
Field operator	$\mathbf{A}$	$4\pi\mathbf{A}$

5.2 COVARIANCE OF THE CLASSICAL THEORY

We postulate that the potentials form a 4-vector

$$A^\mu = (\phi, \mathbf{A}) \tag{5.12}$$

and this determines how Maxwell's equations transform under a Lorentz transformation. Since we always impose a gauge condition, a Lorentz transformation must be accompanied by a gauge transformation

$$A^\mu \rightarrow A^\mu - \partial^\mu\chi \tag{5.13}$$

in order to maintain the gauge condition. Under an infinitesimal Lorentz transformation, therefore, the vector potential transforms according to

$$A'^\mu = A^\mu - \omega^\mu_\nu A^\nu - \partial^\mu\chi \tag{5.14}$$

where χ is such that A'^{μ} satisfies the gauge condition.

The Lorentz gauge $\partial_\mu A^\mu = 0$ is covariant, and the equations of motion take the form (5.7), which are manifestly covariant

$$\Box^2 A^\mu = j^\mu \tag{5.15}$$

where j^μ is the conserved 4-vector current density

$$j^\mu = (\rho, \mathbf{j}) \qquad \partial_\mu j^\mu = 0 \tag{5.16}$$

In this gauge, however, the physical degrees of freedom are not manifest.

In Coulomb gauge, where physical degree freedom are made explicit, the equations of motion (5.9) are not manifestly covariant; but they actually are, because there always exists a gauge transformation to maintain the appearance of the equations in all Lorentz frames. One has to choose between manifest covariance with Lorentz gauge, or manifest transversality with Coulomb gauge, and we choose the latter.

The electric and magnetic fields are components of the antisymmetric field tensor

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{5.17}$$

which is gauge-invariant. The dual field tensor is defined as

$$\tilde{F}^{\mu\nu} = \tfrac{1}{2}\epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta} \tag{5.18}$$

In terms of the electric and magnetic fields, we have

$$\begin{aligned} F^{k0} &= E^k \qquad F^{ij} = -\epsilon^{ijk} B^k \\ B^k &= -\tfrac{1}{2}\epsilon^{kij} F^{ij} \end{aligned} \tag{5.19}$$

The components of the field tensor and its dual can be displayed as matrices:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix} \tag{5.20}$$

$$\tilde{F}^{\mu\nu} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & E^3 & -E^2 \\ B^2 & -E^3 & 0 & E^1 \\ B^3 & E^2 & -E^1 & 0 \end{pmatrix} \tag{5.21}$$

We see that $\tilde{F}^{\mu\nu}$ is obtainable from $F^{\mu\nu}$ through the duality transformation

$$\{\mathbf{E}, \mathbf{B}\} \rightarrow \{\mathbf{B}, -\mathbf{E}\} \tag{5.22}$$

From the field tensors we can form two independent Lorentz invariants:

$$\begin{aligned} \tfrac{1}{4} F^{\mu\nu}F_{\mu\nu} &= \tfrac{1}{2}(\mathbf{B}^2 - \mathbf{E}^2) \qquad \text{(scalar)} \\ \tfrac{1}{4} \tilde{F}^{\mu\nu}F_{\mu\nu} &= -\mathbf{B}\cdot\mathbf{E} \qquad \text{(pseudoscalar)} \end{aligned} \tag{5.23}$$

In terms of the field tensors, Maxwell's equations read

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= j^\nu \\ \partial_\mu \tilde{F}^{\mu\nu} &= 0 \end{aligned} \tag{5.24}$$

which are gauge-invariant and Lorentz-covariant, and are invariant under the duality transformation when $j^\nu = 0$. Since $F^{\mu\nu} = -F^{\nu\mu}$, the first equation is consistent only if $\partial_\mu j^\mu = 0$. The second equation is identically satisfied by putting $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.

5.3 CANONICAL FORMALISM

The Lagrangian density of the free electromagnetic field is

$$\mathcal{L} = -\tfrac{1}{4} F^{\mu\nu}F_{\mu\nu} = \tfrac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2) \tag{5.25}$$

Apart from an overall factor, this is uniquely determined by the requirement that it be Lorentz- and gauge-invariant, and does not contain higher derivatives of A^μ than first derivatives. The minus sign in front is chosen to give a positive energy density for the free field, and the factor $\frac{1}{4}$ sets the normalization of the fields. To obtain the equations of motion from the action principle, we must use the potential A^n as the field variable. The Lagrangian density then reads

$$\mathcal{L} = -\tfrac{1}{2}(\partial^\mu A^\nu - \partial^\nu A^\mu)\partial_\mu A_\nu \tag{5.26}$$

from which we obtain

$$\pi^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu} \tag{5.27}$$

The equation of motion is

$$\partial_\mu \pi^{\mu\nu} - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0 \tag{5.28}$$

Since the last term on the left side is zero, we have

$$\partial_\mu F^{\mu\nu} = 0 \tag{5.29}$$

The use of the potential makes $\partial_\mu \tilde{F}^{\mu\nu} \equiv 0$. Thus, we correctly recover Maxwell's equations for free fields.

The canonical conjugate to A^ν is

$$\pi^{0\nu} = -F^{0\nu} \tag{5.30}$$

which vanishes identically for $\nu = 0$, indicating that A^0 is not a dynamical variable. The dynamical fields are A^k, with canonical conjugate $-F^{0k} = E^k$. However, the longitudinal part of **A** has no physical significance, because it can be changed at will through a gauge transformation (see Table 5.1.) The only dynamical degrees of freedom are the two transverse components of **A**, and we can go to the Coulomb gauge to make this explicit. In Coulomb gauge A^0 satisfies the Poisson equation, and is determined by the external charges.

The canonical energy–momentum tensor is, according to (4.26),

$$T_c^{\mu\alpha} = \pi^{\mu\nu}\partial^\alpha A_\nu - g^{\mu\alpha}\mathcal{L} = -F^{\mu\nu}\partial^\alpha A_\nu - g^{\mu\alpha}\mathcal{L} \tag{5.31}$$

which can be rewritten using the equation of motion:

$$T_{c\ \alpha}^{\mu} = -F^{\mu\nu}F_{\alpha\nu} - g_\alpha^\mu \mathcal{L} - \partial_\nu(F^{\mu\nu}A_\alpha) \tag{5.32}$$

The last term is not symmetric in μ and α, and not gauge-invariant. However, it is a total 4-divergence antisymmetric in μ and ν, and is conserved because $\partial_\mu\partial_\nu(F^{\mu\nu}A^\alpha) \equiv 0$. As discussed in Section 4.5, such a term has no effect on the conservation law and the definitions of total energy and total momentum, and may therefore be omitted. Thus we take as energy–momentum tensor the symmetric and gauge-invariant tensor

$$T_\alpha^\mu = -F^{\mu\nu}F_{\alpha\nu} - g_\alpha^\mu \mathcal{L} \tag{5.33}$$

TABLE 5.1 Fields and Canonical Conjugates

Field	*Canonical Conjugate*	*Remark*
A^0	0	Not dynamical variable
A	**–E**	Only transverse part physical

which satisfies

$$\partial_\mu T^\mu_{\ \alpha} = 0 \tag{5.34}$$

The trace of the tensor vanishes:

$$T^\mu_{\ \mu} = 0 \tag{5.35}$$

It is now straightforward to obtain the Hamiltonian density:

$$\mathcal{H} = T^{00} = -F^{0k}F_{0k} - \mathcal{L} = \tfrac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) \tag{5.36}$$

The momentum density (the Poynting vector) is

$$S^k = T^{0k} = F^{0j}F^{kj} = \epsilon^{kij}E^iB^j \tag{5.37}$$

The total Hamiltonian H and total momentum $\mathbf{P}$ are given by

$$H = \tfrac{1}{2}\int d^3r(\mathbf{E}^2 + \mathbf{B}^2)$$

$$\mathbf{P} = \int d^3r\, \mathbf{E} \times \mathbf{B} \tag{5.38}$$

The conservation of energy and momentum correspond to the statements

$$\frac{\partial}{\partial t}\mathcal{H} + \nabla\cdot\mathbf{S} = 0$$

$$\frac{\partial}{\partial t}S^k + \partial_j T^{jk} = 0 \tag{5.39}$$

where

$$T^{jk} = -(E^jE^k + B^jB^k) + \tfrac{1}{2}\delta_{jk}(\mathbf{E}^2 + \mathbf{B}^2) \tag{5.40}$$

is the stress tensor.

According to (4.47), the generalized angular momentum tensor is

$$M^{\mu\alpha\beta} = x^\alpha T^{\mu\beta} - x^\beta T^{\mu\alpha} \tag{5.41}$$

which satisfies the conservation law

$$\partial_\mu M^{\mu\alpha\beta} = 0 \tag{5.42}$$

It follows that the total angular momentum $\mathbf{J}$ and the Lorentz boost $\mathbf{K}$ are given by

$$\mathbf{J} = \int d^3r\, \mathbf{r} \times (\mathbf{E} \times \mathbf{B})$$

$$\mathbf{K} = t\mathbf{P} - \mathbf{r}H \tag{5.43}$$

5.4 QUANTIZATION IN COULOMB GAUGE

To quantize the electromagnetic field, we must first eliminate all unphysical degrees of freedom by fixing the gauge, and in the following we shall use Coulomb gauge.[1] In the absence of external charges, we can set $A^0 = 0$, and write the Hamiltonian in the form

$$H = \tfrac{1}{2} \int d^3r\, (\mathbf{E}^2 + |\nabla \times \mathbf{A}|^2) \qquad (\nabla \cdot \mathbf{A} = 0) \tag{5.44}$$

The canonical conjugate to $\mathbf{A}$ is $-\mathbf{E} = \partial\, \mathbf{A}/\partial t$, and we would normally impose the equal-time commutation relation $[E^j(\mathbf{r}, t), A^k(\mathbf{r}', t)] = i\delta_{jk}\delta^3(\mathbf{r} - \mathbf{r}')$. But this is incorrect here, because the right side is not consistent with $\nabla \cdot \mathbf{A} = 0$, nor with one of Maxwell's equations $\nabla \cdot \mathbf{E} = 0$. We therefore replace $\delta_{jk}\delta^3(\mathbf{r} - \mathbf{r}')$ by its transverse projection, and take

$$[E^j(\mathbf{r}, t), A^k(\mathbf{r}', t)] = i\delta^{\mathrm{T}}_{jk}(\mathbf{r} - \mathbf{r}') \tag{5.45}$$

where the transverse delta function $\delta^{\mathrm{T}}_{jk}(\mathbf{r} - \mathbf{r}')$ is defined by

$$\delta^{\mathrm{T}}_{jk}(\mathbf{r}) = \int \frac{d^3k}{(2\pi)^3} \left(\delta_{jk} - \frac{k^j k^j}{|\mathbf{k}|^2} \right) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{5.46}$$

and satisfies

$$\partial^j \delta^{\mathrm{T}}_{jk}(\mathbf{r}) = \partial^k \delta^{\mathrm{T}}_{jk}(\mathbf{r}) = 0 \tag{5.47}$$

A complete set of solutions to Maxwell's equations in a periodic box of volume Ω are the transverse plane waves

$$\boldsymbol{\epsilon}(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} \qquad \omega_k = |\mathbf{k}|$$

where $\boldsymbol{\epsilon}(\mathbf{k})$ is a unit polarization vector normal to $\mathbf{k}$. For each $\mathbf{k}$, there are two independent polarization vectors $\boldsymbol{\epsilon}_1(\mathbf{k})$ and $\boldsymbol{\epsilon}_2(\mathbf{k})$; and $\boldsymbol{\epsilon}_1(\mathbf{k})$, $\boldsymbol{\epsilon}_2(\mathbf{k})$, $\hat{\mathbf{k}}$ together form a right-handed coordinate system:

$$\mathbf{k} \cdot \boldsymbol{\epsilon}_1(\mathbf{k}) = \mathbf{k} \cdot \boldsymbol{\epsilon}_2(\mathbf{k}) = 0$$

[1] For quantization in other gauges, see Huang [1].

$$\boldsymbol{\epsilon}_s(\mathbf{k})\cdot\boldsymbol{\epsilon}_r(\mathbf{k}) = \delta_{sr}$$
$$\boldsymbol{\epsilon}_1(\mathbf{k}) \times \boldsymbol{\epsilon}_2(\mathbf{k}) = \frac{\mathbf{k}}{|\mathbf{k}|} \tag{5.48}$$

Having chosen $\boldsymbol{\epsilon}_1(\mathbf{k})$, $\boldsymbol{\epsilon}_2(\mathbf{k})$, there is still arbitrariness in the choice of $\boldsymbol{\epsilon}_1(-\mathbf{k})$, $\boldsymbol{\epsilon}_2(-\mathbf{k})$. By convention, we choose

$$\boldsymbol{\epsilon}_1(-\mathbf{k}) = \boldsymbol{\epsilon}_1(\mathbf{k})$$
$$\boldsymbol{\epsilon}_2(-\mathbf{k}) = -\boldsymbol{\epsilon}_2(\mathbf{k}) \tag{5.49}$$

as illustrated in Fig. 5.1. The following sum over polarizations results in the transverse projection operator (see Problem 5.2):

$$I^{ij}(\mathbf{k}) \equiv \sum_{s=1}^{2} \epsilon_s^i(\mathbf{k})\epsilon_s^j(\mathbf{k}) = \delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2} \tag{5.50}$$

We now expand the field in terms of the transverse plane waves:

$$A^j(\mathbf{r}, 0) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k \Omega}} \sum_{s=1}^{2} \epsilon_s^j(\mathbf{k})[a_s(\mathbf{k})e^{i\mathbf{k}\mathbf{r}} + a_s^\dagger(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}}]$$
$$\dot{A}^j(\mathbf{r}, 0) = -i\sum_{\mathbf{k}} \sqrt{\frac{\omega_k}{2\Omega}} \sum_{s=1}^{2} \epsilon_s^j(\mathbf{k})[a_s(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}} - a_s^\dagger(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{r}}] \tag{5.51}$$

The commutation relations (5.45) are satisfied by imposing the commutation relations

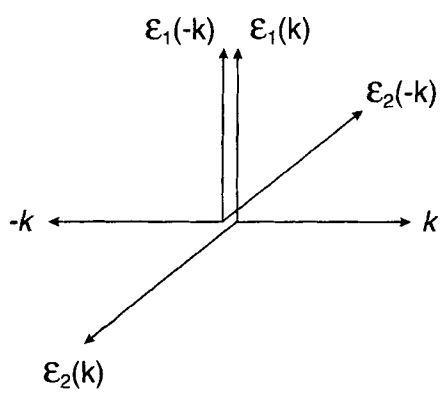


Figure 5.1 Polarization vectors of an electromagnetic wave.

$$[a_s(\mathbf{k}), a_r^\dagger(\mathbf{p})] = \delta_{sr}\delta_{\mathbf{kp}}$$

$$[a_s(\mathbf{k}), a_r(\mathbf{p})] = 0 \tag{5.52}$$

where $a_s(\mathbf{k})$ is the annihilation operator of a photon—a field quantum of momentum $\mathbf{k}$ and linear polarization s. In the free-field case, the time-dependent operator $A^j(\mathbf{r}, t)$ is simply obtained by replacing ($\mathbf{k}\cdot\mathbf{r}$ by ($\mathbf{k}\cdot\mathbf{r} - \omega_k t$) in the exponents in (5.51), because

$$e^{iHt}a_s(\mathbf{p})e^{-iHt} = a_s(\mathbf{p})e^{-i\omega_p t} \tag{5.53}$$

The Hamiltonian and the total momentum operator of the electromagnetic field are given respectively by

$$H = \tfrac{1}{2}\sum_{\mathbf{k},s} |\mathbf{k}|[a_s^\dagger(\mathbf{k})a_s(\mathbf{k}) + \tfrac{1}{2}]$$

$$\mathbf{P} = \sum_{\mathbf{k},s} \mathbf{k}\, a_s^\dagger(\mathbf{k})a_s(\mathbf{k}) \tag{5.54}$$

These equations show that photons are boson with energy–momentum relation $\omega_k = |\mathbf{k}|$. The vacuum state $|0\rangle$ is the state with no photons. All other states of the system can be generated by applying creation operators repeatedly to the vacuum state.

In the limit $\Omega \to \infty$, the expansion (5.51) becomes a Fourier integral:

$$A^j(\mathbf{r}, 0) = \int \frac{d^3k}{(2\pi)^3\sqrt{2\omega_k}} \sum_{s=1}^{2} \epsilon(\mathbf{k}, s)[a(\mathbf{k}, s)e^{i\mathbf{k}\cdot\mathbf{r}} + a^\dagger(\mathbf{k}, s)e^{-i\mathbf{k}\cdot\mathbf{r}}] \tag{5.55}$$

where we write $\epsilon(\mathbf{k}, s) = \epsilon_s(\mathbf{k})$ for consistency in notation. The continuum form of the annihilation operator is

$$a(\mathbf{k}, s) = \sqrt{\Omega}a_s(\mathbf{k}) \tag{5.56}$$

which obeys the commutation relations

$$[a(\mathbf{k}, s), a^\dagger(\mathbf{k}', s')] = (2\pi)^3\delta_{ss'}\delta^3(\mathbf{k} - \mathbf{k}') \tag{5.57}$$

The Hamiltonian and total momentum now take the forms

$$H = \frac{1}{2}\sum_{s=1}^{2}\int \frac{d^3k}{(2\pi)^3}|\mathbf{k}|a^\dagger(\mathbf{k}, s)a(\mathbf{k}, s)$$

$$\mathbf{P} = \sum_{s=1}^{2}\int \frac{d^3k}{(2\pi)^3}\mathbf{k}\, a^\dagger(\mathbf{k}, s)a(\mathbf{k}, s) \tag{5.58}$$

5.5 SPIN ANGULAR MOMENTUM

According to (5.43), the angular momentum density is

$$\mathbf{u} = \mathbf{r} \times [\mathbf{E} \times (\nabla \times \mathbf{A})] \tag{5.59}$$

We define the spin density to be the part that is independent of the origin of $\mathbf{r}$. To find it, let us first rewrite the preceding in component form:

$$u^i = \epsilon^{ijk}\epsilon^{klm}\epsilon^{mnq}x^j E^l \partial_n A^q \tag{5.60}$$

Now combine the last two ϵ symbols according to the rule

$$\epsilon^{klm}\epsilon^{nqm} = \delta_{kn}\delta_{lq} - \delta_{kq}\delta_{ln} \tag{5.61}$$

We then obtain

$$\begin{aligned} u^i &= \epsilon^{ijk}(x^j E^q \partial_k A^q - x^j E^n \partial_n A^k) \\ &= \epsilon^{ijk}[x^j E^q \partial_k A^q - \partial_n(x^j E^n A^k) + x^j(\partial_n E^n)A^k + E^n A^k] \end{aligned} \tag{5.62}$$

The factor x^j in the first term cannot be removed by manipulations involving ∂_k, because $j \neq k$. The second term is a total 3-divergence, and can be ignored. The third term vanishes because $\nabla \cdot \mathbf{E} = 0$. The last term is independent of $\mathbf{r}$, and is identified as the spin density:

$$\mathbf{s} = \mathbf{E} \times \mathbf{A} \tag{5.63}$$

This is the spatial part of the tensor

$$s^{\mu\nu} = A^\alpha \Sigma^{\mu\nu}_{\alpha\beta} \pi^\beta \tag{5.64}$$

where $\Sigma^{\mu\nu}_{\alpha\beta}$ is the spin matrix given in (3.49). The spin angular momentum is given by

$$\mathbf{S} = \int d^3r\, \mathbf{E} \times \mathbf{A} \tag{5.65}$$

Using the expansion (5.51), we obtain

$$\mathbf{S} = i\sum_{\mathbf{k}} \hat{\mathbf{k}}[a_2(\mathbf{k})^\dagger a_1(\mathbf{k}) - a_1^\dagger(\mathbf{k})a_2(\mathbf{k})] \tag{5.66}$$

where $\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|$. To diagonalize this, we make a linear transform to circularly polarized photons.

The polarization vectors for circularly polarized photons are

$$\boldsymbol{\epsilon}_{\pm} = \frac{1}{\sqrt{2}}(\boldsymbol{\epsilon}_1 \pm i\boldsymbol{\epsilon}_2) \tag{5.67}$$

where the label $\mathbf{k}$ has been suppressed. As one can easily verify, the plane wave

$$\mathrm{Re}[\boldsymbol{\epsilon}_+ e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}] \tag{5.68}$$

represents a traveling wave whose polarization vector rotates in a right-handed sense about $\mathbf{k}$. This is called a *left-circularly polarized wave,* because an observer facing the incoming wave would see the polarization rotating to the left. Similarly, $\boldsymbol{\epsilon}_-$ corresponds to a right-circularly polarized wave. The annihilation operators for a circularly polarized photons are given by

$$\text{Left-circularly polarized:} \qquad a_+(\mathbf{k}) = \frac{1}{\sqrt{2}}[a_1(\mathbf{k}) - ia_2(\mathbf{k})]$$

$$\text{Right-circularly polarized:} \qquad a_-(\mathbf{k}) = \frac{1}{\sqrt{2}}[a_1(\mathbf{k}) + ia_2(\mathbf{k})] \tag{5.69}$$

The commutation relations are

$$\begin{aligned} [a_\pm(\mathbf{k}), a_\pm^\dagger(\mathbf{p})] &= \delta_{\mathbf{kp}} \\ [a_+(\mathbf{k}), a_-(\mathbf{p})] &= [a_+(\mathbf{k}), a_-^\dagger(\mathbf{p})] = 0 \end{aligned} \tag{5.70}$$

In terms of these, the spin operator becomes diagonal:

$$\mathbf{S} = \sum_{\mathbf{k}} \hat{\mathbf{k}}[a_+(\mathbf{k})^\dagger a_+(\mathbf{k}) - a_-^\dagger(\mathbf{k})a_-(\mathbf{k})] \tag{5.71}$$

This shows that the photon has spin 1, but there are only two helicity states. The helicity +1 corresponds to left-circular polarization, and −1 corresponds to right-circular polarization:

$$a_\pm(\mathbf{k}) \text{ annihilates helicity state } \pm 1 \tag{5.72}$$

In terms of circular polarization, the field operator has the expansion

$$\begin{aligned} \mathbf{A}(\mathbf{r}, 0) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k \Omega}} \{ &[\boldsymbol{\epsilon}_+(\mathbf{k})a_+(\mathbf{k}) + \boldsymbol{\epsilon}_-(\mathbf{k})a_-(\mathbf{k})]e^{i\mathbf{k}\cdot\mathbf{r}} \\ &+ [\boldsymbol{\epsilon}_+^*(\mathbf{k})a_+^\dagger(\mathbf{k}) + \boldsymbol{\epsilon}_-^*(\mathbf{k})a_-^\dagger(\mathbf{k})]\, e^{-i\mathbf{k}\cdot\mathbf{r}}\} \end{aligned} \tag{5.73}$$

In the convention (5.49) the sense of the circular polarizations remains unchanged when $\mathbf{k} \to -\mathbf{k}$:

$$\boldsymbol{\epsilon}_+(-\mathbf{k}) = i\boldsymbol{\epsilon}_+(\mathbf{k})$$
$$\boldsymbol{\epsilon}_-(-\mathbf{k}) = -i\boldsymbol{\epsilon}_-(\mathbf{k}) \tag{5.74}$$

5.6 INTRINSIC PARITY

Let us make a coordinate transformation $\mathbf{r} \rightarrow \mathbf{r}'$, with the transformation law

$$x^j = \sum_{j=1}^{3} \rho^{jk} x'^k \tag{5.75}$$

Since A^k is a vector field, this induces the unitary transformation U according to (3.54):

$$UA^j(\mathbf{r})U^{-1} = \sum_{j=1}^{3} \rho^{jk} A^k(\mathbf{r}') \tag{5.76}$$

where $A^j(\mathbf{r}) = A^j(\mathbf{r}, 0)$. For spatial reflection $\mathbf{r}' = -\mathbf{r}$, we denote the unitary transformation by $\mathcal{P}$:

$$\mathcal{P}A^k(\mathbf{r})\mathcal{P}^{-1} = -A^k(-\mathbf{r}) \tag{5.77}$$

This establishes the fact that the electromagnetic field has odd intrinsic parity.

To investigate how photon states transform, we substitute into the preceding the expansion (5.51). Using the abbreviation

$$\mathbf{a}(\mathbf{k}) = \sum_{s=1}^{2} \boldsymbol{\epsilon}_s(\mathbf{k}) a_s(\mathbf{k}) \tag{5.78}$$

we have

$$\begin{aligned}
\mathcal{P}\mathbf{A}(\mathbf{r})\mathcal{P}^{-1} &= \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k \Omega}} [\mathcal{P}\mathbf{a}(\mathbf{k})\mathcal{P}^{-1} e^{i\mathbf{k}\cdot\mathbf{r}} + \mathcal{P}\mathbf{a}^\dagger(\mathbf{k})\mathcal{P}^{-1} e^{-i\mathbf{k}\cdot\mathbf{r}}] \\
&= -\sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k \Omega}} [\mathbf{a}(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}} + \mathbf{a}^\dagger(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}] \\
&= -\sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k \Omega}} [\mathbf{a}(-\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} + \mathbf{a}^\dagger(-\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}}]
\end{aligned} \tag{5.79}$$

where the last relation is obtained by changing the summation variable from $\mathbf{k}$ to $-\mathbf{k}$. Thus

$$\mathcal{P}\mathbf{a}(\mathbf{k})\mathcal{P}^{-1} = -\mathbf{a}(-\mathbf{k}) \tag{5.80}$$

which gives

$$\mathcal{P}a_1(\mathbf{k})\mathcal{P}^{-1} = -a_1(-\mathbf{k})$$

$$\mathcal{P}a_2(\mathbf{k})\mathcal{P}^{-1} = a_2(-\mathbf{k}) \tag{5.81}$$

In terms of circular polarization, we obtain

$$\mathcal{P}a_\pm(\mathbf{k})\mathcal{P}^{-1} = -a_\mp(-\mathbf{k}) \tag{5.82}$$

A one-photon state of momentum $\mathbf{k}$, linear polarization s, is defined by

$$|\mathbf{k}, s\rangle = a_s^\dagger(\mathbf{k})|0\rangle \tag{5.83}$$

States with circular polarization are given by

$$|\mathbf{k}, \pm\rangle = a_\pm^\dagger(\mathbf{k})|0\rangle \tag{5.84}$$

which are linearly superpositions of states with linear polarizations:

$$|\mathbf{k}, \pm\rangle = \frac{1}{\sqrt{2}}[|\mathbf{k},1\rangle \pm i|\mathbf{k},2\rangle] \tag{5.85}$$

Assuming that the vacuum state is invariant under reflection, we have

$$\mathcal{P}|\mathbf{k}, \pm\rangle = \mathcal{P}a_\pm^\dagger(\mathbf{k})\mathcal{P}^{-1}|0\rangle = -|\mathbf{k}, \mp\rangle \tag{5.86}$$

Thus, under spatial reflection, left and right are interchanged, and the state vector changes sign.

5.7 TRANSVERSE PROPAGATOR

We now calculate the photon propagator in Coulomb gauge:

$$D_{\mathrm{T}}^{ij}(x) = -i\langle 0|\mathrm{T}A^i(x)A^j(0)|0\rangle \tag{5.87}$$

where the subscript "T" reminds us that the field is transverse: $\partial^k A^k = 0$. Expanding the field in creation and annihilation operators, we have

$$D_{\mathrm{T}}^{ij}(x) = -\frac{i}{(2\pi)^6}\int \frac{d^3k d^3k'}{\sqrt{4\omega\omega'}} \sum_{ss'} \begin{cases} \langle 0|a^i(\mathbf{k}, s)a^{j\dagger}(\mathbf{k}', s')|0\rangle e^{-ik\cdot x} & (x_0 > 0) \\ \langle 0|a^j(\mathbf{k}', s')a^{i\dagger}(\mathbf{k}, s)|0\rangle e^{ik\cdot x} & (x_0 > 0) \end{cases} \tag{5.88}$$

where

$$\omega = k_0 \tag{5.89}$$

and we use the abbreviation

$$a^i(\mathbf{k}, s) \equiv \epsilon^i(\mathbf{k}, s)a(\mathbf{k}, s) \tag{5.90}$$

The vacuum expectation values are easily calculated:

$$\begin{aligned}\langle 0|a^i(\mathbf{k}, s)a^{j\dagger}(\mathbf{k}', s')|0\rangle &= \langle 0|a^j(\mathbf{k}', s')a^{i\dagger}(\mathbf{k}, s)|0\rangle \\ &= (2\pi)^3\delta_{ss'}\delta^3(\mathbf{k}-\mathbf{k}')\epsilon^i(\mathbf{k}, s)\epsilon^j(\mathbf{k}', s')\end{aligned} \tag{5.91}$$

Therefore

$$D_{\mathrm{T}}^{ij}(x) = -\frac{i}{(2\pi)^3}\int\frac{d^3k}{2\omega}e^{i\mathbf{k}\cdot\mathbf{x}-\omega|t|}I^{ij}(\mathbf{k}) \tag{5.92}$$

where I^{ij} is defined in (5.50).This can be rewritten as a four-dimensional Fourier integral, with the help of the identity

$$\frac{e^{-i\omega|t|}}{2\omega} = -\frac{1}{2\pi i}\int_{-\infty}^{\infty}du\frac{e^{-iut}}{\omega^2-u^2-i\eta} \qquad (\eta\to 0^+) \tag{5.93}$$

The final form is

$$D_{\mathrm{T}}^{ij}(x) = \int\frac{d^4k}{(2\pi)^4}\frac{e^{ik\cdot x}}{k^2+i\eta}I^{ij}(\mathbf{k}) \tag{5.94}$$

where $k^2 = k_0{}^2 - |\mathbf{k}|^2$. The Fourier transform is

$$\tilde{D}_{\mathrm{T}}^{ij}(k) = \frac{1}{k^2+i\eta}\left(\delta_{ij}-\frac{k^ik^j}{|\mathbf{k}|^2}\right) \tag{5.95}$$

This is not Lorentz-covariant, for it is in Coulomb gauge. To prove that the quantized field theory is covariant, we should exhibit the gauge transformation that will maintain the form of the transverse propagator under Lorentz transformations. However, this is unnecessary, as we shall show in Chapter 11. The point is that noncovariant part of the propagator is physically irrelevant, because, owing to current conservation, it does not contribute to the scattering amplitude.

5.8 VACUUM FLUCTUATIONS

The vacuum state is neither an eigenstate of **E** nor **B**, since these operators annihilate or create photons singly. Although the fields average to zero, their mean-square

fluctuations are large. This can be shown via direct calculation, as in Problem 2.8. We can also demonstrate it through the following argument. The energy density in the vacuum state is

$$\tfrac{1}{2}\langle 0|\mathbf{E}^2 + \mathbf{B}^2|0\rangle = \langle 0|\mathbf{E}^2|0\rangle \tag{5.96}$$

for the free-field theory is invariant under the duality transformation. Equating this with the zero-point energy per unit volume in (5.54), we have

$$\langle 0|\mathbf{E}^2|0\rangle = \frac{1}{2\Omega}\sum_{\mathbf{k}s}\omega_k = \frac{1}{(2\pi)^3}\int d^3k|\mathbf{k}| \tag{5.97}$$

which diverges because of the short-wavelength modes. This divergence is harmless, since only energy differences have physical significance; but the long-wavelength part of the fluctuations gives rise to observable effects, including the Casimir effect.

We illustrate the essence of the Casimir effect in a simple one-dimensional example, leaving for the next section a more detailed treatment. Consider the modes of a harmonic oscillator in a box of length L. The zero-point energy is

$$E_0(L) = \frac{1}{2}\sum_{\omega}\omega f(\omega)$$

$$\omega = \frac{\pi n}{L} \qquad (n = 1,2,\ldots\infty) \tag{5.98}$$

where we have introduced a cutoff function $f(\omega)$, with the properties

$$f(0) = 1$$

$$f(\omega) \underset{\omega\to\infty}{\to} 0 \tag{5.99}$$

There is a cutoff frequency ω_c, above which $f(\omega)$ decreases rapidly to zero, and we take the limit $\omega_c \to \infty$ eventually. Suppose that a partition is inserted, such that normal modes are required to have a node at the wall. The modes near the cutoff frequency are hardly affected, because their wavelengths are vanishingly small. Therefore, there are now fewer normal modes below the cutoff, as illustrated in Fig. 5.2, and the zero-point energy decreases.

For definiteness, choose the cutoff function to be

$$f(\omega) = e^{-\omega/\omega_c} \tag{5.100}$$

The zero-point energy for a box without partitions can be easily calculated, with the result

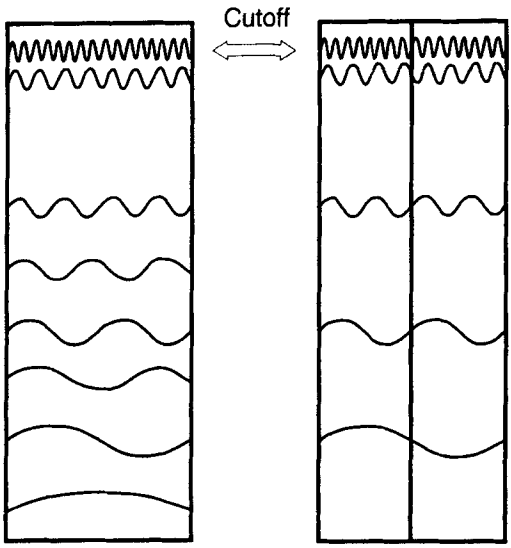


Figure 5.2 When a wall is inserted into a box, those normal modes that do not have a node at the wall are suppressed. Consequently, the number of modes below a fixed frequency decreases, and the zero-point energy is lowered.

$$E_0(L) = \frac{\pi}{8L}\frac{1}{\sinh^2(\pi/\omega_c L)} \xrightarrow[\omega_c\to\infty]{} \frac{L\omega_c^2}{2\pi^2} - \frac{\pi}{24L} + O(\omega_c^{-2}) \tag{5.101}$$

Now insert two partitions centered about the midpoint, separated by distance a. The box is divided into three compartments—one with length a and the others with length $(L-a)/2$—and the zero-point energy becomes

$$E(a) = E_0(a) + 2E_0\frac{L-a}{2} \xrightarrow[\omega_c\to\infty]{} \frac{L\omega_c^2}{2\pi^2} - \frac{\pi}{6(L-a)} - \frac{\pi}{24a} \tag{5.102}$$

In the limit $L \to \infty$, the attractive force between the walls is given by

$$-\frac{\partial E(a)}{\partial a} = -\frac{\pi}{24a^2} \tag{5.103}$$

which is independent of the cutoff.

5.9 THE CASIMIR EFFECT

We now calculate the force between two metallic plates in the electrodynamic vacuum. The first task is to obtain the normal modes of the electromagnetic field in a perfectly conducting box of size $a \times b \times c$. We choose one corner of the box as origin, and use Coulomb gauge. On each face of the box, the boundary condition is

$$\mathbf{E}_{\parallel} = 0 \qquad \mathbf{B}_{\perp} = 0 \tag{5.104}$$

where the subscripts $\parallel$ and $\perp$ denote respectively the tangential and normal components. We put $\mathbf{B} = \nabla \times \mathbf{A}$, $\mathbf{E} = -\dot{\mathbf{A}}$, to obtain

$$\mathbf{A}_{\parallel} = 0 \qquad (\nabla \times \mathbf{A})_{\perp} = 0 \tag{5.105}$$

On the y–z plane, for example, the boundary conditions are

$$A_y = A_z = 0 \qquad \partial_y A_z - \partial_z A_y = 0$$

The first says that $\mathbf{A}$ is normal to the surface, and therefore the second condition is automatically satisfied. We must, however, satisfy the gauge condition

$$\partial_x A_x + \partial_y A_y + \partial_z A_z = 0 \tag{5.106}$$

which leads to

$$\partial_x A_x = 0 \tag{5.107}$$

Thus, the boundary conditions in Coulomb gauge are

$$\mathbf{A}_{\parallel} = 0 \qquad \frac{\partial}{\partial n}\mathbf{A}_{\perp} = 0 \tag{5.108}$$

For A_x, for example, the conditions are

$$\begin{aligned} [\partial_x A_x(x, y, z)]_{x=0} &= [\partial_x A_x(x, y, z)]_{x=a} = 0 \\ A_x(x, y, 0) &= A_x(x, y, a) = 0 \\ A_x(x, y, 0) &= A_x(x, y, c) = 0 \end{aligned} \tag{5.109}$$

A complete set of solutions to the wave equation is given by

$$A_x = \pm \cos(k_x x) \sin(k_y y) \sin(k_z z)$$

$$A_y = \sin(k_x x)\cos(k_y y)\sin(k_z z)$$
$$A_z = \sin(k_x x)\sin(k_y y)\cos(k_z z) \tag{5.110}$$

where

$$k_x = \frac{\pi n_x}{a} \qquad k_y = \frac{\pi n_y}{b} \qquad k_z = \frac{\pi n_z}{c} \tag{5.111}$$

with $n_i = 1,2,\ldots,\infty$. The frequency is given by

$$\omega_k = \sqrt{k_x^2 + k_y^2 + k_z^2} \tag{5.112}$$

If all three components of **k** are nonzero, there are two independent solutions corresponding to the ± signs in A_x. If any component of **k** vanishes, there is only one solution. For example, if $k_x = 0$, then $A_x = A_z = 0$, and the ± sign does not make any difference. We can now obtain the zero-point energy:

$$E_0(a,b,c) = \tfrac{1}{2}\sum_{k_y,k_z} \sqrt{k_y^2 + k_z^2}\, F\left(\sqrt{k_y^2 + k_z^2}\right) \tag{5.113}$$

$$+ \sum_{k_x,k_y,k_z} \sqrt{k_x^2 + k_y^2 + k_z^2}\, F\left(\sqrt{k_x^2 + k_y^2 + k_z^2}\right) \tag{5.114}$$

where $F(k)$ is a cutoff function.

Consider now a large cubicle box of edge L, which is divided into three compartments as shown in Fig. 5.3, with two parallel metallic plates inserted normal to the x axis, separated by a distance a, symmetric about the midpoint. The zero-point energy is the sum of those of the compartments. That of the middle compartment, of dimensions $L \times L \times a$, with $L \to \infty$, is given by

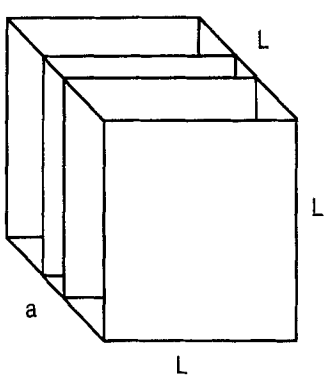


Figure 5.3 Two metallic plates separated by distance a in the electrodynamic vacuum, which is represented by a cube of edge $L \to \infty$.

$$
\begin{aligned}
U(a) &= E_0(a, L, L) \\
&= \frac{L^2}{2(2\pi)^2}\int d^2k\left[k\,F(k) + 2\sum_{n=1}^{\infty}\sqrt{k^2 + \frac{\pi^2 n^2}{a^2 F}}\left(\sqrt{k^2 + \frac{\pi^2 n^2}{a^2}}\right)\right] \\
&= \frac{L^2}{4\pi}\int d^2k k\,F(k) + \frac{\pi^2 L^2}{4a^3}\sum_{n=1}^{\infty} G(n)
\end{aligned}
\tag{5.115}
$$

where

$$
G(n) = \int_{n^2}^{\infty} dy\,\sqrt{y}\,F\left(\frac{\pi\sqrt{y}}{a}\right) \tag{5.116}
$$

We can rewrite the n-sum using the Euler–MacLaurin formula [2]

$$
\sum_{n=1}^{\infty} G(n) = \int_0^{\infty} dn\,G(n) + \frac{1}{2}G(0) - \frac{B_2}{2!}G'(0) - \frac{B_4}{4!}G'''(0) + \cdots \tag{5.117}
$$

where $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$. Using $G'(0) = 0$, $G'''(0) = -4$, and the fact that all higher derivatives vanish at $n = 0$, we obtain

$$
\sum_{n=1}^{\infty} G(n) = \int_0^{\infty} dn\,G(n) + \frac{1}{2}\int_0^{\infty} dy\sqrt{y}F\left(\frac{\pi\sqrt{y}}{a}\right) - \frac{4}{4!30}
$$

This leads to

$$
U(a) = L^2\left[C_1 a + C_2 - \frac{\pi^2}{720a^3}\right] \tag{5.118}
$$

where

$$
\begin{aligned}
C_1 &= \frac{1}{4\pi^2}\int_0^{\infty} d\tau \int_0^{\infty} dk\,k\sqrt{k^2 + \tau^2}\,F(k) \\
C_2 &= \frac{1}{8\pi}\int_0^{\infty} dk\,k^2 F(k)
\end{aligned}
\tag{5.119}
$$

The zero-point energy in the box in Fig. 5.3 is given by

$$
\begin{aligned}
E(a) &= U(a) + 2U((L-a)/2) \\
&= L^2\left[C_1 + 2C_2 - \frac{\pi^2}{720a^3}\right] + O(L^{-1})
\end{aligned}
$$

This gives an attractive force per unit area between the plates:

$$f = -\frac{1}{L^2}\frac{\partial E(a)}{\partial a} = -\frac{\pi^2}{240a^4} \tag{5.120}$$

or, in practical units,

$$f = -\frac{\pi^2 hc}{240a^4} = -\frac{0.013}{a^4}\ \text{dyn/cm}^2 \tag{5.121}$$

where a is in micrometers. Figure 5.4 compares this result and early measurements [3], with reasonable agreement. More recent measurements of a similar force between a plate and a sphere have achieved much greater experimental accuracy [4].

5.10 THE GAUGE PRINCIPLE

We now discuss how the electromagnetic field should be coupled to charged fields. A nonrelativistic charged particle obeys the Schrödinger equation

$$\left[-\frac{1}{2m}(\nabla + ie\mathbf{A})^2 + e\phi\right]\psi(\mathbf{r}, t) = i\frac{\partial}{\partial t}\psi(\mathbf{r}, t) \tag{5.122}$$

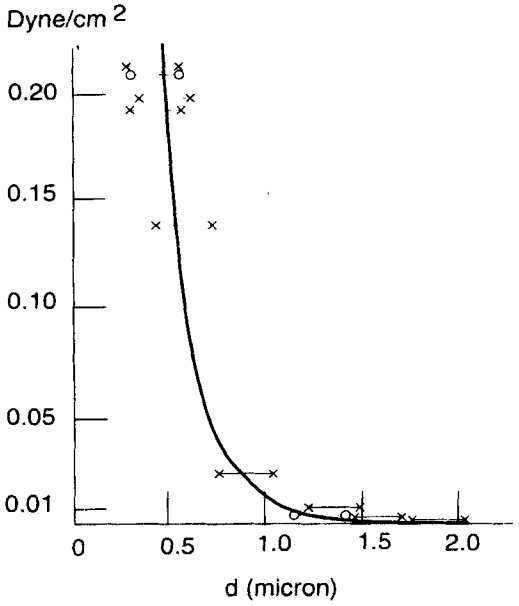


Figure 5.4 The Casimir attractive force between two metallic plates in the vacuum: x—chromium steel, o—chromium; solid line—theory. [Data from M. J. Sparnay, *Physica* **24**, 751 (1958).]

where $\mathbf{A}$, ϕ are respectively the vector and scalar potentials of an external electromagnetic field and e is the charge of the particle. We can derive the form of the interaction as follows. In the absence of external fields, the Schrödinger equation is invariant under a global gauge transformation

$$\psi(\mathbf{r}, t) \rightarrow e^{-i\omega}\psi(\mathbf{r}, t) \tag{5.123}$$

where ω is an arbitrary real constant. The invariance depends on the fact that $\partial^\mu\psi$ transforms in the same manner as ψ. If we make a local gauge transformation, with ω dependent on $\mathbf{r}$, t, this condition will not hold, for we have

$$\partial^\mu\psi \rightarrow e^{-i\omega}[\partial^\mu\psi - i(\partial^\mu\omega)\psi] \tag{5.124}$$

To make the equation invariant, we must cancel the terms involving $\partial^\mu\omega$. This can done by introducing the fields $A^\mu = (\mathbf{A}, \phi)$ through the replacement

$$\frac{\partial}{\partial x_\mu}\psi(x) \rightarrow \left[\frac{\partial}{\partial x_\mu} + ieA^\mu(x)\right]\psi(x) \tag{5.125}$$

The Schrödinger equation is now invariant under the local gauge transformation

$$\begin{aligned} A^\mu(x) &\rightarrow A^\mu(\mathbf{x}) + \partial^\mu\chi(x) \\ \psi(x) &\rightarrow \exp[-ie\chi(x)]\psi(x) \end{aligned} \tag{5.126}$$

The quantity

$$D^\mu\psi(x) = [\partial^\mu + ieA^\mu(x)]\psi(x) \tag{5.127}$$

is called the *covariant derivative*, A^μ is called the *gauge field*, and the recipe for replacing ∂^μ by D^μ is called the *gauge principle*.

Actually, the gauge principle works only for a fully relativistic theory. For the nonrelativistic Schrödinger equation, it fails to produce magnetic moment terms of the form $-\boldsymbol{\mu}\cdot\nabla \times \mathbf{A}$, which has to be put in by hand, with $\boldsymbol{\mu}$ arbitrary. In the relativistic Dirac equation discussed in the next chapter, the gauge principle gives the full electromagnetic interaction of the electron, with a completely determined magnetic moment.

As a relativistic example, consider the complex scalar field with Lagrangian density

$$\mathcal{L}_0(x) = \partial^\mu\psi^*\partial^\mu\psi - m^2\psi^*\psi \tag{5.128}$$

which is invariant under the global gauge transformation

$$\psi(x) \rightarrow e^{-i\omega}\psi(x) \tag{5.129}$$

where ω is a constant; but it is not invariant when ω depends on x. To extend the symmetry to local gauge invariance, we make the replacement

$$D^\mu = \partial^\mu + ieA^\mu(x) \tag{5.130}$$

where e is the electric charge. The Lagrangian density is generalized to

$$\mathcal{L}(x) = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + (D^\mu\psi)^* D^\mu\psi - m^2\psi^*\psi \tag{5.131}$$

which is invariant under the local gauge transformation

$$\begin{aligned}\psi(x) &\to e^{-ie\chi(x)}\psi(x)\\ A^\mu(x) &\to A^\mu(x) + \partial^\mu\chi(x)\end{aligned} \tag{5.132}$$

where $\chi(x)$ is an arbitrary space–time function. The Lagrangian density of the free electromagnetic field is included to make the system self-contained dynamically.

PROBLEMS

5.1 The Lagrangian density for the electromagnetic field in the presence of an external current density j^μ is

$$\mathcal{L} = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} - j^\mu A_\mu$$

What is the condition on j^μ for this to be gauge-invariant?

5.2 Consider the symmetric tensor

$$I^{ij} = \sum_{a=1}^{2} \epsilon^i_{\mathbf{k}a}\epsilon^j_{\mathbf{k}a}$$

(a) Show that $k_i T^{ij} = k_j T^{ij} = 0$, and $T^{ii} = 2$.

(b) Using the preceding conditions show the statement in (5.50):

$$I^{ij}(\mathbf{k}) = \delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2}$$

5.3 Verify that the field operators (5.51) satisfy the commutation relations (5.45). Show, in particular, that the transverse delta function arises from the transversality of the polarization vectors expressed by (5.50).

5.4 **Rotations** Apply the transformation law (5.76) to rotations. In particular, let R be a rotations of the coordinate system about z axis through φ, and ζ be that about the x axis through π:

$$R:\begin{cases} x = x'\cos\varphi + y'\sin\varphi \\ y = -x'\sin\varphi + y'\cos\varphi \\ z = z' \end{cases} \qquad \zeta:\begin{cases} x = x' \\ y = -y' \\ z = -z' \end{cases}$$

Show that the creation operators transform as indicated in the following:

$$Ra_+^\dagger(\mathbf{k})R^{-1} = e^{-i\varphi}a_+^\dagger(\mathbf{k}) \qquad \xi a_+^\dagger(\mathbf{k})\xi^{-1} = a_+^\dagger(-\mathbf{k})$$
$$Ra_+^\dagger(-\mathbf{k})R^{-1} = e^{i\varphi}a_+^\dagger(-\mathbf{k}) \qquad \xi a_+^\dagger(-\mathbf{k})\xi^{-1} = a_+^\dagger(\mathbf{k})$$
$$Ra_-^\dagger(\mathbf{k})R^{-1} = e^{i\varphi}a_-^\dagger(\mathbf{k}) \qquad \xi a_-^\dagger(\mathbf{k})\xi^{-1} = a_-^\dagger(-\mathbf{k})$$
$$Ra_-^\dagger(-\mathbf{k})R^{-1} = e^{-i\varphi}a_-^\dagger(-\mathbf{k}) \qquad \xi a_-^\dagger(-\mathbf{k})\xi^{-1} = a_-^\dagger(\mathbf{k})$$

5.5 Two-Photon States [5] We can obtain interesting information about a state of two photons by examining its behavior under rotations and reflection. Consider two photons with momenta $\mathbf{k}$ and $-\mathbf{k}$. There are four independent states of polarization, which can be classified according to circular polarizations:

$$|{+}{+}\rangle = a_+^\dagger(\mathbf{k})a_+^\dagger(-\mathbf{k})|0\rangle$$
$$|{+}{-}\rangle = a_+^\dagger(\mathbf{k})a_-^\dagger(-\mathbf{k})|0\rangle$$
$$|{-}{+}\rangle = a_-^\dagger(\mathbf{k})a_+^\dagger(-\mathbf{k})|0\rangle$$
$$|{-}{-}\rangle = a_-^\dagger(\mathbf{k})a_-^\dagger(-\mathbf{k})|0\rangle$$

(a) Verify that, in terms of states with linear polarization,

$$|{+}{+}\rangle + |{-}{-}\rangle = [a_1^\dagger(\mathbf{k})a_1^\dagger(-\mathbf{k}) - a_2^\dagger(\mathbf{k})a_2^\dagger(-\mathbf{k})]\,|0\rangle$$
$$|{+}{+}\rangle - |{-}{-}\rangle = [a_1^\dagger(\mathbf{k})a_2^\dagger(-\mathbf{k}) + a_2^\dagger(\mathbf{k})a_1^\dagger(-\mathbf{k})]\,|0\rangle$$
$$|{+}{-}\rangle = [a_1^\dagger(\mathbf{k})a_1^\dagger(-\mathbf{k}) + a_2^\dagger(\mathbf{k})a_2^\dagger(-\mathbf{k}) + ia_1^\dagger(\mathbf{k})a_2^\dagger(-\mathbf{k}) - ia_2^\dagger(\mathbf{k})a_1^\dagger(-\mathbf{k})]\,|0\rangle$$
$$|{-}{+}\rangle = [a_1^\dagger(\mathbf{k})a_1^\dagger(-\mathbf{k}) + a_2^\dagger(\mathbf{k})a_2^\dagger(-\mathbf{k}) - ia_1^\dagger(\mathbf{k})a_2^\dagger(-\mathbf{k}) + ia_2^\dagger(\mathbf{k})a_1^\dagger(-\mathbf{k})]\,|0\rangle$$

From this, note that the polarization of the two photons are correlated:

- In the state $|{+}{+}\rangle + |{-}{-}\rangle$ the planes are parallel.
- In the state $|{+}{+}\rangle - |{-}{-}\rangle$ the planes are orthogonal.
- In the states $|{+}{-}\rangle$ and $|{-}{+}\rangle$, the planes have equal probability of being parallel or orthogonal.

(b) Work out the transformation laws for the four polarization states under R, ξ, P, using results of the last problem, and the fact that the vacuum state is invariant. Verify the results summarized in the table of eigenvalues (listed whenever the state in question is an eigenstate of the operation indicated):

	$\|{+}{+}\rangle + \|{-}{-}\rangle$	$\|{+}{+}\rangle - \|{-}{-}\rangle$	$\|{+}{-}\rangle$	$\|{-}{+}\rangle$
R	1	1	$e^{-2i\varphi}$	$e^{2i\varphi}$
ξ	1	1		
P	1	−1	1	1

(c) From the preceding table, verify the following quantum numbers for a two-photon state:

- The only state with odd parity is $|{+}{+}\rangle - |{-}{-}\rangle$. There are three states with even parity: $|{+}{+}\rangle + |{-}{-}\rangle$, $|{+}{-}\rangle$, and $|{-}{+}\rangle$.
- For odd total angular momentum $J = 1,3,5,\ldots$, the only possible states are $|{+}{-}\rangle$ and $|{-}{+}\rangle$. The reason is as follows. The other two states are both eigenstates of R and ξ with eigenvalue 1. However, an initial state that is an eigenstate of R with eigenvalue 1 must have the rotation properties of the spherical harmonic Y_J^0, and therefore changes sign under ξ for $J = 1,3,5,\ldots$
- For total angular momentum $J = 0,1$, the only possible states are $|{+}{+}\rangle + |{-}{-}\rangle$ and $|{+}{+}\rangle - |{-}{-}\rangle$, because the other two states have spin projections ± 2 along the z axis, values that are too large for $J = 0,1$.

(d) Verify that a two-photon state cannot have $J = 1$.

This gives Yang's selection rule [6]: *A spin 1 particle cannot decay into two photons.* For example, just by observing that the π^0 meson decays into two photons, we can conclude that its spin cannot be 1. (It is, in fact, a spin 0 particle.)

5.6 Dirac Monopole A magnetic monopole has a magnetic field $\mathbf{B}_{\text{pole}} = g\hat{\mathbf{r}}/r^2$, with total magnetic flux $4\pi g$. Accommodate such a magnetic field into Maxwell's equations in the following manner. To keep $\nabla \cdot \mathbf{B} = 0$, postulate that there is a return flux $-4\pi g$ concentrated in an infinitely thin string attached to the monopole. The vector potential then consists of a part due to the monopole, and a part due to the string:

$$\mathbf{A} = \mathbf{A}_{\text{pole}} + \mathbf{A}_{\text{string}}$$

where $\mathbf{A}_{\text{pole}}$ is any vector potential that satisfies $\nabla \times \mathbf{A}_{\text{pole}} = \mathbf{B}_{\text{string}}$, and is, of course, determined only up to a gauge transformation.

(a) Give one solution for $\mathbf{A}_{\text{pole}}$.

(b) The shape of the string can be changed through a gauge transformation. For a straight-line string leading from the monopole to infinity, show that the vector potential of the string is of the pure-gauge form

$$\mathbf{A}_{\text{string}} = -2g\,\nabla\theta$$

where θ is the azimuthal angle around the string.

(c) Consider a quantum-mechanical particle of electric charge e in the field of the monopole, with wave function ψ. Show that the string can be transformed away through a gauge transformation

$$\psi \rightarrow e^{-2ige\theta}\psi$$

(d) Since ψ has to be single-value, the coefficient of θ in the exponent must be an integer n, and thus

$$ge = n/2$$

This is the Dirac quantization condition. The mere possibility that a monopole can exist quantizes the electric charge.

(e) Show that the total angular momentum of the system consisting of a charge e and a monopole g points from the charge to the monopole, and has the magnitude ge. Obtain the Dirac quantization condition by quantizing the angular momentum.

5.7 Cutoff Functions

(a) Calculate the vacuum energy (5.98) for a one-dimensional system using a sharp cutoff, which corresponds to $f(\omega) = \theta(\omega_c - \omega)$, and show

$$E(a) = \frac{L\omega_c^2}{4\pi} - \frac{3\omega_c}{2} \qquad \text{(sharp cutoff)}$$

Since this is independent of a, there will be no force between inserted walls.

(b) Show, on the other hand, that any continuous cutoff function will have a nonzero cutoff-independent force. To do this, write

$$E_0(L) = \frac{\pi}{2L} \sum_{n=1}^{\infty} n f\left(\frac{\pi b}{\omega_c L}\right)$$

Since the argument of f approaches a continuous variable in the limit $\omega_c \to \infty$, we can approximate the sum by an integral, using the Euler-MacLaurin formula (5.117):

$$E_0(L) = \frac{\pi}{2L} \int dn\, n f\frac{\pi n}{\omega_c L} - \frac{\pi B_2}{4L} + \cdots$$

$$= \frac{L\omega_c^2}{2\pi} \int dy\, y f(y) - \frac{\pi}{24L} + O(\omega_c^{-2})$$

The cutoff-independent term is the same as that in (5.101).

REFERENCES

1. K. Huang, *Quarks, Leptons, and Gauge Fields,* 2nd ed., World Scientific, Singapore, 1992, Chapter 8.
2. M. Abramowitz and I. A, Stegun, Eds., *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, DC, 1964, p. 806.
3. M. J. Sparney, *Physica* **24,** 751 (1958).
4. S. K. Lamoreaux, *Phys. Rev. Lett.* **78,** 5 (1997).
5. C. N. Yang, *Phys. Rev.* **77,** 242 (1950).
6. C. N. Yang, *Phys. Rev.* **77,** 242 (1950); L. D. Landau, *Dokl. Akad. Nauk* (USSR) **60,** 207 (1948).

CHAPTER SIX

Dirac Equation

6.1 DIRAC ALGEBRA

A relativistic wave equation must treat space and time on the same footing. The Klein–Gordon equation does that, but it involves second time derivatives, a feature responsible for its failure as a one-particle equation. Dirac tries to remedy this by proposing a first-order differential equation. To obtain a equation for the wave function ψ that is linear in the space–time derivatives $\partial_\mu \psi$, Dirac writes

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0 \tag{6.1}$$

where the γ^μ are numerical coefficients, so far undetermined. To satisfy the relativistic kinematics, $\psi(x)$ must also satisfy the Klein–Gordon equation. Multiplying from the left by $(i\gamma^\mu \partial_\mu + m)$, we have

$$\begin{aligned} 0 &= (i\gamma^\mu \partial_\mu + m)(i\gamma^\mu \partial_\mu - m)\psi(x) \\ &= -(\gamma^\mu \gamma^n \partial_\mu \partial_\nu + m^2)\psi(x) \\ &= -[\tfrac{1}{2}(\gamma^\mu \gamma^n + \gamma^n \gamma^\mu)\partial_\mu \partial_\nu + m^2]\ \psi(x) \end{aligned} \tag{6.2}$$

This reduces to the Klein–Gordon equation

$$(\Box^2 + m^2)\psi(x) = 0 \tag{6.3}$$

if and only if

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \tag{6.4}$$

This algebraic relation defines four objects γ^μ, which anticommute with one another, with

$$(\gamma^0)^2 = 1$$
$$(\gamma^k)^2 = -1 \tag{6.5}$$

Clearly γ^μ cannot be numbers. They can be represented by matrices, called *Dirac matrices*.

According to (6.4) $\gamma^\mu\gamma^\nu$ should be a Hermitian matrix. Thus γ^μ is either Hermitian or anti-Hermitian. Putting $\mu \neq \nu$ and taking the trace of both sides in (6.4), we obtain

$$\text{Tr } \gamma^\mu = 0 \qquad (\mu = 0,1,2,3) \tag{6.6}$$

This condition immediately rules out matrices of odd dimension. It also rules out dimension 2, for there are only three independent traceless 2×2 matrices—the Pauli matrices. Therefore, the dimension must be at least 4. That a 4×4 representation exists can be shown by explicit construction.

Define the following 4×4 Hermitian matrices:

$$\alpha^k = \begin{pmatrix} 0 & \sigma^k \\ \sigma^k & 0 \end{pmatrix} \qquad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{6.7}$$

where 1 stands for the 2×2 unit matrix, and σ^k are the 2×2 Pauli matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{6.8}$$

We shall not use different notations for 2×2 and 4×4 matrices, since the context usually makes the meaning clear. It follows from the definitions that

$$(\alpha^k)^2 = \beta^2 = 1$$
$$\{\alpha^k, \beta\} = 0 \tag{6.9}$$

A standard representation for the Dirac matrices is

$$\gamma^0 = \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{6.10}$$

$$\gamma^k = \gamma^0\alpha^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} \tag{6.11}$$

The matrix γ^0 is Hermitian with $(\gamma^0)^2 = 1$, and γ^k is anti-Hermitian, with $(\gamma^k)^2 = -1$:

$$\gamma^0 = (\gamma^0)^\dagger \qquad (\gamma^0)^2 = 1$$

$$\gamma^k = -(\gamma^k)^\dagger \quad (\gamma^k)^2 = 1 \tag{6.12}$$

From these we can show

$$\gamma^0(\gamma^\mu)^\dagger\gamma^0 = \gamma^\mu \tag{6.13}$$

The representation given here is not unique. A unitary transformation $S\gamma^\mu S^{-1}$ gives an equally acceptable set of matrices, since such a transformation obviously preserves (6.4).

The γ^μ and their products, together with the unit matrix, generate a set of 16 independent 4×4 matrices, in terms of which any 4×4 matrix can be expanded. We introduce special symbols for some of their products:

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$$

$$\sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) \tag{6.14}$$

The "fifth" Dirac matrix γ_5 is Hermitian, with square 1, and anticommutes with all four γ^μ:

$$(\gamma_5)^\dagger = \gamma_5$$
$$(\gamma_5)^2 = 1$$
$$\{\gamma_5, \gamma^\mu\} = 0 \tag{6.15}$$

In our standard representation it has the form

$$\gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{6.16}$$

The generalized Pauli matrices $\sigma^{\mu\nu} = -\sigma^{\mu\nu}$ have six independent members:

$$\sigma^{0k} = i\alpha^k$$
$$\sigma^{ij} = \epsilon^{ijk}\sigma^k \tag{6.17}$$

where σ^k denotes the matrix of 2×2 blocks made up of Pauli matrices along the diagonal. It is straightforward to show that

$$\sigma^k = \gamma_5\alpha^k$$
$$\gamma^0\sigma^{\mu\nu\dagger}\gamma^0 = \sigma^{\mu\nu} \tag{6.18}$$

A complete set of 16 independent 4×4 matrices Γ_n is given in Table 6.1. By definition, we take $\Gamma_0 \equiv 1$. All the Γ_n are traceless except for Γ_0:

TABLE 6.1 Matrices of Dirac Algebra

Γ_n	*Number*
1	1
γ^μ	4
$\gamma_5\gamma^\mu$	4
$\sigma^{\mu\nu}$	6
γ_5	1
Total	16

$$\mathrm{Tr}\Gamma_n = 0 \qquad (n \neq 0) \tag{6.19}$$

The set is closed under multiplication and commutation, and is called *Dirac algebra*. The commutators are given in Table 6.2.

An arbitrary 4×4 matrix M can be expanded in the form

$$M = \frac{15}{n=0} c_n \Gamma_n \tag{6.20}$$

where

$$c_0 = \frac{1}{4}\mathrm{Tr}\, M$$

$$c_n = \frac{\mathrm{Tr}(M\Gamma_n)}{\mathrm{Tr}(\Gamma_n)^2} \tag{6.21}$$

TABLE 6.2 Commutators of Dirac Algebra

$$[\gamma^\lambda, \gamma^\mu] = 2\gamma^\lambda\gamma^\mu - 2g^{\lambda\mu}$$
$$[\gamma_5, \gamma^\mu] = 2\gamma_5\gamma^\mu$$
$$[\gamma_5\gamma^\lambda, \gamma^\mu] = 0$$
$$[\gamma_5\gamma^\lambda, \gamma_5] = 2\gamma^\lambda$$
$$[\sigma^{\lambda\rho}, \gamma^\mu] = 2i(\gamma^\lambda g^{\rho\mu} - \gamma^\rho g^{\lambda\mu})$$
$$[\sigma^{\lambda\rho}, \gamma_5] = 0$$
$$[\sigma^{\lambda\rho}, \gamma_5\gamma^\mu] = 0$$
$$[\sigma^{\lambda\rho}, \sigma^{\mu\nu}] = 2i(\sigma^{\lambda\mu}g^{\rho\nu} - \sigma^{\lambda\nu}g^{\rho\mu} + \sigma^{\rho\nu}g^{\lambda\mu} - \sigma^{\rho\mu}g^{\lambda\nu})$$

6.2 WAVE FUNCTIONS AND CURRENT DENSITY

Rewriting the Dirac equation in a 3-vector notation, we have

$$(-i\boldsymbol{\alpha}\cdot\nabla+\beta m)\,\psi(x)=i\frac{\partial\psi(x)}{\partial t} \tag{6.22}$$

This looks like a single-particle wave equation with Hamiltonian

$$H=\boldsymbol{\alpha}\cdot\mathbf{p}+\beta m \tag{6.23}$$

where $\mathbf{p}$ is the momentum operator. The wave function $\psi(x)$ is a four-component column vector called a *Dirac spinor:*

$$\psi=\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix} \tag{6.24}$$

where ψ_i are complex numbers. The complex conjugate is the column vector

$$\psi^*=\begin{pmatrix}\psi_1^*\\ \psi_2^*\\ \psi_3^*\\ \psi_4^*\end{pmatrix} \tag{6.25}$$

and there are other types of conjugates:

$$\begin{aligned}\text{Hermitian adjoint:}\quad & \psi^\dagger=(\psi_1^* \quad \psi_2^* \quad \psi_3^* \quad \psi_4^*)\\ \text{Pauli adjoint:}\quad & \overline{\psi}=\psi^*\gamma^0=(\psi_1^* \quad \psi_2^* \quad -\psi_3^* \quad -\psi_4^*)\end{aligned} \tag{6.26}$$

The Hermitian conjugate of (6.1) reads

$$-i[\partial_\mu\psi^\dagger(x)]\gamma^{\mu\dagger}-m\psi^\dagger(x)=0 \tag{6.27}$$

Now write $\psi^\dagger=\overline{\psi}\gamma^0$, and use $\gamma^0\gamma^{\mu\dagger}\gamma^0=\gamma^\mu$ to obtain the equation for the Pauli adjoint:

$$(\partial_\mu\overline{\psi})i\gamma^\mu+m\overline{\psi}=0 \tag{6.28}$$

Another way of writing this is

$$\overline{\psi}(i\gamma^\mu\overleftarrow{\partial}_\mu+m)=0 \tag{6.29}$$

where the overhead arrow on ∂_μ indicates that it acts to the left.

The conserved density current is given by

$$j^\mu = \overline{\psi}\gamma^\mu\psi \tag{6.30}$$

It is easy to see, with the help of (6.1) and (6.29), that

$$\begin{aligned}\partial_\mu j^\mu &= (\partial_\mu\psi)\gamma^\mu\psi + \overline{\psi}\gamma^\mu\partial_\mu\psi \\ &= -m\overline{\psi}\psi + m\overline{\psi}\psi = 0\end{aligned} \tag{6.31}$$

Note that j^0 is positive-definite:

$$j^0 = \psi^\dagger\psi = \psi_1^*\psi_1 + \psi_2^*\psi_2 + \psi_3^*\psi_3 + \psi_4^*\psi_4 \tag{6.32}$$

As opposed to

$$\overline{\psi}\psi = \psi_1^*\psi_1 + \psi_2^*\psi_2 - \psi_3^*\psi_3 - \psi_4^*\psi_4 \tag{6.33}$$

The current j^μ can therefore serve as a particle current density. As we shall see, however, the Dirac equation fails to qualify as a single-particle equation for a different reason; namely, the energy spectrum is not bounded from below. As we shall discuss in Section 6.9, the remedy is a redefinition of the vacuum state known as "hole theory," which makes the system a many-particle system. With this modification, j^0 will become an operator, whose expectation values are no longer positive-definite, but can be interpreted as charge density.

6.3 PLANE WAVES

Plane-wave solutions to the Dirac equation can be constructed by putting

$$\psi(x) = e^{-ip\cdot x}u(\mathbf{p}) \tag{6.34}$$

where $p^\mu = (p^0, \mathbf{p})$, and $u(\mathbf{p})$ is a column vector called a *Dirac spinor:*

$$u(\mathbf{p}) = \begin{pmatrix} u_1(\mathbf{p}) \\ u_2(\mathbf{p}) \\ u_3(\mathbf{p}) \\ u_4(\mathbf{p}) \end{pmatrix} \tag{6.35}$$

Since $\psi(x)$ satisfies the Klein–Gordon equation, we have

$$p_0^2 - \mathbf{p}^2 - m^2 = 0$$

For given momentum, there are two roots for the energy p^0, with opposite signs:

$$p_0 = \pm E \tag{6.36}$$

where E is defined as the positive quantity

$$E = + \sqrt{\mathbf{p}^2 + m^2} \tag{6.37}$$

The Dirac equation now takes the form

$$(\not{p} - m)u(\mathbf{p}) = 0 \tag{6.38}$$

where $\not{p}$ is a 4×4 matrix defined by

$$\not{p} = \gamma^\mu p_\mu = \gamma^0 p^0 - \gamma^k p^k \tag{6.39}$$

It has the property

$$\not{p}\not{q} + \not{p}\not{q} = 2p \cdot q \tag{6.40}$$

which follows from (6.4).

To find explicit solutions, we note that

$$(\not{p} - m)(\not{p} + m) = p^2 - m^2 = 0 \tag{6.41}$$

Thus, each column of the matrix $(\not{p} + m)$ satisfies the Dirac equation. The explicit form of the matrix is

$$\not{p} + m = \begin{pmatrix} m + p^0 & 0 & -p^3 & -p_- \\ 0 & m + p^0 & -p_+ & p^3 \\ p^3 & p_- & m - p^0 & 0 \\ p_+ & -p^3 & 0 & m - p^0 \end{pmatrix} \tag{6.42}$$

where

$$p_\pm = p^1 \pm p^2 \tag{6.43}$$

The number of independent columns can be found by letting $p^k \to 0$, since the matrix is a continuous function of p^k. In that limit $p^0 = \pm m$, and the matrix becomes proportional to

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{for } p^0 > 0 \qquad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{for } p^0 < 0$$

This shows that columns 1 and 2 are independent for $p^0 > 0$, while columns 3 and 4 are independent for $p^0 < 0$. The independent solutions are then columns 1 and 2 of (6.42) for $p^0 = E$, and columns 3 and 4 for $p^0 = -E$. We designate them as $u(\mathbf{p}, s)$. The explicit solutions for $p^0 = E$ are

$$u(\mathbf{p}, 1) = C\begin{pmatrix} 1 \\ 0 \\ \dfrac{p^3}{m+E} \\ \dfrac{p_+}{m+E} \end{pmatrix} \qquad u(\mathbf{p}, 2) = C\begin{pmatrix} 0 \\ 1 \\ \dfrac{p_-}{m+E} \\ \dfrac{-p^3}{m+E} \end{pmatrix} \tag{6.44}$$

The solutions for $p^0 = -E$ are

$$u(\mathbf{p}, 3) = C\begin{pmatrix} \dfrac{-p^3}{m+E} \\ \dfrac{-p_+}{m+E} \\ 1 \\ 0 \end{pmatrix} \qquad u(\mathbf{p}, 4) = C\begin{pmatrix} \dfrac{-p_-}{m+E} \\ \dfrac{p^3}{m+E} \\ 0 \\ 1 \end{pmatrix} \tag{6.45}$$

where

$$E = +\sqrt{\mathbf{p}^2 + m^2}$$

$$C = \sqrt{\frac{m+E}{2m}} \tag{6.46}$$

For a given $\mathbf{p}$, these solutions form an orthogonal set:

$$u^\dagger(\mathbf{p}, s)u(\mathbf{p}, s') = \frac{E}{m}\delta_{ss'} \tag{6.47}$$

For a given energy, the wave functions above resemble those of a nonrelativistic particle of spin $\frac{1}{2}$, and it is natural to regard s as a spin label. We shall see that this is a correct interpretation.

Taking the Hermitian conjugate of (6.38), we have

$$u^\dagger(\mathbf{p}, s)(\not{p}^\dagger - m) = 0 \tag{6.48}$$

Multiplying the equation from the right by γ^0, and using the identity

$$\gamma^0 \not{p}^\dagger \gamma^0 = \not{p} \tag{6.49}$$

we find the Pauli-adjoint equation

$$\bar{u}(\mathbf{p}, s)(\not{p} - m) = 0 \tag{6.50}$$

Multiplying (6.48) from the right by u, and (6.38) from the left by $u^\dagger$, and writing out $\not{p}$ more explicitly, we have

$$\begin{aligned} u^\dagger(\gamma^0 p^0 - \gamma^k p_k - m)u = 0 \\ u^\dagger(\gamma^0 p^0 + \gamma^k p_k - m)u = 0 \end{aligned} \tag{6.51}$$

Adding the two equations leads to the relation

$$u^\dagger u = \frac{p^0}{m}\bar{u}u \tag{6.52}$$

We can restate the orthonormality of the solutions in the form

$$\bar{u}(\mathbf{p}, s)u(\mathbf{p}, s') = \pm\, \delta_{ss'} \tag{6.53}$$

where the plus sign applies for the positive-energy solutions, corresponding to s = 1,2, and the minus sign is used for the negative-energy solution with s = 3,4.

6.4 LORENTZ TRANSFORMATIONS

Under a Lorentz transformation $x' = \Lambda x$, the Dirac equation in the new frame reads

$$(i\gamma^\mu \partial'_\mu - m)\psi'(x') = 0 \tag{6.54}$$

Note that γ^μ remains unchanged, because it is just a numerical matrix. We relate the new wave function to the old through a linear unitary transformation:

$$\psi'(x') = S\psi(x) \qquad S^\dagger S = 1 \tag{6.55}$$

where S is a nonsingular 4×4 unitary matrix. To demonstrate Lorentz covariance, we shall show that there exists a nonsingular transformation on γ^μ that will restore the Dirac equation to the old form. Putting $\partial'_\mu = \Lambda^\nu_\mu \partial_\nu$ and multiplying the equation by S^{-1} from the left, we obtain

$$S^{-1}(i\gamma^\mu \Lambda^\nu_\mu \partial_\nu - m)S\psi(x) = 0$$

which reduces to the original equation if $(S^{-1}\gamma^\mu S)\Lambda^\nu_\mu = \gamma^\nu$ or

$$S^{-1}\gamma^\mu S = \Lambda^\mu_\nu \gamma^\nu \tag{6.56}$$

The existence of S will be demonstrated by explicit construction.

It suffices to consider an infinitesimal Lorentz transformation

$$x^{\mu\prime} = x^\mu + \omega^\mu_\nu x^\nu \tag{6.57}$$

where ω^μ_ν contains six infinitesimal parameters, the three rotations θ^k and the three boosts v^k of the coordinate frame:

$$\begin{aligned} \omega^{0k} &= \omega^{k0} = v^k \\ \omega^{ij} &= -\epsilon^{ijk}\theta^k \end{aligned} \tag{6.58}$$

We put

$$S = 1 + iR \tag{6.59}$$

where R is an infinitesimal Hermitian matrix linear in ω^μ_ν. To first order in ω^μ_ν, the condition for covariance is

$$\omega^\nu_\mu \gamma^\mu = i[\gamma^\nu, R] \tag{6.60}$$

We can write R as a linear combination of the Γ_n of Table 6.1, whose commutators are listed in Table 6.2. As we can see, only $\sigma^{\lambda\rho}$ can contribute to R. Thus R must have the form $R = C\omega_{\mu\nu}\sigma^{\mu\nu}$, and a short calculation determines $C = -\frac{1}{4}$. Therefore

$$S = 1 - \frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu} \tag{6.61}$$

This demonstrates the Lorentz covariance of the Dirac equation. Comparison with (3.42) shows that the spin operator is

$$\Sigma^{\mu\nu} = -\frac{i}{2}\sigma^{\mu\nu} \tag{6.62}$$

Writing S in terms of the parameters of the infinitesimal Lorentz transformation, we have

$$S = 1 - \frac{1}{2}\mathbf{v}\cdot\boldsymbol{\alpha} + \frac{i}{2}\boldsymbol{\theta}\cdot\boldsymbol{\sigma} \tag{6.63}$$

A finite rotation about a given axis can be built up from successive infinitesimal rotations about the same axis:

$$S = e^{i\boldsymbol{\theta}\cdot\boldsymbol{\sigma}/2} = \cos\frac{\theta}{2} + i\hat{\boldsymbol{\theta}}\cdot\boldsymbol{\sigma}\sin\frac{\theta}{2} \tag{6.64}$$

where $\boldsymbol{\theta}$ is a vector whose direction is the axis of rotation, and whose magnitude is the angle of rotation. Similarly, for a Lorentz boost of the reference frame with finite velocity v corresponds to

$$S = e^{-\boldsymbol{\phi}\cdot\boldsymbol{\alpha}/2} = \cosh\frac{\phi}{2} - \hat{\boldsymbol{\phi}}\cdot\boldsymbol{\alpha}\sinh\frac{\phi}{2} \tag{6.65}$$

where

$$\boldsymbol{\phi} = \hat{\mathbf{v}}\tanh^{-1}v \tag{6.66}$$

The advantage of using $\boldsymbol{\phi}$ is that, unlike v, it is additive for successive boosts. We note that S is unitary for rotations, but not for Lorentz boosts.

The Dirac equation realizes a finite-dimensional representation of the Lorentz group. As we have seen in Chapter 3, the smallest faithful representation of the Lorentz group is of dimension 2. Here the dimension is doubled in order to represent spatial reflections.

Consider the total reflection of the spatial coordinate system

$$\mathbf{r}' = -\mathbf{r} \qquad t' = t \tag{6.67}$$

In the transformed frame, the Dirac equation reads

$$\left(i\gamma^0\frac{\partial}{\partial t'} + i\gamma^k\frac{\partial}{\partial x'^k} - m\right)\psi'(x') = 0$$

Putting $\psi'(x') = S\psi(x)$, and multiplying the equation from the left by S^{-1}, we have

$$\left[i(S^{-1}\gamma^0 S)\frac{\partial}{\partial t} - i(S^{-1}\gamma^k S)\frac{\partial}{\partial x^k} - m\right]\psi(x) = 0$$

which exhibits covariance if there exist S that commutes with γ^0, and anticommutes with γ^k, with $S^2 = 1$. An obvious choice is

$$S = \gamma^0 \qquad \text{(spatial reflection)} \tag{6.68}$$

which shows the necessity for 4×4 matrices.

For total space–time reflection $x' = -x$, we have

$$(-i\gamma^\mu\partial_\mu - m)\psi'(-x) = 0$$

with complex conjugate

$$(i\gamma^{*\mu}\partial_\mu - m)\psi'^*(-x) = 0$$

Accordingly, we put

$$\psi'(-x) = S\psi^*(x)$$

To restore the original equation, we seek S that anticommutes with γ^μ, and with $S^2 = 1$, and an obvious solution is

$$S = \gamma_5 \qquad \text{(space–time reflection)} \tag{6.69}$$

For time reversal, which is the product of space–time reflection with spatial reflection, we have

$$S = \gamma^0\gamma_5 \qquad \text{(time reversal)} \tag{6.70}$$

This is an algebraic transformation that preserves the form of the Dirac equation when t is replaced by $-t$; but it is not the operation that governs physical states, which must be taken as states in quantum field theory. In the next chapter, we shall see that physical time reversal must involve complex conjugation of the state.

For a plane-wave state we have

$$\begin{aligned} \psi(x) &= e^{-ip\cdot x}u(\mathbf{p}, s) \\ \psi'(x') &= e^{-ip'\cdot x'}u'(\mathbf{p}', s') \end{aligned} \tag{6.71}$$

where p' is the 4-momentum with respect to the new frame and s' labels the new solutions. Since $p\cdot x$ is invariant, we have

$$\begin{aligned} u'(\mathbf{p}', s') &= Su(\mathbf{p}, s) \\ \bar{u}'(\mathbf{p}', s') &= \bar{u}(\mathbf{p}, s)S^{-1} \end{aligned} \tag{6.72}$$

In general, we can reshuffle the four solutions in the new frame; but since a Lorentz transformation preserves the sign of the energy, the mixing of solutions can occur only among s = 1,2, and separately among s = 3,4. With this freedom understood, we set $s' = s$ by convention. It is straightforward to show (Problem 6.2) that

$$\gamma^0 u(-\mathbf{p}, s) = \begin{cases} u(-\mathbf{p}, s) & (s = 1,2) \\ -u(\mathbf{p}, s) & (s = 3,4) \end{cases} \tag{6.73}$$

which indicates that positive-energy and negative-energy states have opposite parity.

We see from (6.72) that $\bar{u}u$ is invariant under a Lorentz transformation, and $\bar{u}\gamma^\mu u$ transforms like a 4-vector. More generally, the transformation properties of $\bar{u}\Gamma^\mu u$ are listed in Table 6.3.

TABLE 6.3 Transformation Properties

$\bar{u}u$	Scalar
$\bar{u}\gamma^\mu u$	Vector
$\bar{u}\gamma_5\gamma^\mu u$	Pseudovector
$\bar{u}\sigma^{\mu\nu}u$	Tensor
$\bar{u}\gamma_5 u$	Pseudoscalar

6.5 INTERPRETATION OF DIRAC MATRICES

The Dirac matrices α^k are velocity components in the sense

$$\frac{dx^k}{dt} = i[H, x^k] = \alpha^k \tag{6.74}$$

Individually, each component has eigenvalues ± 1, but the components do not commute with one another. This conjures up the picture that the electron performs a curious dance at the speed of light about its average motion called the *zitterbewegung*. But this motion is not directly observable. (See Problem 6.5.) Expectation values of the Dirac matrices, on the other hand, have physical significance.

Let $|E\rangle$ be a normalized energy eigenstate:

$$\begin{aligned} H|E\rangle &= E|E\rangle \\ \langle E|E\rangle &= 1 \end{aligned} \tag{6.75}$$

We have the trivial identity

$$\langle E|(HO - OH)|E\rangle = 0 \tag{6.76}$$

which can be rewritten in the form

$$\langle E|(HO + OH - 2OH)|E\rangle = 0 \tag{6.77}$$

This gives

$$\langle O\rangle = \frac{1}{2E}\langle\{H, O\}\rangle \tag{6.78}$$

where $\langle O\rangle \equiv \langle E|O|E\rangle/\langle E|E\rangle$. Choosing for O the Dirac matrix β, we find

$$\langle\beta\rangle = \frac{m}{E} = \sqrt{1 - v^2} \tag{6.79}$$

As we have seen, $\boldsymbol{\alpha}$ is the velocity. Hence the magnetic moment is

$$\boldsymbol{\mu} = \frac{e}{2}\mathbf{r} \times \boldsymbol{\alpha} \tag{6.80}$$

where e is the electric charge. Now put $O = \boldsymbol{\mu}$ in (6.78). A straightforward calculation gives

$$\langle \boldsymbol{\mu} \rangle = \frac{e}{2E}\langle \mathbf{L} + 2\mathbf{S} \rangle \tag{6.81}$$

where $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is the orbital angular momentum and $\mathbf{S} = \boldsymbol{\sigma}/2$ is the spin operator. This shows that the gyromagnetic ratio associated with the intrinsic spin is 2, but that associated with orbital motion, as in spin–orbit coupling, is 1. We shall verify these properties, and give physical interpretations.

6.6 EXTERNAL ELECTROMAGNETIC FIELD

The Dirac equation in the presence of an external electromagnetic field $A^{\mu}(x)$ is

$$[i\gamma^{\mu}D_{\mu} - m]\,\psi(x) = 0 \tag{6.82}$$

with

$$D_{\mu} = \partial_{\mu} + ieA_{\mu}(x) \tag{6.83}$$

The Hamiltonian has the form

$$H = \boldsymbol{\alpha} \cdot \boldsymbol{\pi} + \beta m + e\phi \tag{6.84}$$

where $\phi(x) = A^{0}(x)$ and $\boldsymbol{\pi}$ is the kinetic momentum

$$\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}(x) \tag{6.85}$$

where $\mathbf{p}$ is the one-particle momentum operator defined by $[p^{j}, x^{k}] = -i\delta_{jk}$.

It is straightforward to verify the equations

$$\frac{d\boldsymbol{\pi}}{dt} = e(\mathbf{E} + \boldsymbol{\alpha} \times \mathbf{B})$$

$$\frac{d\boldsymbol{\sigma}}{dt} = 2\boldsymbol{\pi} \times \boldsymbol{\alpha} \tag{6.86}$$

where $\mathbf{E}$ and $\mathbf{B}$ are respectively the external electric and magnetic fields. The first

equation is the analog of the Lorentz-force equation, and the second describes spin precession. Taking the dot product of the second equation with $\boldsymbol{\pi}$, we have

$$\boldsymbol{\pi} \cdot \frac{d\boldsymbol{\sigma}}{dt} = 2\boldsymbol{\pi} \cdot \boldsymbol{\pi} \times \boldsymbol{\alpha} = 2\boldsymbol{\pi} \times \boldsymbol{\pi} \cdot \boldsymbol{\alpha} \tag{6.87}$$

Noting that

$$\boldsymbol{\pi} \times \boldsymbol{\pi} = -e(\mathbf{A} \times \mathbf{p} + \mathbf{p} \times \mathbf{A}) = -ie\mathbf{B} \tag{6.88}$$

we can write

$$\boldsymbol{\pi} \cdot \frac{d\boldsymbol{\sigma}}{dt} = -2ie\boldsymbol{\alpha} \cdot \mathbf{B} \tag{6.89}$$

Combining this with the equation for $d\boldsymbol{\pi}/dt$, we obtain

$$\frac{d(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})}{dt} = e\boldsymbol{\sigma} \cdot \mathbf{E} \tag{6.90}$$

In arriving at this equation, we used the identity $\boldsymbol{\sigma} \times \boldsymbol{\alpha} = 2i\boldsymbol{\alpha}$, which can be obtained from $\boldsymbol{\sigma} \times \boldsymbol{\sigma} = 2i\boldsymbol{\sigma}$ by observing that $\boldsymbol{\alpha} = \gamma_5\boldsymbol{\sigma}$.

The equations (6.89) and (6.90) have interesting physical consequences [1]. For $\mathbf{E} = 0$, (6.90) states

$$\frac{d(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})}{dt} = 0 \qquad \text{(in pure magnetic field)} \tag{6.91}$$

That is, the spin projection along the velocity, is a constant of the motion in a pure magnetic field. This means that the precessional frequency of the spin is exactly equal to the orbital frequency. Deviations from this law measure corrections due to quantum field effects.

For $\mathbf{B} = \mathbf{0}$, (6.89) states

$$\boldsymbol{\pi} \cdot \frac{d\boldsymbol{\sigma}}{dt} = 0 \qquad \text{(in pure electric field)} \tag{6.92}$$

Suppose that a longitudinally polarized particle moves along the z axis, loses energy, and stops. This equation says $d\sigma_z/dt = 0$, or that the particle is not depolarized. This result is crucial to the interpretation of the classic experiment [2] that established parity violation in the decay process $\pi^- \rightarrow \mu^- + \bar{\nu}$. The μ mesons were observed after being stopped in matter, and found to be longitudinally polarized. The preceding result shows that they had the same longitudinal polarization at the moment of decay.

6.7 NONRELATIVISTIC LIMIT

We shall study the nonrelativistic limit of (6.82), by first putting it in second-order form. Multiplying it from the left by $i\gamma^\nu D_\nu + m$, we obtain

$$(\gamma^\nu\gamma^\mu D_\nu D_\mu + m^2)\psi = 0 \tag{6.93}$$

Writing the first term as half the symmetric part plus the antisymmetric part with respect to the labels μ and ν, we can show

$$\gamma^\nu\gamma^\mu D_\nu D_\mu = g^{\nu\mu} D_\nu D_\mu + \tfrac{1}{2}\gamma^\mu\gamma^\nu[D_\mu, D_\nu] \tag{6.94}$$

A straightforward calculation gives

$$[D_\mu, D_\nu] = ieF_{\mu\nu} \tag{6.95}$$

We thus arrive at the second-order equation

$$\left(D^\mu D_\mu + \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu} + m^2\right)\psi = 0 \tag{6.96}$$

Consider a stationary solution of energy E, with $\partial\psi/\partial t = -iE\psi$. We can rewrite the equation in the form

$$[(\mathbf{p} - e\mathbf{A})^2 - e\boldsymbol{\sigma}\cdot\mathbf{B} + ie\boldsymbol{\alpha}\cdot\mathbf{E} - m^2]\psi = (E - e\phi)^2\psi \tag{6.97}$$

where we have used the relation

$$\tfrac{1}{2}\sigma^{\mu\nu}F_{\mu\nu} = -\boldsymbol{\sigma}\cdot B + i\boldsymbol{\alpha}\cdot\mathbf{E} \tag{6.98}$$

The equation displays a magnetic-moment term $\boldsymbol{\sigma}\cdot\mathbf{B}$, with electric-moment term $\boldsymbol{\alpha}\cdot\mathbf{E}$ generated by the moving magnetic moment.

In the nonrelativistic limit the components ψ_3 and ψ_4 are small, and it is convenient to rewrite the above in two-component form by putting

$$\psi = \begin{pmatrix}\chi\\ \xi\end{pmatrix} \tag{6.99}$$

where χ and ξ are two-component column vectors. Substituting this into the Dirac equation (6.82), we obtain the coupled equations

$$\begin{aligned}(E - e\phi - m)\chi - \boldsymbol{\sigma}\cdot(\mathbf{p} - e\mathbf{A})\xi = 0\\ (E - e\phi + m)\xi - \boldsymbol{\sigma}\cdot\mathbf{p} - e\mathbf{A})\chi = 0\end{aligned} \tag{6.100}$$

Solving for the "small" component ξ, we have

$$\xi = \frac{\boldsymbol{\sigma} \cdot (\mathbf{p} - e\mathbf{A})\chi}{E - e\phi + m} \tag{6.101}$$

which shows that it is of order $|\mathbf{p}|/E$ compared to the "large" component χ.

The second-order equation can be rewritten in the block form

$$\begin{pmatrix} \boldsymbol{\pi}^2 - e\boldsymbol{\sigma} \cdot \mathbf{B} + m^2 & ie\boldsymbol{\sigma} \cdot \mathbf{E} \\ ie\boldsymbol{\sigma} \cdot \mathbf{E} & ie\boldsymbol{\sigma} \cdot \mathbf{E} \end{pmatrix} \begin{pmatrix} \chi \\ \xi \end{pmatrix} = (E - e\phi)^2 \begin{pmatrix} \chi \\ \xi \end{pmatrix} \tag{6.102}$$

where $\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$. We write the equation for χ, and eliminate ς with the help of (6.101):

$$\left[\boldsymbol{\pi}^2 - e\boldsymbol{\sigma} \cdot \mathbf{B} + m^2 + ie\frac{(\boldsymbol{\sigma} \cdot \mathbf{E})(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})}{E - e\phi + m} \right] \chi = (E - e\phi)^2 \chi \tag{6.103}$$

which is an exact equation. We go to the nonrelativistic limit by putting

$$E = m + \epsilon \tag{6.104}$$

and assume $\epsilon \ll m$ and $e\phi \ll m$. Keeping only terms to first-order in ϵ and $e\phi$, we obtain

$$\left[\frac{1}{2m} \boldsymbol{\pi}^2 + e\phi - \frac{e}{2m} \boldsymbol{\sigma} \cdot \mathbf{B} - \frac{e}{4m^2} \boldsymbol{\sigma} \cdot \mathbf{E} \times \boldsymbol{\pi} + \frac{ie}{4m^2} \mathbf{E} \cdot \boldsymbol{\pi} \right] \chi = \epsilon \chi \tag{6.105}$$

This has the form of a Schrödinger equation except for the non-Hermitian $i\mathbf{E} \cdot \boldsymbol{\pi}$ term, which reflects the fact that $\chi^\dagger \chi$ is not conserved, due to the existence of negative-energy states. A one-particle interpretation is consistent only when this term can be neglected.

The magnetic-moment term has the form

$$-\frac{e}{2m} \boldsymbol{\sigma} \cdot \mathbf{B} = -\frac{e}{m} \mathbf{S} \cdot \mathbf{B} \tag{6.106}$$

where $\mathbf{S} = \boldsymbol{\sigma}/2$ is the spin operator. The coefficient of the magnetic field gives the magnetic moment:

$$\boldsymbol{\mu} = \frac{e}{2m} g\mathbf{S} \tag{6.107}$$

where

$$g = 2 \tag{6.108}$$

is the gyromagnetic ratio. There is a small correction to this value, known as the *anomalous magnetic moment,* due to vacuum fluctuations of the quantum fields. We shall calculate the latter in Section 12.3.

The spin–orbit interaction is contained in the term

$$H_1 = -\frac{e}{4m^2}\boldsymbol{\sigma}\cdot\mathbf{E}\times\boldsymbol{\pi} \tag{6.109}$$

In a central electrostatic field, with $\mathbf{A} = 0$ and

$$\mathbf{E} = -\hat{\mathbf{r}}\phi'(r) \tag{6.110}$$

we can write

$$H_1 = \frac{e}{4m^2}\frac{\phi'}{r}\boldsymbol{\sigma}\cdot\mathbf{r}\times\mathbf{p} = \frac{e}{2m^2}\frac{\phi'}{r}\mathbf{S}\cdot\mathbf{L} \tag{6.111}$$

where $\mathbf{L} = \mathbf{r}\times\mathbf{p}$. Viewed from a frame moving with the electron, this can be interpreted as the energy $-\boldsymbol{\mu}'\cdot\mathbf{B}'$ of a magnetic moment $\boldsymbol{\mu}'$ in the magnetic field

$$\mathbf{B}' = \mathbf{E}\times\mathbf{v} = -\mathbf{r}\times\mathbf{v}\frac{\phi'}{r} \tag{6.112}$$

where $\mathbf{v} = \mathbf{p}/m$. This gives

$$\boldsymbol{\mu}' = g'\frac{e}{2m}\mathbf{S} \tag{6.113}$$

with

$$g' = 1 \tag{6.114}$$

which is in agreement with experiments on the fine structure of atomic spectral lines.

6.8 THOMAS PRECESSION

The electron exhibits two different gyromagnetic ratios: $g = 2$ with respect to an external magnetic field, and $g' = 1$ with respect to a magnetic field generated by its orbital motion in a central electrostatic potential. To understand these results, consider first an electron in an external magnetic field B. It moves in a circular orbit with the cyclotron frequency

$$\omega_{\text{cyclotron}} = \frac{e}{m}B \tag{6.115}$$

The spin precesses about the magnetic field according to the equation $d\boldsymbol{\mu}/dt = \boldsymbol{\mu} \times \mathbf{B}$. With $\boldsymbol{\mu} = g\mathbf{S}$, this gives a precession frequency

$$\omega_{\text{precession}} = \frac{e}{2m}gB \tag{6.116}$$

Thus

$$\omega_{\text{precession}} - \omega_{\text{cyclotron}} = \frac{e}{2m}(g-2)B \tag{6.117}$$

If $g = 2$, as implied by the Dirac equation, the orbital motion and the spin precession are synchronized, as stated previously in (6.91). In reality, g deviates slightly from 2, due to vacuum fluctuations, and $g - 2$ can be measured to very high accuracy by observing the slippage between orbital motion and spin precession.

Consider now an electron moving in an electrostatic potential in a circular orbit. In the frame moving with the electron, there is a magnetic field $\mathbf{B}'$, about which the spin precesses with frequency

$$\omega_{\text{rest}} = \frac{e}{2m}gB' \tag{6.118}$$

with $g = 2$. However, this is not the precession frequency observed in the laboratory frame, due to the *Thomas precession*. This arises because the electron frame has a precession relative to the lab frame, due to the nonadditivity of velocities in successive Lorentz transformations. To see this, make a Lorentz transformation from the laboratory frame (which is presumed to be an inertial frame) to the instantaneous rest frame of the electron at time t. Since the orbital velocity $\mathbf{v}$ is small in the nonrelativistic limit, the electron spinor undergoes an infinitesimal Lorentz transformation:

$$S_1 = 1 - \tfrac{1}{2}\boldsymbol{\alpha} \cdot \mathbf{v} \tag{6.119}$$

At time $t + dt$, the velocity becomes $\mathbf{v} + \mathbf{a}dt$, where $\mathbf{a}$ is the instantaneous acceleration, and we make a successive Lorentz transformation

$$S_2 = 1 - \tfrac{1}{2}\boldsymbol{\alpha} \cdot \mathbf{a}dt \tag{6.120}$$

Over the time interval dt, we have made the overall transformation

$$\begin{aligned} S = S_2 S_1 &= (1 - \tfrac{1}{2}\boldsymbol{\alpha} \cdot \mathbf{a}dt)\,(1 - \tfrac{1}{2}\boldsymbol{\alpha} \cdot \mathbf{v}) \\ &= 1 - \tfrac{1}{2}\boldsymbol{\alpha} \cdot (\mathbf{v} + \mathbf{a}dt) + \tfrac{1}{4}(\boldsymbol{\alpha} \cdot \mathbf{a})(\boldsymbol{\alpha} \cdot \mathbf{v})dt \end{aligned} \tag{6.121}$$

with last term causing the nonadditivity of velocities. Now use the identity

$$(\boldsymbol{\alpha}\cdot\mathbf{A})(\boldsymbol{\alpha}\cdot\mathbf{B}) = (\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot\mathbf{A}\times\mathbf{B} \tag{6.122}$$

and write

$$(\boldsymbol{\alpha}\cdot\mathbf{a})(\boldsymbol{\alpha}\cdot\mathbf{v})dt = i\boldsymbol{\sigma}\cdot\mathbf{a}\times\mathbf{v}dt \tag{6.123}$$

where we have used the fact that $\mathbf{a}\cdot\mathbf{v} = 0$ for circular motion. Thus

$$S = 1 - \frac{1}{2}\boldsymbol{\alpha}\cdot(\mathbf{v} + \mathbf{a}dt) + \frac{i}{4}\boldsymbol{\sigma}\cdot\mathbf{a}\times\mathbf{v}dt \tag{6.124}$$

The last term is a pure rotation—an example of the Wigner rotation discussed in Section 3.6. It leads to the Thomas precession, with angular frequency

$$\boldsymbol{\omega}_T = \tfrac{1}{2}\mathbf{a}\times\mathbf{v} \tag{6.125}$$

As illustrated in Fig. 6.1, its sense is opposite that of ω_{rest}, and the magnitude is given by

$$\omega_T = \frac{e}{2m}B' \tag{6.126}$$

The spin precession frequency in the lab frame is therefore

$$\omega_{\text{lab}} = \omega_{\text{rest}} - \omega_T = \frac{e}{2m}(g-1)B' \tag{6.127}$$

which leads to the result $g' = g - 1 = 1$.

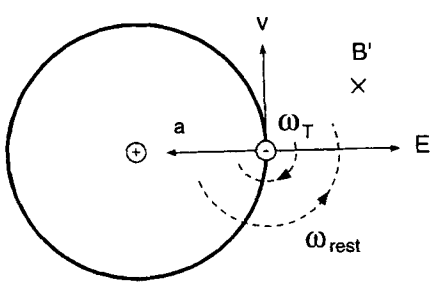


Figure 6.1 Thomas precession.

6.9 HOLE THEORY

The energy spectrum of the Dirac equation consists of a positive continuum extending from m to ∞, and a negative continuum from $-m$ to $-\infty$. The negative-energy levels cannot be ignored, because they are required by relativistic kinematics. Their existence destabilizes the theory, for, if there are any interactions at all, a particle can lose energy and fall down the bottomless pit of negative-energy states. Thus, no stable particles of positive energy can exist.

Imagine that the negative spectrum is cut off at some large but finite depth. If the particles obey Fermi statistics, the avalanche will stop when all negative-energy levels are filled with two particles (of opposite spin). The filled "negative-energy sea" will be the state of lowest energy: the vacuum state. Removing a particle in this sea will create a hole that appears as an "antiparticle," in the sense that it can be annihilated with a particle falling into the hole. Redefining the vacuum state in this manner results in "hole theory," as illustrated in Fig. 6.2. This redefinition of the vacuum state stabilizes the theory, and the important points are that

- It is possible only if spin-$\frac{1}{2}$ particles obey Fermi statistics.
- It makes the system into a many-particle system—a quantized field.

In hole theory, the absence of a negative-energy particle corresponds to the presence of an antiparticle with positive energy. This is expressed through the fact that the

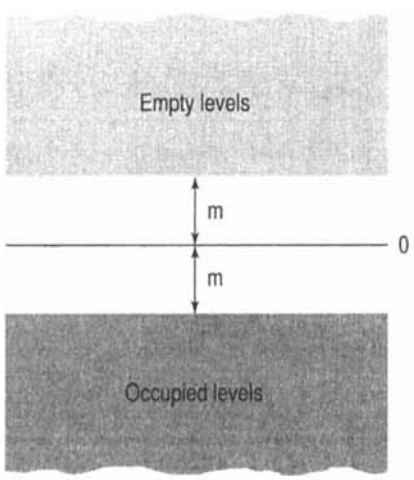


Figure 6.2 Energy spectrum in hole theory.

wave function of an antiparticle of energy E and momentum $\mathbf{p}$ is the complex conjugate of that of a particle of energy $-E$ and momentum $-\mathbf{p}$. Accordingly, we shall redefine the solutions to the Dirac equation by reversing the momentum for negative-energy states. Let the positive-energy solutions be redesignated $u_+(\mathbf{p}, 1)$, $u_+(\mathbf{p}, 2)$, and the negative-energy solutions $u_-(\mathbf{p}, 1)$, $u_-(\mathbf{p}, 2)$. We choose as a complete set the following:

$$u(\mathbf{p}, s) = C_0 u_+(\mathbf{p}, s)$$
$$v(\mathbf{p}, s) = C_0 u_-(-\mathbf{p}, s) \tag{6.128}$$

where

$$C_0 = \sqrt{\frac{m + E}{2m}}$$
$$E = +\sqrt{\mathbf{p}^2 + m^2} \tag{6.129}$$

They satisfy the equations

$$(\not{p} - m)u(\mathbf{p}, s) = 0$$
$$(\not{p} + m)v(\mathbf{p}, s) = 0 \tag{6.130}$$

with adjoint equations

$$u(\mathbf{p}, s)(\not{p} - m) = 0$$
$$v(\mathbf{p}, s)(\not{p} + m) = 0 \tag{6.131}$$

In these equations, p^μ is defined such that $p^0 > 0$:

$$p^\mu \equiv (E, \mathbf{p}) \tag{6.132}$$

In 3-vector form, the equations read

$$(\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m)u(\mathbf{p}, s) = Eu(\mathbf{p}, s)$$
$$(-\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m)v(\mathbf{p}, s) = -Ev(\mathbf{p}, s) \tag{6.133}$$

Note that the energy of $v(\mathbf{p}, s)$ is still negative, for all we did was reverse the momentum, and write its energy as $-E$ (with $E > 0$).

The orthonormality of the solutions is expressed by the relations

$$\bar{u}(\mathbf{p}, s)u(\mathbf{p}, s') = \delta_{ss'}$$
$$\bar{v}(\mathbf{p}, s)v(\mathbf{p}, s') = -\delta_{ss'}$$
$$\bar{v}(\mathbf{p}, s)u(\mathbf{p}, s') = 0 \tag{6.134}$$

which are equivalent to

$$u^\dagger(\mathbf{p}, s)u(\mathbf{p}, s') = \frac{E}{m}\delta_{ss'}$$
$$v^\dagger(\mathbf{p}, s)v(\mathbf{p}, s') = \frac{E}{m}\delta_{ss'}$$
$$u^\dagger(\mathbf{p}, s)v(-\mathbf{p}, s') = 0 \tag{6.135}$$

The completeness of the solution is stated as

$$\sum_{s=1}^{2}[u_a(\mathbf{p}, s)u_b^*(\mathbf{p}, s) + v_a(\mathbf{p}, s)v_b^*(\mathbf{p}, s)] = \frac{E}{m}\delta_{ab} \tag{6.136}$$

where a and b are spinor indices. This is equivalent to the matrix equation

$$\sum_{s=1}^{2}[u(\mathbf{p}, s)u(\mathbf{p}, s) - v(\mathbf{p}, s)\bar{v}(\mathbf{p}, s)] = 1 \tag{6.137}$$

The terms above are respectively projection operators onto positive-energy and negative-energy states:

$$\Lambda_+(\mathbf{p}) = \sum_{s=1}^{2}u(\mathbf{p}, s)\bar{u}(\mathbf{p}, s) = \frac{m + \not{p}}{2m}$$
$$\Lambda_-(\mathbf{p}) = -\sum_{s=1}^{2}v(\mathbf{p}, s)\bar{v}(\mathbf{p}, s) = \frac{m - \not{p}}{2m} \tag{6.138}$$

which have the properties

$$[\Lambda_\pm(\mathbf{p})]^2 = \Lambda_\pm(\mathbf{p}) \qquad \Lambda_+(\mathbf{p}) + \Lambda_-(\mathbf{p}) = 1 \tag{6.139}$$

Note that the 4-vector p^μ in $\not{p}$ is defined to have positive time component $p^0 = E$.

The Dirac equation cannot be a one-particle equation, but it furnishes a finite-dimensional representation of the full Lorentz group. As such, it provides a complete set of one-particle wave functions, in terms of which we can analyze the operator of a spin-$\frac{1}{2}$ field, as we shall do in the next chapter.

6.10 CHARGE CONJUGATION

An antiparticle should have opposite charge to a particle, since it represents the absence of a particle in the negative-energy sea. This is intuitively obvious; but let us make certain that the formalism gives this result. In the presence of an external electromagnetic field $A^\mu(x)$, the Dirac equation is as given by (6.82). We denote the wave function as $\psi(x)$ for positive-energy plane wave states, and $\psi^c(x)$ for negative-energy plane-wave states:

$$\begin{aligned} \psi(x) &\equiv e^{-iEt+i\mathbf{p}\cdot\mathbf{x}}u(\mathbf{p}, s) \\ \psi^c(x) &\equiv e^{iEt-i\mathbf{p}\cdot\mathbf{x}}v(\mathbf{p}, s) \end{aligned} \tag{6.140}$$

where $E = +\sqrt{\mathbf{p}^2 + m^2}$. Then (6.82) can be rewritten

$$\begin{aligned} [i\gamma^\mu(\partial_\mu + ieA_\mu) - m]\ \psi(x) &= 0 \\ [i\gamma^\mu(\partial_\mu - ieA_\mu) - m]\ \psi^c(x) &= 0 \end{aligned} \tag{6.141}$$

which show that the charge indeed has opposite signs for particle and antiparticle. The two equations above can transformed into each other through "charge conjugation," or "particle–antiparticle conjugation." To change the sign of the coupling term in the first equation, we take the complex conjugate:

$$[-i\gamma^{*\mu}(\partial_\mu - ieA_\mu) - m]\ \psi^*(x) = 0 \tag{6.142}$$

We then make a unitary transformation to bring it to the form of the second equation. Thus

$$\psi^c(x) = \eta\psi^*(x) \tag{6.143}$$

where η is a 4×4 matrix with the properties

$$\begin{aligned} \eta^2 &= 1 \\ \eta^{-1}(\gamma^\mu)^*\eta &= -\gamma^\mu \end{aligned} \tag{6.144}$$

The solution is, in our standard representation of the Dirac matrices,

$$\eta = i\gamma^2 \tag{6.145}$$

(where γ^2 is the second Dirac matrix). In terms of the spinors, charge conjugation corresponds to the transformation

$$v(\mathbf{p}, s) = \eta u^*(\mathbf{p}, s) \tag{6.146}$$

Since $\{\gamma^2, \gamma^0\} = 0$. This shows that particles and antiparticles have opposite parity.

Like the time reversal discussed earlier, the charge conjugation here is an operation on Dirac wave functions, and not on physical states, which are defined in quantum field theory. The operation is relevant because we expand the quantum field operators in terms of Dirac wave functions.

6.11 MASSLESS PARTICLES

For a massless Dirac particle, with $m = 0$, the equation for the Dirac spinor reduces to $\not{p}u(\mathbf{p}) = 0$, or

$$\boldsymbol{\alpha}\cdot\mathbf{p}u(\mathbf{p}) = p_0 u(\mathbf{p}) \tag{6.147}$$

where

$$p_0 = \pm E \qquad E \equiv |\mathbf{p}| \tag{6.148}$$

Since $[\alpha^k, \gamma_5] = 0$, we can diagonalize γ_5, whose eigenvalue ± 1 is called "chirality." The solution with chirality + 1 is called "right-handed," denoted u_R; one with chirality -1 is called "left-handed," denoted u_L:

$$\begin{aligned}\gamma_5 u_R(\mathbf{p}) &= u_R(\mathbf{p}) \\ \gamma_5 u_L(\mathbf{p}) &= -u_L(\mathbf{p})\end{aligned} \tag{6.149}$$

Using the relation

$$\gamma_5\boldsymbol{\alpha} = \boldsymbol{\sigma} \tag{6.150}$$

we have

$$\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}u(\mathbf{p}) = \frac{p_0}{E}\gamma_5 u(\mathbf{p}) \tag{6.151}$$

which states that the helicity $\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}$ is the chirality time the sign of the energy. Thus, for a right-handed particle, the helicity is correlated with the sign of the energy, and for a left-handed particle it is *anti*-correlated. For a given momentum $\mathbf{p}$, the four independent solutions are $u_C(\mathbf{p}, s)$, where $C = R, L$ denotes chirality and $s = \pm 1$ denotes helicity. Explicit solutions can be obtained from (6.45) by putting $m = 0$; but obviously we cannot normalize them according to (13.105). Instead, we put

$$u_C^\dagger(\mathbf{p}, s)u_{C'}(\mathbf{p}, s') = \delta_{CC'}\delta_{ss'}2E \tag{6.152}$$

It is easy to show $\bar{u}_C(\mathbf{p}, s)u_C(\mathbf{p}, s) = 0$, it follows that for, since $\{\gamma_5, \gamma^0\} = 0$, it fol-

lows that $\bar{u}$ and u have opposite chirality. The one-particle states $|\mathbf{p}\rangle$ have the properties

$$\langle \mathbf{p}'|\mathbf{p}\rangle = 2E(2\pi)^3\delta^3(\mathbf{p}-\mathbf{p}')$$

$$\int \frac{d^3p}{(2\pi)^3 2E}|\mathbf{p}\rangle\langle|\mathbf{p}| = 1 \tag{6.153}$$

We must, of course, define the vacuum using hole theory. In analogy with the massive case, we define antiparticle spinors:

$$v_C(\mathbf{p}, s) = u_C(-\mathbf{p}, -s) \tag{6.154}$$

For a given $\mathbf{p}$, the independent solutions can be taken to be $u_R(\mathbf{p}, 1)$, $v_R(\mathbf{p}, -1)$, $u_L(\mathbf{p}, -1)$, $v_L(\mathbf{p}, 1)$. Thus, a right-handed particle is a right-handed screw, and a left-handed particle is a left-handed screw. The correlation between handedness and helicity is reversed for antiparticles.

PROBLEMS

6.1 Lorentz Boost

(a) The transformation matrix for an infinitesimal Lorentz transformation is of the form $S = 1 + iR$, where R satisfies (6.60). Review the argument leading to the form $R = C\omega_{\mu\nu}\sigma^{\mu\nu}$, and show that $C = \frac{1}{4}$.

(b) Using the identity $\gamma^0\{\sigma^{\mu\nu}\}^\dagger\gamma^0 = \sigma^{\mu\nu}$, show that

$$\gamma^0 S^\dagger \gamma^0 = S^{-1}$$

With this, verify the transformation law for $\bar{u}$ given in (6.72).

(c) Obtain the free-particle solutions $u(\mathbf{p}, s)$ to the Dirac equation by applying a Lorentz boost to the solutions in the rest frame:

$$u(\mathbf{p}, s) = \left[\cosh\frac{\phi}{2} + \hat{p}\cdot\boldsymbol{\alpha}\sinh\frac{\phi}{2}\right]b_s$$

where $\phi = \tanh^{-1}v$, and

$$b_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \quad b_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix} \quad b_3 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix} \quad b_4 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

6.2 Intrinsic Parity Show (6.73) that

$$\gamma^0 u(-\mathbf{p}, s) \begin{cases} u(\mathbf{p}, s) & (s = 1,2) \\ -u(\mathbf{p}, s) & (s = 3,4) \end{cases}$$

and therefore

$$\gamma^0 u(-\mathbf{p}, s) = u(\mathbf{p}, s) \qquad (s = 1,2)$$
$$\gamma^0 v(-\mathbf{p}, s) = -v(\mathbf{p}, s) \qquad (s = 1,2)$$

These relations indicate that particles and antiparticles have opposite intrinsic parity.

6.3 Pauli Term The Dirac equation describes a particle with $g = 2$. Physical particles have g factors different from 2 because of interactions, which give rise to an "anomalous" magnetic moment. The electron acquire the anomalous moment through interactions with the quantized electromagnetic field. That for the proton and neutron are dominated by the strong interactions. Suppose that the g factor is $2 + \kappa$. Show that this can be accommodated by taking the Dirac equation in external electromagnetic field to be

$$\left[i\gamma^\mu(\partial_\mu + ieA_\mu) - \frac{e\kappa}{4m}\sigma^{\mu\nu}F_{\mu\nu} - m \right] \psi(x) = 0$$

The extra term is called the "Pauli term." For the proton and the neutron, the experimental values are $\kappa_P = 1.79$, $\kappa_N = -1.91$, respectively.

6.4 Chiral Current The *chiral current density* is defined by

$$j_5^\mu(x) = \overline{\psi}(x)\gamma^\mu\gamma_5\psi(x)$$

Using the Dirac equation, show that

$$\partial_\mu j_5^\mu = 2m\overline{\psi}\gamma_5\psi$$

The chiral current becomes conserved in the massless limit $m \to 0$. In quantum field theory with electromagnetic interactions turned on, this property is destroyed by the axial anomaly [3].

6.5 Zitterbewegung The *zitterbewegung* [4] is a kinematic property of the spin-$\frac{1}{2}$ representation of the Lorentz group, the "clockwork" of the Dirac equation. To exhibit this motion, construct a wave packet for a Dirac particle:

$$\psi(\mathbf{r}, t) = \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{r}}[w_+(\mathbf{p})e^{-iEt} + w_-(\mathbf{p})e^{iEt}]$$

where $E = +\sqrt{\mathbf{p}^2 + m^2}$, and $w_\pm(\mathbf{p})$ are linear combinations of Dirac spinors with positive (negative) energies $\pm E$. Calculate the expectation value of the velocity $\langle \mathbf{v} \rangle = \int d^3r\psi^\dagger\boldsymbol{\alpha}\psi$. Show that

$$\langle \mathbf{v} \rangle = \mathbf{v}_0 + 2\mathrm{Re}\int \frac{d^3p}{(2\pi)^3}(w_+^\dagger \boldsymbol{\alpha} w_-)e^{2iEt}$$

where $\mathbf{v}_0 = (2\pi)^{-3}\int d^3p(w_+^\dagger \boldsymbol{\alpha} w_+ + w_-^\dagger \boldsymbol{\alpha} w_-)$. Integrate this to obtain the average position

$$\langle \mathbf{r} \rangle = \mathbf{r}_0 + \mathbf{v}_0 t + \mathrm{Im}\int \frac{d^3p}{(2\pi)^3 E} e^{2iEt} w_+^\dagger \boldsymbol{\alpha} w_-$$

The last term is the *zitterbewegung*, which arises from an interference between positive- and negative-energy states. On dimensions grounds, we can conclude that the amplitude of this oscillatory motion is of the order of the Compton wavelength $1/m$, and therefore unobservable. In the hole theory, when all negative energy state are filled, the *zitterbewegung* becomes part of the vacuum fluctuations of the Dirac field, for it can happen only when holes are momentarily created as a result of fluctuations.

6.6 Gordon Decomposition

(a) From the definition of the Dirac matrices, show that

$$\gamma^\mu \gamma^\nu = g^{\mu\nu} - i\sigma^{\mu\nu}$$

(b) Multiply the equation $(i\gamma^\mu \partial_\mu - m)\psi = 0$ from the left by $\bar{\psi}\gamma^\nu$, and use the identity to rewrite the result in the form

$$\bar{\psi}\gamma^\mu \psi = \frac{1}{2m}\{i[\bar{\psi}(\partial^\mu \psi) - (\partial^\mu \bar{\psi})\psi] + \partial_\nu(\bar{\psi}\sigma^{\mu\nu}\psi)\}$$

with spatial components

$$\bar{\psi}\boldsymbol{\gamma}\psi = \frac{1}{2m}\left\{i[(\nabla\bar{\psi})\psi - \bar{\psi}(\nabla\psi)] + \nabla \times (\bar{\psi}\boldsymbol{\sigma}\psi) - i\frac{\partial}{\partial t}(\bar{\psi}\boldsymbol{\alpha}\psi)\right\}$$

This is the *Gordon decomposition*, which splits the current density into a "convection" part, plus contributions from the spin. It suggests that the spin is the orbital angular momentum of the *zitterbewegung*.

(c) Let $u_i = u(\mathbf{p}_i, s_i)$, $(i = 1,2)$, be two Dirac spinors. Let

$$P^\mu = p_1^\mu + p_2^\mu$$
$$k^\mu = p_1^\mu - p_2^\mu$$

Show that

$$\bar{u}_2 \gamma^\mu u_1 = \frac{1}{2m}\bar{u}_2(P^\mu + i\sigma^{\mu\nu}k_\nu)u_1$$

6.7 Massless Particles Consider massless Dirac particles.

(a) Show $\bar{u}_C(\mathbf{p}, s)u_C(\mathbf{p}, s) = 0$.

(b) Show that the projection operators $\Lambda_\pm(\mathbf{p})$ for positive and negative energies have the properties

$$\Lambda_{\pm}(\mathbf{p}) = E \pm \boldsymbol{\alpha}\cdot\mathbf{p}$$

$$\Lambda_{\pm}^2(\mathbf{p}) = E\Lambda_{\pm}(\mathbf{p})$$

$$\Lambda_{+}(\mathbf{p}) + \Lambda_{-}(\mathbf{p}) = E$$

REFERENCES

1. K. M. Case, *Phys. Rev.* **106,** 177 (1957).
2. R. L. Garwin, L. M. Lederman, and M. Weinrich, *Phys. Rev.* **105,** 1415 (1957).
3. K. Huang, *Quark, Leptons, and Gauge Fields*, 2nd ed. World Scientific, Singapore, 1992, Chapter 11.
4. K. Huang, *Am. J. Phys.* **20,** 479 (1952).

CHAPTER SEVEN

The Dirac Field

7.1 QUANTIZATION OF THE DIRAC FIELD

In hole theory, the Dirac equation describes a many-fermion system, and thus the Dirac "wave function" $\psi(x)$ should be regarded as a classical field to be quantized according to Fermi statistics. To carry out the quantization in the canonical formalism, we take as classical Lagrangian density

$$\mathcal{L}(x) = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \tag{7.1}$$

where $\psi(x)$ is a four-component spinor and the independent field variables are the components $\psi_a(x)$. We note that $\mathcal{L}(x)$ is Lorentz-invariant, and globally gauge-invariant. This is a first-order Lagrangian density, involving first instead of second derivatives with respect to time. We have illustrated the self-consistency of the canonical formalism in this case in Problem 4.5. Therefore, following strict canonical procedures, we calculate

$$\pi_a^\mu \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_a)} = i\bar{\psi}_b(\gamma^\mu)_{ba} \qquad \frac{\partial \mathcal{L}}{\partial \psi_a} = -m\bar{\psi} \tag{7.2}$$

The equation of motion is

$$i\partial_\mu \bar{\psi}\gamma^\mu + m\bar{\psi} = 0 \tag{7.3}$$

which correctly gives the Dirac equation in Hermitian-conjugate form. The Lagrangian density vanishes for fields satisfying the equation of motion:

$$\mathcal{L}(x) = 0 \qquad \text{(for fields satisfying equation of motion)} \tag{7.4}$$

The canonical conjugate to ψ_a is $i\psi_a^*$, since

$$\pi_a = \pi_a^0 = i\psi_a^* \tag{7.5}$$

The Lagrangian does not depend on $\dot{\psi}_a^*$, and therefore ψ_a^* has no conjugate. One must resist the temptation to make the Lagrangian more "symmetric" by replacing $i\bar{\psi}\gamma^\mu\partial_\mu\psi$ with $(i/2)[\bar{\psi}\gamma^\mu(\partial_\mu\psi) + (\partial_\mu\bar{\psi})\gamma^\mu\psi]$. This would be akin to "adding feet when drawing a snake," as a Chinese saying goes.

The canonical quantization rules lead to the following anticomutation relations:

$$\begin{aligned}\{\psi_a(\mathbf{r}, t), \psi_b^\dagger(\mathbf{r}', t)\} &= \delta_{ab}\delta^3(\mathbf{r} - \mathbf{r}') \\ \{\psi_a(\mathbf{r}, t), \psi_b(\mathbf{r}', t)\} &= 0\end{aligned} \tag{7.6}$$

where a and b denote spinor indices. The anticommutators serve as initial conditions for the Dirac equation. They also fix the normalization left arbitrary in the Dirac equation.

The Lagrangian density is invariant under the global gauge transformation $\psi \rightarrow e^{-i\omega}\psi$, where ω is a constant. The associated Noether current is

$$j^\mu(x) = \psi^\dagger(x)\gamma^\mu\psi(x) \tag{7.7}$$

which is conserved:

$$\partial_\mu j^\mu(x) = 0 \tag{7.8}$$

The canonical energy–momentum tensor, which is associated with translational invariance, is given by

$$T_c^{\mu\alpha}(x) = i\bar{\psi}(x)\gamma^\mu\partial^\alpha\psi(x) \tag{7.9}$$

with conservation law

$$\partial_\mu T_c^{\mu\alpha}(x) = 0 \tag{7.10}$$

The energy and momentum densities are respectively

$$\begin{aligned}T_c^{00} &= \psi^\dagger i\partial^0\psi = \psi^\dagger(-i\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi \\ T_c^{0k} &= \psi^\dagger i\partial^k\psi\end{aligned} \tag{7.11}$$

When integrate over space, they give the Hamiltonian H and total momentum $\mathbf{P}$:

$$\begin{aligned}H &= \int d^3r\, \psi^\dagger(\mathbf{r}, t)(-i\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi(\mathbf{r}, t) \\ \mathbf{P} &= -i\int d^3r\, \psi^\dagger(\mathbf{r}, t)\nabla\, \psi(\mathbf{r}, t)\end{aligned} \tag{7.12}$$

The generalized angular momentum tensor is

$$M_c^{\mu\alpha\beta}(x) = \psi(\bar{x})[i(x^\alpha\gamma^\mu\partial^\beta - x^\beta\gamma^\mu\partial^\alpha) + \tfrac{1}{2}\gamma^\mu\sigma^{\alpha\beta}]\psi(x) \tag{7.13}$$

and the angular momentum and the boost operators are respectively

$$J^j = \epsilon^{jkl}\int d^3r\, M_c^{0kl} = i\epsilon^{jkl}\int d^3r\psi^\dagger[(x^k\partial^l - x^l\partial^k) - \frac{i}{2}\sigma^{kl}]\,\psi$$

$$K^j = \int d^3r\, M_c^{00j} = i\int d^3r\ \psi^\dagger(x^0\partial^j - x^j\partial^0 + \frac{1}{2}\alpha^j)\,\psi \tag{7.14}$$

From the angular momentum, we can read off the spin operator:

$$\Sigma^{\mu\nu} = -\frac{i}{2}\sigma^{\mu\nu} \tag{7.15}$$

in agreement with what we found in the last chapter.

The one-particle solutions obtained in Chapter 6 constitute a basis in terms of which the field operators may be expanded. We normalize the wave functions in a large periodic box of volume Ω, and write

$$\psi(\mathbf{r}, t) = \sum_{\mathbf{p}s}\sqrt{\frac{m}{\Omega E_p}}[a_{\mathbf{p}s}e^{i(\mathbf{p}\cdot\mathbf{r}-E_pt)}u(\mathbf{p}, s) + b_{\mathbf{p}s}{}^\dagger e^{-i(\mathbf{p}\cdot\mathbf{r}-E_pt)}v(\mathbf{p}, s)]$$

$$\psi^\dagger(\mathbf{r}, t) = \sum_{\mathbf{p}s}\sqrt{\frac{m}{\Omega E_p}}[a_{\mathbf{p}s}{}^\dagger e^{-i(\mathbf{p}\cdot\mathbf{r}-E_pt)}u^\dagger(\mathbf{p}, s) + b_{\mathbf{p}s}e^{i(\mathbf{p}\cdot\mathbf{r}-E_pt)}v^\dagger(\mathbf{p}, s)] \tag{7.16}$$

where

$$E_p = +\sqrt{\mathbf{p}^2 + m^2} \tag{7.17}$$

The factor $\sqrt{m/E_p}$ appears because the Dirac spinors are normalized according to (6.133) and (6.134):

$$\bar{u}(\mathbf{p}, s)u(\mathbf{p}, s') = \frac{m}{E_p}u^\dagger(\mathbf{p}, s)u(\mathbf{p}, s') = \delta_{ss'}$$

$$\bar{v}(\mathbf{p}, s)v(\mathbf{p}, s') = -\frac{m}{E_p}v^\dagger(\mathbf{p}, s)v(\mathbf{p}, s') = -\delta_{ss'}$$

$$\bar{v}(\mathbf{p}, s)u(\mathbf{p}, s') = 0 \tag{7.18}$$

With this factor taken out, we have simple anticommutation rules

$$\{a_{\mathbf{p}s}, a^\dagger_{\mathbf{p}'s'}\} = \{b_{\mathbf{p}s}, b^\dagger_{\mathbf{p}'s'}\} = \delta_{ss'}\delta_{\mathbf{p}\mathbf{p}'}$$

$$\{a_{\mathbf{p}s}, a_{\mathbf{p}'s'}\} = \{b_{\mathbf{p}s}, b_{\mathbf{p}'s'}\} = \{a_{\mathbf{p}s}, b_{\mathbf{p}'s'}\} = 0 \tag{7.19}$$

which lead to the interpretation that $a_{\mathbf{p}s}$ annihilates a particle whose wave function is $u(\mathbf{p}, s)$ and $b_{\mathbf{p}s}$ annihilates an antiparticle whose wave function is $v^\dagger(\mathbf{p}, s)$.

Hole theory is implemented through the statement

$$a_{\mathbf{p}s}|0\rangle = b_{\mathbf{p}s}|0\rangle = 0 \qquad \text{for all } \mathbf{p},s \tag{7.20}$$

This implies that there are neither particles nor antiparticles in the vacuum state $|0\rangle$. In terms of the annihilation and creation operators we have

$$H = \sum_{\mathbf{p}s} E_p(a^\dagger_{\mathbf{p}s}a_{\mathbf{p}s} - b_{\mathbf{p}s}b^\dagger_{\mathbf{p}s}) = \sum_{\mathbf{p}s} E_p(a^\dagger_{\mathbf{p}s}a_{\mathbf{p}s} + b^\dagger_{\mathbf{p}s}b_{\mathbf{p}s} - 1)$$

$$\mathbf{P} = \sum_{\mathbf{p}s} \mathbf{p}(a^\dagger_{\mathbf{p}s}a_{\mathbf{p}s} - b_{\mathbf{p}s}b^\dagger_{\mathbf{p}s}) = \sum_{\mathbf{p}s} \mathbf{p}(a^\dagger_{\mathbf{p}s}a_{\mathbf{p}s} + b^\dagger_{\mathbf{p}s}b_{\mathbf{p}s}) \tag{7.21}$$

If we had not used hole theory, $b_{\mathbf{p}s}$ would be creation instead of annihilation operator, and $b_{\mathbf{p}s}b_{\mathbf{p}s}{}^\dagger$ would have eigenvalues 0,1. Consequently, the Hamiltonian would not be bounded from below. The sign reversal that makes the Hamiltonian positive-definite, of course, depends on the fact that we quantized the system according to Fermi statistics.

The charge operator is given by

$$Q = \int d^3r\, \bar{\psi}\gamma^0\psi = \int d^3r\, \psi^\dagger\psi$$

$$= \sum_{\mathbf{p}s} (a^\dagger_{\mathbf{p}s}a_{\mathbf{p}s} - b_{\mathbf{p}s}b^\dagger_{\mathbf{p}s} + 1) \tag{7.22}$$

which shows that particles and antiparticles have opposite charge. The minus sign above arises through rewriting $b_{\mathbf{p}s}b^\dagger_{\mathbf{p}s}$ as $-b^\dagger_{\mathbf{p}s}b_{\mathbf{p}s} + 1$. This is dictated by the fact that $b^\dagger_{\mathbf{p}s}b_{\mathbf{p}s}$ has positive eigenvalues in hole theory. The normalization of Q is arbitrary, for the magnitude of the charge is determined only when there is interaction with the electromagnetic field.

7.2 FEYNMAN PROPAGATOR

The Feynman propagator for the Dirac field is a 4×4 matrix

$$iS_F(x) = \langle 0|T\psi(x)\bar{\psi}(0)|0\rangle \tag{7.23}$$

where the time-ordering operator T is defined to include a sign change when two fermion operators A and B are interchanged:

$$T[A(t_1)B(t_2)] = \begin{cases} A(t_1)B(t_2) & \text{if } t_1 > t_2 \\ -B(t_2)A(t_1) & \text{if } t_2 > t_1 \end{cases} \tag{7.24}$$

The propagator can be calculated straightforwardly, using the expansions (7.16). Let $x = (t, \mathbf{r})$. For $t > 0$, the only contribution comes from terms in the expansion of the form $aa^\dagger$:

$$i[S_F(x)]_{ab} = \langle 0|\psi_a(\mathbf{r}, t)\bar{\psi}_b(0)|0\rangle = \frac{1}{\Omega}\sum_{\mathbf{p}}\frac{m}{E_p}e^{-i(E_p t-\mathbf{p}\cdot\mathbf{r})}\sum_{s=1}^{2}u_a(\mathbf{p}, s)\bar{u}_b(\mathbf{p}, s) \tag{7.25}$$

For $t < 0$ we need only keep terms of the form $bb^\dagger$:

$$i[S_F(x)]_{ab} = -\langle 0)\bar{\psi}_b(0)\psi_a(\mathbf{r}, t)|0\rangle = -\frac{1}{\Omega}\sum_{\mathbf{p}}\frac{m}{E_p}e^{i(E_p t-\mathbf{p}\cdot\mathbf{r})}\sum_{s=1}^{2}v_a(\mathbf{p}, s)\bar{v}_b(\mathbf{p}, s) \tag{7.26}$$

The sum over spin states results in the projection operators given in (6.137). Suppressing the spinor indices and going to the limit $\Omega \to \infty$, we have

$$iS_F(x) = \int\frac{d^3p}{(2\pi)^3}\frac{1}{2E_p}\begin{cases}(m + \not{p})e^{-i(E_p t-\mathbf{p}\cdot\mathbf{r})} & (t > 0)\\ (m - \not{p})e^{i(E_p t-\mathbf{p}\cdot\mathbf{r})} & (t < 0)\end{cases} \tag{7.27}$$

We can make the replacement

$$m + \not{p} = m + \gamma^0 E_p - \gamma^k p^k \to m + i\gamma^0\frac{\partial}{\partial t} - \gamma^k p^k \tag{7.28}$$

because this operator acts on the exponential factor. For $t < 0$, make the change of variables $\mathbf{p} \to -\mathbf{p}$. Then we have

$$iS_F(x) = \int\frac{d^3p}{(2\pi)^3}\frac{1}{2E_p}\left(m + i\gamma^0\frac{\partial}{\partial t} - \gamma^k p^k\right)e^{-iE_p|t|}e^{i\mathbf{p}\cdot\mathbf{r}} \tag{7.29}$$

Now use the representation

$$e^{-iE_p|t|} = \frac{E_p}{i\pi}\int_{-\infty}^{\infty}dp_0\frac{e^{-ip_0 t}}{E_p^2 - p_0^2 - i\eta} = \frac{iE_p}{\pi}\int_{-\infty}^{\infty}dp_0\frac{e^{-ip_0 t}}{p^2 - m^2 + i\eta} \tag{7.30}$$

where $\eta \to 0^+$. Then, the operator $i\partial/\partial t$ in the previous formula can be replaced by p_0. The final result is

$$S_F(x) = \int\frac{d^4p}{(2\pi)^4}e^{-ip\cdot x}\frac{\not{p} + m}{p^2 - m^2 + i\eta} \tag{7.31}$$

We leave it as an exercise to show that

$$S_F(x) = (i\gamma^\mu\partial_\mu + m)\Delta_F(x) \tag{7.32}$$

where $\Delta_F(x)$ is the Feynman propagator for a free scalar field of mass m.

The Fourier transform of the propagator is

$$\tilde{S}_F(p) = \frac{\not{p} + m}{p^2 - m^2 + i\eta} = \frac{1}{\not{p} - \mathrm{m} + i\eta} \tag{7.33}$$

where the right side is the inverse of a 4×4 matrix. The 4-vector p is arbitrary, with either $p^0 > 0$ or $p^0 > 0$. For $p^0 > 0$, we have, according to (6.137),

$$\frac{(\not{p} + m)}{2m} = \sum_{s=1}^{2} u(\mathbf{p}, s)\bar{u}(\mathbf{p}, s) \qquad (p^0 > 0) \tag{7.34}$$

For $p^0 > 0$, let us define $q^\mu \equiv -p^\mu$, Then according to (6.137), we obtain

$$\frac{(\not{p} + m)}{2m} = \frac{(-\not{q} + m)}{2m} = -\sum_{s=1}^{2} v(\mathbf{q}, s)\bar{v}(\mathbf{q}, s) \qquad (p^0 > 0) \tag{7.35}$$

This shows that the residue at the mass-shell pole at $p^2 = m^2$ contains the wave functions of an electron of momentum $\mathbf{p}$, or those of a positron of momentum $\mathbf{q} = -\mathbf{p}$.

7.3 NORMAL ORDERING

Both H and Q contain divergent contributions from the zero-point energy and charge of the vacuum state. These terms have no physical relevance since energy and charge are measured relative to those of the vacuum state. They can be eliminated by redefining the reference points, and this can be achieved by arranging the order of operators appropriately.

We first introduce the notion of *normal ordering*. Suppose that O is a product of creation and annihilation operators. The corresponding *normal product* $: O :$ is defined as that obtained from O by rearranging the order of the factors, if necessary, such that all creation operators stand to the left of all annihilation operators. In the rearrangement process, an interchange of two fermion operators gives rise to a factor -1. As an example:

$$\begin{aligned} &: a^\dagger_{\mathbf{p}s} a_{\mathbf{p}'s'} : = a^\dagger_{\mathbf{p}s} a_{\mathbf{p}'s'} \\ &: a_{\mathbf{p}'s'} a^\dagger_{\mathbf{p}s} : = -a^\dagger_{\mathbf{p}s} a_{\mathbf{p}'s'} \end{aligned} \tag{7.36}$$

Normal ordering can be naturally extended to a sum of products:

$$: O_1 + O_2 : = : O_1 : + : O_2 : \tag{7.37}$$

We now redefine the Hamiltonian and the current as

$$H = : \int d^3 r \psi^\dagger(\mathbf{r}, t)(-i\boldsymbol{\alpha} \cdot \nabla + \beta m)\psi(\mathbf{r}, t) :$$

$$j^{\mu}(x) = :\psi^{\dagger}(x)\gamma^{\mu}\psi(x): \tag{7.38}$$

It is clear that these operators give zero when operating on the vacuum state, because annihilation operators stand to the far right. This is just a formal way of stating that the zero-point energy and currents are to be omitted. As the notation is somewhat cumbersome, we shall not explicit indicate normal ordering unless necessary.

7.4 ELECTROMAGNETIC INTERACTIONS

We consider systems of interacting fields with a Lagrangian density consisting of the sum of the free Lagrangian densities of the participating fields, plus an interaction Lagrangian density that couple the fields together. This is not the most general case conceivable, but it is what we can handle mathematically. We illustrate the types of interactions commonly encountered. Consider a Dirac field, a complex scalar field, and the electromagnetic field, which have free Lagrangian densities given by

$$\begin{aligned}
\mathcal{L}_{\text{Dirac}} &= \overline{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi \\
\mathcal{L}_{\text{scalar}} &= \partial^{\mu}\phi^{*}\partial_{\mu}\phi - \kappa^{2}\phi^{*}\phi \\
\mathcal{L}_{\text{em}} &= -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu}
\end{aligned} \tag{7.39}$$

According to the gauge principle, the matter fields can be coupled to the electromagnetic (em) field by replacing ∂_{μ} by the covariant derivative

$$D^{\mu} = \partial_{\mu} + ieA^{\mu}(x) \tag{7.40}$$

where e is the electric charge. Assuming that both the Dirac field and the scalar field have the same charge e, the electrodynamic Lagrangian density is

$$\begin{aligned}
\mathcal{L} &= -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi + [D^{\mu}\phi]^{*}D_{\mu}\phi - \kappa^{2}\phi^{*}\phi \\
&= \mathcal{L}_{\text{em}} + \mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{scalar}} + \mathcal{L}_{\text{int}}
\end{aligned} \tag{7.41}$$

where

$$\begin{aligned}
\mathcal{L}_{\text{int}} &= (j^{\mu} + J^{\mu})A_{\mu} \\
j^{\mu} &= ie\overline{\psi}\gamma^{\mu}\psi \\
J^{\mu} &= -ie[\phi^{*}(\partial^{\mu}\phi) - (\partial^{\mu}\phi^{*})\phi] + e^{2}\phi^{*}\phi A^{\mu}
\end{aligned} \tag{7.42}$$

The matter fields are coupled through conserved currents, which are the Noether

currents associated with global gauge invariance. For the scalar field, the current has an e^2 term proportional to A^μ. This becomes a mass term for the photon when $\phi^*\phi$ develops a vacuum expectation value, in spontaneous symmetry breaking. (See Problem 15.5.)

The electromagnetic field couples to all charged fields through the gauge principle, and is universal in this sense. The vacuum fluctuations of the electromagnetic field include the momentary creation of virtual particle–antiparticle pairs and their subsequent annihilation. The temporary charge separation makes the vacuum into a dielectriclike medium, and all charged fields of the world participate in this "vacuum polarization," as their contributions being determined solely by charge and mass.

7.5 ISOSPIN

The Dirac field can be used in a phenomenological description of protons and neutrons, which are really made of quarks. The effective theory is useful in describing the "charge-independent" pion–nucleon interactions at low energies. It is based on the fact that proton and neutron are almost identical, and so are the three π mesons, and the strong nuclear forces respect the identities. By ignoring the electromagnetic and weak interactions, we can regard the proton and neutron as different states of a particle called the *nucleon,* and the π mesons as different states of the *pion.*

The nucleon field is represented by a two-component Dirac field

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix} \tag{7.43}$$

where $i = 1$ corresponds to proton, and $i = 2$ to neutron. Each ψ_i is a four-component Dirac spinor field. Writing out all the indices, we have eight complex fields $\psi_{ia}(x)$, with $a = 1, \ldots, 4$ and $i = 1, 2$. By analogy with spin angular momentum, we define the isospin $\vec{\tau}/2$ as generators of rotations in the two-dimensional internal space spanned by ψ_1 and ψ_2:

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{7.44}$$

The proton and neutron states are eigenstates of $\tau_3/2$ with respective eigenvalues $+\frac{1}{2}$ and $-\frac{1}{2}$:

$$|\mathrm{p}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |\mathrm{n}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{7.45}$$

which can be created from the vacuum by applying $\psi_i^\dagger(x)$. We use a shorthand notation in which the spinor and internal indices are suppressed. For example,

$$\bar{\psi}\gamma^{\mu}\vec{\tau}\psi \equiv \bar{\psi}_{ia}(\gamma^{\mu})_{ab}(\vec{\tau})_{ij}\psi_{bj} \tag{7.46}$$

where a, b are summed from 1 to 4 and i, j are summed from 1 to 2.

More generally we define *isospin* as an internal symmetry group whose generators $\vec{I}$ obey angular momentum commutation relations (Lie algebra):

$$[I_k, I_l] = i\epsilon^{klm}I_m$$

Thus one can simultaneously diagonalize $\vec{I}^2 = I(I+1)$ and I_3, and denote isospin eigenstates by $|I, I_3\rangle$. The nucleon belongs to the fundamental representation with $I = \frac{1}{2}$, in which $\vec{I} = \vec{\tau}/2$. The overhead arrow denotes a vector in isospin space, which has three components because that is the number of generators of the group.

The pion field has $I = 1$, and is described by a three-component real field

$$\vec{\phi}(x) = \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \end{pmatrix} \tag{7.47}$$

This the "adjoint representation" of the group, which has the same dimension as the number of generators, and in which the generators are represented by matrices $\vec{T}$ taken directly from Lie algebra:

$$(T_k)_{lm} = -i\epsilon^{klm} \tag{7.48}$$

Experimental evidence dictates that $\phi_i(x)$ be pseudoscalar, that is, that it change sign under spatial reflection. We note that I_3 is not diagonal. The physical pion fields, which are eigenstates of I^3, are related to ϕ_k through

$$\begin{aligned} \pi^{+}(x) &= \frac{1}{\sqrt{2}}[\phi_1(x) + i\phi_2(x)] \\ \pi^{-}(x) &= \frac{1}{\sqrt{2}}[\phi_1(x) - i\phi_2(x)] \\ \pi^{0}(x) &= \phi_3(x) \end{aligned} \tag{7.49}$$

These operators create states with $I_3 = +1, -1, 0$ respectively, when they operate on the vacuum state,

"Charge independence" in the pion–nucleon system means that the interaction conserves isospin. A Lorentz-invariant effective Lagrangian density, known as "pseudoscalar coupling," is given by

$$\mathcal{L}(x) = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - M)\psi + \tfrac{1}{2}[\partial^{\mu}\vec{\phi}\cdot\partial^{\mu}\vec{\phi} - m^2\vec{\phi}\cdot\vec{\phi}] + g(\bar{\psi}\gamma_5\vec{\tau}\psi)\cdot\vec{\phi} \tag{7.50}$$

The vector notation makes manifest the rotational invariance in isospin space. A competing model is the "pseudovector coupling" model, with

$$\mathcal{L}(x) = \bar{\psi}(i\gamma^\mu\partial_\mu - M)\psi + \tfrac{1}{2}[\partial_\mu\vec{\phi}\cdot\partial_\mu\vec{\phi} - m^2\vec{\phi}\cdot\vec{\phi}] + g'(\bar{\psi}\gamma_5\gamma^\mu\vec{\tau}\psi)\cdot\partial_\mu\vec{\phi} \quad (7.51)$$

Some consequences of isospin invariance are explored in Problem 7.4.

7.6 PARITY

We discuss the discrete symmetries, using as an example the electromagnetic coupling as contained in

$$\mathcal{L}(x) = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}[i\gamma^\mu(\partial_\mu + ieA_\mu) - m]\,\psi \quad (7.52)$$

Under a Lorentz transformation $x \to \Lambda x$, the field operators $\phi_a(x)$ undergoe a unitary transformation U given by (3.54):

$$U\phi_a(x)U^{-1} = S_{ab}\phi_b(\Lambda^{-1}x) \quad (7.53)$$

where S_{ab}. This can be extended to spatial reflection $\mathbf{x} \to \mathbf{x}$, $t \to t$, for which the unitary operator U is denoted by $\mathcal{P}$. For the Dirac field, we have $S = \gamma^0$ according to (6.69) and (6.70), and thus

$$\begin{aligned}\mathcal{P}\psi(\mathbf{r}, t)\mathcal{P}^{-1} &= \gamma^0\psi(-\mathbf{r}, t)\\ \mathcal{P}\bar{\psi}(\mathbf{r}, t)\mathcal{P}^{-1} &= \bar{\psi}(-\mathbf{r}, t)\gamma^0\end{aligned} \quad (7.54)$$

Since A^μ transforms like a vector,

$$\begin{aligned}\mathcal{P}A^k(\mathbf{r}, t)\mathcal{P}^{-1} &= -A^k(-\mathbf{r}, t)\\ \mathcal{P}A^0(\mathbf{r}, t)\mathcal{P}^{-1} &= A^0(-\mathbf{r}, t)\end{aligned} \quad (7.55)$$

Thus we have

$$\mathcal{P}\mathcal{L}(\mathbf{r}, t)\mathcal{P}^{-1} = \mathcal{L}(-\mathbf{r}, t) \quad (7.56)$$

which show that the Lagrangian $L = \int d^3x\,\mathcal{L}(x)$ is invariant.

From the expansion (7.16) at $t = 0$:

$$\psi(\mathbf{r}) = \sum_{\mathbf{p}s}\sqrt{\frac{m}{\Omega E_p}}[a_{\mathbf{p}s}e^{i\mathbf{p}\cdot\mathbf{r}}u(\mathbf{p}, s) + b^\dagger_{\mathbf{p}s}e^{-i\mathbf{p}\cdot\mathbf{r}}v(\mathbf{p}, s)] \quad (7.57)$$

we have

$$\mathcal{P}\psi(\mathbf{r})\mathcal{P}^{-1} = \sum_{\mathbf{p}s}\sqrt{\frac{m}{\Omega E_p}}[\mathcal{P}a_{\mathbf{p}s}\mathcal{P}^{-1}e^{i\mathbf{p}\cdot\mathbf{r}}u(\mathbf{p}, s) + \mathcal{P}b^\dagger_{\mathbf{p}s}\mathcal{P}^{-1}e^{-i\mathbf{p}\cdot\mathbf{r}}v(\mathbf{p}, s)]$$

$$\gamma^0\psi(-\mathbf{r}) = \sum_{\mathbf{p}s} \sqrt{\frac{m}{\Omega E_p}} [a_{\mathbf{p}s} e^{-i\mathbf{p}\cdot\mathbf{r}} \gamma^0 u(\mathbf{p}, s) + b^\dagger_{\mathbf{p}s} e^{i\mathbf{p}\cdot\mathbf{r}} \gamma^0 v(\mathbf{p}, s)] \tag{7.58}$$

Using the relations (Problem 6.2)

$$\begin{aligned} \gamma^0 u(-\mathbf{p}, s) &= u(\mathbf{p}, s) \\ \gamma^0 v(-\mathbf{p}, s) &= -v(\mathbf{p}, s) \end{aligned} \tag{7.59}$$

we obtain the statement that particles and antiparticles in Dirac theory have opposite intrinsic parity:

$$\begin{aligned} \mathcal{P} a_{\mathbf{p}s} \mathcal{P}^{-1} &= a_{-\mathbf{p}s} \\ \mathcal{P} b_{\mathbf{p}s} \mathcal{P}^{-1} &= -b_{-\mathbf{p}s} \end{aligned} \tag{7.60}$$

The transformation $\mathcal{P}$ may be accompanied by a rotation in spin space with respect to the index s, as is clear from (7.58); but we leave it out for simplicity.

7.7 CHARGE CONJUGATION

Charge conjugation, or particle–antiparticle conjugation, is defined as a unitary operation C on the Hilbert space that interchanges particle and antiparticle, and reverses the sign of the electromagnetic field:

$$\begin{aligned} C a_{\mathbf{p}s} C^{-1} &= b_{\mathbf{p}s} \\ C b_{\mathbf{p}s} C^{-1} &= a_{\mathbf{p}s} \\ C A^k(x) C^{-1} &= -A^k(x) \end{aligned} \tag{7.61}$$

The transformation of $A^0(x)$ is not specified independently, because in Coulomb gauge it is not an independent field. It is clear that $\mathcal{L}(x)$ is invariant under this transformation, because the free-field Lagrangian densities are invariant, and the Dirac field is coupled to the electromagnetic field through the current density, which changes sign.

To find how the Dirac field operator transforms, let us compare the following expansions:

$$\psi(\mathbf{r}) = \sum_{\mathbf{p}s} \sqrt{\frac{m}{\Omega E_p}} [a_{\mathbf{p}s} e^{i\mathbf{p}\cdot\mathbf{r}} u(\mathbf{p}, s) + b^\dagger_{\mathbf{p}s} e^{-i\mathbf{p}\cdot\mathbf{r}} v(\mathbf{p}, s)]$$

$$\psi^\dagger(\mathbf{r}) = \sum_{\mathbf{p}s} \sqrt{\frac{m}{\Omega E_p}} [a^\dagger_{\mathbf{p}s} e^{-i\mathbf{p}\cdot\mathbf{r}} u^*(\mathbf{p}, s) + b_{\mathbf{p}s} e^{i\mathbf{p}\cdot\mathbf{r}} v^*(\mathbf{p}, s)] \tag{7.62}$$

The expansion coefficients satisfy (6.146):

$$v(\mathbf{p}, s) = \eta u^*(\mathbf{p}, s) \tag{7.63}$$

where $\eta = i\gamma^2$ is a real 4×4 matrix. Therefore

$$\eta\psi^\dagger(\mathbf{r}) = \sum_{\mathbf{p}s} \sqrt{\frac{m}{\Omega E_p}} [b_{\mathbf{p}s} e^{i\mathbf{p}\cdot\mathbf{r}} u(\mathbf{p}, s) + a^\dagger_{\mathbf{p}s} e^{-i\mathbf{p}\cdot\mathbf{r}} v(\mathbf{p}, s)] \tag{7.64}$$

which shows

$$C\psi(\mathbf{r})C^{-1} = \eta\psi^\dagger(\mathbf{r}) \tag{7.65}$$

Note that the Dirac wave functions undergo complex conjugation, which is a nonlinear operation, because $(\lambda u)^* = \lambda^* u^*$. The field operator, however, undergoes a linear transformation, because $C(\lambda\psi)C^{-1} = \lambda\, C\psi\, C^{-1}$. The difference can be traced to the fact that in the Dirac equation we have to change the sign of the coupling to an external electromagnetic field, whereas in the field theory, the electromagnetic field is part of the system, and changes sign under charge conjugation.

7.8 TIME REVERSAL

Time reversal is the operation of interchanging past and future, represented by a operator $\mathcal{T}$ on Hilbert space. Suppose that Ψ_a is a member of a complete set of state in Hilbert space, where a stands for quantum numbers, such as momentum $\mathbf{p}$ and spin projection s on a fixed axis. The time-reversed state $\mathcal{T}\Psi_a$ must be a member of the same set:

$$\mathcal{T}\Psi_a = \Psi_{\bar{a}} \tag{7.66}$$

where $\bar{a}$ are the time-reversed quantum numbers, defined by correspondence with classical mechanics:

$$\bar{\mathbf{p}} = -\mathbf{p} \qquad \bar{s} = -s \tag{7.67}$$

and the helicity is invariant. The basic property of $\mathcal{T}$ is

$$(\mathcal{T}\Psi_a, \mathcal{T}\Psi_b) = (\Psi_b, \Psi_a) \tag{7.68}$$

that is, it interchanges initial and final states. This can be rewritten

$$(\mathcal{T}\Psi_a, \mathcal{T}\Psi_b) = (\Psi_a, \Psi_b)^* \tag{7.69}$$

Replacing Ψ_b by $\lambda\Psi_b$, where λ is a complex number, we have

$$(\mathcal{T}\Psi_a, \mathcal{T}(\lambda\Psi_b)) = \lambda^*(\Psi_a, \Psi_b)^* \tag{7.70}$$

Therefore

$$\mathcal{T}(\lambda\Psi_b) = \lambda^*\mathcal{T}\Psi_b \tag{7.71}$$

Thus, when acting on a number, $\mathcal{T}$ takes its complex conjugate. This makes $\mathcal{T}$ nonlinear. More specifically, it is called an. "antilinear" operator. A general representation of $\mathcal{T}$ is complex conjugation followed by a unitary transformation:

$$\begin{aligned} \mathcal{T} &= U* \\ \mathcal{T}^{-1} &= U^{-1}* \end{aligned} \tag{7.72}$$

where it is assumed that U commutes with complex conjugation. For the Schrödinger equation

$$H\Psi = i\frac{\partial\Psi}{\partial t} \tag{7.73}$$

time reversal means

$$H(\mathcal{T}\Psi) = -i\frac{\partial(\mathcal{T}\Psi)}{\partial t} \tag{7.74}$$

The system is invariant under time reversal if the time-reversed equation is equivalent to the original. Taking the complex conjugate, we have

$$H^*(U\Psi) = i\frac{\partial(U\Psi)}{\partial t} \tag{7.75}$$

Thus, the system is invariant under time reversal if the Hamiltonian is real:

$$H = H^* \tag{7.76}$$

which implies that the Lagrangian must be real.

Without going through all the details, we can conclude that

$$\begin{aligned} \mathcal{T}A^k(\mathbf{r})\mathcal{T}^{-1} &= -A^k(\mathbf{r}) \\ \mathcal{T}\psi(\mathbf{r})\mathcal{T}^{-1} &= \gamma^0\gamma_5\psi(\mathbf{r}) \end{aligned} \tag{7.77}$$

The first equation follows from the requirement that A^k transform like the current density, which must change sign, because classically it is a velocity. The second follows from the fact that $\gamma^0\gamma_5$ is the transformation that preserves the Dirac equation

under time reversal, as shown in (6.70). It is straightforward to verify that the Lagrangian is invariant, if the charge e is real.

There is a theorem known as the *PCT* theorem, which states that a local field theory that is Lorentz invariant is automatically invariant under the product *PCT*, even though it may be separately invariant under *P*,*C*,*T* separately. We refer the reader elsewhere [1] for proof.

PROBLEMS

7.1 Energy–Momentum Tensor The canonical energy–momentum tensor $T_c^{\mu\alpha}$ for the Dirac field is not symmetric in $\mu\alpha$. According to Section 4.5, we can construct an equivalent symmetric tensor $T^{\mu\alpha} = T_c^{\mu\alpha} + \frac{1}{2}\partial_\lambda X^{\lambda\mu\alpha}$. Find $X^{\lambda\mu\alpha}$.

7.2 Propagator Show that the propagator for the Dirac field is related to that of the scalar field through

$$S_F(x) = (i\gamma^\mu \partial_\mu + m)\Delta_F(x)$$

7.3 Neutrinos Neutrinos are massless Dirac particles. Using the convention for wave function given in Section 6.11, expand the field operator in terms of annihilation and creation operators.

7.4 Isospin Transformations

(a) Show that under an infinitesimal isospin transformation, the nucleon field, and the pion field transform according to

$$\psi \rightarrow (1 - \frac{i}{2}\vec{\omega}\cdot\vec{\tau})\,\psi$$

$$\vec{\phi} \rightarrow \vec{\phi} + \vec{\omega} \times \vec{\phi}$$

where the components of $\vec{\omega}$ are arbitrary infinitesimal real parameters.

(b) Let $\vec{V} = \psi\Gamma\,\vec{\tau}\psi$, where γ is a 4×4 Dirac matrix. Show that $\vec{V}$ transforms like a vector in isospin space:

$$\vec{V} \rightarrow \vec{V} + \vec{\omega} \times \vec{V}$$

7.5 Pion–Nucleon Scattering As far as isospin properties are concerned, the pion and nucleon states can be labeled by I and I_3:

$$|\pi^+\rangle = |1, 1\rangle \qquad |\pi^-\rangle = |1, -1\rangle \qquad |\pi^0\rangle = |1, 0\rangle \}$$

$$|p\rangle = |\tfrac{1}{2}, \tfrac{1}{2}\rangle \qquad |n\rangle = |\tfrac{1}{2}, -1\backslash\rangle$$

(a) A state containing a pion and a nucleon is a direct product in isospin space, as, for example, $|\pi^+ n\rangle = |1, 1\rangle \times |\frac{1}{2}, \frac{1}{2}\rangle$. However, this is not an eigenstate of total isospin and thus not an eigenstate of the Hamiltonian of the system. Show that eigenstates of the isospin are the following:

$I = \frac{3}{2}$:

$$|\tfrac{3}{2}, \tfrac{3}{2}\rangle = |p\pi^+\rangle$$

$$|\tfrac{3}{2}, \tfrac{1}{2}\rangle = \sqrt{\tfrac{1}{3}}|n\pi^+\rangle + \sqrt{\tfrac{2}{3}}|p\pi^0\rangle$$

$$|\tfrac{3}{2}, -\tfrac{1}{2}\rangle = \sqrt{\tfrac{1}{3}}|p\pi^-\rangle + \sqrt{\tfrac{2}{3}}|n\pi^0\rangle$$

$$|\tfrac{3}{2}, -\tfrac{3}{2}\rangle = |n\pi^0\rangle$$

$I = \frac{1}{2}$:

$$|\tfrac{1}{2}, \tfrac{1}{2}\rangle = \sqrt{\tfrac{2}{3}}|n\pi^+\rangle - \sqrt{\tfrac{1}{3}}|p\pi^0\rangle$$

$$|\tfrac{1}{2}, -\tfrac{1}{2}\rangle = \sqrt{\tfrac{1}{3}}|n\pi^0\rangle - \sqrt{\tfrac{2}{3}}|p\pi^-\rangle$$

(b) For interactions that conserve isospin, the pion–nucleon scattering amplitude depends only on total isospin and not on I_3 (for the same reason that atomic energy levels are independent of the magnetic quantum number, i.e., the Wigner–Eckart theorem.) Denote the transition amplitudes by $a_{3/2}$ and $a_{1/2}$, and show that they have the form

$$\langle \mathrm{p}\pi^+|T|p\pi^+\rangle = a_{3/2}$$

$$\langle \mathrm{p}\pi^-|T|p\pi^-\rangle = \tfrac{1}{3}a_{3/2} + \tfrac{2}{3}a_{1/2}$$

$$\langle \mathrm{n}\pi^0|T|p\pi^-\rangle = \frac{\sqrt{2}}{3}a_{3/2} - \frac{\sqrt{2}}{3}a_{1/2}$$

where T is the transition operator. The corresponding scattering cross sections are proportional to the squares of these amplitudes.

(c) Neglect $a_{1/2}$ compared to $a_{3/2}$, and show that pion–nucleon scattering cross sections bear the ratio

$$\sigma(p\pi^+) : \sigma(p\pi^-) : \sigma(n\pi^0 \to \mathrm{p}\pi^-) = 9 : 1 : 2$$

This is verified experimentally at low energies. The reason that $a_{3/2}$ dominates is the resonance scattering $\pi + N \to \Delta \to \pi + N$, where Δ is a particle of spin $\frac{3}{2}$, isopin $\frac{3}{2}$, with mass 1232 MeV, known at one time as the "33 resonance."

REFERENCE

1. R. Streater and A. Wightman, *PCT, Spin and Statistics, and All That*, Benjamin, New York, 1964.

CHAPTER EIGHT

Dynamics of Interacting Fields

8.1 TIME EVOLUTION

The dynamics of a quantum mechanical system is governed by the Hamiltonian H, which generates time translations. One may view the time development from different perspectives. In the *Schrödinger picture,* one regards the operators O_s as time-independent objects, and the state vector Ψ_s changes with time according to the Schrödinger equation

$$i\frac{\partial \Psi_s(t)}{\partial t} = H\Psi_s(t) \tag{8.1}$$

Assuming that H is time-independent, we have the formal solution

$$\Psi_s(t) = e^{-itH}\Psi_s(0) \tag{8.2}$$

The matrix element of an operator O_s evolves in time according to

$$\langle \Phi_s(t)|O_s|\Psi_s(t)\rangle = \langle \Phi_s(0)|e^{itH}O_s e^{-itH}|\Psi_s(0)\rangle \tag{8.3}$$

The subscript "s" identifies states and operators in the Schrödinger picture.

In the *Heisenberg picture,* the state is assumed to be constant in time, but the operators evolve. The matrix elements of an operator must be independent of the picture, and this requirement relates the Heisenberg picture to the Schrödinger picture:

$$\begin{aligned} \Psi_h &= \Psi_s(0) \\ O_h(t) &= e^{itH}O_s e^{-itH} \end{aligned} \tag{8.4}$$

The "h" subscript denotes states and operators in the Heisenberg picture. The two pictures coincide at time t_0, taken to be zero. The equation of motion in the Heisenberg picture is

$$i\frac{\partial O_h(t)}{\partial t} = [O_h(t), H] \tag{8.5}$$

If the Hamiltonian has explicit time dependence, then e^{-itH} will be replaced by a more complicated evolution operator, but Ψ_h is still defined to be constant in time.

8.2 INTERACTION PICTURE

Suppose that the Hamiltonian can be split into a "free" part and an interaction term:

$$H = H_0 + H' \tag{8.6}$$

where H_0 represents an "unperturbed" system that we understand. The two terms above are time-dependent, even if H is independent of time. By convention, the split is made at some fixed time, say, $t = 0$. Thus, H_0 and H' are constant operators by definition, as they are shorthand notations for $H_0(0)$ and $H'(0)$.

We now introduce the *interaction picture,* in which H_0 governs the time evolution of operators, while H' governs that of the state vector. In this picture, the state Ψ_ι and operators O_ι are related to those in the Schrödinger picture by

$$\begin{aligned} \Psi_\iota(t) &= e^{iH_0t}\Psi_s(t) \\ O_\iota(t) &= e^{iH_0t}O_s e^{-iH_0t} \end{aligned} \tag{8.7}$$

Thus, interaction-picture operators are the Heisenberg operators of the unperturbed system.

The equations of motion in the interaction picture can be found by substituting (8.7) into the Schrödinger equation, resulting in

$$\begin{aligned} i\frac{\partial \Psi_\iota(t)}{\partial t} &= H'(t)\Psi_\iota(t) \\ i\frac{\partial O_\iota(t)}{\partial t} &= [O_\iota(t), H_0] \end{aligned} \tag{8.8}$$

where

$$H'(t) = e^{iH_0t}H'e^{-iH_0t} \tag{8.9}$$

We define an evolution operator $U(t_2, t_1)$ through

$$\Psi_I(t) = U(t, t_0)\Psi_I(t_0) \tag{8.10}$$

It is a unitary operator, since H' is Hermitian, and the norm of $\Psi_I(t)$ is conserved:

$$U^\dagger(t, t_0)U(t, t_0) = 1 \tag{8.11}$$

The following properties are easily deduced:

$$\begin{aligned} U(t, t) &= 1 \\ U(t_1, t_0) &= U^{-1}(t_0, t_1) \\ U(t_2, t_0) &= U(t_2, t_1)U(t_1, t_0) \end{aligned} \tag{8.12}$$

The equation of motion can now be cast in the form

$$i\frac{\partial U(t, t_0)}{\partial t} = H'(t)U(t, t_0) \tag{8.13}$$

with the initial condition $U(t, t) = 1$.

It is not trivial to solve for $U(t, t_0)$, because $H'(t_1)$ and $H'(t_2)$ generally do not commute. Let us divide the time interval (t, t_0) into N small steps of duration Δt:

$$\Delta t = \frac{t - t_0}{N} \tag{8.14}$$

To first order in Δt, the equation of motion gives

$$U(t_0 + \Delta t, t_0) = [1 - iH'(t_0)\Delta t]U(t_0, t_0) \tag{8.15}$$

The last factor is, of course, unity. We built up the finite time interval $t - t_0$ from successive infinitesimal ones. Putting

$$t_n = t_0 + n\Delta t \tag{8.16}$$

we have

$$\begin{aligned} U(t, t_0) &= [1 - iH'(t_N)\Delta t][1 - iH'(t_{N-1})\Delta t] \cdots [1 - iH'(t_0)\Delta t] \\ &= 1 - i\sum_{n=0}^{N} H'(t_n) + (-i)^2 \sum_{n<m} H'(t_n)H'(t_m) \\ &\quad + (-i)^3 \sum_{n<m<l} H'(t_n)H'(t_m)H'(t_l) + \cdots \end{aligned} \tag{8.17}$$

Taking the limit $N \to \infty$ gives

$$U(t, t_0) = 1 - i\int_{t_0}^{t} dt_1 H'(t_1) + (-i)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, H'(t_1)H'(t_2) + \cdots$$

$$+ (-i)^n \int_{t_0}^{t} dt_1 \cdots \int_{t_n}^{t_{n-1}} dt_n [H'(t_1)\cdots H'(t_n)] + \cdots \tag{8.18}$$

This is a formal solution to the equation of motion.

To make the integration regions more symmetric, consider, for example, the integral in the third term:

$$I_1 = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, H'(t_1)H'(t_2)$$

where the integrations extend over region I in Fig. 8.1, in which $t_1 \geq t_2$. If we integrate over region II, we would have

$$I_2 = \int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1\, H'(t_2)H'(t_1)$$

where $t_2 > t_1$. The ordering of the operators in the integrand is such that they stand in chronological order from right to left. Evidently $I_1 = I_2$, for they differ only by an exchange of integration variables. We can therefore replace I_1 by $(I_1 + I_2)/2$, which can be rewritten in the form

$$I_1 = \tfrac{1}{2}\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\, \mathrm{T}[H'(t_2)H'(t_1)]$$

where the time-ordering operator T is the same as that defined in (7.23):

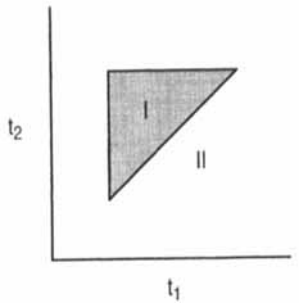


Figure 8.1 Regions of integration.

$$\mathrm{T}[AB\cdots C] = \pm AB\cdots C \qquad \text{(in chronological order from right to left)} \tag{8.19}$$

A factor -1 is attached to the result each time two fermion operators are exchanged in the rearrangement process.

For an integral with n factors in the integrand, there are $n!$ ways to permute the integration variables. Taking $1/n!$ of the sum over all possible permutations, we obtain

$$\int_{t_0}^{t} dt_1 \cdots \int_{t_n}^{t_{n-1}} dt_n [H'(t_1)\cdots H'(t_n)] = \frac{1}{n!}\int_{t_0}^{t} dt_1 \cdots \int_{t_n}^{t} dt_n \,\mathrm{T}[H'(t_1)\cdots H'(t_n)]$$

We can now write the evolution in the final form

$$U(t, t_0) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!}\int_{t_0}^{t} dt_1 \cdots \int_{t_n}^{t} dt_n \,\mathrm{T}[H'(t_1)\cdots H'(t_n)] \tag{8.20}$$

which can also be written in a shorthand notation as

$$U(t, t_0) = \mathrm{T}e^{-i\int_{t_0}^{t} dt_1 \, H'(t_1)} \tag{8.21}$$

In this form, we should expand the exponential in a power series, and then apply the operator T to each term.

8.3 ADIABATIC SWITCHING

The operator $U(t, t_0)$ contains an exponential factor that oscillates rapidly when $t \to \pm\infty$. To make the expression well defined, we assume that the interaction Hamiltonian H' was "switched on" very slowly from the distant past, and will be "switched off" very slowly in the distant future. Specifically, we replace H' with

$$H'_\varepsilon \equiv H' e^{-\varepsilon|t|} \qquad (\varepsilon \to 0^+) \tag{8.22}$$

The limit $\varepsilon \to 0^+$ is to be taken last, after all calculations. The corresponding evolution operator is denoted

$$U_\varepsilon(t, t_0) = \mathrm{T}e^{-i\int_{t_0}^{t} dt_1 \, H'_\varepsilon(t_1)} \tag{8.23}$$

which approaches the S matrix, (or S operator) when $t_0 \to -\infty$, $t \to \infty$:

$$S \equiv \lim_{\varepsilon\to 0} U_\varepsilon(\infty, -\infty) = \lim_{\varepsilon\to 0} \mathrm{T}e^{-i\int_{-\infty}^{\infty} dt \, H'_\varepsilon(t)} \tag{8.24}$$

This is a unitary operator:

$$S^\dagger S = 1 \tag{8.25}$$

and is central to the theory of scattering, which we shall discuss later.

Let us put adiabatic switching to use. We assume that the vacuum state in the Heisenberg picture $|0\rangle$ is unique and normalizable:

$$\langle 0|0\rangle = 1 \tag{8.26}$$

The interaction-picture vacuum state $|\Psi_0(t)\rangle$ evolves according to the interaction Hamiltonian, with initial condition $|\Psi_0(0)\rangle = |0\rangle$, up to a phase factor. By switching off the interactions, we make

$$\lim_{t\to-\infty} |\Psi_0(t)\rangle = |0\rangle \tag{8.27}$$

which defined the phase. In the infinite future we have

$$\lim_{t\to\infty} |\Psi_0(t)\rangle = S|0\rangle = e^{iL}|0\rangle \tag{8.28}$$

where L is a real number. Taking the scalar product of the preceding with $\langle 0|$, we obtain

$$e^{iL} = \langle 0|S|0\rangle \tag{8.29}$$

The relation between the interaction and Heisenberg pictures can be obtained from (8.7) and (8.4):

$$\begin{aligned} \Psi_\iota(t) &= e^{iH_0t}e^{-iH_\varepsilon t}\Psi_\mathrm{h} \\ O_\iota(t) &= e^{iH_0t}e^{-iH_\varepsilon t}O_\mathrm{h}(t)e^{-iH_\varepsilon t}e^{iH_0t} \end{aligned} \tag{8.30}$$

where

$$H_\varepsilon = H_0 + H'e^{-\varepsilon|t|} \tag{8.31}$$

Substituting this into (8.10), we obtain an explicit form for the evolution operator:

$$U_\varepsilon(t, t_0) = e^{iH_0t}e^{-iH_\varepsilon(t-t_0)}e^{-iH_0t_0} \tag{8.32}$$

We can rewrite (8.30) as

$$\begin{aligned} \Psi_\iota(t) &= U_\varepsilon(t, 0)\Psi_\mathrm{h} \\ O_\iota(t) &= U_\varepsilon(t, 0)O_\mathrm{h}(t)U_\varepsilon^{-1}(t, 0) \end{aligned} \tag{8.33}$$

The initial time $t = 0$ can be changed to $t = t_0$ by using the identity $U_\varepsilon(t, 0) = U_\varepsilon(t, t_0)U_\varepsilon(t_0, 0)$ to write

$$\Psi_\iota(t) = U_\varepsilon(t, t_0)U_\varepsilon(t_0, 0)\Psi_\mathrm{h} \tag{8.34}$$

Owing to adiabatic switching, we obtain

$$U_\varepsilon(t_0, 0) = e^{iH_0t_0}e^{-iH_\varepsilon t_0} \underset{t_0\to\infty}{\longrightarrow} 1 \tag{8.35}$$

Therefore

$$\begin{aligned}\Psi_\iota(t) &= U_\varepsilon(t, -\infty)\Psi_\mathrm{h}\\ O_\iota(t) &= U_\varepsilon(t, -\infty)O_\mathrm{h}(t)U_\varepsilon^{-1}(t, -\infty)\end{aligned} \tag{8.36}$$

This formula is more convenient to use than (8.33), because the two pictures coincide at $t \to -\infty$, when the system was being prepared. As we shall see, perturbation theory relies on this relation.

Why is it legitimate to use adiabatic switching? In the theory of scattering, the answer is that it reflects what happens in the laboratory. An incident particle, in the form of a wave packet, approached the target particle, but did not interact with it, until the collision took place. After the collision, the scattered particles eventually leave the scattering region as nonoverlapping wave packets, and there is no interaction after that time. The adiabatic switching is a mathematical simulation of the situation, justified by the fact that the scattering cross section is insensitive to the detailed shapes of the wave packets.

Actually, adiabatic switching need not correspond to any physical process; in general, it serves a more abstract function. When we describe a physical process by saying "A interacts with B," we have to define what is meant by A and B, and to do that, we must conceptually turn off the interaction. In the theory of metals, for example, it is customary to apply adiabatic switching to the electron–phonon interaction, which is, of course, always present. What we are doing is to imagine that the system was "assembled" by putting free electrons into an idealized lattice. As long as the conceptual assemblage took place a long time ago, it should not make any difference whether it actually happened; but we need it to identify the "free parts."

As long as we accept the crucial formula (8.36), adiabatic switching has done its job, and need not be invoked again.

8.4 CORRELATION FUNCTIONS IN THE INTERACTION PICTURE

A vacuum correlation function is the vacuum expectation value of a time-ordered product of Heisenberg field operators:

$$G_n(x_1, \ldots, x_n) = \langle 0|\mathrm{T}A_\mathrm{h}(x_1)B_\mathrm{h}(x_2)\cdots C_\mathrm{h}(x_n)|0\rangle \tag{8.37}$$

where $|0\rangle$ is the exact vacuum state, normalized such that $\langle 0|0\rangle = 1$. They are also

called *vacuum Green's functions*, because they satisfy differential equations with a delta-function source. But "correlation function" more closely describes their role in our applications.

In the interaction picture, we can use adiabatic switching to "assemble" the correlation functions of the interacting system from those of the unperturbed system, and this leads to perturbation theory. Consider a correlation function involving two operators:

$$G(t_1, t_2) = \langle 0|T A_h(t_1)B_h(t_2)|0\rangle \tag{8.38}$$

where we display only the time dependences; the spatial coordinates remain fixed. Our objective is to reexpress the above in terms of interaction-picture operators.

Consider first the time-ordering $t_1 > t_2$. From (8.36), we have

$$A_h(t) = U^{-1}(t, -\infty)A_i(t)U(t, -\infty) = U(-\infty, t)A_i(t)U^{-1}(-\infty, t) \tag{8.39}$$

Thus,

$$\begin{aligned} G(t_1, t_2) &= \langle 0|U(-\infty, t_1)A_i(t_1)U^{-1}(-\infty, t_1)U^{-1}(t_2, -\infty)B_i(t_2)U(t_2, -\infty)|0\rangle \\ &= \langle 0|U(-\infty, t_1)A_i(t_1)U(t_1, t_2)B_i(t_2)U(t_2, -\infty)|0\rangle \end{aligned} \tag{8.40}$$

where we have used the relation $U^{-1}(-\infty, t_1)U^{-1}(t_2, -\infty) = U(t_1, -\infty)U(-\infty, t_2) = U(t_1, t_2)$. The time development is represented in Fig. 8.2*a* by a contour beginning from the far right at $t = -\infty$, going to t_2, then to t_1, and then back to $-\infty$. This contour can be deformed to two straight paths from $-\infty$ to ∞, and from ∞ to $-\infty$, by rewriting

$$U(-\infty, t_1) = U(-\infty, \infty)U(\infty, t_1) = S^{-1}U(\infty, t_1) \tag{8.41}$$

Therefore

$$G(t_1, t_2) = \langle 0|S^{-1}U(\infty, t_1)A_i(t_1)U(t_1, t_2)B_i(t_2)U(t_2, -\infty)|0\rangle \tag{8.42}$$

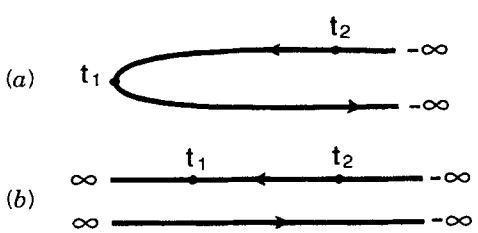


Figure 8.2 (*a*) Contour of time evolution; (*b*) contour deformed into two branches. Because of the uniqueness of the vacuum state, the lower branch gives only a phase factor.

The new time contour is illustrated in Fig. 8.2*b*. The returning branch of the contour, represented by the operator S^{-1}, contributes a phase factor when acting on the vacuum, for, using (8.28) and (8.29), we can write

$$\langle 0|S^{-1} = e^{-iL}\langle 0| = \frac{1}{\langle 0|S|0\rangle}\langle 0| \tag{8.43}$$

Therefore

$$G(t_1, t_2) = \frac{\langle 0|U(\infty, t_1)A_\iota(t_1)U(t_1, t_2)B_\iota(t_2)U(t_2, -\infty)|0\rangle}{\langle 0|S|0\rangle} \tag{8.44}$$

The time contour now consists of the path from $-\infty$ to ∞ only. The trick to straighten the contour works because the vacuum state is unique. If the vacuum expectation value were replaced by an ensemble average, then we would be stuck with a contour that comes from $-\infty$ and returns to $-\infty$. A technique to deal with such a case has been developed by Keldysh [1], but we do not need it here.

We can simplify (8.44) further, by examining more closely the operator

$$\begin{aligned} X &\equiv U(\infty, t_1)A_\iota(t_1)U(t_1, t_2)B_\iota(t_2)U(t_2, -\infty) \\ &= [\mathrm{T}e^{-i\int_{t_1}^{\infty} d\tau\, H'(\tau)}]\, A_\iota(t_1)[\mathrm{T}e^{-i\int_{t_2}^{t_1} d\tau\, H'(\tau)}]\, B_\iota(t_2)[\mathrm{T}e^{-i\int_{-\infty}^{t_2} d\tau\, H'(\tau)}] \end{aligned} \tag{8.45}$$

If all the exponential factors are expanded out, we shall have a sum of terms, each containing products of the H' bunched into three groups, of the form

$$\mathrm{T}\{\mathrm{T}[H'(\tau_1)H'(\tau_2)\cdots]A_\iota(t_1)\mathrm{T}[H'(\tau_1')H'(\tau_2')\cdots]B_\iota(t_2)\mathrm{T}[H'(\tau_1'')H'(\tau_2'')\cdots]\} \tag{8.46}$$

where we have supplied a redundant T in front, not needed because $t_1 > t_2$. With it in place, however, we are free to rearrange *all* the H' factors in arbitrary order. The condition $t_1 > t_2$, which determines the relative position of $A_\iota(t_1)$ and $B_\iota(t_2)$, ensures that the factors cannot wind up in the wrong group when the overall time ordering is enforced. As a shorthand notation, therefore, we can write

$$X = \mathrm{T}e^{-i\int_{-\infty}^{\infty} d\tau\, H'(\tau)}A_\iota(t_1)B_\iota(t_2) \tag{8.47}$$

The correlation function can then be written in a more compact notation:

$$G(t_1, t_2) = \frac{\langle 0|\mathrm{T}SA_\iota(t_1)B_\iota(t_2)|0\rangle}{\langle 0|S|0\rangle} \tag{8.48}$$

Although derived under the assumption $t_1 > t_2$, this result is also valid for $t_1 < t_2$, as one can easily verify by repeating the derivation. The preceding results can be immediately generalized:

$$G_n(x_1, x_2, \ldots, x_n) = \frac{\langle 0|\mathrm{T}SA_{\mathrm{I}}(x_1)B_{\mathrm{I}}(x_2)\cdots C_{\mathrm{I}}(x_n)|0\rangle}{\langle 0|S|0\rangle} \tag{8.49}$$

This is convenient for calculations, because the interaction-picture operators evolve like free fields.

8.5 *S* MATRIX AND SCATTERING

In a scattering experiment, an incident beam of particles A falls on a target particle B, producing a number of outgoing particles $C, \ldots, D$, which are detected by devices that measure their momenta, and possibly other quantum numbers. We indicate the reaction by

$$A + B \rightarrow C + \cdots + D$$

The states of the particles are labeled by momentum, mass, spin, and other single-particle quantum numbers.

The wave functions of A and B are wave packets that do not overlap initially. When they eventually overlap, the reaction takes place, and after a short time the final particles emerge from the interaction region as wave packets that recede from one another, eventually becoming well separated.

The size of the interaction regime is microscopic, and by comparison the wave packet is practically a plane wave. In our theoretical treatment, we idealize the wave packets as plane waves; but only before the interaction began, and after the interaction is over. The is done mathematically via adiabatic switching.

Let us denote the initial and final free-particle states by Φ_{i} and Φ_{f}, respectively. In our idealized scattering experiment, Φ_{i} is the state prepared in the infinite past, and Φ_{f} is the state for which the detectors are set up to detect in the infinite future. The actual state of the system evolves from Φ_{i} into some interaction-picture state at time t_0, which is denoted by

$$\Psi_{\mathrm{i}}^{(+)}(t_0) = U(t_0, -\infty)\Phi_{\mathrm{i}} \tag{8.50}$$

The superscript (+) indicates that eventually final particles will emerge as outgoing spherical waves. In the infinite future, the particle detectors look for plane waves Φ_{f} instead of spherical waves. The state that will become Φ_{f} is, at the time t_0, given by

$$\Psi_{\mathrm{f}}^{(-)}(t_0) = U(t_0, \infty)\Phi_{\mathrm{f}} \tag{8.51}$$

where the superscript (–) indicates that it consists of incoming spherical waves in the past. The overlap between $\Psi_{\mathrm{f}}^{(-)}$ and $\Psi_{\mathrm{i}}^{(+)}$ is the probability amplitude that the scattering process takes place:

$$(\Psi_f^{(-)}, \Psi_i^{(+)}) = (\Phi_i, U^{-1}(t_0, \infty)U(t_0, -\infty)\Phi_i) = (\Phi_i, U(\infty, -\infty)\Phi_i) \tag{8.52}$$

This defines the S matrix:

$$\langle f|S|i\rangle \equiv (\Psi_f^{(-)}, \Psi_i^{(+)}) \tag{8.53}$$

where $|i\rangle$ and $|f\rangle$ are free-particle states. Thus, as anticipated earlier,

$$S = U(\infty, -\infty) = \mathrm{T}e^{-i\int_{-\infty}^{\infty} dt\, H'(t)} \tag{8.54}$$

The process is illustrated in Fig. 8.3.

8.6 SCATTERING CROSS SECTION

If $S = 1$, then the final state is the same as the initial state, and no actual scattering occurs. The probability amplitude that a reaction occurs is therefore the matrix ele-

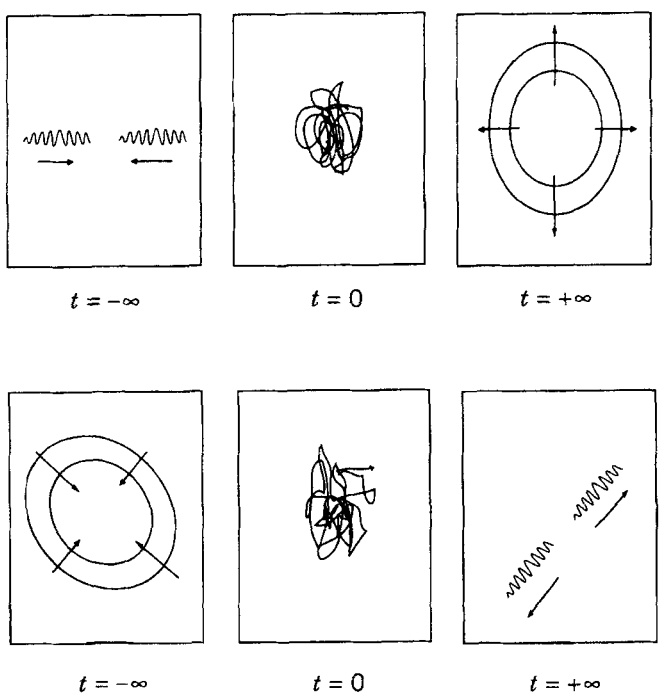


Figure 8.3 Upper panel shows time evolution of $\Psi^{(+)}$ and lower panel, that of $\Psi^{(-)}$. The S matrix is the overlap of the two states at $t = 0$.

ment of $S-1$. By analogy with potential scattering, we define the T matrix through the relation

$$\langle f|(S-1)|i\rangle = -i(2\pi)^4\delta^4(P_f - P_i)\langle f|T|i\rangle \tag{8.55}$$

where P_f and P_i are respectively the total 4-momentum in the final and initial state. The factor $\delta^4(P_f - P_i)$ expresses conservation of total momentum and energy. The energy conserving factor $2\pi\delta(P_f^0 - P_i^0)$ is familiar from potential scattering; it comes from the assumption that the total Hamiltonian is time-independent, and that the process was observed over a long time. The momentum-conserving factor $(2\pi)^3\delta^3(P_f - P_i)$ was extracted from the transition matrix element, under the assumption that the system is invariant under spatial translations.

The transition probability for the reaction $i \to f$ is given by

$$|\langle f|(S-1)|i\rangle|^2 = (2\pi)^4\delta^4(0)(2\pi)^4\delta^4(P_f - P_i)|\langle f|T|i\rangle|^2 \tag{8.56}$$

The factor $(2\pi)^4\delta^4(0)$ should be interpreted as the total volume of space–time, specifically, the limit

$$(2\pi)^4\delta^4(P) = \int d^4x e^{iP\cdot x} \underset{P\to 0}{\to} \int d^4x \tag{8.57}$$

We omit this factor to obtain the transition rate per unit volume.

The final state f, which must lie in a continuum of states, is contained in an infinitesimal neighborhood $d\Omega$ of some state. For example, we may specify that final particles are scattered into specific solid-angle elements. The transition rate per unit volume into $d\Omega$ is given by

$$dR = \sum_{f\in d\Omega} (2\pi)^4\delta^4(P_f - P_i)|\langle f|T|i\rangle|^2 = I d\sigma \tag{8.58}$$

This defines the differential cross section $d\sigma$, with the incident flux I given by

$$I = n_1 n_2 v_{12} \tag{8.59}$$

where v_{12} is the relative velocity of the colliding partners and n_1 and n_2 are their spatial densities. For the decay of an unstable particle from the initial state i, the lifetime τ is given by

$$\frac{1}{\tau} = \sum_f (2\pi)^4\delta^4(P_f - P_i)|\langle f|T|i\rangle|^2 \tag{8.60}$$

where the sum extends over all possible final states f.

The S matrix is unitary: $S^\dagger S = 1$. In terms of the T matrix, this means

$$\langle f|(T-T^\dagger)|i\rangle = 2i\sum_n (2\pi)^4\delta^4(P_i-P_n)\langle f|T^\dagger|n\rangle\langle n|T^\dagger|i\rangle \tag{8.61}$$

Putting $f = i$, we obtain the *optical theorem:*

$$\mathrm{Im}\langle i|(T-T^\dagger)|i\rangle = \sum_n (2\pi)^4\delta^4(P_i-P_n)|\langle i|T^\dagger|n\rangle|^2 = I\sigma_{\mathrm{tot}} \tag{8.62}$$

where σ_{tot} is the total cross section.

The normalization of particle wave functions affects the matrix element $\langle f|T|i\rangle$, phase space volume elements, and the incident flux I; but the cross section is independent of the convention, as long as it is used consistently. A wave function normalized to n particles per unit volume has the form

$$\psi(\mathbf{r}) = \begin{cases} \sqrt{n}e^{i\mathbf{p}\cdot\mathbf{r}} & \text{(boson)} \\ \sqrt{n}e^{i\mathbf{p}\cdot\mathbf{r}}u(\mathbf{p},s) & \text{(fermion)} \end{cases} \tag{8.63}$$

where $\bar{u}u$ is 1 for particle, and -1 for antiparticle. This corresponds to a single-particle state $|\mathbf{p}\rangle$ with the properties

$$\langle \mathbf{p}|\mathbf{p}'\rangle = n(2\pi)^3\delta^3(\mathbf{p}-\mathbf{p}')$$
$$\int \frac{d^3p}{(2\pi)^3 n}|\mathbf{p}\rangle\langle\mathbf{p}| = 1 \tag{8.64}$$

Whatever we choose for n, the convention must be adhered to in the expansion of field operators into creation and annihilation operators.

The convention used in this book is $n = 1$. With this, the creation or annihilation of a boson or fermion is accompanied respectively by a factor $(2E)^{-1/2}$ or $(m/E)^{-1/2}$. This is indicated in the field expansions (2.29) and (7.16), respectively, and originates from the fact that the boson Lagrangian is quadratic in the time derivative, whereas the fermion Lagrangian is linear. If we take

$$n = \begin{cases} 2E & \text{(boson)} \\ \dfrac{E}{m} & \text{(fermion)} \end{cases} \tag{8.65}$$

then these factors disappear in the field expansions, but reappear in the properties of $|\mathbf{p}\rangle$. This convention is call *invariant normalization*, because it has the same form in all Lorentz frames. In practice, it makes little difference which convention we choose, for these factors appear either with the matrix element, or squared in the phase-space volume element.. It is merely a matter of when to include them.

As an illustration, consider a reaction in which two particles go into N particles, with initial and final states given by

$$|i\rangle = |\mathbf{p}_1, \mathbf{p}_2\rangle$$

$$|f\rangle = |\mathbf{p}_1', \mathbf{p}_2' \cdots \mathbf{p}_N'\rangle \tag{8.66}$$

With the convention $n = 1$, the incident flux is

$$I = \left| \frac{\mathbf{p}_1}{E_1} - \frac{\mathbf{p}_2}{E_2} \right| = \frac{1}{E_1 E_2} \sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2} \tag{8.67}$$

where p denotes 4-momentum and E and m denote energy and mass, respectively. The momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ are assumed to be collinear. The differential cross section $d\sigma$ is given by

$$Id\sigma = \prod_{j=1}^{N} \frac{d^3 p_j'}{(2\pi)^3} (2\pi)^4 \delta^4(P_i - P_f) |\langle f|T|i\rangle|^2 \tag{8.68}$$

where P_i and P_f are respectively the total 4-momentum of the initial and final state. The wave function factors are contained in $\langle f|T|i\rangle$. With invariant normalization, we take them out of the matrix element, and put their squares under $d^3 p_j'$. The transition rate $Id\sigma$ is not changed.

8.7 POTENTIAL SCATTERING

To help understand scattering in quantum field theory, we give a parallel description of potential scattering in nonrelativistic quantum mechanics. The experimental situation is that a steady beam of particles is scattered by a potential well, and detectors are set up very far away to register scattered particles of definite momenta. The incident beam is represented by a stream of nonoverlapping wave packets, which can be considered one at a time. A wave packet must be small enough that it does not overlap the scattering center initially, but large enough that its momentum can be well defined. In this sense, we can replace the wave packet by a plane wave.

In the plane-wave limit, the overall scattering process is described by the stationary wave function $\psi(\mathbf{r})$ of the particle, which satisfies the Schrödinger equation with outgoing-wave boundary condition:

$$\left[-\frac{1}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi(r) = E\psi(\mathbf{r})$$

$$\psi(r) \xrightarrow[r \to \infty]{} e^{i\mathbf{k}_0 \cdot \mathbf{r}} + f(\Omega) \frac{e^{ik_0 r}}{r} \tag{8.69}$$

where $V(\mathbf{r})$ is the scattering potential and Ω denotes the scattering angles. We have written $r = |\mathbf{r}|$, $k_0 = |\mathbf{k}_0|$, and $E = k_0^2/2m$. The incident wave corresponds to $\exp(i\mathbf{k}_0 \cdot \mathbf{r})$, and the outgoing spherical wave corresponds to $\exp(ik_0 r)/r$. We can calculate the number of particles scattered per unit time into solid angle $d\Omega$ from the current density of the scattered wave. The differential cross section $d\sigma/d\Omega$ is the preceding rate

per unit solid angle, divided by the incident flux. This procedure yields the familiar result

$$\frac{d\sigma}{d\Omega} = |f(\Omega)|^2 \tag{8.70}$$

where $f(\Omega)$ is called the *scattering amplitude.* The wave packet nature was taken into account through the neglect of interference between incident and scattered wave.

In the laboratory, scattering is a time-dependent process, which can be described by the stationary wave function $\psi^{(+)}$ through the following construction. Let the initial wave packet be

$$F(\mathbf{r}) = \int d^3k f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}$$

where $f(\mathbf{k})$ is a function peaked about $\mathbf{k} = \mathbf{k}_0$. Then the time-dependent wave function describing the actual scattering process is

$$\Psi(\mathbf{r}, t) = \int d^3k f(\mathbf{k}) \psi_{\mathbf{k}}^{(+)} e^{-iEt}$$

where $E = k^2/2m$. A "motion picture" of this wave function will show the wave packet being scattered by the potential, receding as an outgoing spherical wave front. The stationary wave function obtains in the plane-wave limit $f(\mathbf{k}) \rightarrow \delta^3(\mathbf{k} - \mathbf{k}_0)$.

We now restate the problem in a more formal language, for comparison with the treatment in quantum field theory. Let us write the Hamiltonian as

$$H = H_0 + V \tag{8.71}$$

with $H_0 = -\nabla^2/2m$. Consider the eigenvalue problem with a continuous spectrum E:

$$(H_0 + V)\psi_{\mathbf{k}_0}^{(\pm)}(r) = E\psi_{\mathbf{k}_0}^{(\pm)}(\mathbf{r})$$

$$\psi^{(\pm)}(r) \xrightarrow[r\to\infty]{} e^{i\mathbf{k}_0\cdot\mathbf{r}} + f(\Omega)\frac{e^{\pm ik_0 r}}{r} \tag{8.72}$$

where the superscript $(\pm)$ labels a solution with outgoing (incoming) spherical wave. The unperturbed problem at the same energy E is defined by

$$H_0\phi_{\mathbf{k}_0} = E\phi_{\mathbf{k}_0} \tag{8.73}$$

where

$$\phi_{\mathbf{k}} = e^{i\mathbf{k}\cdot\mathbf{r}} \tag{8.74}$$

The Schrödinger equation and the boundary condition are incorporated in the following integral equation:

$$\psi_{\mathbf{k}_0}^{(\pm)} = \phi_{\mathbf{k}_0} + \frac{1}{E - H_0 \pm i\eta} V\psi_{\mathbf{k}_0}^{(\pm)} \tag{8.75}$$

where $\eta \to 0^+$. The term $\pm i\eta$ prescribes the way to skirt the pole $(E - H_0)^{-1}$, so as to give an outgoing or incoming wave.

We define the T matrix by

$$T\phi_{\mathbf{k}_0} \equiv V\psi_{\mathbf{k}_0}^{(+)} \tag{8.76}$$

It satisfies the Lippmann–Schwinger equation

$$T = V + V\frac{1}{E - H_0 + i\eta}T \tag{8.77}$$

and is related to the scattering amplitude through

$$f(\Omega) = -\frac{m}{2\pi}(\phi_{\mathbf{k}}, V\psi_{\mathbf{k}_0}^{(+)}) = -\frac{m}{2\pi}(\phi_{\mathbf{k}}, \mathrm{T}\phi_{\mathbf{k}_0}) \tag{8.78}$$

where $\mathbf{k}$ is the scattered wave vector, with magnitude k_0 and polar angles Ω.

The set of wave functions with the (+) and (−) boundary conditions separately form complete sets of eigenstates of H (barring bound states). The unitary transformation matrix relating the two sets is called the *S matrix:*

$$\langle \mathbf{k}|S|\mathbf{k}_0\rangle \equiv (\psi_{\mathbf{k}_0}^{(-)}, \psi_{\mathbf{k}_0}^{(+)}) \tag{8.79}$$

To find the relation between the S matrix and the T matrix, rewrite (8.75) in the more explicit form

$$\psi_{\mathbf{k}_0}^{(\pm)} = \left[1 + \frac{1}{E - H \pm i\eta} V\right]\phi_{\mathbf{k}_0} \tag{8.80}$$

which can be obtained by iterating (8.75), and resumming. In the limit $\eta \to 0^+$ we have

$$\frac{1}{E - H \pm i\eta} = \frac{P}{E - H} \mp i\pi\delta(E - H) \tag{8.81}$$

where P denotes principal part. We therefore have

$$\psi^{(+)}_{\mathbf{k}_0} - \psi^{(-)}_{\mathbf{k}_0} = -2\pi i \delta(E - H) V \phi_{\mathbf{k}_0} \tag{8.82}$$

Now take the scalar product of both sides with respect to $\psi^{(+)}_k$, noting that

$$(\psi^{(+)}_k, \psi^{(+)}_{k_0}) = \langle \mathbf{k} | \mathbf{k}_0 \rangle = (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}_0) \tag{8.83}$$

This leads to

$$\langle \mathbf{k} | S | \mathbf{k}_0 \rangle = \langle \mathbf{k} | [1 - 2\pi i \delta(E_k - E_{k0}) T] \, | \mathbf{k}_0 \rangle \tag{8.84}$$

Energy conservation is enforced in the S matrix, but not in the T matrix. The latter can be defined for arbitrary E, including complex values, through the Lippmann–Schwinger equation (8.77).

We can write as an operator relation

$$S = 1 - 2\pi i \delta(E - H) T \tag{8.85}$$

The unitarity condition $S^\dagger S = 1$ implies that

$$T^\dagger - T = 2\pi i T^\dagger \delta(E - H) T \tag{8.86}$$

The diagonal matrix element of this relation gives the optical theorem:

$$\sigma_{\text{tot}} = \frac{4\pi}{k} \operatorname{Im} f(0) \tag{8.87}$$

where σ_{tot} is the total cross section for incident wave number k and $f(0)$ is the scattering amplitude in the forward direction.

8.8 ADIABATIC THEOREM

The meaning of the S matrix rests on the statements (8.27) and (8.28), specifically that the vacuum state in the interaction picture approaches that of the Heisenberg picture in the infinite past and in the infinite future. The two limiting states can differ only by a phase factor, which defines the S matrix. The assumption was that the system did not get excited from the ground state. A formal statement is provided by the *adiabatic theorem*, which can be loosely stated as follows:

A system being perturbed in its ground state will remain in the ground state, as long as the perturbation varies sufficiently slowly with time.

A precise mathematical statement of the theorem will be given later. We write the Hamiltonian in the form

$$H = H_0 + \lambda H' \tag{8.88}$$

where H' is the interaction to be turned on and off and λ is a parameter introduced for convenience, to be set to unity at the end of the calculations. The split of H into the two terms is made at time $t = 0$. We enclose the system in a large but finite box, and assume that the eigenvalues of H are discrete and nondegenerate.

Suppose that Ψ is an eigenstate of H, and Φ the corresponding unperturbed state:

$$\begin{aligned} H\Psi &= E\Psi \\ H_0\Phi &= \epsilon\Phi \\ (\Phi, \Phi) &= 1 \end{aligned} \tag{8.89}$$

In stationary perturbation theory, the relation between Ψ and Φ can be found as follows. We write

$$\Psi = \Phi + \chi \tag{8.90}$$

where $(\chi, \Phi) = 0$. Note that Ψ is not normalized to 1. It simplifies the analysis to take the coefficient of Φ to be unity. Substituting this equation into $H\Psi = E\Psi$, we obtain, after some rearrangement,

$$(E - \epsilon)\Phi = (H_0 - E)\chi - \lambda H'\Psi \tag{8.91}$$

Taking the scalar product of this equation with Φ, we obtain an expression for the perturbed energy:

$$E = \epsilon + \lambda (\Phi, H'\Psi) \tag{8.92}$$

The perturbed state Ψ is found by solving for χ from (8.91). To ensure that the result is orthogonal to Φ, we first multiply both sides of (8.91) by the projection operator Q onto the space orthogonal to Φ:

$$\begin{aligned} Q\Phi &= 0 \\ Q\chi &= \chi \end{aligned} \tag{8.93}$$

We then obtain the result

$$\Psi = \Phi - \lambda \frac{1}{H_0 - E} QH'\Psi \tag{8.94}$$

The following is an equivalent form of the equation with the unperturbed energy ϵ in the denominator:

$$\Psi = \Phi - \lambda \frac{1}{H_0 - \epsilon} Q(H' + \epsilon - E)\Psi \tag{8.95}$$

The formulas for E and Ψ are valid for any state, as long as the energy levels do not cross when λ varies from 0 to 1. In particular, we use them for the vacuum state, which is assumed to be unique for all λ.

With adiabatic switching, the Hamiltonian is taken to be

$$H_\varepsilon = H_0 + \lambda\, e^{-\varepsilon|t|} H' \qquad (\varepsilon \to 0^+) \tag{8.96}$$

and the time-development operator in the interaction picture is given by

$$U_\varepsilon \equiv U_\varepsilon(0, -\infty) = \sum_{n=0}^{\infty} \frac{(-i\lambda)^n}{n!} \int_{-\infty}^{0} dt_1 \cdots \int_{-\infty}^{0} dt_n\, e^{\varepsilon(t_1 + \cdots + t_n)} \mathrm{T}[H'(t_1) \cdots H'(t_n)] \tag{8.97}$$

Let us calculate $[H_0, U_\varepsilon]$. For the commutator $[H_0, H'(t)]$, we use the formula

$$[H_0, H'(t)] = -i \frac{\partial H'}{\partial t} \tag{8.98}$$

which can be deduced from (8.9). Thus

$$\begin{aligned}
[H_0, U_\varepsilon] &= -i \sum_{n=0}^{\infty} \frac{(-i\lambda)^n}{n!} \int_{-\infty}^{0} dt_1 \cdots \int_{-\infty}^{0} dt_n\, e^{\varepsilon(t_1 + \cdots + t_n)} \left(\frac{\partial}{\partial t_1} + \cdots + \frac{\partial}{\partial t_n} \right) \\
&\quad \times \mathrm{T}[H'(t_1) \cdots H'(t_n)] \\
&= -i \sum_{n=1}^{\infty} \frac{(-i\lambda)^n n}{n!} \int_{-\infty}^{0} dt_1 \cdots \int_{-\infty}^{0} dt_n\, e^{\varepsilon(t_1 + \cdots + t_n)} \frac{\partial}{\partial t_1} \mathrm{T}[H'(t_1) \cdots H'(t_n)] \\
&= -H'(0)U + \varepsilon\lambda \sum_{n=1}^{\infty} \frac{(-i\lambda)^{n-1}}{(n-1)!} \int_{-\infty}^{0} dt_1 \cdots \int_{-\infty}^{0} dt_n\, e^{\varepsilon(t_1 + \cdots + t_n)} \\
&\quad \times \mathrm{T}[H'(t_1) \cdots H'(t_n)]
\end{aligned} \tag{8.99}$$

where, in the last step, we made a partial integration with respect to t_1. The above can be rewritten as

$$[H_0, U_\varepsilon] = -H'(0)U_\varepsilon + i\varepsilon\lambda \frac{\partial U_\varepsilon}{\partial \lambda} \tag{8.100}$$

or

$$U_\varepsilon H_0 U_\varepsilon^{-1} = H_\varepsilon - i\varepsilon\lambda \frac{\partial U_\varepsilon}{\partial \lambda} U_\varepsilon^{-1} \tag{8.101}$$

This shows that U_ε transforms the unperturbed Hamiltonian into the perturbed one. We can now give a more precise statement of the adiabatic theorem [2].

■ **Adiabatic Theorem** *Let*

$$X_\varepsilon \equiv \frac{U_\varepsilon \Phi}{(\Phi, U_\varepsilon \Phi)} \tag{8.102}$$

where $\mathrm{H}_0\Phi = \epsilon\Phi$. *Then*

$$\lim_{\varepsilon \to 0} X_\varepsilon = c\Psi \tag{8.103}$$

where Ψ *is the eigenstate of the Hamiltonian corresponding to the unperturbed state* Φ, *and* c *is a normalization constant.*

Proof Let

$$\Psi_\varepsilon = U_\varepsilon \Phi \tag{8.104}$$

Consider

$$(H_0 - \epsilon)\Psi_\varepsilon = (H_0 - \epsilon)U_\varepsilon \Phi = [H_0, U_\varepsilon]\Phi \tag{8.105}$$

Using (8.100), we obtain

$$(H_\varepsilon - \epsilon)\Psi_\varepsilon = i\varepsilon\lambda \frac{\partial \Psi_\varepsilon}{\partial \lambda} \tag{8.106}$$

Dividing both sides by $(\Phi, U_\varepsilon \Phi)$, we can rearrange the preceding to read

$$\left(H_\varepsilon - \epsilon - i\varepsilon\lambda \frac{\partial}{\partial \lambda}\right) X_\varepsilon = i\varepsilon X_\varepsilon \lambda \frac{\partial \ln(\Phi, \Psi_\varepsilon)}{\partial \lambda} \tag{8.107}$$

In the limit $\varepsilon \to 0$ this becomes

$$HX_0 = (\epsilon + \Delta E)X_0 \tag{8.108}$$

where

$$\Delta E = \lim_{\varepsilon \to 0} i\varepsilon\lambda \frac{\partial \ln(\Phi, \Psi_\varepsilon)}{\partial \lambda} \tag{8.109}$$

Thus, X_0 is an eigenstate of H. To show that it has the correct eigenvalue, we take the scalar product of (8.106) with Φ to obtain

$$(\Phi, (H_\varepsilon - \epsilon)\Psi) = i\varepsilon\lambda \frac{\partial(\Phi, \Psi_\varepsilon)}{\partial\lambda} \tag{8.110}$$

Dividing both sides by $(\Phi, U\Phi)$, we can rewrite this in the form

$$i\varepsilon\lambda \frac{\partial \ln(\Phi, \Psi_\varepsilon)}{\partial\lambda} = (\Phi, (H_\varepsilon - \epsilon)X_\varepsilon) = \lambda\,(\Phi, H'X_\varepsilon)e^{-\varepsilon|t|} \tag{8.111}$$

Comparison with (8.109) shows that

$$\Delta E = \lambda(\Phi, H'X_0) \tag{8.112}$$

This completes the proof that X_0 is the eigenstate of H with the correct eigenvalue (8.92). ∎

PROBLEMS

8.1 Show that the term $\pm i\eta$ in the operator $(E - H_0 \pm i\eta)^{-1}$ in (8.75) corresponds respectively to an outgoing (+) and incoming (−) spherical wave, as defined in (8.72).

8.2 **(a)** Consider the scattering of a charged particle by the Coulomb potential of an atomic nucleus, as in Rutherford scattering. The scattering amplitude involves the Fourier transform of $1/r$, which is ambiguous because of the lack of convergence at large r. Show that by introducing a screening factor $e^{-\alpha r}$, one obtains a definite scattering cross section in the limit $\alpha \to 0$.

(b) The screening is a mathematical device like adiabatic switching. Why does it make sense physically in this case?

(c) Can you think of circumstances in which the screening device should not be used, because it would correspond to the wrong physics? (*Hint:* Would you get plasma oscillations if the Coulomb potential were screened?)

8.3 The time-reversal operation $\mathcal{T}$ defined in Section 7.8 interchanges initial and final state. Show that we must have

$$\mathcal{T}S\mathcal{T} = S^{-1} = S^\dagger$$

and this implies

$$(\Psi_b^{(-)}, \Psi_a^{(+)}) = (\Psi_{\bar{a}}^{(+)}, \Psi_{\bar{b}}^{(-)})$$

where $\bar{a}$ denotes the time-reversed quantum numbers corresponding to a.

8.4 From (8.32), we have the relation $U(t, 0) = e^{iH_0t}e^{-iHt}$, or

$$e^{-iHt} = e^{-iH_0t}\mathrm{T}\, e^{-i\int_0^t dt_1 H'(t_1)}$$

where $H = H_0 + H'$, and $H'(t) = e^{iH_0t}H'e^{-iH_0t}$. Obtain from this the expansion formula

$$e^{A+B} = e^{A} + \int_0^1 d\lambda_1\, e^{(1-\lambda_1)A} B\, e^{\lambda_1 A} + \frac{1}{2!}\int_0^1 d\lambda_1 \int_0^1 d\lambda_2\, e^{(1-\lambda_1)A} B\, e^{(\lambda_1-\lambda_2)A} B\, e^{\lambda_2 A} + \cdots$$

8.5 **(a)** From (8.101) show that the ground-state energy is given by

$$E = \epsilon + i\lambda\left(\Phi, U^{-1}\frac{\partial U}{\partial \lambda}\Phi\right)$$

where the symbols are as defined in Section 8.8 and we have set the adiabatic switching parameter $\varepsilon = 0$.

(b) Expand E in a perturbation series in powers of λ. (Eventually $\lambda \to 1$.)

(c) Under what conditions can this formula be used for an excited state?

REFERENCES

1. L. *V.* Keldysh, *Zh. Eksp. Teor. Fiz.***47,** 1515 (1964) [*Sov. Phys. JETP* **20,** 235 (1965)].
2. M. Gell-Mann and F. E. Low, *Phys. Rev.* **84** 350 (1951).

CHAPTER NINE

Feynman Graphs

9.1 PERTURBATION THEORY

We have obtained the S matrix as an expansion in powers of the interaction Hamiltonian. Perturbation theory is based on this expansion, considered as an asymptotic expansion, whose first few terms can give very good approximations. Whether the expansion actually converges is usually unknown. The value of perturbation theory, however, goes beyond obtaining good numerical answers, for the following reasons:

- The formal sum of the perturbation series, or even a partial sum, can reveal interesting properties of the theory, as, for example, the renormalizability of quantum electrodynamics.
- Divergence of the perturbation series usually signals "nonperturbative" effects, such as the formation of bound states or solitons, and the onset of a phase transition.

In this chapter, we illustrate the use of perturbation theory in the calculation of S matrix elements in the ϕ^4 theory. The Lagrangian density of the theory is given by

$$\mathcal{L}(x) = \frac{1}{2}\partial^\mu\phi(x)\partial^\mu\phi(x) - \frac{m_0^2}{2}\phi^2(x) - \frac{\lambda_0}{4!}\phi^4(x) \tag{9.1}$$

where the mass m_0 and coupling constant λ_0 are called "bare" or "unrenormalized" constants. They differ from the physical mass and coupling constant, which are "renormalized" by the effects of interactions.

The unperturbed system is taken to be the free field of mass m_0, and the corresponding interaction Lagrangian density is

$$\mathcal{L}_{\text{int}}(x) = -\frac{\lambda_0}{4!}\phi^4(x) \tag{9.2}$$

The interaction Hamiltonian differs only by a sign:

$$\mathcal{H}_{\text{int}}(x) = -\frac{\lambda_0}{4!}\phi^4(x) \tag{9.3}$$

We use the interaction picture, in which the S matrix can be written as the following power series in the bare coupling constant:

$$S = \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{i\lambda_0}{4!}\right)^n \int d^4x_1 \cdots \int d^4x_n \, \mathrm{T}[\phi^4(x_1) \cdots \phi^4(x_n)] \tag{9.4}$$

where $\phi(x)$ evolves in time like a free field. Our goal is to calculate the matrix element of S between given initial and final free-particle states.

The field operator $\phi(x)$ contains two terms: a positive-frequency part that annihilates a particle, and a negative-frequency part that creates a particle. The nth order term in the S matrix is a sum of terms, each of which is a product of creation and annihilation operators. To obtain the matrix element of such a product, we first try to rearrange the order of the factors in normal order—with all annihilation operators standing to the right of all creation operators. In general, the factors involved do not commute with one another, and we will generate other terms in the rearrangement process. This is, however, a finite process, and in the end we will obtain the nth-order S matrix as a finite sum of normal products. Once this is done, the matrix elements can be read off, because a normal product has nonvanishing matrix element only between an initial state containing the particles to be annihilated, and a final state containing those to be created. In the following section we develop some mathematical tools to implement this task.

9.2 TIME-ORDERED AND NORMAL PRODUCTS

In this section, let capital letters, such as A, denote either a creation or annihilation operator for fermions or bosons. The commutator between any two boson operators, or the anticommutator between any two fermion operators, is a c-number. The same is true of the commutator between a boson and a fermion operator. Such a commutator or anticommutator may be replaced by its vacuum expectation value. We restate the definitions of time-ordered and normal-ordered products in a more formal way.

A time-ordered product (T-product) is defined as

$$\mathrm{T}(A_1 A_2 \cdots A_n) \equiv \eta_{AB} B_1 B_2 \cdots B_n \tag{9.5}$$

where $\{B_i\}$ is the same set of operators as $\{A_i\}$, except possibly in different order, such that an operator with a later time label appears to the left of one with an earlier label. The quantity η_{AB} is a signature factor:

$$\eta_{AB} = \begin{cases} -1 & \text{if } \{A\} \text{ and } \{B\} \text{ differ by an odd permutation of fermion operators} \\ +1 & \text{otherwise} \end{cases} \tag{9.6}$$

Time ordering is distributive:

$$\mathrm{T}[(ABC\cdots)+(DEF\cdots)] = \mathrm{T}(ABC\cdots)+\mathrm{T}(DEF\cdots) \tag{9.7}$$

A normal product (N-product) is defined as

$$:A_1\cdots A_n: \equiv \eta_{AC} C_1\cdots C_n \tag{9.8}$$

where $\{C_i\}$ is the same set as $\{A\}$, except possibly in different order, such that in the set $\{C_i\}$ all creation operators appear to the left of all annihilation operators. The signature factor η_{AC} has been defined in (9.6). Normal ordering is distributive:

$$:[(ABC\cdots)+(DEF\cdots)]: = :(ABC\cdots): + :(DEF\cdots): \tag{9.9}$$

We define the *contraction* between two operators as the vacuum expectation value of their product:

$$A^{\bullet}B^{\bullet} \equiv \langle 0|AB|0\rangle \tag{9.10}$$

If there are intervening operators between A and B, then

$$A^{\bullet}CB^{\bullet} \equiv \pm\langle 0|AB|0\rangle C \tag{9.11}$$

where the sign is $\pm$ depending on whether an even or odd number of fermion operators exchange position when B is pulled across C to the left. If there is more than one contraction in a product, we distinguish the different contractions using multiple dots:

$$A^{\bullet}B^{\bullet\bullet}C^{\bullet\bullet\bullet}DE^{\bullet}F^{\bullet\bullet\bullet}G^{\bullet\bullet} = \pm\langle 0|AE|0\rangle\langle 0|BG|0\rangle\langle 0|CF|0\rangle D \tag{9.12}$$

9.3 WICK'S THEOREM

The operator form of Wick's theorem tells us how to expand a T-product into N-products. A T-product orders operators according to time labels, regardless of whether they are creation or annihilation operators. An N-product, on the other hand, orders operators according to whether they are creation or annihilation operators, regardless of time label. Thus, it is sufficient to learn how to expand an ordinary product into N-products, because at fixed times a T-product is just an ordinary product, and the N-product does not care about time labels. Since a single operator

is its own N-product, the nontrivial cases start with a product of two operators. The general result will be obtained by induction on the number of factors in the product.

LEMMA 1

$$A_1A_2 = :A_1A_2: + A_1^{\bullet}A_2^{\bullet} \tag{9.13}$$

Proof

(a) For two boson operators, we have

$$:A_1A_2: = \begin{cases} A_1A_2 & \text{if } A_2 \text{ is annihilation operator} \\ A_2A_1 & \text{if } A_2 \text{ is creation operator} \end{cases} \tag{9.14}$$

In the latter case we write

$$\begin{aligned} A_2A_1 &= A_1A_2 + [A_2, A_1] = A_1A_2 + \langle 0|[A_2, A_1]|0\rangle \\ &= A_1A_2 - \langle 0|A_1A_2|0\rangle \end{aligned} \tag{9.15}$$

where the last step follows because $\langle 0|A_2 = 0$. Therefore, if A_2 is a creation operator, we have

$$A_1A_2 = :A_1A_2: + \langle 0|A_1A_2|0\rangle \tag{9.16}$$

This is also true if A_2 is an annihilation operator, for then the second term vanishes.

(b) For two fermion operators, we have

$$:A_1A_2: = \begin{cases} A_1A_2 & \text{if } A_2 \text{ is annihilation operator} \\ -A_2A_1 & \text{if } A_2 \text{ is creation operator} \end{cases} \tag{9.17}$$

In the latter case we write

$$\begin{aligned} -A_2A_1 &= A_1A_2 - \{A_2, A_1\} = A_1A_2 - \langle 0|\{A_2, A_1\}|0\rangle \\ &= A_1A_2 - \langle 0|A_1A_2|0\rangle \end{aligned} \tag{9.18}$$

which leads to the same result as in the boson case.

(c) If one operator is boson, and the other fermion, then the boson calculations apply, for the operators commute. ■

LEMMA 2

$$:A_1 \cdots A_n:C = :A_1 \cdots A_nC: + :A_1^{\bullet} \cdots A_nC^{\bullet}: + \cdots + :A_1 \cdots A_n^{\bullet}C^{\bullet}: \tag{9.19}$$

Proof. If C is an annihilation operator, then the lemma is trivially true, for C is already in normal order, and all contractions vanish. Hence assume that C is a creation operator. To normal-order the given product, we commute (or anticommute) C all the way to the far left. First write

$$:A_1 \cdots A_n:C = \eta_{AB} B_1 \cdots B_n C \tag{9.20}$$

where $\{B_1 \cdots B_n\}$ is the set $\{A_1 \cdots A_n\}$ in normal order. Now interchange B_n and C:

$$B_n C = \eta_{nC} C B_n + B_n^{\bullet} C^{\bullet} \tag{9.21}$$

where η_{nC} is -1 if both B_n and C are fermion operators, and $+1$ otherwise. Thus

$$:A_1 \cdots A_n:C = \eta_{AB}\eta_{nC} B_1 \cdots C B_n + \eta_{AB} B_1 \cdots B_n^{\bullet} C^{\bullet} \tag{9.22}$$

Continuing the process, we obtain

$$\begin{aligned} :A_1 \cdots A_n:C &= \eta_{nC} + \eta_{AB}(B_1 \cdots B_n^{\bullet} C^{\bullet}) + \eta_{AB}(B_1 \cdots B_{n-1}^{\bullet} B_n C^{\bullet}) \\ &= \cdots \\ &= :A_1 \cdots A_n C: + :A_1^{\bullet} \cdots A_n C^{\bullet}: + \cdots + :A_1 \cdots A_n^{\bullet} C^{\bullet}: \end{aligned} \tag{9.23}$$

■

■ LEMMA 3

$$A_1 A_2 \cdots A_n = :A_1 A_2 \cdots A_n: + :A_1^{\bullet} A_2^{\bullet} \cdots A_n: + :A_1^{\bullet} A_2 A_3^{\bullet} \cdots A_n: + \cdots$$

(*all possible contractions*) (9.24)

Proof. The statement is valid for $n = 2$, as demonstrated in Lemma 1. Assume that it is also valid for n. We prove it for $n + 1$ by multiplying the preceding equation by A_{n+1} from the left, and applying Lemma 2 term by term. ■

The extension of Lemma 3 to a T-product is straightforward, because for any fixed time ordering the T-product is an ordinary product. However, we have to redefine the contraction symbol to take into account time ordering:

$$\underbrace{AB} \equiv \langle 0|TAB|0\rangle \tag{9.25}$$

Then we have Wick's theorem.

■ Wick's Theorem

$$\mathrm{T}(A_1 A_2 \cdots A_n) = :A_1 A_2 \cdots A_n: + \underbrace{A_1 A_2} A_3 \cdots A_n: + :\underbrace{A_1 A_2} \underbrace{A_3} \cdots A_n: + \cdots$$

(*all possible contractions*) (9.26)

Corollary *Wick's theorem also holds when any* A_i *is replaced by a linear combination of creation and annihilation operators.*

A weaker form of Wick's theorem, which is simpler to state and prove, deals with vacuum expectation values, and is given in Section 10.2.

9.4 FEYNMAN RULES FOR SCALAR THEORY

Suppose we want to calculate the S matrix element for the scattering process $\mathbf{p}_1 + \mathbf{p}_2 \rightarrow \mathbf{p}_3 + \mathbf{p}_4$ to second order in λ_0. A systematic way is to make a normal-product expansion of the S matrix

$$S = 1 + S^1 + S^{(2)} + \cdots$$

$$S^1 = -\frac{i\lambda_0}{4!}\int d^4x \phi^4(x)$$

$$S^{(2)} = \frac{1}{2!}\left(\frac{-i\lambda_0}{4!}\right)^2 \int d^4x_1 \int d^4x_2 \, \mathrm{T}[\phi^4(x_1)\phi^4(x_2)] \tag{9.27}$$

It is convenient to represent the procedure graphically as follows. The basic interaction consists of a product of four field operators at the same space–time point, which we call a "vertex." It may be depicted as follows:

$$\begin{matrix} \phi & \phi & \phi & \phi \\ | & | & | & | \end{matrix} \tag{9.28}$$

We draw a line sticking out from each field operator, to represent a possible incoming or outgoing particle. These lines are distinguishable, since each corresponds to a factor $\phi(x)$ at a definite position in the product.

A contraction according to Wick's theorem will be indicated by joining the lines of the operators being paired. This produces an "internal line" representing a propagating virtual particle:

$$\underbrace{\phi(x)\phi(y)} = \langle 0|\mathrm{T}\phi(x)\phi(y)|0\rangle$$

$$= i\Delta_F(x_2 - x_1) = i\int \frac{d^4k}{(2\pi)^4} \frac{e^{ik\cdot x}}{k^2 - m^2 + i\eta} \qquad (\eta \rightarrow 0^+) \tag{9.29}$$

An uncontracted line is an "external line" associated with a particle in the initial state or final state, whose wave function depends on the normalization convention. We use the continuum normalization (2.44), with

$$\langle 0|\phi(x)|\mathbf{p}\rangle = \frac{e^{-ip\cdot x}}{\sqrt{2\omega_{\mathbf{p}}}} \tag{9.30}$$

We can represent the contractions in the form of a *Feynman graph*. The simplest, without any contraction, is the "vertex graph" in Fig. 9.1*a*. We have four external lines, and external momenta can be assigned to them in 4! different ways. This combinatorial factor 4! cancels the one in the coupling constant $\lambda_0/4!$. Thus, the vertex graph corresponds to the matrix element

$$\mathcal{G}_{\text{vertex}} = -i\lambda_0 \int d^4x \frac{e^{ip3\cdot x}}{\sqrt{2\omega_3}} \frac{e^{ip4\cdot x}}{\sqrt{2\omega_4}} \frac{e^{ip1\cdot x}}{\sqrt{2\omega_1}} \frac{e^{ip2\cdot x}}{\sqrt{2\omega_2}} \equiv -i\lambda_0 K_{1234} \tag{9.31}$$

where

$$\omega_i = +\sqrt{\mathbf{p}_i^2 + m_0^2} \tag{9.32}$$

and

$$K_{1234} = \frac{(2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4)}{\sqrt{(2\omega_1)(2\omega_2)(2\omega_3)(2\omega_4)}} \tag{9.33}$$

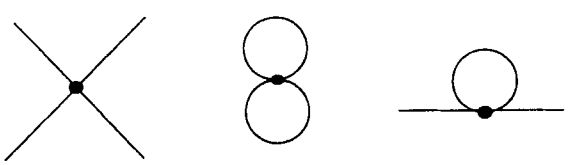
This factor expresses the conservation of total 4-momentum, and the normalization convention (2.44).

Terms with contractions fall into the following patterns:

$$\phi\phi\phi\phi \qquad \phi\,\phi\phi\,\phi \tag{9.34}$$

which correspond to the Feynman graphs in Figs. 9.1*b* and 9.1*c*. However, they do not contribute to the scattering process under consideration, since they do not have four external lines. Their significance will be discussed in the next section. The first-order S matrix consists of the vertex graph only:

$$\langle \mathbf{p}_3,\mathbf{p}_4|S^{(1)}|\mathbf{p}_1,\mathbf{p}_2\rangle = \mathcal{G}_{\text{vertex}} \tag{9.35}$$



(*a*) (*b*) (*c*)

Figure 9.1 (*a*) Vertex graph; (*b*) vacuum graph; (*c*) self-energy graph.

Let us now consider the second-order S matrix. There are two interaction vertices, as depicted below:

$$T[\phi_1\phi_1\phi_1\phi_1\phi_2\phi_2\phi_2\phi_2] \tag{9.36}$$

where $\phi_i = \phi(x_i)$. We must designate four lines as external particles, and contract the remaining lines. A possible pattern of contraction is the following:

$$\phi_1\phi_1\phi_1\overbrace{\phi_1\phi_2}\phi_2\phi_2\phi_2 \tag{9.37}$$

According to Wick's theorem, we must make all possible distinct choices of contraction schemes, normal-order the results, and add their contributions. In graphical terms, to normal-order is to assign external momenta to uncontracted lines. This procedure gives rise to the Feynman graphs shown in Figs. 9.2 and 9.3. The pattern (9.37) gives rise to the three graphs in Fig. 9.2, which differ only in the assignment of external momenta to the external lines. The number of terms in the Wick expansion corresponding to each of these Feynman graphs is

$$2\frac{(4!)^2}{2} = (4!)^2 \tag{9.38}$$

The first factor 2 comes from interchanging the positions x_1 and x_2, which are to be integrated over the same domain. The numerator in the next factor comes from permuting the four lines of each vertex among themselves, but this overcounts by a factor 2, because of the following symmetry: interchanging the two internal lines in graph 1 of Fig. 9.2 does not lead to a distinct situation. Thus we divide by 2. The 4! in both coupling constants are again canceled. If it were not for the symmetry in the graph, the factor 1/2! in front of the second-order S matrix would also be canceled. The correction factor for overcounting is called the *symmetry number* of the graph, which we shall discuss later in more general terms.

Graph 1 of Fig. 9.2 represents the following matrix element:

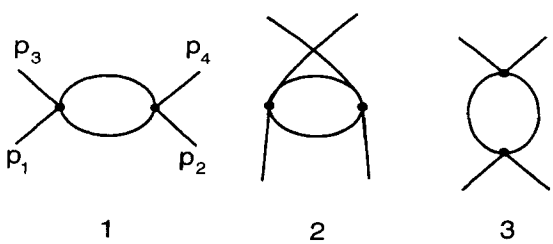


Figure 9.2 Second-order graphs for a scattering process.

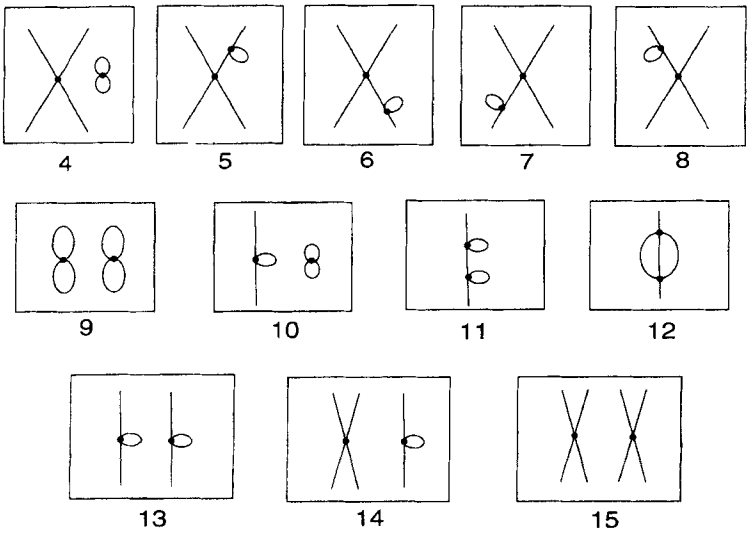


Figure 9.3 The rest of the second-order graphs, in addition to those in Fig. 9.2.

$$\mathcal{G}_1 = \frac{(-i\lambda_0)^2}{2}\int d^4x_1\, d^4x_2 \frac{e^{ip_3\cdot x_1}}{\sqrt{2\omega_3}} \frac{e^{ip_4\cdot x_1}}{\sqrt{2\omega_4}} [i\Delta_F(x_2 - x_1)]^2 \frac{e^{-ip_1\cdot x_2}}{\sqrt{2\omega_1}} \frac{e^{-ip_2\cdot x_2}}{\sqrt{2\omega_2}} \tag{9.39}$$

Substituting the Fourier integral for the propagator into the matrix element, we find

$$\mathcal{G}_1 = K_{1234} \frac{(-i\lambda_0)^2}{2}\int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m_0^2 + i\eta} \frac{i}{(p_1 - p_3 - k)^2 - m_0^2 + i\eta} \tag{9.40}$$

where $p_i = (\omega_i, \mathbf{p}_i)$ are 4-vectors. This shows that total 4-momentum is conserved at each vertex.

In nonrelativistic perturbation theory, virtual transitions conserve momentum but not energy. That is, virtual particles go "off energy shell." In contrast, virtual transitions here conserve both momentum and energy, but the squared mass becomes unphysical, as it is k^2 instead of the fixed value m^2. We say that a virtual particle propagates "off mass shell."

It is evident from (9.39) that an incoming particle of 4-momentum p^μ is associated with factor exp $(-ip\cdot x)$ and an outgoing one, with exp $(ip\cdot x)$. We can therefore convert an initial particle to a final one, or vice versa, by simply reversing the sign of its 4-momentum. This property is called *crossing symmetry.*

From $\mathcal{G}_1$, we can obtain the contributions $\mathcal{G}_2$ and $\mathcal{G}_3$ of graphs 2 and 3 in Fig. 9.2 by interchanging momentum labels. To get $\mathcal{G}_2$, we interchange p_3 and p_4; and to get $\mathcal{G}_3$, we interchange p_2 and $-p_3$. The last operation is an illustration of crossing symmetry. The sum of the three graphs gives

$$\mathcal{G}_1 + \mathcal{G}_2 + \mathcal{G}_3 = \frac{(-i\lambda_0)^2}{2}[I(p_1, p_2, p_3, p_4) + I(p_1, p_2, p_4, p_3) + I(p_1, -p_3, -p_2, p_4)]K_{1234} \tag{9.41}$$

where

$$I(p_1, p_2, p_3, p_4) = \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m_0^2 + i\eta} \frac{i}{(p_1 - p_3 - k)^2 - m_0^2 + i\eta} \tag{9.42}$$

Other second-order Feynman graphs are shown in Fig. 9.3, where only graphs 4–8 contribute to our process, because they are the only ones with four external lines. Thus, the complete second-order S matrix is

$$\langle \mathbf{p}_3, \mathbf{p}_4 | S^{(2)} | \mathbf{p}_1, \mathbf{p}_2 \rangle = \sum_{i=1}^{8} \mathcal{G}_i \tag{9.43}$$

The other graphs in Fig. 9.3, which do not contribute to the matrix element here, will be considered in the next section. Even among the graphs included, we only need to keep $\mathcal{G}_1 + \mathcal{G}_2 + \mathcal{G}_3$ for practical purposes, as we shall explain in the next section.

The nth order S matrix is given by

$$S^{(n)} = \frac{1}{n!}\left(\frac{-i\lambda_0}{4!}\right)^n \int d^4x_1 \cdots \int d^4x_n \, \mathrm{T}[(\phi_1\phi_1\phi_1\phi_1) \cdots (\phi_n\phi_n\phi_n\phi_n)] \tag{9.44}$$

In the following, we state the rules to generate all Feynman graphs of this order. First, draw n vertices, with four lines emanating from each vertex. We then contract the $4n$ lines attached to the vertices in all possible manners, including no contraction. Each distinct contraction scheme gives a graph. We assign external momenta to the uncontracted lines (the external lines.) for either incoming or outgoing particles. Each distinct assignment gives a Feynman graph. The matrix element corresponding to a Feynman graph can be obtained by inspection, through the use of the *Feynman rules*. For real ϕ^4 theory, they are as follows:

- An nth-order Feynman graph consists of n vertices where four lines meet. Some of these are internal lines carrying a internal 4-momentum. Others are external lines identified with incoming or outgoing particles of definite momenta.
- Each vertex contributes a factor $-i\lambda_0$. The net 4-momentum flowing into each vertex is zero.
- An external line of 4-momentum p contributes a factor $(2\omega_p)^{-1/2}$. It is an incoming particle if $p_0 > 0$, outgoing if $p_0 < 0$.
- An internal line of 4-momentum k^μ contributes a factor

$$i\Delta_F(k) = \frac{i}{k^2 - m_0^2 + i\eta} \qquad (\eta \to 0^+)$$

Not all internal momenta are independent because of 4-momentum conservation at vertices. The independent ones are integrated with measure $(2\pi)^{-4} d^4k$.

- Overall conservation of total 4-momentum is enforced through a factor

$$(2\pi)^4 \delta^4(P_f - P_i)$$

where P_f and P_i are respectively the total 4-momentum of the final and initial states.

- The graph is divided by a symmetry number S.

In general, the symmetry number has to be worked out for each graph. However, in ϕ^4 theory, a rule can be stated for connected nonvacuum graphs (graphs with external lines, with no disjoint subgraphs). In such a graph, k internal lines are said to form an equivalent set, if they all share the same vertices at both end. If there are more than one such set, containing respectively k_1, k_2, ... internal lines, then the symmetry number of the graph is (see Huang [1])

$$S = \prod_i k_i! \tag{9.45}$$

Vacuum graphs do not follow this rule. An example is given in Section 9.5.

There are topological relations among graph elements. Consider the more general ϕ^K theory, in which K lines meet at a vertex. We may say that an external line "uses up" $1/K$ vertex, while an internal line uses up $2/K$ vertex. Therefore, for a graph with n vertices, L_e external lines, and L_i internal lines, we have the relation

$$L_e + 2L_i = Kn \tag{9.46}$$

In the case under discussion, $K = 4$.

The 4-momenta of internal lines are not independent of one another, because of 4-momentum conservation at each of the n vertices. There are thus n conservation conditions, one of which is taken into account through total 4-momentum conservation. This leaves $n - 1$ constraints on the L_i internal 4-momenta. The number of independent internal 4-momentum is accordingly

$$N_i = L_i - (n - 1) \tag{9.47}$$

This is the number of $\int d^4k$ integrations in a graph.

9.5 TYPES OF FEYNMAN GRAPHS

9.5.1 Vacuum Graph

Figure 9.1*b* represents a vacuum graph, which describes a vacuum fluctuation involving the creation and annihilation of two virtual particle–antiparticle pairs. The matrix element is

$$\begin{aligned}\mathcal{G}_{\text{vac}} &= \langle 0|S^1|0\rangle = -i\lambda_0 \int d^4x_1 \langle 0|\mathrm{T}\ \phi_1\phi_1|0\rangle\ \langle 0|\mathrm{T}\phi_1\phi_1|0\rangle \\ &= -i\lambda_0 \int d^4x_1 [i\Delta_F(0)]^2 \\ &= -i\lambda_0 (2\pi)^4 \delta^4(0) \left[\int \frac{d^4k}{(2\pi)^4}\ \frac{i}{k^2 - m_0^2 + i\eta}\right]^2 \end{aligned} \tag{9.48}$$

The final form can be written down directly using the Feynman rules. The factor $(2\pi)^4\delta^4(0)$ represents the integral $\int d^4x$, which should be interpreted as the total volume of space–time. Vacuum processes such as this one occur with uniform probability over all space–time, and they can accompany any reaction we consider. Thus, the sum of all vacuum processes $\langle 0|S|0\rangle$ occurs as a factor in any S-matrix element. As we shall show below, this factor is a phase $\exp(i\Phi_0)$, and so does not affect transition probabilities. From a practical point of view, therefore, vacuum graphs may be ignored.

9.5.2 Self-Energy Graph

An example of a self-energy graph is Fig. 9.1*c*, which describes a particle interacting with itself while propagating in the vacuum. It does this by emitting and reabsorbing a virtual particle at the same point. Alternatively, we can say that the particle creates a virtual pair, which eventually annihilates. The matrix element is

$$\begin{aligned}\mathcal{G}_{\text{self}} &= \langle \mathbf{p}_2|S^1|\mathbf{p}_1\rangle = -i\lambda_0 \int d^4x_1 \langle \mathbf{p}_2|\phi_1|0\rangle\ \langle 0|\mathrm{T}\ \phi_1\phi_1|0\rangle\ \langle 0|\phi_1|\mathbf{p}_1\rangle \\ &= -i\lambda_0 \int d^4x_1 \frac{e^{ip_2\cdot x_1}}{\sqrt{2\omega_2}} i\Delta_F(0) \frac{e^{-ip_1\cdot x_1}}{\sqrt{2\omega_1}} \\ &= -i\lambda_0 \frac{(2\pi)^4\delta^4(p_1 - p_2)}{\sqrt{(2\omega_1)(2\omega_2)}} \int \frac{d^4k}{(2\pi)^4}\ \frac{i}{k^2 - m_0^2 + i\eta} \end{aligned} \tag{9.49}$$

Again, we could have obtained this directly from the Feynman rules.

Such a self-energy "bubble" can be inserted into any line, external or internal, any number of times. Some examples are shown in Fig. 9.3. We shall discuss such insertions systematically in Section 13.2. At this point, we merely mention that the effect of all possible self-energy insertions is to replace the free propagator $\Delta_F(k)$ by a "full propagator" $\Delta_F'(k)$, in which the mass pole at $k^2 = m_0^2$ is shifted in position,

and the residue is also changed. The shifted position corresponds to a "renormalized mass," and the changed residue corresponds to a change in normalization of the wave function, by a factor conventionally designated as $Z^{-1/2}$, the "wave function renormalization."

9.5.3 Connected Graph

A connected graph does not contain disjoint subgraphs. The converse is a disconnected graph, illustrated by graph 4 of Fig. 9.3, which corresponds to the matrix element

$$\mathcal{G}_4 = \mathcal{G}_{\text{vertex}}\mathcal{G}_{\text{vac}} \tag{9.50}$$

This describes a vacuum fluctuation during scattering. As mentioned before, vacuum components of graphs can be ignored.

Graphs 13–15 of Fig. 9.3 are disconnected graphs describing separate uncorrelated events. (One takes place on Earth; the other, on Mars 1000 years later, perhaps.) For example, graph 13 corresponds to two independent self-energy interactions; graph 14 represents a scattering event, with a spectator particle interacting with itself; and graph 15 describes two independent uncorrelated reactions. Disconnected graphs do not require separate calculations, because they are products of lower-order connected graphs. Therefore, we only need to consider connected graphs.

9.6 WICK ROTATION

In calculating a Feynman graph, we generally have to integrate over internal 4-momenta. Consider the self-energy graph $\mathcal{G}_{\text{self}}$, where we encounter the integral

$$I = \int d^4k \frac{1}{k^2 - m_0^2 + i\eta} \tag{9.51}$$

Putting aside the question of convergence, let us first describe how the "mass-shell" singularity at $k^2 = m_0^2$ should be handled. Written more explicitly, the integral reads

$$I = \int d^3k \int_{-\infty}^{\infty} dk_0 \frac{1}{k_0^2 - (\mathbf{k}^2 + m_0^2) - i\eta} \tag{9.52}$$

The term $i\eta$ in the denominator displaces the mass-shell poles away from the path of integration, as shown in Fig. 9.4 (left). Since the integrand has no other singularities in the k_0 plane, we can rotate the k_0 contour counterclockwise, until it lies along the imaginary k_0 axis, as indicated in Fig. 9.4 (right). This corresponds to making k_0 pure-imaginary:

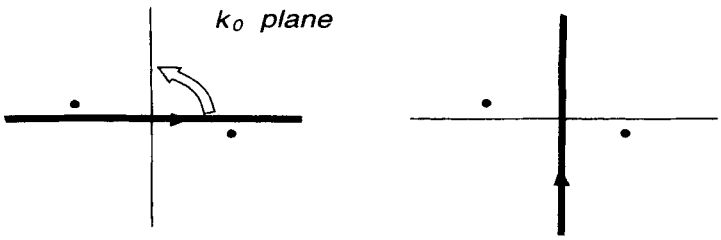


Figure 9.4 Wick rotation of contour to imaginary axis. The contour never crosses the poles of the Feynman propagator shown by the dots.

$$k_0 = ik_4 \tag{9.53}$$

where k_4 is real. We now have an integral over a four-dimensional Euclidean space:

$$I = -i\int d^3k \int_{-\infty}^{\infty} dk_4 \frac{1}{k_4^2 + \mathbf{k}^2 + m_0^2} \tag{9.54}$$

where we have taken the limit $\eta \to 0$. Called a *Wick rotation*, the rotation to the imaginary k_0 axis can be done in all Feynman integrals, because the integrand does not contain singularities other than those in the Feynman propagators. A virtual particle has Euclidean 4-momentum, instead of Minskowskian, because of the $i\eta$ in the Feynman propagators.

9.7 REGULARIZATION SCHEMES

Having defined the contour of integration, we are faced with the problem that I has an "ultraviolet divergence" at the upper limit of integrations. We must render it finite by introducing a high-momentum (or small-distance) cutoff. Eventually, when all graphs are added up, we hope to "renormalize" the theory, by reexpressing the cutoff in terms of experimentally measurable parameters.

Cutting off the high-momentum contributions means modifying the asymptotic behavior of the Feynman propagator, and the way to do this is not unique. The simplest scheme introduces a sharp cutoff Λ, by replacing I with

$$I_\Lambda = -i \int_{|\mathbf{k}|<\Lambda} d^3k \int_{-\infty}^{\infty} dk_4 \frac{1}{k_4^2 + \mathbf{k}^2 + m_0^2} \tag{9.55}$$

The k_4 integration can be performed through contour integration:

$$\int_{-\infty}^{\infty} dk_4 \frac{1}{k_4^2 + \mathbf{k}^2 + m_0^2} = \frac{\pi}{\sqrt{\mathbf{k}^2 + m_0^2}} \tag{9.56}$$

Thus we have

$$I_\Lambda = -4i\pi^2 \int_0^\Lambda dk \frac{k^2}{\sqrt{k^2 + m_0^2}} \xrightarrow[\Lambda \to \infty]{} -2i\pi^2 \Lambda^2 \tag{9.57}$$

The quadratic divergence when $\Lambda \to \infty$ is typical of a boson self-energy.

The sharp cutoff is simple, but not Lorentz-invariant. When it is important to keep the theory invariant, we can use the *Pauli–Villars regularization*, which replaces the propagator in the following manner:

$$\frac{1}{k^2 - m_0^2} \to \frac{1}{k^2 - m_0^2} - \frac{1}{k^2 - \Lambda^2} \tag{9.58}$$

The cutoff Λ appears as the mass of a fictitious particle, whose propagator has the sign opposite that of a physical particle. Since, according to (2.76), the residue at the mass pole is the squared modulus of the wave function, the fictitious particle here has negative probability, and therefore has no physical meaning. In the example considered, the Pauli–Villars regulator gives the same Λ –dependent term as in the sharp-cutoff case.

Another Lorentz-invariant cutoff scheme is *dimensional regularization*, which is based on analytic continuation of the space–time dimensionality. To illustrate the technique, rewrite I as a d–dimensional integral:

$$I = -i \int \frac{d^d \kappa}{\kappa^2 + m_0^2} = -i \int_0^\infty du \int d^d k \, e^{-(k^2 + m_0^2)u} \tag{9.59}$$

We are interested in its value near $d = 4$. Using the formula

$$\int d^d k \, e^{-uk^2} = \left(\frac{\pi}{u}\right)^{d/2} \tag{9.60}$$

we write

$$I = -i\pi^{d/2} \int_0^\infty du \, u^{-d/2} e^{-m_0^2 u} = -i\pi^{d/2} m_0^{d-2} \Gamma\left(1 - \frac{d}{2}\right) \tag{9.61}$$

which can be continued to complex d. We then put $d = 4 + \epsilon$, and obtain

$$I \xrightarrow[\epsilon \to 0]{} -2i\pi^2 m_0^2 \epsilon^{-1} \tag{9.62}$$

This gives the same result as the sharp cutoff if we identify $\Lambda^2 = m_0^2 \epsilon^{-1}$.

Still another cutoff scheme is *lattice regularization*, in which continuous space is replaced by a discrete lattice. The advantage of this scheme is that local gauge invariance can be preserved, and that it is well suited for Monte Carlo simulations. An example of this is given in Section 16.3.

9.8 LINKED-CLUSTER THEOREM

We distinguished between connected and disconnected graphs, because the latter are composed of connected subgraphs. If we know all the connected graphs, that is sufficient to generate all graphs. The linked-cluster theorem tell us exactly how to do this:

■ Linked-Cluster Theorem

$$\exp(\textit{sum of all connected graphs}) = (\textit{sum of all graphs}) \tag{9.63}$$

Proof. *Let* Γ_i denote a connected graph, so that the set of all connected graphs is $\{\Gamma_1, \Gamma_2, \ldots\}$. The general graph $\mathcal{G}$ contains m_i copies of Γ_i, and may be represented in the form

$$\mathcal{G} = \frac{\Gamma_1^{m_1}}{m_1!}\frac{\Gamma_2^{m_2}}{m_2!}\cdots \tag{9.64}$$

The factors $m_i!$ account for the fact that the copies are indistinguishable, and, as we shall show later, arise as symmetry numbers. Summing over all possible choices of subgraphs gives the sum of all possible graphs:

$$\sum \mathcal{G} = \prod_i \sum_{m=0}^{\infty} \frac{\Gamma_i^m}{m!} = e^{\sum_i \Gamma_i} \tag{9.65}$$

This the desired result; but it remains to derive the symmetry numbers. Consider the graph composed of m Γ's. If Γ is of order n, the graph is of order nm, and has the form

$$\text{Graph} \sim \frac{1}{(nm)!}\int d^4x_1 \cdots d^4x_{nm}\langle f|\mathrm{T}\,\mathcal{H}(x_1)\cdots\mathcal{H}(x_{nm})|i\rangle \tag{9.66}$$

A permutation the nm vertices has no effect on the integral, but a permutation that bodily interchanges two subgraphs does not give a distinct term in the Wick expansion. Thus, the number of distinct permutations is $(nm)!/m!$, and the symmetry number is $m!$. This completes the proof. ■

9.9 VACUUM GRAPHS

First we show that a connected vacuum graph is a pure imaginary number. Consider ϕ^K theory, in which an nth-order connected graph is of the form

$$\Gamma \sim (-i\lambda_0)^n \int (d^4k)^{Ni} \left(\frac{i}{k^2 - m_0^2} \right)^{Li} \tag{9.67}$$

where L_i is the number of internal lines. For a vacuum graph there are no external lines, and so, according to (9.46), we have

$$L_i = \frac{Kn}{2} \tag{9.68}$$

By (9.47), the number of 4-momentum integrations is given by

$$N_i = L_i - n + 1 = \frac{Kn}{2} - n + 1 \tag{9.69}$$

After making Wick rotations in all the k_0 integrations, the integral gives a real number; but the rotations produce a factor i^{Ni}. Thus a vacuum graph is of the form

$$i^{n+N_i+L_i} \times (\text{real number}) = i^{1+Kn} \times (\text{real number}) \tag{9.70}$$

For K odd, the theory is unphysical, because the Hamiltonian is not bounded from below. For K even, the graph is pure-imaginary. ∎

It follows from the linked-cluster theorem that the sum of all vacuum graphs is a phase factor:

$$\text{Sum of all vacuum graphs} = \langle 0|S|0\rangle = e^{i\Phi} \tag{9.71}$$

Since vacuum graphs have no external lines, we can freely add them to any graph with a fixed number of external lines. Any graph is therefore multiplied by $e^{i\Phi}$, but this has no effect on the scattering cross section. Therefore we can ignore all vacuum subgraphs.

PROBLEMS

9.1 Complex Field

(a) For a complex scalar field $\psi(x)$, show that the basic contractions are

$$\overbrace{\psi(x)\psi^\dagger(y)} = \langle 0|T\psi(x)\psi^\dagger(y)|0\rangle = i\Delta_F(x-y)$$

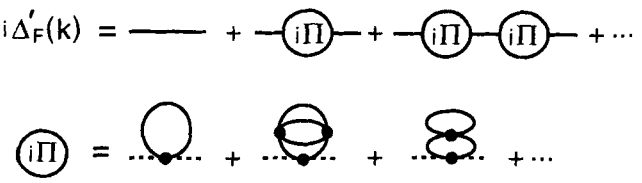
$$\overbrace{\psi(x)\psi(y)} = 0$$

Explain why the second of these is zero. A line in a Feynman graph now has a direction, represented by an arrow pointing along the flow of charge.

(b) Assume an interaction Lagrangian density $-(\lambda_0/4!)(\psi^\dagger\psi)^2$. State the modified Feynman rules. What restrictions follow from charge conservation?

(c) Consider graph 1 in Fig. 9.2 for a complex field. Choose a particular way to place arrows on the lines, and calculate the graph.

9.2 Self-Energy Consider the scalar ϕ^4 theory discussed in the text. To any internal line in a Feynman graph, we can always add any self-energy graph. The sum of all such additions replaces the Feynman propagator $i\Delta_F(k)$ with a full propagator $i\Delta_F'(k)$, as illustrated in the accompanying figure. Any graph for the full propagator can be dissected into "one-particle irreducible" components, which are graphs that cannot be made disconnected by cutting one line. The sum of all irreducible graphs, *with external lines omitted*, is denoted $i\Pi(k^2)$, where $\Pi(k^2)$ is called the "proper self-energy part." It is also known as "vacuum polarization," because it describes virtual pair creation and annihilation in the vacuum, and for a complex field this creates a fluctuation of dipole-moment density.



(a) Show that the full propagator has the form

$$\Delta_F'(k) = \frac{1}{k^2 - m_0^2 + \Pi(k^2) + i\eta}$$

(b) Calculate $\Pi(k^2)$ to the lowest order, namely the simple bubble graph. Introduce a cutoff Λ to make it finite.

(c) Show that the mass of the particle is shifted from m_0 to a renormalized value m. Calculate m to lowest order in the coupling constant λ_0.

9.3 Scalar ϕ^3 Theory A real scalar field $\phi(x)$ has Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - \frac{1}{2}m_0^2\phi^2 - \frac{1}{3!}g_0\phi^3$$

This theory is unphysical, because the Hamiltonian has no lower bound, but we can use it to illustrate Feynman graphs.

(a) State the Feynman rules.

(b) Calculate the vacuum polarization $\Pi(k^2)$ to lowest nonvanishing order, using a sharp momentum cutoff. Obtain the renormalized mass.

(c) Consider a two-particle scattering process. Draw all Feynman graphs for the S matrix up to order g_0^3. Give the matrix elements in the center-of-mass system, with incident momenta $\mathbf{p}$, $-\mathbf{p}$, and scattering angle θ.

(d) Calculate the differential cross section.

9.4 Pauli–Villars Regularization In Problem 9.3 the vacuum polarization yielded an integral proportional to

$$I(k^2) = \int \frac{d^4p}{(2\pi)^4} \frac{1}{(p^2 + m_0^2)[(p-k)^2 + m_0^2]}$$

where a Wick rotation has been made to Euclidean space. Calculate this integral using Pauli–Villars regularization, as follows.

(a) Since the integral is logarithmic divergent, we need to regulate only one of the propagators. Replace $(p^2 + m_0^2)^{-1}$ by $(p^2 + m_0^2)^{-1} - (p^2 + M^2)^{-1}$.

(b) Rewrite the propagators using the representation

$$\frac{1}{D} = \int_0^\infty d\alpha\, e^{-\alpha D}$$

From the two denominators, you get two integrals of the form $\int_0^\infty d\alpha \int_0^\infty d\beta\, e^{-\alpha D - \beta D'}$. Make the substitution of variables $\alpha = \lambda x$, $\beta = \lambda(1 - x)$. Then $\int_0^\infty d\alpha \int_0^\infty d\beta = \int_0^1 dx \int_0^\infty d\lambda \lambda$.

(c) Perform the integration over λ, and obtain $I(k^2)$ as an integral over x. Find the asymptotic behavior when $M \to \infty$.

9.5 Nonrelativistic Electron Gas Consider a gas of nonrelativistic free spin-$\frac{1}{2}$ fermions with fixed density n_0. The Hamiltonian is

$$H_0 = \frac{1}{2m} \int d^3x \sum_\sigma \psi_\sigma^+(\mathbf{x})(-i\nabla)^2 \psi_\sigma(\mathbf{x})$$

where $\psi_\sigma(x)$ is the annihilation operator for spin σ, which satisfies

$$[\psi_\sigma(\mathbf{x}), \psi_{\sigma'}^+(\mathbf{x}')] = (2\pi)^3 \delta^3(\mathbf{x} - \mathbf{x}')\delta_{\sigma\sigma'}$$

We expand it in terms of momentum eigenstates:

$$\psi_\sigma(x) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} c_{\mathbf{k}\sigma}$$

The ground state is denoted $|0\rangle$. It is like a vacuum state in hole theory, with all momentum states $k < k_F$ occupied, and the rest empty. The free propagator is defined by

$$G_{\alpha\beta}^0(\mathbf{x}, t; \mathbf{x}', t') = -i\langle 0|T\psi_\alpha(\mathbf{x}, t)\psi_\beta^+(\mathbf{x}', t')|0\rangle$$

where $\psi_\alpha(\mathbf{x}, t)$ are Heisenberg operators. Because of translational invariance, we can define

$$G_{\alpha\beta}^0(\mathbf{k}, \omega) = \int d^3x\, dt\, e^{i\omega t - i\mathbf{k}\cdot\mathbf{x}} G_{\alpha\beta}^0(x, t; 0, 0)$$

(a) Show that the propagator is given by a diagonal matrix in spin space:

$$G^0(\mathbf{k}, \omega) = \frac{1}{\omega - \epsilon_k + i\eta}$$

$$\eta = \begin{cases} +0 & (\epsilon_k > \epsilon_F) \\ -0 & (\epsilon_k < \epsilon_F) \end{cases}$$

where $\epsilon_k = k^2/2m$ and $\epsilon_F = k_F^2/2m$ is the Fermi energy. More explicitly,

$$G^0_{\alpha\beta}(\mathbf{k}, \omega) = \delta_{\alpha\beta}\left(\frac{\theta(\epsilon_k - \epsilon_F)}{\omega - \epsilon_k + i\eta} + \frac{\theta(\epsilon_F - \epsilon_k)}{\omega - \epsilon_k - i\eta}\right)$$

where θ is the step function: $\theta(x) = 1$ for $x > 0, \theta(x) = 0$ for $x < 0$.

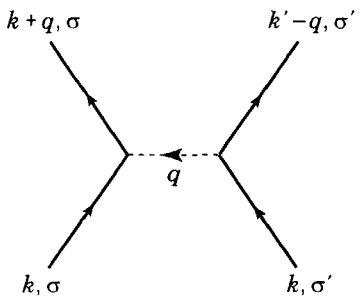
(b) Next, include Coulomb repulsion between the fermions. The interaction Hamiltonian takes the form

$$H' = \frac{e^2}{2}\sum_{\sigma\sigma'}\int d^3x\, d^3x'\ \psi^+_\sigma(\mathbf{x})\psi^+_{\sigma'}(\mathbf{x}')\frac{1}{|\mathbf{x} - \mathbf{x}'|}\psi_{\sigma'}(\mathbf{x}')\psi_\sigma(\mathbf{x})$$

To ensure charge neutrality, we add an interaction with a uniform background with opposite charge density. As shown in Problem 1.5, this is done by omitting the $k = 0$ part of the Fourier transform of the Coulomb potential. Show that H_1 can be written

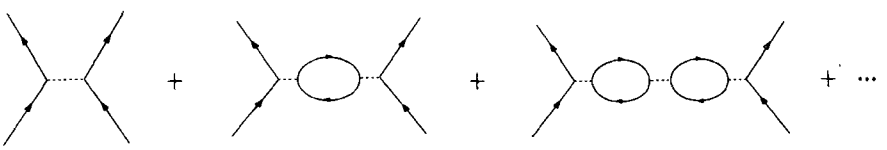
$$H' = \frac{e^2}{2}\sum_{\sigma\sigma'}\int \frac{d^3k\, d^3k'\, d^3q}{(2\pi)^9} c^+_{k+q,\sigma}c_{k',\sigma'}u(q)c_{k'-q,\sigma'}c_{k,\sigma}$$

where $u(q) = 4\pi e^2/q^2$. Represent the interaction by the Feynman graph



Write down a set of Feynman rules in momentum space.

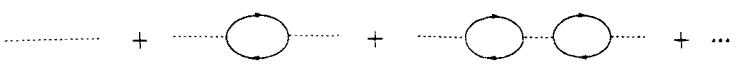
(c) The interaction between two particles is screened by "vacuum" polarization, described by the summation of the following diagrams.



Define the polarization function by

$$\Pi_0(\mathbf{q}, \omega) = -i\sum_\sigma \int \frac{d\nu}{2\pi}\frac{d^3k}{(2\pi)^3} G^0_{\sigma\sigma}(\mathbf{k} + \mathbf{q}, \omega + \nu)G^0_{\sigma\sigma}(\mathbf{k}, \omega)$$

The effective interaction $u_{\text{eff}}(q, \omega)$ is defined by the sum



We can define a dielectric function $\epsilon(q, \omega)$ by

$$u_{\text{eff}}(q, \omega) = \frac{e^2}{\epsilon(q, \omega)} \frac{4\pi}{q^2}$$

Show that

$$u_{\text{eff}}(q, \omega) = \frac{u(q)}{1 - u(q)\Pi_0(q, \omega)}$$

(d) Show that the polarization function is given by

$$\Pi_0(\mathbf{q}, \omega) = 2 \int \frac{d^3k}{(2\pi)^3} \theta(\epsilon_{k+q} - \epsilon_F)\theta(\epsilon_k - \epsilon_F) \times \left(\frac{1}{\omega + \epsilon_k - \epsilon_{k+q} + i\eta} - \frac{1}{\omega + \epsilon_{\epsilon k+q} - \epsilon_k - i\eta} \right)$$

(e) Now we concentrate on the real part of Π_0. Show that

$$\text{Re}\Pi_0(\mathbf{q}, \omega) = 2\text{P} \int \frac{d^3k}{(2\pi)^3} \frac{f(\mathbf{k}) - f(\mathbf{k} + \mathbf{q})}{\omega + \epsilon_k - \epsilon_{k+q}}$$

where $f(\mathbf{k}) = \theta(\epsilon_F - \epsilon_k)$ is the Fermi function at zero temperature. The sum of simple bubbles is known as the "random-phase approximation" for the dielectric function.

(f) Consider the limit $q \to 0$ at finite ω. Show that

$$\text{Re}\Pi_0(\mathbf{q}, \omega) \underset{q \to 0}{\to} \frac{n_0 q^2}{m\omega^2}$$

To do this, it might be helpful to rewrite the result of (e) as

$$\text{Re}\Pi_0(\mathbf{q}, \omega) = 2\text{P} \int \frac{d^3k}{(2\pi)^3} \frac{f(\mathbf{k})(\epsilon_{k+q} - \epsilon_k)}{\omega^2 - (\epsilon_k - \epsilon_{k+q})^2}$$

Note that

$$u_{\text{eff}}(q, \omega) = \frac{4\pi e^2}{q^2} \frac{1}{1 - (\omega_0/\omega)^2}$$

now contains a pole at the plasma frequency $\omega_0 = 4\pi e^2 n_0/m$.

(g) Consider the opposite limit $\omega = 0$, $q \to 0$. This is the appropriate limit for the screening of a static external test charge. Show that

$$\lim_{q \to 0} \text{Re}\Pi_0(\mathbf{q}, 0) = -\rho_0$$

where ρ_0 is the density of states at the Fermi energy:

$$\rho_0 = 2\int \frac{d^3k}{(2\pi)^3}\,\delta(\epsilon_k - \epsilon_{k'}) = \frac{mk_F}{\pi^2}$$

For this limit the result of (e) is useful.

(h) Show that

$$u_{\text{eff}}(q, 0) = \frac{4\pi e^2}{q^2 + \kappa^2}$$

where $\kappa^2 = 4\pi e^2 \rho_0$ is the Thomas–Fermi screening constant. This gives the static screened Coulomb potential

$$u_{\text{eff}}(r) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} u_{\text{eff}}(\mathbf{k})$$

Show that it takes the form of a Yukawa potential:

$$u_{\text{eff}}(r) = \frac{e^2}{r} e^{-\kappa r}$$

REFERENCE

1. K. Huang, *Quarks, Leptons, and Gauge Fields,* 2nd ed., World Scientific Publishers, Singapore, 1992, Section 7.7.

CHAPTER TEN

Vacuum Correlation Functions

10.1 FEYNMAN RULES

The n-point vacuum correlation function is defined by

$$G_n(x_1, \ldots, x_n) \equiv \langle 0|\mathrm{T}\ \phi_h(x_1) \cdots \phi_h(x_n)|0\rangle \tag{10.1}$$

where $\phi_h(x)$ is the field operator in the Heisenberg picture and $|0\rangle$ is the exact vacuum state, with $\langle 0|0\rangle = 1$. The Fourier transform is denoted by

$$\tilde{G}_n(k_1, \ldots, k_n) \equiv \int d^4x_1 \cdots d^4x_n e^{-i(k_1 \cdot x_1 + \cdots k_n \cdot x_n)} G_n(x_1, \ldots, x_n) \tag{10.2}$$

These functions give a complete description of the system. As we shall see, they determine the S matrix. From (8.49), we have the representation in the interaction picture:

$$G_n(x_1, \ldots, x_n) = \frac{\langle 0|\mathrm{T}\ S\phi(x_1) \cdots \phi(x_n)|0\rangle}{\langle 0|S|0\rangle} \tag{10.3}$$

where $\phi(x)$ is the field operator in the interaction picture and S is the S matrix. More explicitly,

$$G_n(x_1, \ldots, x_n) = \frac{1}{\langle 0|S|0\rangle} \sum_{k=0}^{\infty} \frac{(-i)^m}{m!} \int d^4y_1 \cdots d^4y_m \langle 0|\mathrm{T}\ \mathcal{H}(y_1) \cdots \mathcal{H}(y_m)\phi(x_1) \cdots \phi(x_n)|0\rangle \tag{10.4}$$

where $\mathcal{H}(x)$ is the interaction Hamiltonian density. The numerator may be expanded into sums of normal products according to Wick's theorem. Since we are taking the vacuum expectation value here, *all* factors in a normal product must be contracted. This leads to an expansion of the numerator in terms of Feynman graphs. Just as in the case of the S matrix, each graph is multiplied by vacuum subgraphs, and the

sum of all vacuum graphs occurs as a common factor to all graphs. This factor is $\langle 0|S|0\rangle$, which cancel the denominator. We may therefore replace the denominator by 1, and at the same time ignore all vacuum graphs.

The Feynman rules can be derived directly by using the following theorem.

■ **Wick's Theorem for Vacuum Expectations** *Let $a_1 \cdots a_{2p}$ be a product of an even number of creation or annihilation operators, and let*

$$\langle a_1 \cdots a_{2p}\rangle \equiv \langle 0|\mathrm{T}\, a_1 \cdots a_{2p}|0\rangle \tag{10.5}$$

Then

$$\begin{aligned}\langle a_1 \cdots a_{2\mathrm{p}}\rangle = {} & \langle a_1 a_2\rangle \langle a_3 a_4\rangle \cdots \langle a_{2p-1} a_{2p}\rangle \\ & + \delta_{\mathrm{P}} \langle a_1 a_3\rangle \langle a_2 a_4\rangle \cdots \langle a_{2p-1} a_{2p}\rangle \\ & + \cdots (\textit{all possible pairings with signature factor})\end{aligned} \tag{10.6}$$

where the signature factor δ_{P} is defined as follows. For bosons $\delta_{\mathrm{P}} = 1$. For fermions, δ_{P} is the signature of the permutation that will bring the list of operators in the term that it multiplies to the standard order $a_1 \cdots a_{2p}$.

The theorem, of course, follows from the operator form given in Section 9.3, but a direct proof is simple: The case $p = 1$ of this theorem is trivial. The general case can be proven by induction on p, with the help of the following

■ **LEMMA**

$$\begin{aligned}\langle a_1 \cdots a_{2p}\rangle = {} & \delta_1 \langle a_1 a_2\rangle \langle a_3 \cdots a_{2p}\rangle + \delta_2 \langle a_1 a_3\rangle \langle a_2 a_4 \cdots a_{2p}\rangle \\ & + \cdots + \delta_{2p-1} \langle a_1 a_{2p}\rangle \langle a_2 \cdots a_{2p-1}\rangle\end{aligned} \tag{10.7}$$

Proof. Let $b_1 \cdots b_{2p}$ be the set $a_1 \cdots a_{2p}$ in chronological order. We may assume that b_1 is an annihilation operator; otherwise, both sides trivially vanish. We commute b_1 all the way to the right, where it annihilates the vacuum state. Each time it commutes across some b_n, we write

$$b_1 b_n = \begin{cases} b_n b_1 + [b_1, b_n] & \text{(bosons)} \\ -b_n b_1 + \{b_1, b_n\} & \text{(fermions)} \end{cases} \tag{10.8}$$

The commutator or anticommutator above is a c-number, and may be replaced by its vacuum expectation value. Further, in the vacuum expectation value, it may be replaced by $b_1 b_n$, since $b_1|0\rangle = 0$. Then, $b_1 b_n$ may be replaced by $\mathrm{T} b_1 b_n$, since the operators are in chronological order. This proves that a typical term in the expansion is generated, whenever b_1 is commuted across some b_n. The signature factor supplies the appropriate sign. ■

As an example, consider the two-point correlation function G_2 in ϕ^4 theory:

$$G_2(x, y) = \langle 0|\mathrm{T}\ \phi(x)\phi(y)|0\rangle - \frac{i\lambda_0}{4!}\int d^4x_1\langle 0|\mathrm{T}\ \phi^4(x_1)\phi(x)\phi(y)|0\rangle + \cdots \quad (10.9)$$

with the understanding that vacuum graphs are to be ignored. The first term is the basic pairing, which gives $i\Delta_F(x-y)$. This is represent graphically by drawing a line between the points x and y, as shown in Fig. 10.1*a*. To expand the second term using Wick's theorem, write out the time-ordered product as follows:

$$T\ \phi_1\phi_1\phi_1\phi_1\phi_x\phi_y \quad (10.10)$$

where $\phi(x)$ and $\phi(y)$ are distinguished by heavier and longer lines, because the points x and y are "external" points not integrated over, and do not correspond to vertices. There are two distinct patterns of contraction:

$$\phi_1\phi_1\phi_1\phi_1\phi_x\phi_y \qquad \phi_1\phi_1\phi_1\phi_1\phi_x\phi_y \quad (10.11)$$

which correspond to graphs *b* and *c* of Fig. 10.1. We are to ignore *b* because it is the same as *a* when the vacuum subgraph is omitted. To order λ, we have

$$G_2(x, y) = i\Delta_F(x-y) + (-i\lambda_0)\int d^4x_1 i\Delta_F(x-x_1)[i\Delta_F(0)]i\Delta_F(x_1-y) \quad (10.12)$$

with Fourier transform

$$\tilde{G}_2(k_1, k_2) = (2\pi)^4\delta^4(k_1+k_2)\{i\tilde{\Delta}_F(k_1) + (-i\lambda_0)i\tilde{\Delta}_F(k_1)[i\Delta_F(0)]i\tilde{\Delta}_F(k_2)\} \quad (10.13)$$

Compared with Feynman graphs for the S matrix with external lines, there are only two differences:

- The external 4-momenta are arbitrary, not necessarily on the mass shell.
- Each external line contributes a propagator $i\tilde{\Delta}_F(k)$, instead of wave function $(2\omega_k)^{-1/2}$.

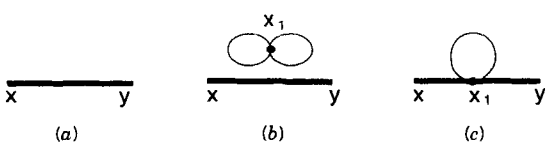


Figure 10.1 Graphs for a two-point correlation function.

Other than these exceptions, the Feynman rules are the same as those for the S matrix.

10.2 REDUCTION FORMULA

The converse of the rules stated above is as follows. A Feynman graph for the S matrix can be obtained from a corresponding one for the correlation function, by performing the following operations for each external line, of 4-momentum k:

- Replace the propagator $i\tilde{\Delta}_F(k)$ with a wave function $(2\omega_k)^{-1/2}$.
- Put k on mass shell.

While this procedure is correct graph by graph, a neater rule applies to the sum of all graphs. The sum of all Feynman graphs can be regrouped, such that all free propagators $i\tilde{\Delta}_F(k)$ are replaced by full propagators $i\tilde{\Delta}'_F(k)$, in which the pole in k^2 is displaced from the bare mass to the renormalized mass and the residue of the pole acquires a factor $Z^{-1/2}$ from wave function renormalization. Thus, to obtain the S matrix from the correlation function, we replace $i\tilde{\Delta}'_F(k)$ by $(2\omega_k Z)^{-1/2}$, and then go on the renormalized mass shell. This rule is the content of the *reduction formula* [1], as illustrated schematically in Fig. 10.2.

To formally derive the reduction formula, we consider external particles whose wave functions are finite wave packets. We shall let the wave packets approach plane waves, but only in the final formula. A wave packet $f(x)$ is defined as a normalizable solution of the Klein–Gordon equation:

$$(\Box^2 + m^2)f(x) = 0$$

$$i\int d^3x\, f^* \overleftrightarrow{\partial_0} f < \infty \tag{10.14}$$

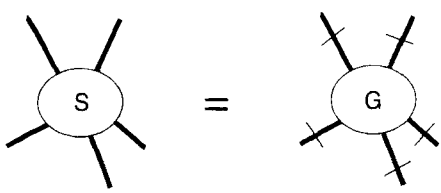


Figure 10.2 Reduction formula in pictures. The S matrix can be obtained from the correlation function by replacing external full propagators with normalization constants of wave functions.

where

$$f_1 \overleftrightarrow{\partial_0} f_2 \equiv f_1 \frac{\partial f_2}{\partial x_0} - \frac{f_2 \partial f_1}{\partial x_0} \tag{10.15}$$

Reduction Theorem *For $f(x)$ a normalizable solution of the Klein–Gordon equation,*

$$\int d^4x f(x)(\Box^2 + m^2)g(x) = \left[\lim_{x_0 \to \infty} - \lim_{x_0 \to -\infty}\right] \int d^3x f(x) \overleftrightarrow{\partial_0} g(x) \tag{10.16}$$

Proof. On the left side, make partial integrations to transfer the operator $(\Box^2 + m^2)$ to $f(x)$, which it annihilates. One is left with surface integrals, which give the result. ∎

Now consider a two-particle scattering process symbolically denoted as

$$3 + 4 \longrightarrow 1 + 2 \tag{10.17}$$

Let the Heisenberg field operator of the ith particle be $\phi_i \equiv \phi_i(x_i)$, and its normalizable wave packet be $f_i \equiv f(x_i)$. The annihilation operator for the ith wave packet can be defined as

$$a_i(x_0) \equiv i \int d^3x f_i^*(x) \overleftrightarrow{\partial_0} \phi_i(x) \tag{10.18}$$

which is a Heisenberg operator. As $x_0 \to \pm\infty$, the wave packets of different particles diverge from each other, and eventually the particles will behave like free particles. Thus, we may assume

$$\begin{aligned} a_i(x_0) &\xrightarrow[x_0 \to -\infty]{} \sqrt{Z_i} a_i^{\text{in}} \\ &\xrightarrow[x_0 \to \infty]{} \sqrt{Z_i} a_i^{\text{out}} \end{aligned} \tag{10.19}$$

where a_i^{in} and a_i^{out} are free-field annihilation operators of the plane-wave state i, in the limit $\alpha \to 0$. The factor $\sqrt{Z_i}$ accounts for wave function renormalization. This is called the *adiabatic condition,* which embodies adiabatic switching in the present context.

The free-field operators are defined by the commutation relations

$$\begin{aligned} [a_i^{\text{in}}, \bar{a}_j^{\text{in}}] &= \delta_{ij} \\ [a_i^{\text{out}}, \bar{a}_j^{\text{out}}] &= \delta_{ij} \end{aligned} \tag{10.20}$$

where, to avoid too many superscripts, we use a bar to denote Hermitian conjugate. These define two equivalent set of operators that may differ by a phase factor, which forms the S matrix. The initial and final states are defined by

$$|3, 4 \text{ in}\rangle \equiv \bar{a}_3^{\text{in}} \bar{a}_4^{\text{in}} |0\rangle$$

$$|1, 2 \text{ out}\rangle \equiv \bar{a}_1^{\text{out}} \bar{a}_2^{\text{out}} |0\rangle \tag{10.21}$$

where $|0\rangle$ is the physical vacuum state. The S matrix element is

$$\langle 1, 2|S|3, 4\rangle = \langle 1, 2 \text{ out}|3, 4 \text{ in}\rangle \tag{10.22}$$

Consider now the vacuum correlation function

$$G(1, 2; 3, 4) \equiv \langle 0|T\, \phi_1 \phi_2 \bar{\phi}_3 \bar{\phi}_4 |0\rangle \tag{10.23}$$

The projection onto wave packet states is defined by

$$\tilde{G}(t_1, t_2; t_3, t_4) = \int d^3x_1 \cdots d^3x_4 f_1^* f_2^* f_3 f_4 \overleftrightarrow{\partial}_{10} \overleftrightarrow{\partial}_{20} \overleftrightarrow{\partial}_{30} \overleftrightarrow{\partial}_{40} G(1,2;3,4) \tag{10.24}$$

$$= \langle 0|\mathrm{T}\, a_1 a_2 \bar{a}_3 \bar{a}_4 |0\rangle \tag{10.25}$$

By the asymptotic assumption we have, as $t_1, t_2 \to \infty$, and $t_3, t_4 \to -\infty$,

$$\tilde{G}(t_1, t_2; t_3, t_4) \to \sqrt{Z_1 Z_2 Z_3 Z_4} \langle 0|a_1^{\text{out}}\, a_2^{\text{out}} \bar{a}_3^{\text{in}} \bar{a}_4^{\text{in}} |0\rangle \tag{10.26}$$

where time-ordering is unnecessary because the operators involved are independent of time. Therefore

$$\tilde{G}(t_1, t_2; t_3, t_4) \to \sqrt{Z_1 Z_2 Z_3 Z_4} \langle 1, 2\, |S|3, 4\rangle \tag{10.27}$$

We now calculate the left side using the reduction theorem. Go to (10.24), and perform the indicated operation with respect to particle 1, in the limit $t_1 \to \infty$. This is called "reducing particle 1":

$$\lim_{t_1 \to \infty} i \int d^3x_1 f_1^* \overleftrightarrow{\partial}_{10} \langle 0|\mathrm{T}\, \phi_1 \phi_2 \phi_3 \phi_4 |0\rangle$$

$$= \lim_{t_1 \to -\infty} i \int d^3x_1 f_1^* \overleftrightarrow{\partial}_{10} \langle 0|\mathrm{T}\, \phi_1 \phi_2 \phi_3 \phi_4 |0\rangle$$

$$+ \int d^4x_1 f_1^* (\Box_1^2 + m_1^2) G(1, 2; 3, 4) \tag{10.28}$$

where the first term on the right side vanishes, as it is equal to $\langle 0|\mathrm{T}\, (\phi_1 \phi_2 \bar{\phi}_3) a_4 |0\rangle$. We then reduce the other particles in a similar manner. The final result is

$$\langle 1, 2\, |S|3, 4\rangle = \int d^4x_1 \cdots d^4x_4 \frac{f_1^* f_2^* f_3 f_4}{\sqrt{Z_1 Z_2 Z_3 Z_4}} ((\Box_1^2 + m_1^2) \cdots (\Box_4^2 + m_4^2) G(1, 2; 3, 4) \tag{10.29}$$

This is the reduction formula.

When we go to the plane-wave limits of the wave packets, the operator $(\Box_1^2 + m_1^2)$ can be replaced by $(-k_1^2 + m_1^2)$, where k_1 is the 4-momentum of particle 1. Thus it vanishes, unless canceled by its inverse from the correlation function. The effect is to cut off the external leg in the correlation function, forcing it to go on mass shell, and multiplying it with $(2k_{10}Z_1)^{-1/2}$.

10.3 THE GENERATING FUNCTIONAL

The vacuum correlation functions may be considered to be the response of the system to an external source $J(x)$ coupled to the field, which is turned on and off adiabatically:

$$J(x) \xrightarrow[|x_0|\to\infty]{} 0 \tag{10.30}$$

The Lagrangian density in the presence of the source is

$$\mathcal{L}_J(x) = \mathcal{L}(x) - J(x)\phi(x) \tag{10.31}$$

where $\mathcal{L}(x)$ is the Lagrangian density without source. We assume that the vacuum state $|0\rangle$ in the absence of source is unique, with $\langle 0|0\rangle = 1$. When the source is turned on and off, it remains unique according to the adiabatic theorem of Section 8.8. We denote the vacuum in the infinite past by $|0^-\rangle_J$, and in the infinite future by $|0^+\rangle_J$. These state vectors describe the same state as $|0\rangle$, but may differ by a phase $W[J]$:

$$\langle 0^+|0^-\rangle_J = e^{iW[J]} \tag{10.32}$$

Let us go to the interaction picture with respect to the source interaction:

$$H' = \int d^3x J(x)\phi(x) \tag{10.33}$$

The field operator in this picture is just the Heisenberg operator in the absence of source. According to the adiabatic theorem, we have

$$|0^-\rangle_J = \mathrm{T}e^{-i\int_{-\infty}^{0} dt\, H'(t)}|0\rangle$$

$$|0^+\rangle_J = \mathrm{T}e^{-i\int_{\infty}^{0} dt\, H'(t)}|0\rangle \tag{10.34}$$

where $|0\rangle$ is the vacuum state vector at $t = 0$. Thus

$$\langle 0^+|0^-\rangle_J = \langle 0|\mathrm{T}\, e^{-i\int_{-\infty}^{\infty} dt\, H'(t)}|0\rangle = \langle 0|\mathrm{T}\, e^{-i\int d^4x\, J(x)\phi(x)}|0\rangle \tag{10.35}$$

Taking $\delta/\delta J(x)$ of the right side brings down a factor $-i\phi(x)$ from the exponent. Hence

$$\frac{\delta\langle 0^+|0^-\rangle_J}{\delta J(x_1)\cdots\delta J(x_n)}\bigg|_{J=0} = (-i)^n\langle 0|\mathrm{T}\,\phi(x_1)\cdots\phi(x_n)|0\rangle = (-i)^n G_n(x_1,\ldots,x_n) \quad (10.36)$$

and we have the expansion

$$e^{iW[J]} = \langle 0^+|0^-\rangle_J = \sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\int d^4x_1\cdots d^4x_1 G_n(x_1,\ldots,x_n)J(x_1)\cdots J(x_n) \quad (10.37)$$

This shows that $\langle 0^+|0^-\rangle_J$ is the generating functional of vacuum correlation functions.

10.4 CONNECTED CORRELATION FUNCTIONS

In terms of Feynman graphs, we can state

$$G_n(x_1,\ldots,x_n) = \text{sum of all Feynman graphs with } n \text{ external lines} \quad (10.38)$$

where the sum includes both connected and disconnected graphs. (It is understood that vacuum components of graphs are to be omitted.) Since disconnected graphs are made up of lower-order connected graphs, it is useful to separate out the connected ones:

$$I_n(x_1,\ldots,x_n) = \text{sum of all connected Feynman graphs with } n \text{ external lines} \quad (10.39)$$

A general G_n consists of a number of disconnected components, which we can enumerate by giving the "occupation numbers" $\{\sigma_1, \sigma_2, \cdots\}$, such that there are σ_k copies of I_k. This is indicated by the formula

$$G_n(x_1,\ldots,x_n) = \sum_{\{\sigma_k\}}\sum_{P} P[\underbrace{I_1(*)\cdots I_1(*)}_{\sigma 1 \text{ factors}}][\underbrace{I_2(*)\cdots I_2(**)}_{\sigma 2 \text{ factors}}]\cdots \quad (10.40)$$

where $\{\sigma_1,\ldots,\sigma_n\}$ is a partition of the integer n, such that

$$n = \sigma_1 + 2\sigma_2 + \cdots + n\sigma_n \quad (10.41)$$

In each term of the sum over $\{\sigma_k\}$, there appear n asterisks representing the coordinates $x_1,\ldots,x_n$ in some fixed order, and P denotes a distinct permutation of these coordinates. The number of such permutations is

$$\frac{n!}{(\sigma_1!\ldots\sigma_n!)[(1!)^{\sigma_1}\cdots(n!)^{\sigma_n}]} \quad (10.42)$$

When (10.40) is substituted into (10.37), terms in the P sum give the same contribution on integration over $x_1, \ldots, x_n$. Thus we have

$$\langle 0^+|0^-\rangle = \sum_{n=0}^{\infty} \sum_{\{\sigma_k\}} (-i)^n \frac{\left[\int d^4x I_1(x)J(x)\right]^{\sigma_1}}{\sigma_1!(1!)^{\sigma_1}} \frac{\left[\int d^4x d^4y I_2(x,y)J(x)J(y)\right]^{\sigma_2}}{\sigma_2!(2!)^{\sigma_2}} \cdots \tag{10.43}$$

The double sum above is equivalent to a sum over $\{\sigma_k\}$ with no restriction. Thus, each σ_k is independently summed from 0 to ∞:

$$\langle 0^+|0^-\rangle = \sum_{\sigma_1=0}^{\infty} \frac{1}{\sigma_1!}\left[\frac{(-i)}{1!}\int d^4x I_1(x)J(x)\right]^{\sigma_1} \sum_{\sigma_2=0}^{\infty}\left[\frac{(-i)^2}{2!}\int d^4x d^4y I_2(x,y)J(x)J(y)\right]^{\sigma_1} \cdots$$

$$= \exp \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \cdots d^4x_n \, I_n(x_1, \ldots, x_n)J(x_1)\cdots J(x_n) \tag{10.44}$$

Thus $\ln\langle 0^+|0^-\rangle$ is the generating functional of connected correlation functions:

$$iW[J] = \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \cdots d^4x_n I_n(x_1, \ldots, x_n)J(x_1)\cdots J(x_n) \tag{10.45}$$

This is a form of the linked-cluster theorem.

10.5 LEHMANN REPRESENTATION

The correlation functions "know" about the mass spectrum of the field theory. For two-point correlation functions, the dependence on the mass spectrum is made explicit in the Lehmann representation [2]. We consider a variety of two-point correlation functions, which for a real scalar field are conventionally designated as follows:

$$\begin{aligned} i\Delta_F'(x) &= \langle 0|T\phi(x)\phi(0)|0\rangle \\ i\Delta'(x) &= \langle 0|[\phi(x), \phi(0)]|0\rangle \\ i\Delta^{(+)'}(x) &= \langle 0|\phi(x)\phi(0)|0\rangle \\ i\Delta^{(-)'}(x) &= \langle 0|\phi(0)\phi(x)|0\rangle \end{aligned} \tag{10.46}$$

where $\phi(x)$ is the Heisenberg operator and $|0\rangle$ is the vacuum state of the interacting theory. The corresponding correlation functions for the free field, denoted without a prime, have the following Fourier representations:

$$\Delta_F(x, m^2) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik\cdot x}}{k^2 - m^2 + i\eta} \qquad (\eta \to 0^+)$$

$$\Delta(x, m^2) = -\int \frac{d^3k}{(2\pi)^3} \frac{\sin(\omega_k t)}{\omega_k} e^{i\mathbf{k}\cdot\mathbf{x}}$$

$$\Delta^{(+)}(x, m^2) = \frac{1}{2\pi i} \int \frac{d^4k}{(2\pi)^4} \theta(k^0)\delta(k^2 - m^2)\, e^{ik\cdot x}$$

$$\Delta^{(-)}(x, m^2) = -\frac{1}{2\pi i} \int \frac{d^4k}{(2\pi)^4} \theta(-k^0)\delta(k^2 - m^2)\, e^{ik\cdot x} \tag{10.47}$$

where m is the free-particle mass.

Consider the function $\Delta^{(+)\prime}$. By writing $\phi(x) = e^{-iP\cdot x}\phi e^{iP\cdot x}$, where P^μ is the 4-momentum operator, and $\phi \equiv \phi(0)$, we obtain

$$i\Delta^{(+)\prime}(x) = \sum_n \langle 0|e^{-iP\cdot x}\phi\, e^{iP\cdot x}|n\rangle \langle n|\phi|0\rangle = \sum_n e^{iP_n\cdot x}|\langle 0|\phi|n\rangle|^2 \tag{10.48}$$

The state $|n\rangle$, with 4-momentum P_n^μ, is a "single-particle state" in the sense that $\langle 0|\phi\,|n\rangle \neq 0$. We assume

$$P_n^0 > 0 \qquad \text{(positive energy)} \tag{10.49}$$

$$P_n^2 > 0 \qquad \text{(positive invariant mass)}$$

Thus

$$\int d^4k\ \theta(k^0)\theta(k^2)\delta^4(k - P_n) = 1 \tag{10.50}$$

when inserted into (10.48), it yields the integral representation

$$i\Delta^{(+)\prime}(x) = \int d^4k\ \theta(k^0)\ \theta(k^2)\, e^{ik\cdot x}\rho(k^2) \tag{10.51}$$

where the mass spectral function $\rho(k^2)$ is defined by

$$\rho(k^2) \equiv \sum_n \delta^4(k - P_n)|\langle 0|\phi|n\rangle|^2 \tag{10.52}$$

which is real, positive-definite, and depends on k^2 only (by Lorentz invariance). We now integrate over k^μ, keeping $k^2 = m^2$, by writing

$$\int d^4k = \int dm^2 \int d^4k\ \delta(k^2 - m^2) \tag{10.53}$$

The result is the Lehmann representation for $\Delta^{(+)\prime}$:

$$\Delta^{(+)\prime}(x) = \int_0^\infty dm^2 \rho(k^2) \Delta^{(+)}(x, m^2) \tag{10.54}$$

A similar result holds for other correlation function: A correlation function is the spectrally weighted integral of the corresponding free function, over all possible values of the mass. In particular, the full Feynman propagator has the representation

$$\Delta_F'(x) = \int_0^\infty dm^2 \rho(m^2) \Delta_F(x, m^2) \tag{10.55}$$

whose Fourier transform gives

$$\tilde{\Delta}_F'(k) = \int_0^\infty dm^2 \frac{\rho(m^2)}{k^2 - m^2 - i\eta} \tag{10.56}$$

This immediately implies that for a free field of mass m_0 we have

$$\rho(m^2) = \delta(m^2 - m_0^2) \qquad \text{(free field)} \tag{10.57}$$

Consider the function $\Delta'(x)$ defined in (10.46). Writing $x = (\mathbf{x}, t)$, and taking the time derivative at $t = 0$, we have

$$\dot{\Delta}'(\mathbf{x}, 0) = -i\langle 0|[\dot{\phi}(\mathbf{x}, 0), \phi(0, 0)]|0\rangle = -\delta^3(\mathbf{x}) \tag{10.58}$$

This is also equal to $\dot{\Delta}(\mathbf{x}, 0)$, since it depends only on the equal-time commutation relation. Thus, performing this operation on both sides of the representation $\Delta'(x) = \int_0^\infty dm^2 \rho(m^2)\, \Delta(x, m^2)$, we obtain

$$\int_0^\infty dm^2 \rho(m^2) = 1 \tag{10.59}$$

An "elementary particle," defined as a one-particle state $|1\rangle$ with definite mass m_1, corresponds to a delta-function term in the spectral function:

$$\rho(m^2) = Z\delta(m^2 - m_1^2) + \sigma(m^2) \tag{10.60}$$

where $Z = |\langle 1|\phi|0\rangle|^2$ is the wave function renormalization constant. The condition (15.43) implies

$$0 \leq Z \leq 1 \tag{10.61}$$

We can see from (10.56) that the particle corresponds to a pole in the full propagator.

In an interacting theory there may exist bound states, which are states of definite mass connected to the vacuum through a product of field operators, but not

through a single-field operator. They correspond to poles not in the full propagator, but in a correlation function involving four or more fields. We shall discuss this in a later section..

10.6 DYSON–SCHWINGER EQUATIONS

The Dyson–Schwinger equations are integral equations for vacuum correlation functions. We shall illustrate them in a field theory of interacting fermions and bosons. The fermion field is denoted $\psi(x)$ with adjoint $\overline{\psi}(x)$, and the real boson field is denoted by $\phi(x)$. The Hamiltonian density is

$$\begin{aligned} \mathcal{H}(x) &= \mathcal{H}_F + \mathcal{H}_B + \mathcal{H}' \\ \mathcal{H}'(x) &= g\overline{\psi}(x)\phi(x)\psi(x) \end{aligned} \tag{10.62}$$

where $\mathcal{H}_F$ and $\mathcal{H}_B$ are respectively the free fermion and boson Hamiltonian density and g is a coupling constant. Our discussion will concentrate on two- and four-fermion correlation functions, with the boson field relagated to the background role of mediating the interaction between fermions.

We assume that ψ is a column vector and $\overline{\psi}$, either the Hermitian or Pauli adjoint; but the number of components are not specified. Similarly, ϕ may be multicomponent, and each component may be a matrix on the fermion internal vector space. The free Hamiltonians, which determine the propagators in Feynman graphs, need not be specified in detail. The general form of the Hamiltonian covers a nonrelativistic electron gas interacting through phonons, or quantum electrodynamics, with $\phi = \gamma^\mu A_\mu$. Our treatment will be based on general properties of Feynman graphs, and detailed specifications are purposely avoided, in order to focus on the relevant algebraic structure. A property we explicitly assume is that the fermion number is conserved.

We consider the two- and four-point fermion correlation functions:

$$\begin{aligned} G_2(x, y) &= \langle 0|\mathrm{T}\ \psi(x)\overline{\psi}(y)|0\rangle \\ G_4(x_1, x_2; y_1, y_2) &= \langle 0|\mathrm{T}\ \psi(x_1)\psi(x_2)\overline{\psi}(y_1)\overline{\psi}(y_2)|0\rangle \end{aligned} \tag{10.63}$$

where all operators are in the Heisenberg picture and the spin indices are suppressed. Consider first the two-point function, with Fourier transform

$$G_2(x, y) = \int \frac{d^4p}{(2\pi)^4} G_2(p)\ e^{-ip\cdot(x-y)} \tag{10.64}$$

The full fermion propagator $S(p)$ is given by

$$iS(p) = G_2(p) \tag{10.65}$$

This a matrix on the fermion internal space. The Feynman graph expansion is shown in Fig. 10.3*a*. Owing to fermion number conservation, fermion lines cannot terminate, and therefore a single fermion line runs through a graph for $S(p)$. Since vacuum subgraphs can be omitted, all graphs for $S(p)$ are connected. Denoting the free propagator by S_0, and leaving the momentum argument p understood, we can write Fig. 10.3*a* in algebraic form:

$$\begin{aligned} iS &= iS_0 + iS_0 i\Sigma i S_0 + iS_0 i\Sigma i S_0 i\Sigma i S_0 + \cdots \\ &= iS_0 + iS_0 i\Sigma (iS_0 + iS_0 i\Sigma i S + \cdots) \end{aligned} \tag{10.66}$$

where $i\Sigma(p)$ is the sum of all "irreducible" graphs, with external leg amputated, as shown in Fig. 10.3*b*. Here, an irreducible graph is a connected graph that cannot be made disconnected by cutting one fermion line. Sometimes the name "one-particle irreducible" graph is used.

The Dyson–Schwinger equation for G_2 is obtained by rewriting (10.66) in the form

$$S = S_0 - S_0 \Sigma S \tag{10.67}$$

When transformed to coordinate space, this becomes an integral equation. The kernel Σ is variously called the "proper self-energy, " or "mass operator." The explicit solution is

$$S = (S_0 + \Sigma)^{-1} \tag{10.68}$$

This has a simple appearance, but all the complications are buried in Σ.

We now turn to the four-point function. In terms of interaction picture operators, we have (without bothering to change notation)

$$G_4(x_1, x_2; y_1, y_2) = \langle 0|\mathrm{T}\ \psi(x_1)\psi(x_2)\overline{\psi}(y_1)\overline{\psi}(y_2)\, e^{-i\int d^4x\, \mathcal{H}'(x)}|0\rangle \tag{10.69}$$

with the instruction that all vacuum graphs are to be omitted in the Feynman graph expansion. The zeroth-order graphs are shown in Fig. 10.4*a* and the second-order

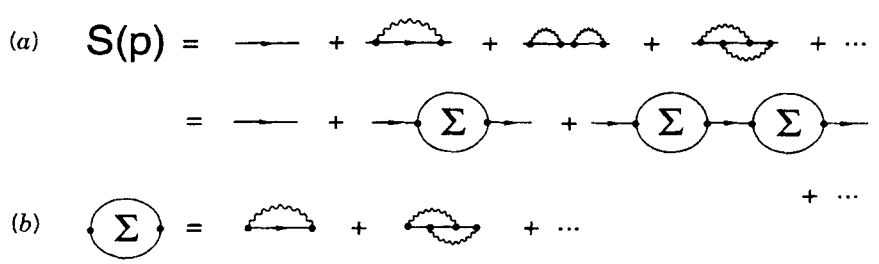


Figure 10.3 (*a*) Full propagator $S(p)$; (*b*) irreducible component $\Sigma(p)$.

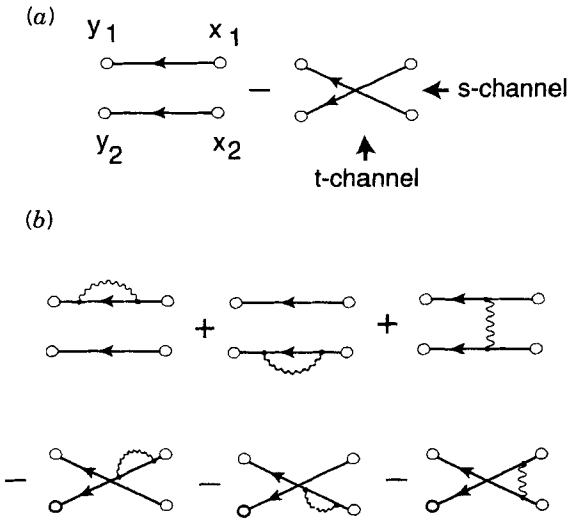


Figure 10.4 Graphs for four-point correlation function for fermions: (*a*) two basic patterns of free graphs, differing by fermion exchange; (*b*) interacting graphs consists of putting "adornments" on the two free patterns.

graphs, in Fig. 10.4*b*. Any graph has two fermion lines running through it, and they can be drawn either parallel or crossed. The interactions merely "adorn" the basic pattern. Each graph with the parallel pattern is in one–one correspondence with one having the crossed pattern, in which x_1, x_2 are interchanged, with a change of sign.

The Fourier transform is denoted by

$$\langle q_1, q_2|G|p_1, p_2\rangle = \int G_4(x_1, x_2; y_1, y_2) e^{i(p_1\cdot x_1 + p_2\cdot x_2) - i(q_1\cdot y_1 + q_2\cdot y_2)} \qquad (10.70)$$

where the measure $d^4x_1\, d^4x_2\, d^4y_1\, d^4y$ is left understood . The same set of Feynman graphs describes three possible channels of scattering (with off-mass-shell momenta) :

- s channel: fermion–fermion scattering with center-of mass energy squared $s = (p_1 + p_2)^2$.
- t channel: fermion–antifermion scattering with center-of mass energy squared $t = (p_1 - q_1)^2$.
- u channel: fermion–antifermion scattering with center-of mass energy squared $u = (p_1 - q_2)^2$.

We are interested in deriving an integral equation for the correlation function, and the kernel of the integral equation will depend on the channel. We shall consider the s channel from now on.

Let $D(x_1, x_2; y_1, y_2)$ be the sum of all "direct" graphs with parallel fermion lines, with Fourier transform $\langle q_1, q_2|D|p_1, p_2\rangle$. Then we can write

$$G_4(x_1, x_2; y_1, y_2) = (1 - \mathcal{A}_{12})D(x_1, x_2; y_1, y_2)$$
$$\langle q_1, q_2|G|p_1, p_2\rangle = (1 - \mathcal{A}_{12})\langle q_1, q_2|D|p_1, p_2\rangle \tag{10.71}$$

where the operation $\mathcal{A}_{12}$ interchanges the labels 1 and 2 in either the initial or final state. Graphs for $\langle q_1, q_2|D|p_1, p_2\rangle$ are shown in Fig. 10.5. There is only one disconnect graph, the product of two full fermion propagators. We denote it by

$$\langle q_1, q_2|I|p_1, p_2\rangle \equiv (2\pi)^4\delta^4(p_1 - q_1)(2\pi)^4\delta^4(p_2 - q_2)iS(p_1)iS(p_2) \tag{10.72}$$

The connected graphs can be decomposed into two-particle irreducible components, that is, connected graphs that cannot be made disconnected by cutting two fermion lines. As indicated in Fig. 10.5, the sum of all such irreducible components, with external line omitted, is denoted by $\langle q_1, q_2|\Gamma|p_1, p_2\rangle$.

We introduce a matrix notation by regarding $|p_1, p_2\rangle$ as a vector. with the properties

$$|p_1, p_2\rangle = |p_1\rangle \otimes |p_2\rangle$$
$$\langle p_1|p_2\rangle = (2\pi)^4\delta^4(p_1 - {}_2)$$
$$\int \frac{d^4p}{(2\pi)^4}|p\rangle\langle p| = 1 \tag{10.73}$$

It should be noted that $|p_1\rangle \otimes |p_2\rangle \neq |p_2\rangle \otimes |p_1\rangle$. The vectors in the basis are mathematical constructs without physical significance. In this notation we can write $D = I + I\Gamma I + I\Gamma I\Gamma I + \cdots$, which gives the integral equation

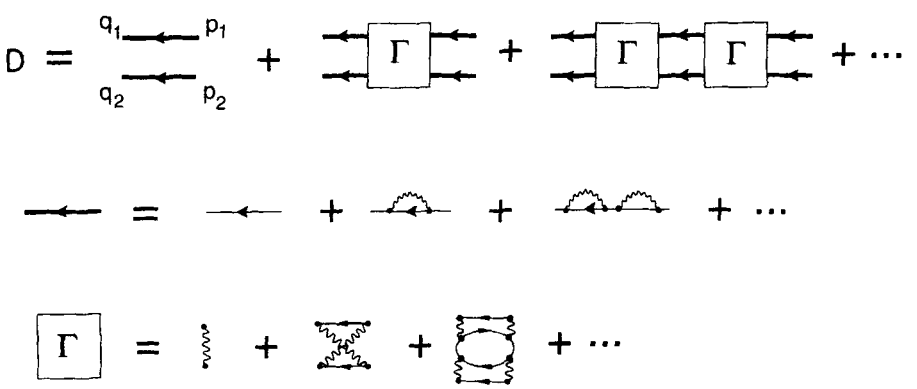


Figure 10.5 The connected graphs of the four-point function can be expressed in terms of an irreducible kernel Γ.

$$D = I + I\Gamma D \tag{10.74}$$

We can obtain $\langle q_1, q_2|G|p_1, p_2\rangle$ by antisymmetrizing $\langle q_1, q_2|D|p_1, p_2\rangle$ with respect to either p_1, p_2, or q_1, q_2. Defining the antisymmetrizing operator $\mathcal{A}$ by $\mathcal{A}|p_1, p_2\rangle = |p_2, p_1\rangle$, we can write the integral equation for G in matrix notation:

$$\begin{aligned} G &= K + K\Gamma G \\ K &\equiv (1 - \mathcal{A})I = I(1 - \mathcal{A}) \end{aligned} \tag{10.75}$$

Further analysis of the integral equation (10.74) is left to the exercises.

A t-channel process can be described through the continuation $\langle q_1, -q_2|G|p_1, -p_2\rangle$. Although such a continuation of (10.74) yields an expression for the correlation function, it is not in the form of an integral equation. To get an integral equation for the t channel, we have to go back to the Feynman graphs to define a different kernel. We refer the reader elsewhere [3] for a more general discussion, as well as derivation of integral equations for higher correlation functions.

10.7 BOUND STATES

The distinction between elementary and composite particles is purely theoretical. It depends on the model we use to describe the particles. For example, nucleons were once regarded as "elementary," but are now considered bound states of quarks. Similarly, the electron is considered elementary because so far it is adequate to describe it by a basic field. Within a given quantum field theory, there is a clear distinction between elementary and composite particles.

Like any single-particle state, a bound state should have definite mass and spin. In addition, it must be orthogonal to any elementary particle state. For illustration, let us consider a bound state in the s channel, with fermion number 2. Suppose that x_1, x_2 have times earlier than those of y_1, y_2. We insert a complete set of states into the four-point function

$$\langle 0|\psi(x_1)\psi(x_2)\bar{\psi}(y_1)\bar{\psi}(y_2)|0\rangle = \sum_n \langle 0|T\psi(x_1)\psi(x_2)|n\rangle \langle n|\bar{\psi}(y_1)\bar{\psi}(y_2)|0\rangle \tag{10.76}$$

Only states of fermion number 2 can occur, and the bound state will be among them, if it exists. Thus we expect the bound state to show up as a pole of the Fourier transform, in an appropriate momentum variable. The residue of that pole can be regarded as the wave function of the bound state.

We introduce total and relative coordinates:

$$X = \frac{x_1 + x_2}{2} \qquad x = x_1 - x_2$$

$$Y = \frac{y_1 + y_2}{2} \qquad y = y_1 - y_2 \tag{10.77}$$

and consider times such that x_1, x_2 are well ahead of y_1, y_2. This means that $X^0 - Y^0 \geqslant a$, where a is some sufficiently large number dependent on x^0 and y^0, as illustrated in Fig. 10.6. We separate out the term fulfilling this condition by writing

$$G_4 = \theta(X^0 - Y^0 - a)\,\langle 0|[\mathrm{T}\psi(X + x/2)\psi(X - x/2)][\mathrm{T}\bar{\psi}(Y + y/2)\bar{\psi}(Y - y/2)]|0\rangle + \mathcal{R} \tag{10.78}$$

where $\mathcal{R}$ denotes the rest. Now insert a complete set of states between the two T-products, using the completeness relation

$$1 = \int \frac{d^3P}{(2\pi)^3 2E}|B\rangle\,\langle B| + \cdots \tag{10.79}$$

where $|B\rangle$ denotes a bound state and the dots denote contribution from other states. The bound state has fermion number 2, and energy–momentum $(E, \mathbf{P})$. In the relativistic case $E = \sqrt{\mathbf{P}^2 + M^2}$. We have opted to normalize the bound state covariantly, and left understood that $|B\rangle$ depends on $\mathbf{P}$.

Splitting off the contribution from the bound state, we have

$$G_4(x_1, x_2; y_1, y_2) = G_B(X, Y; x, y) + \mathcal{R}' \tag{10.80}$$

where

$$G_B(X, Y; x, y) = \theta(X^0 - Y^0 - a)\int \frac{d^3P}{(2\pi)^3 2E}\left\langle 0\left|\left[\mathrm{T}\psi\left(X + \frac{x}{2}\right)\psi\left(X - \frac{x}{2}\right)\right]\right|B\right\rangle$$

$$\times \left\langle B\left|\left[\mathrm{T}\bar{\psi}\left(Y + \frac{y}{2}\right)\bar{\psi}\left(Y - \frac{y}{2}\right)\right]\right|0\right\rangle \tag{10.81}$$

We can take out the dependence on X and Y by using the translation operation

$$\psi\left(X + \frac{x}{2}\right) = e^{iP\cdot X}\psi\left(\frac{x}{2}\right)e^{-iP\cdot X} \tag{10.82}$$

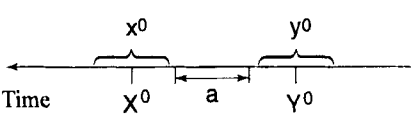


Figure 10.6 Choice of time ordering to exhibit a bound-state pole in the appropriate correlation function.

where P^μ is the total 4-momentum operator. Then

$$G_B(X, Y; x, y) = \theta(X^0 - Y^0 - a)\int \frac{d^3P}{(2\pi)^3 2E} U_{\mathbf{P}}(x)\bar{U}_{\mathbf{P}}(y)\, e^{-iK\cdot(X-Y)} \tag{10.83}$$

where $U_{\mathbf{P}}(x)$ is the relative wave function of the bound state:

$$[U_{\mathbf{P}}(x)]_{\alpha\beta} = \left\langle 0\left|\left[\mathrm{T}\,\psi_\alpha\left(\frac{x}{2}\right)\psi_\beta\left(\frac{-x}{2}\right)\right]\right|B\right\rangle \tag{10.84}$$

where we have displayed the spin indices.

Now substitute G_B into the Fourier transform of G_4 in (10.70):

$$\langle q_1, q_2|G|p_1, p_2\rangle = \langle q_1, q_2|B|p_1, p_2\rangle + \mathcal{R} \tag{10.85}$$

where $\mathcal{R}$ does not contain the bound state, and

$$\langle q_1, q_2|B|p_1, p_2\rangle = \int G_B(X, Y; x, y)e^{i(P\cdot X + p\cdot x) - i(Q\cdot Y + q\cdot y)} \tag{10.86}$$

where the integration extends over X, Y, x, y, and

$$\begin{aligned} P &= p_1 + p_2 \qquad p = \frac{p_1 - p_2}{2} \\ Q &= q_1 + q_2 \qquad q = \frac{q_1 - q_2}{2} \end{aligned} \tag{10.87}$$

Using the representation

$$\theta(t) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} d\omega \frac{e^{i\omega t}}{\omega - i\eta} \qquad (\eta \to 0^+) \tag{10.88}$$

we obtain the result

$$\langle q_1, q_2|B|p_1, p_2\rangle = (2\pi)^4\delta^4(P - Q)\frac{e^{i(E-P0)a}}{2iE(E - P^0 + i\eta)} u_{\mathbf{P}}(p)\bar{u}_{\mathbf{P}}(q) \tag{10.89}$$

where

$$\begin{aligned} [u_{\mathbf{P}}(p)]_{\alpha\beta} &= \int d^4x\, e^{ip\cdot x}[U_{\mathbf{P}}(x)]_{\alpha\beta} \\ &= \int d^4x\, e^{ip\cdot x}\left\langle 0\left|\mathrm{T}\psi_\alpha\left(\frac{x}{2}\right)\psi_\beta\left(-\frac{x}{2}\right)\right|B\right\rangle \end{aligned} \tag{10.90}$$

We can rewrite

$$\frac{1}{2E(E-P^0)} = -\frac{E+P^0}{2E}\frac{1}{(P^2-M^2)} \tag{10.91}$$

Thus $\langle q_1, q_2|B|p_1, p_2\rangle$ has a pole at $P^2 = M^2$:

$$\langle q_1, q_2|B|p_1, p_2\rangle \xrightarrow[P^2\to M^2]{} i(2\pi)^4\delta^4(P-Q)\frac{u_{\mathbf{P}}(p)\bar{u}_{\mathbf{P}}(q)}{P^2-M^2+i\eta} \tag{10.92}$$

Since a pole is absent in the term $\mathcal{R}$ we have

$$\langle p_1, p_2|G|q_1, q_2\rangle \xrightarrow[P^2\to M^2]{} (2\pi)^4\delta^4(P-Q)\frac{u_{\mathbf{P}}(p)\bar{u}_{\mathbf{P}}(q)}{P^2-M^2+i\eta} \tag{10.93}$$

This shows that, like an elementary particle, the bound state occurs as a mass pole in the appropriate correlation function. The residue of the pole gives the bound state wave function. In matrix notation, we can abbreviate the preceding as

$$G \to \frac{i|B\rangle\langle B|}{P^2-M^2+i\eta} \tag{10.94}$$

10.8 BETHE–SALPETER EQUATION

Using (10.94) in the Dyson–Schwinger equation $G = K + K\Gamma G$, we obtain the *Bethe–Salpeter equation* [4]

$$|B\rangle = K\Gamma|B\rangle \tag{10.95}$$

More explicitly,

$$u(p_1, p_2) = iS(p_1)iS(p_2)\times$$

$$\int \frac{d^4p_1'd^4p_2'}{(2\pi)^8}[\langle p_1, p_2|\Gamma|p_1', p_2'\rangle - \langle p_2, p_1|\Gamma|p_1', p_2'\rangle]u(p_1', p_2') \tag{10.96}$$

where for convenience of notation we have written

$$u(p_1, p_2) = u_{\mathbf{P}}(p) \tag{10.97}$$

This is not a wave function in the nonrelativistic sense, since the two particles involved have different time coordinates. But it occurs in the S matrix for bound state scattering [3], and is in this sense a natural generalization of the wave function of an elementary particle.

The normalization of the wave function is not fixed by the Bethe–Salpeter equation, but may be determined as follows. Consider the Dyson–Schwinger equation $G = K + K\Sigma G$, which can be rewritten

$$(K^{-1} - \Sigma)G = 1 \tag{10.98}$$

At given $\mathbf{P}$, the bound state occurs as a pole in G at $P^2 = M^2$:

$$G = \frac{i|B\rangle\langle B|}{P^2 - M^2 + i\eta} + \sum_n C_n|\chi_n\rangle\langle\chi_n| \tag{10.99}$$

where $\langle\chi_n|B\rangle = 0$. Substituting this into the Dyson–Schwinger equation and multiplying through by $P^2 - M^2$, we have

$$(K^{-1} - \Sigma)\left\{i|B\rangle\langle B| + (P^2 - M^2)\sum_n C_n|\chi_n\rangle\langle\chi_n|\right\} = P^2 - M^2 \tag{10.100}$$

Now differentiate with respect to P^μ, and then put $P^2 = M^2$. Noting $(K^{-1} - \Sigma)|B\rangle = 0$ at $P^2 = M^2$, we have

$$i\left[\frac{\partial}{\partial P^\mu}(K^{-1} - \Sigma)|B\rangle\right]\langle B| + 2P^\mu\sum_n C_n(I^{-1} - K)|\chi_n\rangle\langle\chi_n| = 2P^\mu$$

Sandwiching this between $\langle B|$ and $|B\rangle$ leads to the normalization condition

$$i\left\langle B\left|\frac{\partial}{\partial P^\mu}(K^{-1} - \Sigma)\right|B\right\rangle = 2P^\mu \tag{10.101}$$

PROBLEMS

10.1 Consider a free real scalar field.

(a) Show that the only connected vacuum correlation function is

$$I_2(x_1, x_2) = \langle 0|T\,\phi(x_1)\phi(x_2)|0\rangle = i\Delta_F(x_1 - x_2)$$

(b) Hence show that

$$W[J] = \tfrac{1}{2}\int d^4x\, d^4x_2\, J(x_1)\Delta_F(x_1 - x_2)J(x_2)$$

10.2 From the Lehmann representation (10.56) and discussion of vacuum polarization in Problem 9.2, we have alternative representations for the full propagator:

$$\tilde{\Delta}'_F(k) = \int_0^\infty dm^2 \frac{\rho(m^2)}{k^2 - m^2 - i\eta} = \frac{1}{k^2 - m_0^2 + \Pi(k^2)}$$

The $i\eta$ term can be omitted from the second form, because $\Pi(k^2)$ is generally complex.

Assume that the theory has only one stable particle of mass M.

(a) Show that $\rho(m^2)$ has the general form

$$\rho(m^2) = Z\delta(m^2 - M^2) + \theta(m^2 - 4M^2)\sigma(m^2)$$

(b) Show that

$$\Pi(M^2) = m_0^2$$

$$\Pi'(M^2) = \frac{1}{Z} - 1$$

(c) Express $\sigma(m^2)$ in terms of $\Pi(k^2)$. (*Hint:* Equate the imaginary parts of the two representations.)

10.3 To further analyze the Dyson–Schwinger equation (10.74), factor out from D and Γ factors common to all Feynman graphs:

$$\langle q_1, q_2|D|p_1, p_2\rangle = (2\pi)^4\delta^4(p_1 + p_2 - q_1 - q_2)S(q_1)S(q_2)\bar{D}(p_1, p_2; k)S(p_1)S(p_2)$$

$$\langle q_1, q_2|\Gamma|p_1, p_2\rangle = (2\pi)^4\delta^4(p_1 + p_2 - q_1 - q_2)\bar{\Gamma}(p_1, p_2; k)$$

where $k = q_1 - p_1$ and $\bar{D}$ and $\bar{\Gamma}$ are matrices in internal space.

(a) Show that the Dyson–Schwinger equation becomes

$$\bar{D}(p_1, p_2; k) = (2\pi)^4\delta^4(k)[S(p_1)S(p_2)]^{-1} + \int \frac{d^4k'}{(2\pi)^4}\bar{\Gamma}(p_1 + k, p_2 - k; k - k')\bar{D}(p_1, p_2; k')$$

(b) Take the kernel from the lowest-order Feynman graph:

$$\bar{\Gamma}(p_1, p_2; k) = g^2\Delta(k)$$

where $\Delta(k)$ is the free boson propagator. This is known as the "ladder approximation," because D is the sum of graphs that resemble ladders with increasing number of rungs. Put

$$\bar{D}(p_1, p_2; k) = [S(p_1)S(p_2)]^{-1}F(k)$$

and show that

$$F(k) = (2\pi)^4\delta^4(k) + g^2\int \frac{d^4k'}{(2\pi)^4}\Delta(k - k')F(k')$$

(c) Solve the preceding equation. Obtain $\bar{D}(p_1, p_2; k)$ in the ladder approximation, whence $\langle q_1, q_2|D|p_1, p_2\rangle$ and $\langle q_1, q_2|G|p_1, p_2\rangle$.

10.4 **(a)** Show that in the ladder approximation the Bethe–Salpeter equation (10.96) is

$$[u_{\mathbf{P}}(p)]_{\alpha\beta} = -g^2[S(p_1)]_{\alpha\alpha'}[S(p_2)]_{\beta\beta'}\int \frac{d^4q}{(2\pi)^4}[\Delta(p - q) - \Delta(p + q)][u_{\mathbf{P}}(q)]_{\alpha'\beta'}$$

where $p_1 = (P/2) + p$, $p_2 = (P/2) - p$.

(b) Consider a bound state of zero total momentum $\mathbf{P} = 0$. For the fermion, use Dirac propagator $S(p) = (\not{p} - m)^{-1}$; for the boson, use massless propagator $\Delta(k) = ik^{-2}$. Put

$$[u_0(p)]_{\alpha\beta} = \frac{\delta_{\alpha\beta}\chi(p)}{(p^2 - m^2)}$$

and show

$$\chi(p) = ig^2 \int \frac{d^4q}{(2\pi)^4} \left[\frac{1}{(p-q)^2} - \frac{1}{(p+q)^2} \right] \frac{\chi(q)}{q^2 - m^2}$$

By ignoring the second term in the kernel (the antisymmetrizing term), one can solve this eigenvalue problem in terms of hypergeometric functions [5].

REFERENCES

1. H. Lehmann, K. Symanzik, and W. Zimmermann, *Nuovo Cimento* **11,** 342 (1954).
2. H. Lehmann, *Nuovo Cimento* **11,** 342 (1954).
3. K. Huang and H. A. Weldon, *Phys. Rev.* **D11,** 257 (1975).
4. H. A. Bethe and E. Salpeter, *Phys. Rev.* **84,** 1232 (1951); M. Gell-Mann and F. E. Low, *Phys. Rev.***84,** 350 (1951).
5. J. Goldstein, *Phys. Rev.* **91,** 1516 (1953).

CHAPTER ELEVEN

Quantum Electrodynamics

11.1 INTERACTION HAMILTONIAN

Quantum electrodynamics (QED) describes the interaction between electrons and photons. The Lagrangian density is

$$\mathcal{L}(x) = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}[i\gamma^{\mu}(\partial_{\mu} + ie_0A_{\mu}) - m_0]\psi \tag{11.1}$$

where $F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}$ is the electromagnetic field tensor, A^{μ} the 4-vector potential, and ψ the Dirac spinor field. We use Coulomb gauge, in which

$$\nabla \cdot \mathbf{A} = 0 \tag{11.2}$$

The fields are quantized according to canonical commutation or anticommutation rules. When there is ambiguity, operators are taken in normal order. The charge e_0 and mass m_0 of the electron are "bare" or "unrenormalized" parameters. The corresponding renormalized or physical parameters have the values

$$\begin{aligned} e &= -1.6 \times 10^{-19}\ \text{C} \\ m &= 9.1 \times 10^{-31}\ \text{kg (kilogram)} \end{aligned} \tag{11.3}$$

We are using rationalized electromagnetic units, in which the dimensionless fine-structure constant is

$$\frac{e^2}{4\pi\hbar c} \approx \frac{1}{137} \tag{11.4}$$

The smallness of this quantity makes us think that we can use perturbation theory.

The total Lagrangian can be put in the form

$$L = \int d^3x[\tfrac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2) + \overline{\psi}(i\gamma^\mu\partial_\mu - m_0)\psi - e_0 j^\mu A_\mu] \tag{11.5}$$

where

$$j^\mu = \overline{\psi}\gamma^\mu\psi \tag{11.6}$$

is the electron current density. The total Hamiltonian reads

$$H = \int d^3x[\tfrac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) + \mathbf{E}\cdot\nabla A_0 + \psi^\dagger(\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m_0)\psi + e_0 j^\mu A_\mu] \tag{11.7}$$

We note that

$$\int d^3x\, \mathbf{E}\cdot\nabla A_0 = -\int d^3x\, A^0\nabla\cdot\mathbf{E} = -e_0\int d^3x\, j^0 A^0 \tag{11.8}$$

which cancels part of the interaction term. Thus

$$H = \int d^3x[\tfrac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) + \psi^\dagger(\boldsymbol{\alpha}\cdot\mathbf{p} + \beta m_0)\psi - e_0\mathbf{j}\cdot\mathbf{A}] \tag{11.9}$$

In Coulomb gauge, A^0 is entirely determined by j^0 through Poisson's equation:

$$\begin{aligned}\nabla^2 A^0 &= -e_0 j^0 \\ A_0(\mathbf{x}) &= e_0\int d^3y \frac{j^0(\mathbf{y})}{4\pi|\mathbf{x}-\mathbf{y}|}\end{aligned} \tag{11.10}$$

Putting $\mathbf{E} = -\partial\mathbf{A}/\partial t - \nabla A^0$, we obtain, after some straightforward calculations,

$$H = H_{\text{em}} + H_{\text{electron}} + H_{\text{int}} \tag{11.11}$$

where

$$\begin{aligned}H_{\text{em}} &= \frac{1}{2}\int d^3x\left[\left(\frac{\partial\mathbf{A}}{\partial t}\right)^2 + (\nabla\times\mathbf{A})^2\right] \\ H_{\text{el}} &= \int d^3x\, \psi^\dagger(-i\boldsymbol{\alpha}\cdot\nabla + \beta m_0)\psi \\ H_{\text{int}} &= \int d^3x\left[-e_0\mathbf{j}\cdot\mathbf{A} + \frac{e_0}{2}j^0A^0\right] \\ &= -e_0\int d^3x\, \mathbf{j}\cdot\mathbf{A} + \frac{e_0^2}{2}\int d^3x\, d^3y\frac{j^0(\mathbf{x})j^0(\mathbf{y})}{4\pi|\mathbf{x}-\mathbf{y}|}\end{aligned} \tag{11.12}$$

In the presence of an external electromagnetic field A^{μ}_{ext}, there will be an additional term in H_{int} given by

$$H_{\text{ext}} = e_0 \int d^3x (j^0 A^0_{\text{ext}} - \mathbf{j}\cdot\mathbf{A}_{\text{ext}}) \tag{11.13}$$

and thus

$$H_{\text{int}} = \int d^3x [-e_0 \mathbf{j}\cdot(\mathbf{A} + \mathbf{A}_{\text{ext}}) + e_0 j^0 (\tfrac{1}{2} A^0 + A^0_{\text{ext}})] \tag{11.14}$$

$$= e_0 \int d^3x [\tfrac{1}{2} j^0 A^0 - j^k A^k + + j_{\mu}(x) A^{\mu}_{\text{ext}}(x)] \tag{11.15}$$

Note that there is a factor $\frac{1}{2}$ in front of A_0, but not A^0_{ext}.

Let us define the "Coulomb propagator" as

$$D_{\text{Coul}}(x-y) = \frac{\delta(x^0 - y^0)}{4\pi|\mathbf{x}-\mathbf{y}|} \tag{11.16}$$

Then the interaction Hamiltonian density can be written in the form

$$\mathcal{H}_{\text{int}}(x) = -e_0 j^k(x) A^k(x) + \frac{e_0^2}{2} \int d^4y D_{\text{Coul}}(x-y) j^0(x) j^0(y)$$

$$+ e_0 j_{\mu}(x) A^{\mu}_{\text{ext}}(x) \tag{11.17}$$

In practice we can replace this with a simpler expression. As shown in the following section, the first two terms can be effectively replaced by $e_0 j_{\mu}(x) A^{\mu}(x)$, because the difference does not contribute to S-matrix elements.

In free propagators and external wave functions, the bare mass m_0 will be replaced by the renormalized mass m when all Feynman graphs are added up. Thus, it is convenient to redefine the unperturbed problem such that renormalized mass occurs in the free propagator. Accordingly, we take the unperturbed electron Hamiltonian density to be

$$\mathcal{H}_{\text{el}} = \psi^{\dagger}(-i\boldsymbol{\alpha}\cdot\nabla + \beta m)\psi \tag{11.18}$$

The difference $(m_0 - m)\psi^{\dagger}\beta\psi$ with the original form, called the "mass counterterm," is considered part of the interaction. With this, the interaction Hamiltonian density of QED will later taken to be

$$\mathcal{H}_{\text{int}}(x) = e_0 : \bar{\psi}(\not{A} + \not{A}_{\text{ext}})\psi : - \delta m : \bar{\psi}\psi : \tag{11.19}$$

where we have used the effective form of the electromagnetic interaction, and

$$\delta m = m - m_0 \tag{11.20}$$

This quantity is considered $O(e_0^2)$ in perturbation theory. In scattering processes calculated to second order in e_0, therefore, the mass counterterm need not be taken into account.

11.2 PHOTON PROPAGATOR

Since the Coulomb gauge is not Lorentz-covariant, neither is the photon propagator $\langle 0|TA^\mu(x)A^\nu(y)|0\rangle$ in that gauge. However, the part of the propagator that contributes to the S matrix is covariant, because of current conservation. We shall show this in the context of electron–electron (ee) scattering, which is indicated schematically as follows:

$$\begin{array}{ccccccc} e & + & e & \rightarrow & e & + & e \\ p_1 & & p_2 & & p_3 & & p_4 \\ s_1 & & s_2 & & s_3 & & s_4 \end{array} \tag{11.21}$$

where p_i and s_i refer respectively to 4-momentum and spin. The initial state $|i\rangle$ and final state $|f\rangle$ are denoted by

$$\begin{aligned} |i\rangle &= |\mathbf{p}_1, s_1; \mathbf{p}_2, s_2\rangle \\ |f\rangle &= |\mathbf{p}_3, s_3; \mathbf{p}_4, s_4\rangle \end{aligned} \tag{11.22}$$

To second order in e_0, the S matrix is given by

$$\langle f|S-1|i\rangle = -i\int d^4x\langle f|\mathcal{H}_{\text{int}}(x)|i\rangle + \frac{(-i)^2}{2}\int d^4x\, d^4y\langle f|\mathrm{T}\mathcal{H}_{\text{int}}(x)\mathcal{H}_{\text{int}}(y)|i\rangle \tag{11.23}$$

We use (11.17) for this calculation, since the whole point is to show that it can be replaced by the simpler interaction (11.19).

The first-order matrix element is

$$-i\int d^4x\langle f|\mathcal{H}_{\text{int}}(x)|i\rangle = -\frac{ie_0^2}{2}\int d^4x\, d^4y\langle f|Tj^0(x)j^0(y)|i\rangle D_{\text{Coul}}(x-y) \tag{11.24}$$

Note that this is actually proportional to e_0^2, and the factor $\frac{1}{2}$ in front makes this similar to a second-order matrix element in Feynman graphs. The second-order matrix element can be put in the form

$$\frac{(-i)^2}{2}\int d^4x\, d^4y\langle f|\mathrm{T}\,\mathcal{H}_{\text{int}}(x)\mathcal{H}_{\text{int}}(y)|i\rangle = -\frac{ie_0^2}{2}\int d^4x\, d^4y\langle f|\mathrm{T}j^r(x)j^s(y)|i\rangle\, D_{\mathrm{T}}^{rs}(x-y) \tag{11.25}$$

where the transverse photon propagator D_{T} is defined as

$$D_{\mathrm{T}}^{ij}(x-y) \equiv -i\langle 0|\mathrm{T}\, A^i(x)A^j(y)|0\rangle \tag{11.26}$$

There is no need to include the square of the Coulomb term, which is proportional to e_0^4.

We can write the transition matrix element to order e_0^2 as

$$\langle f|(S-1)|i\rangle = -\frac{ie_0^2}{2}\int d^4x\, d^4y \langle f|\mathrm{T}j^\mu(x)j^\nu(y)|i\rangle D_{\mu\nu}(x-y) \tag{11.27}$$

where $D_{\mu\nu}(x-y)$ is defined by

$$\begin{aligned} D_{00}(x) &= D_{\mathrm{Coul}}(x) \\ D_{rs}(x) &= D_{\mathrm{T}}^{rs}(x) \\ D_{s0}(x) &= 0 \end{aligned} \tag{11.28}$$

As it is an object in Coulomb gauge, $D^{\mu\nu}$ does not transform in a simple manner under a Lorentz transformation. We shall show, however, that because of current conservation, the part of $D^{\mu\nu}$ that contributes to the transition matrix element is covariant, and that can be taken as the effective photon propagator. The transverse propagator has been calculated in (5.94):

$$D_{\mathrm{T}}^{ij}(x) = \int \frac{d^4k}{(2\pi)^4}\frac{e^{ik\cdot x}}{k^2+i\eta} I^{ij}(\mathbf{k}) \tag{11.29}$$

where $k^2 = k_0{}^2 - |\mathbf{k}|^2$ and

$$I^{ij}(\mathbf{k}) = \delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2} \tag{11.30}$$

The Fourier transform is

$$\tilde{D}_{\mathrm{T}}^{ij}(k) = \frac{1}{k^2+i\eta}\left(\delta_{ij} - \frac{k^i k^j}{|\mathbf{k}|^2}\right) \tag{11.31}$$

The 3-tensor $I^{ij}(\mathbf{k})$ can be extended to a 4-tensor $I^{\mu\nu}(k)$ defined such that it reduces to $I^{ij}(\mathbf{k})$ in the special frame in which the Coulomb gauge is defined, namely, in which $k^\mu = (0, \mathbf{k})$ and $I^{i0} = I^{00} = 0$. Let ξ^μ and $\hat{k}^\mu$ be 4-vectors that reduce in the special frame to

$$\begin{aligned} \xi^\mu &= (1,0,0,0) \\ \hat{k}^\mu &= (0, \hat{\mathbf{k}}) \end{aligned} \tag{11.32}$$

where $\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|$. Then, in a general Lorentz frame, we have

$$I^{\mu\nu}(k) = -g^{\mu\nu} + \xi^\mu \xi^\nu - \hat{k}^\mu \hat{k}^n$$

$$\hat{k}^\mu = \frac{k^\mu - \xi^\mu (k\cdot\xi)}{\sqrt{(k\cdot\xi)^2 - k^2}} \tag{11.33}$$

We can now write down the 4-tensor that reduces to D_{T}^{ij} in the special frame:

$$D_{\mathrm{T}}^{\mu\nu}(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot x}}{k^2 + i\eta} I^{\mu\nu}(k) \tag{11.34}$$

The term $-g^{\mu\nu}$ in $I^{\mu\nu}$ gives the covariant part of the tensor, which we call the Feynman propagator D_{F}:

$$D_{\mathrm{T}}^{\mu\nu}(x) = D_{\mathrm{F}}^{\mu\nu}(x) + D_{\mathrm{I}}^{\mu\nu}(x) \tag{11.35}$$

In terms of Fourier transforms, we have

$$\tilde{D}_{\mathrm{F}}^{\mu\nu}(k) = -\frac{1}{k^2 + i\eta}$$

$$\tilde{D}_{\mathrm{I}}^{\mu\nu}(k) = \frac{1}{k^2 + i\eta}(\xi^\mu \xi^\nu - \hat{k}^\mu \hat{k}^\nu) \tag{11.36}$$

The inverse transform of $\tilde{D}_{\mathrm{I}}$ can be written as

$$D_{\mathrm{I}}^{\mu\nu}(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot x}}{(k\cdot\xi)^2 - k^2} [-\xi^\mu \xi^\nu - k^\mu k^\nu + (k\cdot\xi)(k^\mu \xi^\nu + k^\nu \xi^\mu)] \tag{11.37}$$

The last two terms vanish when contracted with $j^\mu j^\nu$, because $\partial_\mu j^\mu = 0$. In the special frame the first term reduces to

$$-g^{\mu 0} g^{\nu 0} \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik\cdot x}}{|\mathbf{k}|^2} = -g^{\mu 0} g^{\nu 0} D_{\mathrm{Coul}}(x) \tag{11.38}$$

Therefore

$$D_{\mathrm{T}}^{\mu\nu}(x) = D_{\mathrm{F}}^{\mu\nu}(x) - g^{\mu 0} g^{\nu 0} D_{\mathrm{Coul}}(x) + (\text{irrelevant terms}) \tag{11.39}$$

The S matrix can now be represented as

$$\langle f|(S-1)|i\rangle = -\frac{e^2}{2} \int d^4x\, d^4y \langle f|\mathrm{T} j_\mu(x) j_\nu(y)|i\rangle i D_{\mathrm{F}}^{\mu\nu}(x-y) \tag{11.40}$$

To obtain the same result, we can effectively take the interaction Hamiltonian density to be

$$\mathcal{H}_{\text{int}}(x) = e_0 j_\mu(x)[A^\mu(x) + A^\mu_{\text{ext}}(x)] \tag{11.41}$$

with the contraction rule

$$\langle 0|\mathrm{T}\, A^\mu(x)A^\nu(y)|0\rangle = iD_{\mathrm{F}}^{\mu\nu}(x-y)$$

$$D_{\mathrm{F}}^{\mu\nu}(x-y) = \int \frac{d^4k}{(2\pi)^4} e^{-ik\cdot x}\tilde{D}_{\mathrm{F}}^{\mu\nu}(k)$$

$$\tilde{D}_{\mathrm{F}}^{\mu\nu}(k) = -\frac{g^{\mu\nu}}{k^2 + i\eta} \tag{11.42}$$

This replacement applies not only to second-order matrix elements but to higher order ones as well. This justifies the effective interaction Hamiltonian density (11.19), which we shall use from now on, together with (11.42).

The freedom of gauge choice is reflected in the fact that terms in the photon propagator proportional to $k^\mu k^\nu$ have no effect on S-matrix elements. Special choices of such gauge terms can be useful for technical reasons. A popular form of the photon propagator is

$$\tilde{D}_{\mathrm{F}}^{\mu\nu}(k) = \left(g^{\mu\nu} - \lambda\frac{k^\mu k^\nu}{k^2}\right)\frac{-1}{k^2 + i\eta} \tag{11.43}$$

where λ is the gauge parameter. The choice $\lambda = 0$ corresponds to the so-called Feynman gauge and $\lambda = 1$, to Landau gauge.

11.3 FEYNMAN GRAPHS

The S matrix in QED can be effectively taken to be

$$S = \sum_{n=0}^{\infty} \frac{(-ie_0)^n}{n!} \int d^4x_1 \cdots d^4x_n \mathrm{T}[:\bar{\psi}_1 \not{A}_1 \psi_1: \cdots :\bar{\psi}_n \not{A}_n \psi_n:] \tag{11.44}$$

where $\not{A} = \gamma^\mu A_\mu$, $\psi_1 = \psi(x_1)$, and so forth. When there are external fields A^μ_{ext}, we replace $\not{A}$ by $\not{A} + \not{A}_{\text{ext}}$. In (11.44) we have ignored the term $-\delta m:\bar{\psi}\psi:$, which will be taken care of later.

We can represent the basic vertex $-ie_0\bar{\psi}(x)\not{A}(x)\psi(x)$ graphically as shown in Fig. 11.1. The wavy line represents an incoming or outgoing photon, and a directed line whose arrow points along the direction of flow of electron charge (which is negative), represents either an electron propagating along the arrow, or a positron against the arrow. To generate the Wick expansion of the nth-order S matrix, we draw n vertices, and make all possible contractions of the lines. The only nonvanishing contractions are the following:

$$\underbrace{A^\mu(x)A^\nu(y)} = iD_{\mathrm{F}}^{\mu\nu}(x-y)$$

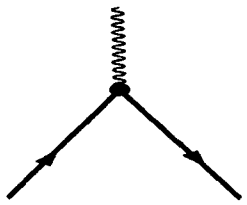


Figure 11.1 At the basic vertex of QED, a photon is emitted or absorbed from an electron line.

$$\underbrace{\psi(x)\psi^{\dagger}(y)} = iS_{\mathrm{F}}(x-y) \tag{11.45}$$

where the electron propagator is given in (7.30). The contraction $\psi^{\dagger}(x)\psi(x)$ does not occur, because the interaction is defined as a normal product. In contrast with scalar ϕ^4 theory, the lines meeting at a vertex are all different. As a consequence, the symmetry number is unity for all connected nonvacuum graphs.

For illustration, we display the Wick expansion of the second-order S matrix:

$$S^{(2)} = \tfrac{1}{2}(-ie_0)^2 \int d^4x_1 d^4x_2 (I_1 + \cdots + I_6) \tag{11.46}$$

The operators $I_1 \cdots I_6$ contribute to different processes, as indicated below:

$$\begin{aligned}
&\text{Disconnected:} && I_1 = :\bar{\psi}_1 \not{A}_1 \psi_1 : :\bar{\psi}_2 \not{A}_2 \psi_2 : \\
&\text{e–e scattering:} && I_2 = :\bar{\psi}_1 \not{A}_1 \underbrace{\psi_1 \bar{\psi}_2} \not{A}_2 \psi_2 : \\
&\text{Compton scattering:} && I_3 = :\bar{\psi}_1 \not{A}_1 \underbrace{\psi_1 \bar{\psi}_2} \not{A}_2 \psi_2 : + :\bar{\psi}_1 \not{A}_1 \underbrace{\psi_1 \bar{\psi}_2} \not{A}_2 \psi_2 : \\
&\text{Electron self-energy:} && I_4 = :\bar{\psi}_1 \not{A}_1 \underbrace{\psi_1 \bar{\psi}_2} \not{A}_2 \psi_2 : + :\bar{\psi}_1 \not{A}_1 \underbrace{\psi_1 \bar{\psi}_2} \not{A}_2 \psi_2 : \\
&\text{Photon self-energy:} && I_5 = :\bar{\psi}_1 \not{A}_1 \underbrace{\psi_1 \bar{\psi}_2} \not{A}_2 \psi_2 : \\
&\text{Vacuum process:} && I_6 = :\bar{\psi}_1 \not{A}_1 \underbrace{\psi_1 \bar{\psi}_2} \not{A}_2 \psi_2 :
\end{aligned} \tag{11.47}$$

We denote the matrix element for e–e scattering by

$$\mathcal{G}_{\mathrm{ee}}(p_1, p_2; p_3, p_4) \equiv \langle f | S^{(2)} | i \rangle \tag{11.48}$$

where $\mathbf{p}_i$ stands for the 4-momentum and the spin of the ith particle and the initial and final states are

$$\begin{aligned}
|i\rangle &= |\mathbf{p}_1, s_1; \mathbf{p}_2, s_2\rangle \\
|f\rangle &= |\mathbf{p}_3, s_3; \mathbf{p}_4, s_4\rangle
\end{aligned} \tag{11.49}$$

Explicitly, the matrix element is

$$\mathcal{G}_{ee}(p_1, p_2; p_3, p_4) = \frac{(-ie_0)^2 m^2}{\sqrt{E_1 E_2 E_3 E_4}} \int d^4x_1\, d^4x_2$$

$$[e^{i(p_3 - p_1)\cdot x1 - i(p_4 - p_2)\cdot x_2}(\bar{u}_3 \gamma^\mu u_1) i D_{F\mu\nu}(x - y)(\bar{u}_4 \gamma^\nu u_2) - (p_3 \rightleftarrows p_4)] \tag{11.50}$$

where $u_1 = u(\mathbf{p}_1, s_1)$, and so on. The two terms above correspond to distinct associations of external lines with external particles. The relative minus sign between them arises from the fact that the two possibilities differ by an interchange of fermion operators, which anticommute. In each case, the matrix elements acquire a factor 2 from the fact that the two vertices can be associated with x_1 and x_2, respectively, or vice versa. This cancels the $\frac{1}{2}$ in front of the second-order S matrix. Performing the space–time integrations leads to

$$G_{ee}(p_1, p_2; p_3, p_4) = -i(2\pi)^4 \delta^4(P_f - P_i) T_{fi}$$

$$T_{fi} = N_{fi} e_0^2 \left[\frac{(\bar{u}_3 \gamma^\mu u_1)(\bar{u}_4 \gamma_\nu u_2)}{(p_3 - p_2)^2 + i\eta} - (p_3 \rightleftarrows p_4) \right]$$

$$N_{fi} = \frac{m^2}{\sqrt{E_1 E_2 E_3 E_4}} \tag{11.51}$$

where P_f and P_i are respectively the total 4-momentum of the final and initial states. The S-matrix element is represented by the two Feynman graphs in Fig. 11.2*a*.

The operator I_2 also contributes to electron–positron ($e\bar{e}$) scattering:

$$\begin{array}{ccccccc} e & + & \bar{e} & \rightarrow & e & + & \bar{e} \\ p_1 & & q_1 & & p_2 & & q_2 \\ s_1 & & \sigma_1 & & s_2 & & \sigma_2 \end{array} \tag{11.52}$$

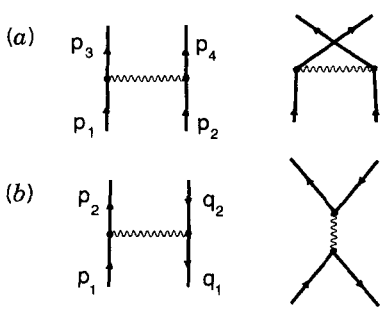


Figure 11.2 (*a*) Electron–electron scattering; (*b*) electron–positron scattering.

where positron spins are denoted by σ. The Feynman graphs are shown in Fig. 11.2*b*. Denoting $|i\rangle = |\mathbf{p}_1, s_1; \mathbf{q}_1, \sigma_1\rangle$ and $|f\rangle = |\mathbf{p}_2, s_2; \mathbf{q}_2, \sigma_2\rangle$, and

$$\mathcal{G}_{e\bar{e}}(p_1, q_1; p_2, q_2) \equiv \langle f|S^{(2)}|i\rangle \tag{11.53}$$

we note the relation

$$\mathcal{G}_{e\bar{e}}(p_1, q_1; p_2, q_2) = \mathcal{G}_{ee}(p_1, -p_2; q_1, -q_2) \tag{11.54}$$

which is a statement of *crossing symmetry*.

It in similar way, crossing symmetry relates Compton scattering

$$\begin{array}{ccccccc} e & + & \gamma & \rightarrow & e & + & \gamma \\ p_1 & & k_1 & & p_2 & & k_2 \\ s_1 & & \epsilon_1 & & s_2 & & \epsilon_2 \end{array} \tag{11.55}$$

and electron–positron annihilation

$$\begin{array}{ccccccc} e & + & \bar{e} & \rightarrow & \gamma & + & \gamma \\ p_1 & & q_1 & & k_1 & & k_2 \\ s_1 & & \sigma_1 & & \epsilon_1 & & \epsilon_2 \end{array} \tag{11.56}$$

where photon polarization is denoted by a 4-vector ϵ^μ, which takes the form

$$\epsilon^\mu = (0, \epsilon^1, \epsilon^2, 0) \tag{11.57}$$

with the 3-direction taken along the photon momentum. The corresponding Feynman graphs are shown in Fig. 11.3, and the matrix elements are given by

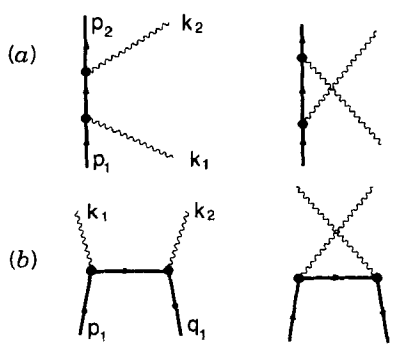


Figure 11.3 (*a*) Compton scattering; (*b*) pair annihilation.

$$\mathcal{G}_{\text{Compton}}(p_1, k_1; p_2, k_2) = -i(2\pi)^4\delta^4(P_f - P_i)T_{fi}$$

$$T_{\text{fi}} = N_{\text{fi}}e_0^2\left[\left(\bar{u}_2\not{\epsilon}_2\frac{1}{\not{p}_1 + \not{k}_1 - m + i\eta}\not{\epsilon}_1 u_1\right) + \left(u_2\not{\epsilon}_1\frac{1}{\not{p}_1 - \not{k}_2 - m + i\eta}\not{\epsilon}_2 u_1\right)\right]$$

$$N_{\text{fi}} = \frac{1}{\sqrt{2\omega_1 2\omega_2}}\frac{m}{\sqrt{E_1E_2}} \tag{11.58}$$

The statement of crossing symmetry in this case is

$$\mathcal{G}_{\text{annihilation}}(p_1, q_1; k_1, k_2) = \mathcal{G}_{\text{Compton}}(p_1, -k_1; -q_1, k_2) \tag{11.59}$$

The electron self-energy corresponds to

$$\langle \mathbf{p}_2, s_2|S^{(2)}|\mathbf{p}_1, s_1\rangle = \frac{(2\pi)^4\delta^4(p_2 - p_1)m}{\sqrt{E_1E_2}}\bar{u}_2\Sigma\, u_1$$

$$\Sigma = (-ie_0)^2\int\frac{d^4k}{(2\pi)^4}\gamma^\mu iS_{\text{F}}(p_1 - k)\gamma^\nu i\tilde{D}_{\text{F}\mu\nu}(k) \tag{11.60}$$

and the photon self-energy corresponds to

$$\langle \mathbf{k}_2, \epsilon_2|S^{(2)}|\mathbf{k}_1, \epsilon_1\rangle = \frac{(2\pi)^4\delta^4(k_2 - k_1)}{\sqrt{2\omega_1 2\omega_2}}\epsilon_{2\mu}\Pi^{\mu\nu}\epsilon_{1\mu}$$

$$\Pi^{\mu\nu} = (-ie_0)^2\int\frac{d^4p}{(2\pi)^4}(-1)\text{Tr}[\gamma^\mu iS_F(p)\gamma^\nu iS_F(k_1 - p)] \tag{11.61}$$

The Feynman graphs are shown in Fig. 11.4. The expression for the photon self-energy needs a little explanation. The contractions indicated in I_5 of (11.47) give, with all spinor indices written out,

$$\begin{aligned}\overline{\psi}_1\gamma^\mu\overline{\psi_1}\overline{\psi}_2\gamma^\nu\psi_2 &= \langle 0|\text{T}\,\overline{\psi}_{1\alpha}\psi_{2\rho}|0\rangle\,\langle 0|\text{T}\,\psi_{1\beta}\overline{\psi}_{2\lambda}|0\rangle(\gamma^\mu)_{\alpha\beta}(\gamma^\nu)_{\lambda\rho}\\ &= iS_{\text{f}}(x_1 - x_2)_{\beta\lambda}(\gamma^\nu)_{\lambda\rho}(-i)S_{\text{F}}(x_2 - x_1)_{\rho\alpha}(\gamma^\mu)_{\alpha\beta}\\ &= -\text{Tr}[iS_{\text{F}}(x_1 - x_2)\gamma^\nu iS_{\text{F}}(x_2 - x_1)\gamma^\mu]\end{aligned} \tag{11.62}$$

Figure 11.4 (*a*) Electron self-energy; (*b*) photon self-energy.

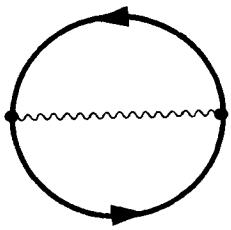


Figure 11.5 A vacuum graph.

Note the minus sign in $(-i)S_F$ on the second line. It arises when we had to rewrite $\bar{\psi}\psi$ as $-\psi\bar{\psi}$. As a rule, a closed fermion loop always contributes a factor -1.

The second-order vacuum graph in Fig. 11.5 corresponds to the matrix element

$$\langle 0|S^{(2)}|0\rangle = (2\pi)^4\delta^4(0)\frac{(-ie_0)^2}{2}\int\frac{d^4k\,d^4p}{(2\pi)^8}(-1)\mathrm{Tr}[\gamma^\mu i\tilde{S}_F(p)\gamma^\nu i\tilde{S}_F(p+k)i\tilde{D}_{F\mu\nu}(k)] \tag{11.63}$$

In this case, interchanging the two vertices does not lead to a new configuration, and the symmetry number is 2. But this is a graph we can ignore.

Consider now the mass counterterm $-\delta m{:}\bar{\psi}\psi{:}$ in (11.19). Because of normal ordering, the two fields occurring in this term do not contract with each other, and must be contracted with other fields. This term therefore gives rise to a vertex where an electron line enters and leaves, with no photon line, and give a factor $(-i)(-\delta m) = i\delta m$. Every internal or external electron line should, in principle, be replaced according to

$$iS_F \to iS_F + iS_F(i\delta m)iS_F \tag{11.64}$$

as indicated in Fig. 11.6; but since δm is considered to be second order in e_0^2, the replacement has no effect on second-order matrix elements. Its effect in higher orders will be discussed in Chapter 12. We now state the Feynman rules for QED.

11.4 FEYNMAN RULES

A Feynman graph is made up of vertices, electron lines, and photon lines. An electron line is directed. It describes the propagation of an electron along the line direc-

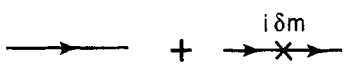


Figure 11.6 The effect of the mass counterterm is to replace all electron lines by the sum of the two graphs above, where a cross denotes a factor $i\delta$.

tion, or a positron against the line direction. A photon line has no direction, and represents either an incoming or outgoing photon.

Each vertex emits an electron line, absorbs another electron line, and emits (or absorbs) a photon line. Thus, an electron line cannot originate or end inside a graph. This means that the "electron number, " the number of electrons minus the number of positrons, is conserved. The number of photons, on the other hand, is not conserved.

A graph consists of a number of electron lines going through the graph, and a number of electron lines in closed loops. On these lines are vertices from which photons are emitted or absorbed.. A closed loop must have at least two vertices.

A disconnected graph is a product of its subgraphs. Thus it suffices to consider connected graphs, whose contribution to the S matrix may be obtained as follows:

- Each vertex with 4-vector index μ contributes a factor $-ie_0\gamma^\mu$. The total 4-momentum flowing into a vertex is zero.
- Each internal photon line of 4-momentum k contributes a factor $i\tilde{D}_{\mathrm{F}\mu\nu}(k)$, where μ and ν are contracted with the indices of the vertices at its ends.
- Each internal electron line of 4-momentum p contributes a factor $i\tilde{S}_{\mathrm{F}}(p)$.
- Each external photon line of momentum k and polarization ϵ contributes a wave function factor $\epsilon^\mu/\sqrt{2\omega_k}$, where μ is contracted with the index of the vertex to which the line is attached.
- Each outgoing external electron line of momentum $\mathbf{p}$ and spin s contribute a wave function factor

$$\sqrt{\frac{m}{E_p}}\bar{u}(\mathbf{p}, s) \quad \text{(if outgoing electron)}$$

$$\sqrt{\frac{m}{E_p}}v(-\mathbf{p}, s) \quad \text{(if incoming positron)}$$

- Each incoming external electron line of momentum p and spin s contribute a wave-function factor

$$\sqrt{\frac{m}{E_p}}u(\mathbf{p}, s) \quad \text{(if incoming electron)}$$

$$\sqrt{\frac{m}{E_p}}v(-\mathbf{p}, s) \quad \text{(if outgoing positron)}$$

- Each closed electron loop contributes a factor -1.
- There is an overall factor $(2\pi)^4\delta^4(P_\mathrm{f} - P_\mathrm{i})$.
- Each independent internal 4-momentum k is integrated over, with measure $d^4k/(2\pi)^4$.
- Connected nonvacuum graphs have symmetry number 1. The symmetry

number of a vacuum graph is generally different from 1, but vacuum graphs should be ignored. (See below.)

- The contribution of a Feynman graph to a correlation function follows the same rules as above, except that the wave function factors of external lines are replaced with propagators.

11.5 PROPERTIES OF FEYNMAN GRAPHS

In a general Feynman graph, let

$$\begin{aligned} n &= \text{number of vertices} \\ E_e &= \text{number of external electron lines} \\ E_i &= \text{number of internal electron lines} \\ P_e &= \text{number of external photon lines} \\ P_i &= \text{number of internal photon lines} \end{aligned} \tag{11.65}$$

Note first that $n - P_e$ is the number of vertices that do not emit external photon lines. Since these vertices must be connected in pairs by internal photon lines, we have $n - P_e = 2P_i$, or

$$P_i = \tfrac{1}{2}(n - P_e) \tag{11.66}$$

Since an external electron line must enter and exit the graph, E_e is an even integer. An external electron line touches only one vertex, while an internal electron lines touches two. Therefore

$$E_i = n - \tfrac{1}{2}E_e \tag{11.67}$$

The number of internal 4-momenta is $P_i + E_i$, but not all are independent because of 4-momentum conservation at the vertices. The overall 4-momentum conservation is not a restriction on internal lines. Therefore there are $n - 1$ constraints, and the number of independent integrations over internal 4-momenta is

$$N = P_i + E_i - n + 1 \tag{11.68}$$

We can now show that, just as in scalar theory, a vacuum graph in QED is a pure imaginary number. In a vacuum graph, n must be even. There are E_i electron propagators of the form

$$\frac{i(\not{p} + m)}{p^2 - m^2 + i\eta}$$

where $(\not{p} + m)$ yield a real number when operating on a Dirac spinor and the de-

nominator is real after a Wick rotation. Hence each electron propagator contributes a pure-imaginary factor. There are P_i photon propagators of the form

$$-\frac{i}{k^2 + i\eta}$$

which is pure-imaginary after a Wick rotation.

Therefore a vacuum graph is of the form

$$\Gamma \sim (-i\lambda_0)^n i^{E_i+P_i+N} \int d^4k_1 \cdots d^4k_N f(k_1, \cdots, k_N) \tag{11.69}$$

where f is real and a factor i^N comes from the Wick rotations. Noting that $(-i)^n$ is real because n is even, we have $\Gamma \sim i^{E_i+P_i+N} \times$ (real number). Since $E_e = P_e = 0$, we have $E_i = n$, $P_i = n/2$, $N = (n/2) + 1$. Thus $E_i + P_i + N = 2n + 1$. Therefore

$$\Gamma \sim i \times \text{(real number)} \tag{11.70}$$ ■

As shown in Section 9.9, this means that the sum of all vacuum graphs gives a phase factor, and consequently we can ignore all vacuum subgraphs.

A useful property, known as *Furry's theorem* [1], is that a graph or subgraph with an odd number of external photons may be ignored, because it is cancelled by similar graphs. This is because such a graph must contain one closed electron loop that emits an odd number of photons. As illustrated in Fig. 11.7, there exist a graph identical in every way except that the sense of the electron is reversed. These two graphs go into each other under charge conjugation, which sends e_0 to $-e_0$. Hence the sum of the two graphs is zero.

Suppose the interaction Hamiltonian density is of the form

$$\mathcal{H}(x) = P(x) + Q(x) \tag{11.71}$$

as, for example, $\mathcal{H}(x) = e_0\bar{\psi}\not{A}\psi + e_0\bar{\psi}\not{A}_{ext}\psi$. The following double series expansion makes it possible to take P into account to order n, and Q into account to order m, independently:

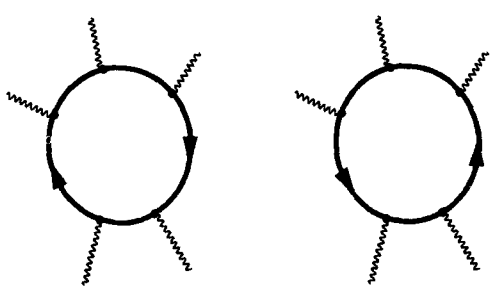


Figure 11.7 According to Furry's theorem, these graphs cancel each other. They have an odd number of external photon lines and no external electrons lines, and differ only in the sense of the closed electron loop.

$$S=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(-1)^n}{n!}\frac{(-1)^m}{m!}\int d^4x_1\cdots d^4x_{n+m}\mathrm{T}[P(x_1)\cdots P(x_n)Q(x_{n+1})\cdots Q(x_{n+m})] \tag{11.72}$$

For example, in an atomic problem, we might want to treat the Coulomb interaction A^{μ}_{ext} exactly, but take radiative corrections into account only to second order.

To derive this expansion, we start with the perturbation series

$$S=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\int d^4x_1\cdots d^4x_k\mathrm{T}\{[(P(x_1)+Q(x_1)]\cdots[(P(x_k)+Q(x_k)]\} \tag{11.73}$$

and group terms by the number of factors of P. Terms in which P occurs n times, regardless of the number of Q factors, have the form

$$\sum_{m=0}^{\infty}\binom{n+m}{n}\int d^4x_1\cdots d^4x_{n+m}\mathrm{T}[(P(x_1)\cdots P(x_n)Q(x_{n+1})\cdots Q(x_{n+m})]$$

To get S, we multiply this by $(-i)^{n+m}/(n+m)!$ and sum over n. Using

$$\frac{1}{(n+m)!}\binom{n+m}{m}=\frac{1}{n!m!}$$

we obtain the desired result. ∎

PROBLEMS

11.1 Write down the matrix elements corresponding to the contractions in (11.47).

11.2 Draw all fourth-order Feynman graphs for electron–electron scattering, that is, graphs with four vertices. Write down the corresponding matrix elements using the Feynman rules. Include *all* graphs, connected or disconnected.

11.3 **(a)** Draw all fourth-order Feynman graphs for photon–photon scattering (scattering of light by light), and write down the corresponding matrix elements.

(b) Adopt the matrix elements obtained above to the scattering of light by an external Coulomb field (Delbrück scattering).

11.4 Adopt the matrix element for Compton scattering to the scattering of an electron by an external Coulomb field, with emission of light (*bremsstrahlung*).

11.5 Consider the scattering of light by a Dirac particle of charge e and mass M, to lowest order in e. Obtain the matrix element in the limit $M\to\infty$, and calculate the differential cross section. The particle can be a proton, except for the neglect of the anomalous magnetic moment.

REFERENCE

1. W. Furry, *Phys. Rev.* **51**, 125 (1937).

CHAPTER TWELVE

Processes in Quantum Electrodynamics

12.1 COMPTON SCATTERING

We shall derive the differential cross section for Compton scattering. The S matrix element has been given in (11.58):

$$\mathcal{G}_{\text{Compton}}(p_1, k_1; p_2, k_2) = -i(2\pi)^4 \delta^4(P_f - P_i) T_{fi}$$

$$T_{fi} = N_{fi} e_0^2 \left[\left(\bar{u}_2 \not{\epsilon}_2 \frac{1}{\not{p}_1 + \not{k}_1 - m + i\eta} \not{\epsilon}_1 u_1 \right) + \left(u_2 \not{\epsilon}_1 \frac{1}{\not{p}_1 - \not{k}_2 - m + i\eta} \not{\epsilon}_2 u_1 \right) \right]$$

$$N_{fi} = \frac{1}{\sqrt{2\omega_1 2\omega_2}} \frac{m}{\sqrt{E_1 E_2}} \tag{12.1}$$

where

$$\omega = k^0$$

$$E_p = +\sqrt{\mathbf{p}^2 + m^2} \tag{12.2}$$

and the photon polarization vectors have the form

$$\epsilon^\mu = (0, \epsilon^1, \epsilon^2, 0) \tag{12.3}$$

with the x^3 axis taken along the photon momentum.

Since the S matrix element is of second order in e_0, we can put $e_0 = e$. We can also drop the $i\eta$ in the propagators, because the internal momenta are fixed, and not integrated over. The transition matrix is thus

$$T_{\rm fi} = N_{\rm fi} e^2 \bar{u}_2 \left[\frac{\not\epsilon_2(\not p_1 - \not k_2 + m)\not\epsilon_1}{(p_1 + k_1)^2 - m^2} + \frac{\not\epsilon_1(\not p_1 - \not k_2 + m)\not\epsilon_2}{(p_1 - k_2)^2 - m^2} \right] u_1 \tag{12.4}$$

We now use the relations

$$\begin{aligned} p_1 \cdot \epsilon_1 &= p_1 \cdot \epsilon_2 = 0 \\ k_1 \cdot \epsilon_1 &= k_2 \cdot \epsilon_2 = 0 \\ k_1^2 &= k_2^2 = 0 \end{aligned} \tag{12.5}$$

The first holds because it is true in the laboratory frame. The second is a statement of transversality, and the third is a statement that the photons are real (i.e., on mass shell). Therefore we have

$$\begin{aligned} (p_1 + k_1)^2 - m^2 &= 2p_1 \cdot k_1 \\ (p_1 - k_2)^2 - m^2 &= -2p_1 \cdot k_2 \end{aligned} \tag{12.6}$$

Using the identity

$$\not A \not B + \not B \not A = 2A \cdot B \tag{12.7}$$

we can write

$$\not p_1 \not\epsilon_i u_1 = -\not\epsilon_i \not p_1 u_1 = -m \not\epsilon_i u_1 \tag{12.8}$$

Hence

$$T_{\rm fi} = \tfrac{1}{2} e^2 N_{\rm fi} (\bar{u}_2 \Gamma u_1) \tag{12.9}$$

where

$$\Gamma = \frac{\not\epsilon_2 \not k_1 \not\epsilon_1}{p_1 \cdot k_1} + \frac{\not\epsilon_1 \not k_2 \not\epsilon_2}{p_1 \cdot k_2} \tag{12.10}$$

The differential cross section is given by

$$I d\sigma = \int \frac{d^3 p_2 d^3 k_2}{(2\pi)^6} (2\pi)^4 \delta^4 (P_{\rm f} - P_{\rm i}) \tfrac{1}{4} \sum_{\rm spins} |T_{\rm fi}|^2 \tag{12.11}$$

where $I = 1$ in the normalization we are using. The final spin states are summed over, because they all contribute to the cross section. The initial spin states are averaged over, because we assume that the incident electron is unpolarized.

We now work out the kinematics of the reaction. In the laboratory coordinate system, shown in Fig. 12.1, we have

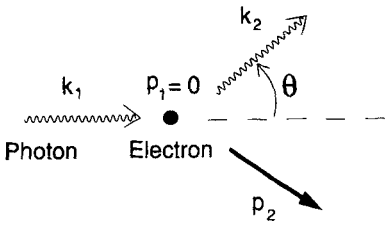


Figure 12.1 Kinematics of Compton scattering.

$$\begin{aligned} \mathbf{p}_1 &= 0 \qquad E_1 = \mathrm{m} \\ \mathbf{p}_2 &= \mathbf{k}_1 - \mathbf{k}_2 \\ |\mathbf{k}_1| &= \omega \qquad |\mathbf{k}_2| = \omega' \end{aligned} \tag{12.12}$$

Let

$$f \equiv E_2 + \omega' - \omega - m \tag{12.13}$$

Energy conservation requires $f = 0$, or

$$\begin{aligned} 0 &= \sqrt{(\mathbf{k}_1 - \mathbf{k}_2)^2 + m^2} + \omega' - \omega - m \\ &= \sqrt{\omega^2 + \omega'^2 - 2\omega\omega'\cos\theta + m^2} + \omega' - \omega - m \end{aligned} \tag{12.14}$$

where θ is the scattering angle. A little algebra yields the relation between the final photon frequency and the scattering angle:

$$\omega' = \frac{\omega}{1 + (\omega/m)(1 - \cos\theta)} \tag{12.15}$$

In the cross section, the integration over the final momenta $\mathbf{p}_2$ is entirely fixed by momentum conservation, while the $\mathbf{k}_2$ integration is limited to a solid-angle element $d\Omega$. The magnitude of $\mathbf{k}_2$ is fixed by energy conservation. The "phase-space" integral is

$$\int \frac{d^3p_2\, d^3k_2}{(2\pi)^6}(2\pi)^4\delta^3(\mathbf{p}_2 + \mathbf{k}_2 - \mathbf{k}_1)\delta(f) = \frac{d\Omega}{(2\pi)^2}\left[\frac{\omega'^2}{d\omega'/df}\right]_{f=0} = \frac{E_2\omega'^3}{m\omega} \tag{12.16}$$

Putting all this together yields

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{16}\left(\frac{\omega'}{\omega}\right)^2 \sum_{\text{spins}} |\bar{u}_2 \Gamma u_1|^2 \tag{12.17}$$

where $\alpha = e^2/4\pi$ is the fine-structure constant.

To perform the sum over the initial and final spin states, we insert projection operators for positive-energy states, and then sum over all states with the given momenta:

$$\sum_{\text{spins}} |\bar{u}_2 \Gamma u_1|^2 = \sum_{\text{spins}} (\bar{u}_2 \Gamma u_1)^* (\bar{u}_2 \Gamma u_1) = \sum_{\text{spins}} (\bar{u}_1 \beta \Gamma^\dagger \beta u_2)(\bar{u}_2 \Gamma u_1)$$

$$= \text{Tr}\left(\beta \Gamma^\dagger \beta \frac{\not{p}_2 + m}{2m} \Gamma \frac{\not{p}_1 + m}{2m} \right) \tag{12.18}$$

Using the relation

$$\beta \Gamma^\dagger \beta = \frac{\not{\varepsilon}_1 \not{k}_1 \not{\varepsilon}_2}{p_1 \cdot k_1} + \frac{\not{\varepsilon}_2 \not{k}_2 \not{\varepsilon}_1}{p_1 \cdot k_2} \tag{12.19}$$

we find, after some rearrangement

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{64m^2} \left(\frac{\omega'}{\omega} \right)^2 \text{Tr}[F(\not{p}_2 + m) G(\not{p}_1 + m)] \tag{12.20}$$

where

$$F = \frac{\not{k}_1 \not{\varepsilon}_1 \not{\varepsilon}_2}{p_1 \cdot k_1} + \frac{\not{k}_2 \not{\varepsilon}_2 \not{\varepsilon}_1}{p_1 \cdot k_2}$$

$$G = \frac{\not{\varepsilon}_2 \not{\varepsilon}_1 \not{k}_1}{p_1 \cdot k_1} + \frac{\not{\varepsilon}_1 \not{\varepsilon}_2 \not{k}_2}{p_1 \cdot k_2} \tag{12.21}$$

The spin traces can be evaluated using the basic identities listed in Table 12.1.

The result for the differential cross section for polarized photons is

TABLE 12.1 Traces of Dirac Matrices

$\text{Tr } 1 = 4$
$\text{Tr } \gamma^\mu = 0$
$\text{Tr } \gamma_5 = 0$
$\text{Tr}(\gamma^\mu \gamma^\nu) = 4g^{\mu\nu}$
$\text{Tr}(\not{p}_1 \not{p}_2) = p_1 \cdot p_2$
$\text{Tr}(\not{p}_1 \not{p}_2 \not{p}_3 \not{p}_4) = (p_1 \cdot p_2)(p_3 \cdot p_4) + (p_1 \cdot p_4)(p_2 \cdot p_3) - (p_1 \cdot p_3)(p_2 \cdot p_4)$
$\text{Tr}(\text{odd number of } \gamma) = 0$

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4m^2}\left(\frac{\omega'}{\omega}\right)^2\left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} + 4(\not{\epsilon}_1\cdot\not{\epsilon}_2)^2 - 2\right] \tag{12.22}$$

Summing over the final photon polarization, and averaging over the initial polarization, we obtain the *Klein–Nishina formula* [1]:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{2m^2}\left(\frac{\omega'}{\omega}\right)^2\left[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} - \sin^2\theta\right] \tag{12.23}$$

In the low-frequency limit $\omega/m \to 0$, this approaches the Thomson cross section

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{Thomson}} = \frac{\alpha^2}{2m^2}\cos^2\theta \tag{12.24}$$

which is symmetric about $\theta = 90°$. As the energy of the incident photon increases, the distribution tends to peak about the forward direction. Graphs for the angular distribution, normalized to unity in the forward direction, are shown in Fig. 12.2, with comparison to experiments [2].

12.2 ELECTROMAGNETIC FORM FACTORS

Consider the scattering of an electron by an external electromagnetic field A_μ^{ext}. If the external field is very weak, we can treat it in lowest-order perturbation theory, but in principle consider radiative corrections to all orders. The S matrix element between the initial state 1 and final state 2 is then, according to (11.72),

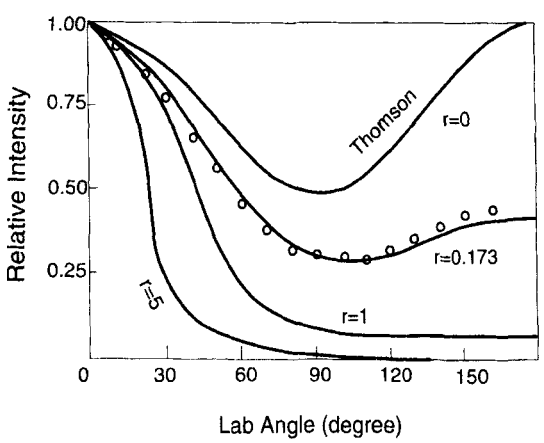


Figure 12.2 Angular distribution in Compton scattering, with $r = \omega/m$. (After W. Heitler, *Quantum Theory of Radiation*, 3rd ed. Oxford Univ. Press, London, 1954).

$$S = -i\sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int d^4x\, d^4y_1 \cdots d^4y_n \langle 2|\mathrm{T}\mathcal{H}^{\mathrm{ext}}(x)\mathcal{H}(y_1)\cdots\mathcal{H}(t_n)|1\rangle \quad (12.25)$$

where

$$\mathcal{H}^{\mathrm{ext}}(x) = e_0\bar{\psi}(x)\not{A}^{\mathrm{ext}}(x)\psi(x) \quad (12.26)$$

For a static external field, we have the following in lowest order ($n = 0$):

$$\begin{aligned} S &= -ie_0 \int d^4x \langle 2|\bar{\psi}(x)\gamma^\mu\psi(x)|1\rangle\, A_\mu^{\mathrm{ext}}(\mathbf{x}) \\ &= 2\pi\delta(E_2 - E_1)\frac{-ie_0 m}{\sqrt{E_2E_1}}(\bar{u}_1)\gamma^\mu u_1 \tilde{A}^{\mathrm{ext}}(\mathbf{p}_2 - \mathbf{p}_1) \end{aligned} \quad (12.27)$$

This is represented by the lowest-order Feynman graph in Fig. 12.3*a*. The transition probability per unit time is formally given by $|S_{21}|^2/2\pi\delta(0)$, and the scattering cross section is given by

$$Id\sigma = \int \frac{d^3p_2}{(2\pi)^3} 2\pi\delta(E_2 - E_1)\frac{e_0^2 m^2}{E_2E_1}|(\bar{u}_1)\gamma^\mu u_1 \tilde{A}^{\mathrm{ext}}(\mathbf{p}_2 - \mathbf{p}_1)|^2 \quad (12.28)$$

The electron wave functions are normalized to unit density, and hence in the laboratory frame we have

$$I = \frac{|\mathbf{p}_1|}{E_1} \quad (12.29)$$

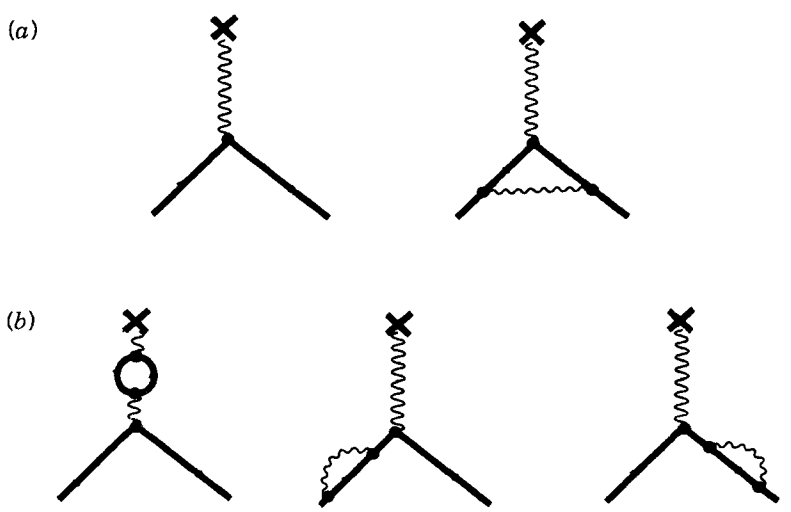


Figure 12.3 (*a*) Proper vertex graphs. (*b*) Vertex graphs with insertions.

which is the velocity of the incident electron.

To include higher-order radiative corrections, we only need to replace $\bar{u}_2\gamma^\mu u_1$ by a more general quantity corresponding to the Feynman graphs in Fig. 12.3, which are call "vertex graphs." These graphs can be separated into two categories: the "proper" ones and "improper" ones. The former, also called "one-particle irreducible," are connected graphs that cannot be made disconnected by cutting one internal line. The improper graphs have insertions on the external lines, and contribute to mass renormalization of the external particles.

The sum of the proper vertex graphs, with external wave functions omitted, defines the "proper vertex part" Λ^μ:

$$\Gamma^\mu(p_2, p_1) = \gamma^\mu + \Lambda^\mu(p_2, p_1) \tag{12.30}$$

where the term γ is the bare vertex, and $\Lambda\ (p_2, p_1)$ is defined by the series of graphs in Fig. 12.3*a*. The scattering cross section to first order in the external field, with all radiative corrections taken into account, is given by (12.28) with $\bar{u}_2\gamma^\mu u_1$ replaced by $\bar{u}_2\Gamma^\mu u_1$.

By Lorentz covariance, $\Gamma^\mu(p_2, p_1)$ must be of the form

$$\Gamma^\mu(p_1, p_2) = C_1 p_1^\mu + C_2 p_2^\mu + C_3\gamma^\mu + C_4\sigma^{\mu\nu}p_{1\nu} + C_5\sigma^{\mu\nu}p_{2\nu} \tag{12.31}$$

where C_i are Lorentz-invariant functions of p_1 and p_2. Let

$$\begin{aligned} P^\mu &= p_1^\mu + p_2^\mu \\ k^\mu &= p_1^\mu - p_2^\mu \end{aligned} \tag{12.32}$$

When the external lines are on mass shell, with $p_1^2 = p_2^2 = m^2$, there is only one independent invariant, which we shall choose to be k^2. Current conservation requires

$$k_\mu \bar{u}_2\Gamma^\mu(p_1, p_2)u_1 = 0 \tag{12.33}$$

This leads to the conditions $C_1 = C_2 = 0$ and $C_4 + C_5 = 0$. Consequently, there are only two independent functions C_3 and C_4, and we can rewrite the vertex operator in terms of two invariant form factors:

$$\Gamma^\mu(p_1, p_2) = F_1(k^2)\gamma^\mu + \frac{1}{2m}F_2(k^2)i\sigma^{\mu\nu}k_\nu \tag{12.34}$$

It should be emphasized that there will be extra form factors when p_1 and p_2 are not on mass shell.

We now use the Gordon decomposition (Problem 6.6)

$$\bar{u}_2\gamma^\mu u_1 = \frac{1}{2m}\bar{u}_2(P^\mu + i\sigma^{\mu\nu}k_\nu)u_1 \tag{12.35}$$

to write

$$\bar{u}_2\Gamma^\mu(p_2, p_1)u_1 = \frac{1}{2m}\bar{u}_2[F_1(P^\mu + i\sigma^{\mu\nu}k_\nu) + F_2 i\sigma^{\mu\nu}k_\nu]u_1$$

$$= \frac{1}{2m}\bar{u}_2[F_1 P^\mu + (F_1 + F_2)i\sigma^{\mu\nu}k_\nu]u_1 \tag{12.36}$$

Two alternative form factors are

$$F_E(k^2) = F_1(k^2)$$
$$F_M(k^2) = F_1(k^2) + F_2(k^2) \tag{12.37}$$

which are respectively the electric and magnetic form factors.

The S matrix element for scattering from a weak external field, with radiative corrections fully taken into account, is given by

$$S = -ie_0 N\int d^4x\, e^{ik\cdot x}\,\bar{u}_2\Gamma^\mu(p_2, p_1)u_1 A_\mu^{\text{ext}}(x) \tag{12.38}$$

where

$$N = \frac{m}{\sqrt{E_1 E_2}} \tag{12.39}$$

For a static field this reduces to

$$S = -iNe_0 2\pi\delta(E_2 - E_1)\int d^3x\, e^{-i\mathbf{k}\cdot\mathbf{x}}\,\bar{u}_2\left[F_1\gamma^\mu + \frac{F_2}{2m}i\sigma^{\mu\nu}\,k_\nu\right]u_1 A_\mu^{\text{ext}}(\mathbf{x}) \tag{12.40}$$

To discover the physical meaning of the form factors, consider forward scattering in the nonrelativistic limit, for which

$$\mathbf{k} \to 0 \qquad P^0 \to 2m \qquad N \to 1 \tag{12.41}$$

First look at the $F_1\gamma^\mu$ term in the vertex part. Using the Gordon decomposition to rewrite $\bar{u}_2\gamma^\mu u_1$, we can write the nonrelativistic limit in the form

$$S_1 \to -\frac{ie_0 F(0)}{2m}2\pi\delta(E_2 - E_1)\int d^3x\, e^{-i\mathbf{k}\cdot\mathbf{x}}\bar{u}_2(P^\mu + i\sigma^{\mu\nu}k_\nu)u_1 A_\mu^{\text{ext}}(\mathbf{x}) \tag{12.42}$$

For an electrostatic potential $A_0 = V(\mathbf{r})$, $A_k = 0$, the second term vanishes when $\mathbf{k} \to 0$, and we have

$$S_1 \to -ie_0 F_1(0)2\pi\delta(E_1 - E_2)\int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}}V(\mathbf{r})\bar{u}_2 u_1 \tag{12.43}$$

This shows that the coupling constant to an electrostatic potential is $e_0F_1(0)$, which is by definition the electric charge of the particle:

$$e \equiv e_0F_1(0) \tag{12.44}$$

Next choose the external field be a static magnetic field **B**, with $A^0 = 0$, $\mathbf{A} = \mathbf{r} \times \mathbf{B}/2$. We have

$$A_\mu^{\text{ext}}(P^\mu + i\sigma^{\mu\nu}k_\nu) = -\mathbf{A}\cdot(\mathbf{p}_1 + \mathbf{p}_2) + iA_i\sigma^{ij}k_j \tag{12.45}$$

The first term gives no contribution in the limit $\mathbf{k} \to 0$, because $\int d^3r\mathbf{A} = 0$. The second term can be worked out as follows:

$$A_i\sigma^{ij}k_j = A_i\epsilon^{ijn}\sigma^n k_j = \boldsymbol{\sigma}\cdot\mathbf{A} \times \mathbf{k}$$
$$\to \boldsymbol{\sigma}\cdot\mathbf{A} \times i\nabla \to \boldsymbol{\sigma}\cdot i\nabla \times \mathbf{A} = i\boldsymbol{\sigma}\cdot\mathbf{B}$$

Thus

$$S_1 \to \frac{ie}{m}2\pi\delta(E_1 - E_2)\int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}}\bar{u}_2\mathbf{S}\cdot\mathbf{B}u_1 \tag{12.46}$$

where $\mathbf{S} = \boldsymbol{\sigma}/2$. This describes the scattering of a particle of magnetic moment

$$\boldsymbol{\mu}_0 = \frac{e}{m}\mathbf{S} \tag{12.47}$$

which is the Dirac moment, with g factor 2.

Consider now the $F_2\sigma^{\mu\nu}$ term in the vertex part. The S-matrix element for forward scattering from an external magnetic field is, in the nonrelativistic limit

$$S_2 \to \frac{e_0}{m}F_2(0)2\pi i\delta(E_1 - E_2)\int d^3r\, e^{-i\mathbf{k}\cdot\mathbf{r}}\bar{u}_2\mathbf{S}\cdot\mathbf{B}u_1 \tag{12.48}$$

which describes the effect of an extra magnetic moment over and above the Dirac moment:

$$\boldsymbol{\mu}_1 = \frac{e_0}{m}F_2(0)\mathbf{S} \tag{12.49}$$

This is called the "anomalous magnetic moment." The total magnetic moment is thus given by

$$\boldsymbol{\mu} = \frac{e}{m}\left[1 + \frac{F_2(0)}{F_1(0)}\right]\mathbf{S} \tag{12.50}$$

The g factor is given by

$$g = 2\left[1 + \frac{F_2(0)}{F_1(0)}\right] \tag{12.51}$$

The factor 2 comes from expressing it in units of the Bohr magneton $e\hbar/2mc$.

12.3 ANOMALOUS MAGNETIC MOMENT

From the second-order Feynman graph in Fig. 12.3*a* we obtain

$$\begin{aligned}\Lambda^{\mu}(p_1, p_2) &= (-ie_0)^2 \int \frac{d^4q}{(2\pi)^4} i\tilde{D}_F(q)\gamma^{\lambda} i\tilde{S}_F(p_1 - q)\gamma^{\mu} i\tilde{S}_F(p_2 - q)\gamma_{\lambda} \\ &= -ie_0^2 \int \frac{d^4q}{(2\pi)^4} \frac{\gamma^{\lambda}(\not{p}_1 - \not{q} + m)\gamma^{\mu}(\not{p}_2 - \not{q} + m)\gamma_{\lambda}}{(q^2 + i\eta)[(p_1 - q)^2 - m^2 + i\eta][(p_2 - q)^2 - m^2 + i\eta]}\end{aligned} \tag{12.52}$$

The integral is divergent both in the ultraviolet ($q \to \infty$) and in the infrared ($q \to 0$). The ultraviolet divergence is logarithmic. We introduce both high-q and low-q cutoffs, and explain later how to dispose of them.

The factors in the denominator of the integrand can be combined, with the help of Feynman's formula [3]:

$$\frac{1}{a_1 \cdots a_n} = (n-1)\int_0^1 dx_1 \cdots dx_n \frac{\delta((x_1 + \cdots + x_n) - 1)}{(x_1 a_1 + \cdots + x_n a_n)^n} \tag{12.53}$$

In particular,

$$\begin{aligned}\frac{1}{ab} &= \int_0^1 dx \frac{1}{[ax + b(1-x)]^2} \\ \frac{1}{abc} &= 2\int_0^1 dx\, dy\, dz \frac{\delta(x + y + z - 1)}{(ax + by + cz)^3}\end{aligned} \tag{12.54}$$

We can now rewrite our integral in the form

$$\Lambda^{\mu}(p_1, p_2) = -2ie_0^2 \int_0^1 dx\, dy\, dz\, \delta(x + y + z - 1) \int \frac{d^4q}{(2\pi)^4} \frac{N^{\mu}}{D^3} \tag{12.55}$$

where

$$\begin{aligned}N^{\mu} &= \gamma^{\lambda}(\not{p}_1 - \not{q} + m)\gamma^{\mu}(\not{p}_2 - \not{q} + m)\gamma_{\lambda} \\ D &= xq^2 + y[(p_1 - q)^2 - m^2] + z[(p_2 - q)^2 - m^2] \\ &= q^2 - 2q\cdot(yp_1 + zp_2)\end{aligned} \tag{12.56}$$

Now change the integration variable from q to $q + yp_1 + zp_2$. Generally, doing this in a divergent integral can introduce ambiguous terms that depend on the cutoff scheme. Here it is safe, because the divergence is only logarithmic. We now have

$$\Lambda^\mu(p_1, p_2) = -2ie_0^2 \int_0^1 dx\, dy\, dz\, \delta(x+y+z-1) \int \frac{d^4q}{(2\pi)^4} \frac{M^\mu}{(q^2 - C + i\eta)^3} \tag{12.57}$$

where

$$\begin{aligned} M^\mu &= \gamma^\lambda(\not{A} + m)\gamma^\mu(\not{B} + m)\gamma_\lambda \\ A &= p_1(1-y) - p_2 z - q \\ B &= p_2(1-z) - p_1 y - q \\ C &= (yp_1 + zp_2)^2 = (1-x)^2 m^2 - yzk^2 \end{aligned} \tag{12.58}$$

where $k^2 = (p_1 - p_2)^2$. Note that $k^2 < 0$, so that C is positive-definite. With the help of the identities listed in Table 12.2, we can rewrite

$$M^\mu = -2\not{B}\gamma^\mu\not{A} + 4m(A^\mu + B^\mu) - 2m^2\gamma^\mu \tag{12.59}$$

Further simplification can be made by noting that terms linear in q^μ may be dropped because they integrate to zero, and that y and z may be interchanged because the rest of the integrand is symmetric in y and z. Since we are on mass shell, we may also use the Gordon decomposition. After some algebraic manipulations, we obtain

$$\begin{aligned} F_1(k^2) &= 1 + 4ie_0^2 \int_0^1 dx\, dy\, dz\, \delta(x+y+z-1) \int \frac{d^4q}{(2\pi)^4} \frac{q^2 + (x-yz)k^2 + m^2(1-4x+x^2)}{(q^2 - C + i\eta)^3} \\ F_2(k^2) &= 16ie_0^2 m^2 \int_0^1 dx\, dy\, dz\, \delta(x+y+z-1)xy \int \frac{d^4q}{(2\pi)^4} \frac{1}{(q^2 - C + i\eta)^3} \end{aligned} \tag{12.60}$$

The form factor F_1 is ultraviolet and infrared-divergent, but F_2 is finite.

To calculate F_2, we need the integral

$$I = \int d^4q \frac{1}{(q^2 - C + i\eta)^3} \tag{12.61}$$

TABLE 12.2 Useful Identities

$\gamma_\lambda \gamma^\alpha \gamma^\lambda = -2\gamma^\alpha$
$\gamma_\lambda \gamma^\alpha \gamma^\beta \gamma^\lambda = 4g^{\alpha\beta}$
$\gamma_\lambda \gamma^{\alpha_1} \gamma^{\alpha_2} \gamma^{\alpha_3} \gamma^\lambda = -2\gamma^{\alpha_1} \gamma^{\alpha_2} \gamma^{\alpha_3}$

We perform a Wick rotation $q_0 \to iq_4$, and use the fact that the surface area of a 4-sphere of radius R is $2\pi^2$ to obtain

$$I = -2\pi^2 i \int_0^\infty dR \frac{R^3}{(R^2 + C)^3} = -\frac{i\pi}{2C} \tag{12.62}$$

Substituting this into (12.60), we obtain

$$F_2(k^2) = \frac{e_0^2 m^2}{2\pi^2} \int_0^1 dx\, dy\, dz\, \delta(x + y + z - 1) \frac{xy}{(1 - x)^2 - yz(k^2/m^2)} \tag{12.63}$$

which gives

$$F_2(0) = \frac{e_0^2}{8\pi^2} \tag{12.64}$$

In the formula (12.51) for the g factor, we may replace $F_2(0)/F_1(0)$ by $F_2(0)$, and e_0 by e, since $F_1(0) = 1 + O(e_0^2)$. Therefore

$$g = 2\left(1 + \frac{\alpha}{2\pi}\right) \tag{12.65}$$

where $\alpha = e^2/4\pi$ is the fine-structure constant. The second term comes from the anomalous magnetic moment, and is known as the *Schwinger correction*.

The anomalous magnetic moment of the electron has been calculated to sixth order in quantum electrodynamics, beyond which the weak interactions should be taken into account. The result is as follows:

$$\frac{1}{2}g = 1 + \frac{\alpha}{2\pi} - 0.32848\left(\frac{\alpha}{\pi}\right)^2 + (1.195 \pm 0.026)\left(\frac{\alpha}{\pi}\right)^3 \tag{12.66}$$

The first term is the prediction of the Dirac equation, dated from 1928, and the second the Schwinger correction [4] arising from one Feynman graph. The third term is the result of summing 18 Feynman graphs [5], while the fourth involves 72 Feynman graphs [6]. The comparison with experiments [7] is as follows:

$$\tfrac{1}{2}g_{\text{theory}} = 1 + (1159651.7 \pm 2.2) \times 10^{-9}$$

$$\tfrac{1}{2}g_{\text{expt}} = 1 + (1159656.7 \pm 3.5) \times 10^{-9} \tag{12.67}$$

The theoretical value was computed using the fine-structure constant [8]

$$\frac{1}{\alpha} = 137.03608(26) \tag{12.68}$$

which was obtained experimentally via the Josephson effect.

The anomalous magnetic moment of the electron arises from the electromagnetic interactions. Nucleons, on the other hand, have strong interactions, which give dominant contributions to their anomalous magnetic moments. The values are relatively large, as can be seen from the following experimental values of the total magnetic moments:

$$\mu = \frac{|e|}{M}\begin{cases} 2.79 & \text{(proton)} \\ -1.91 & \text{(neutron)} \end{cases} \tag{12.69}$$

where M is the nucleon mass.

12.4 CHARGE DISTRIBUTION

The charge form factor can be written in the form

$$F_1(k^2) = 1 + \frac{e_0^2}{2\pi^2}\int_\epsilon^1 dx\, dy\, dz\, \delta(x+y+z-1) \times \int_0^{\Lambda/m} dR \frac{-R^2 + (1-4x+x^2) + (x-yz)(k^2/m^2)}{[R^2 + (1-x)^2 - yz(k^2/m^2)]^3} \tag{12.70}$$

where we have performed a Wick rotation and changed the integration variable to $R = \sqrt{-q^2/m^2}$. As we shall see, the high-momentum cutoff Λ can be absorbed through charge renormalization.

The lower cutoff ϵ is introduced to avoid the "infrared catastrophe" occurring at $k^2 = 0$ and $x = 1$. The divergence occurs because the photon is massless, and an infinite number of channels for multiphoton emission simultaneously open up at the same threshold. Our lower cutoff essentially supplies the photon with a finite mass and, of course, violates gauge invariance. This divergence is real, and must be canceled by the addition of Feynman graphs with soft-photon emission from the external lines. Emission of photons of arbitrarily long wavelength cannot be detected by any conceivable measuring device, and must therefore be included as part of the physical process. The Feynman graphs for the soft-photon processes are themselves infrared-divergent. It can shown, however, that when all relevant graphs are summed up, the infrared divergences cancel (See Bjorken and Drell [9]).

We shall deal only with the ultraviolet divergence here. The form factor at zero momentum transfer is given by

$$F_1(0) = 1 + \frac{e_0^2}{2\pi^2}\int_0^1 dx\, dy\, dz\, \delta(x+y+z-1)\int_\epsilon^{\Lambda/m} dR \frac{-R^2 + (1-4x+x^2)}{[R^2 + (1-x)^2]^3} \tag{12.71}$$

which contains the ultraviolet divergence. To isolate it, we define the finite quantity

$$f(k^2) \equiv 1 + \frac{F_1(k^2) - F_1(0)}{F_1(0)} \tag{12.72}$$

and write

$$F_1(k^2) = F_1(0) f(k^2) \tag{12.73}$$

The divergent factor $F_1(0)$ can now be absorbed through charge renormalization:

$$e_0 F_1(k^2) = e f(k^2) \tag{12.74}$$

The charge distribution in the physical electron is given by the Fourier transform of $f(k^2)$. To second order in e_0, we can take $f(k^2) = 1 + F_1(k^2) - F_1(0)$, and replace e_0 by e, since the errors incurred in so doing are of a higher order.

The charge structure of the electron has not yet been seen in experiments because it has such a small radius. The proton, on the other hand, has stronger interactions, and a bigger charge radius, which has been detected experimentally via electron–proton scattering (see Hofstadter [10]). In this process, the electron produces the virtual photon that probes the charge structure of the proton. Figure 12.4 shows

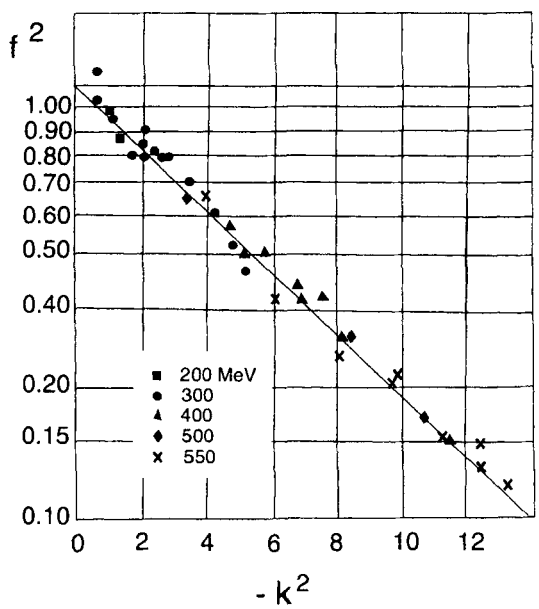


Figure 12.4 Charge form factor of the proton obtained from experiments on electron–proton scattering. [After E. E. Chambers and R. Hofstadter, *Phys. Rev.* **103,** 1454 (1956).]

the results for $f^2(k^2)$ at various incident laboratory energies of the electron [11]. The straight-line fit in the semilogarithmic plot corresponds to a Gaussian distribution

$$f(k^2) = e^{-r_0^2 q/4} \tag{12.75}$$

where $q^2 \equiv -k^2$. This gives the charge radius of the proton as

$$r_0 = 0.70 \times 10^{-13} \text{ cm} \tag{12.76}$$

PROBLEMS

12.1 **(a)** Set up the differential cross section for electron–electron scattering, using the matrix element (11.51). Regard all particles involved as unpolarized. Work out the kinematics, and specify the independent variables. Express the result as spin traces.

(b) Do the same for electron–positron annihilation. Obtain the matrix element from (11.59) for Compton scattering through crossing symmetry.

12.2 Consider Mott scattering, the scattering of an electron by an external Coulomb field. This is the relativistic version of Rutherford scattering. Take $A_k^{\text{ext}} = 0$, and

$$A_0^{\text{ext}} = -\frac{Ze}{4\pi r} \qquad \tilde{A}_0^{\text{ext}}(\mathbf{k}) = \frac{Ze}{k^2}$$

The problem is to calculate the differential cross section

$$\frac{d\sigma}{d\Omega} = \int \frac{d^3p_2}{(2\pi)^3} 2\pi\delta(E_2 - E_1) \frac{1}{2} \sum_{\text{spin}} |\bar{u}_2 \gamma^0 u_1|^2$$

Choose kinematics such that $E_2 = E_1 \equiv E$, $|\mathbf{p}_2| = |\mathbf{p}_1| \equiv p$. The incident velocity is $v = p/E$.

(a) Evaluate the spin sum:

$$\begin{aligned} \frac{1}{2} \sum_{\text{spin}} |\bar{u}_2 \gamma^0 u_1|^2 &= \frac{1}{8m^2} \text{Tr}[\gamma^0(\not{p}_2 + m)\gamma^0(\not{p}_1 + m)] \\ &= \frac{1}{2m^2}(E^2 - p^2 \cos\theta + m^2) \\ &= \frac{E^2}{m^2}\left(1 - v^2 \sin^2 \frac{\theta}{2}\right) \end{aligned}$$

where θ is the scattering angle in the laboratory frame.

(b) Obtain the Mott cross section [12]:

$$\frac{d\sigma}{d\Omega} = \frac{Z^2\alpha^2}{4p^2v^2 \sin^4(\theta/2)}\left(1 - v^2 \sin^2 \frac{\theta}{2}\right)$$

12.3 The proton and neutron have strong interactions, whose contribution to the anomalous magnetic moments dominates over that produced by electromagnetic interactions. A phenomenological way to take this into account is to introduce a "Pauli term" in the Hamiltonian density (see also Problem 6.3)

$$\mathcal{H}_{\text{Pauli}}(x) = \frac{\kappa\mu_0}{2}\sigma^{\mu\nu}F_{\mu\nu}(x)$$

where $\kappa_p = 1.79$, $\kappa_N = -1.91$, and $\mu_0 = |e|/2M$ is the nuclear magneton. The neutron, although electrically neutral, interacts with the electromagnetic field through this term.

(a) State the Feynman rules for the Pauli term.

(b) Obtain the matrix element for electron–neutron scattering.

(c) Obtain the matrix element for the creation of a neutron–antineutron pair by two photons.

REFERENCES

1. O. Klein and Y. Nishina, *Z. Phys.* **52,** 853 (1929).
2. W. Friedrich and G. Golhaber, *Z. Phys.* **44,** 700 (1927).
3. R. P. Feynman, *Phys. Rev.* **76,** 679 (1949).
4. J. Schwinger, *Phys. Rev.* **75,** 651 (1949).
5. C. M. Summerfield, *Ann. Phys.* (NY) **5,** 26 (1958).
6. P. Cvitanovic and T. Kinoshita, *Phys. Rev.* **D10,** 4007 (1974).
7. J. C. Wesley and A. Rich, *Phys. Rev.* **A4,** 1341 (1971).
8. B. N. Taylor, W. H. Parker, and D. N. Langenberg, *Rev. Mod. Phys.* **41,** 375 (1969).
9. J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields*, McGraw-Hill, New York, 1965, Section 17.10; T. D. Lee and M. Nauenberg, *Phys. Rev.* **B133,** 1549 (1964).
10. R. Hofstadter, *Nuclear and Nucleon Structure*, Benjamin, New York, 1963.
11. E. E. Chambers and R. Hofstadter, *Phys. Rev.* **103,** 1454 (1956).
12. N. F. Mott, *Proc. Roy. Soc.* (London) **A124,** 425 (1929).

CHAPTER THIRTEEN

Perturbative Renormalization

13.1 PRIMITIVE DIVERGENCES IN QED

Feynman graphs are generally ultraviolet-divergent, and a high-energy cutoff Λ is needed. The cutoff sets the scale at which the coupling parameters in the Lagrangian are defined. The process of renormalization relates the parameters at one scale to those at another. When this is achieved, we can express the "bare" parameters defined at the cutoff scale to the "renormalized" ones at a lower energy scale. In this chapter we show how this is done in perturbation theory in QED.

A Feynman graph in QED may be represented schematically in the form

$$\text{graph} \sim \int \frac{d^{4N}k}{(k^2)^{P_i} k^{E_i}} \tag{13.1}$$

where P_i is the number of internal photon lines, E_i the number of internal electron lines, and N the number of independent internal 4-momenta. This integral, which is generally divergent, is being cut off at a momentum Λ much larger than any physical momenta in the problem. We define a *primitively divergent graph* as a divergent graph that becomes convergent when *any* one of its internal lines is cut, that is, when *any* integration variable is held fixed. Any divergent graph can be reduced to a primitively divergent one by cutting a sufficient number of internal lines. This is obvious because the graph becomes convergent when all internal lines are cut.

The superficial degree of divergence $\mathcal{D}$ of a primitively divergent graph may be obtained by counting powers:

$$\mathcal{D} = 4N - 2P_i - E_i = 4 - P_e - \tfrac{3}{2}E_e \tag{13.2}$$

where the final result follows by eliminating N, P_i, E_i using (11.66)–(11.68). This number is independent of the number of vertices, and decreases with increasing number of external lines. This property is what makes the theory renormalizable.

It should be noted that the actual degree of divergence may be smaller than $\mathcal{D}$.

Note also that the counting procedure cannot be used for a non–primitively divergent graph. For example, the integral $\int dk\, dp\, k^{-1}p^{-2}$ is logarithmically divergent; but power counting would give $\mathcal{D} = 1$, which suggests incorrectly that it is convergent.

Since $\mathcal{D} \geq 0$ for a divergent graph, and $\mathcal{D}$ decreases with the number of external lines, there exist only a finite number of types of primitively divergent graphs, and they can be classified according to P_e and E_e. There are six possible cases, as shown in Table 13.1. Among these, the vacuum graphs can be ignored. The graphs with three external photon lines can be ignored according to Furry's theorem. The graphs with four external photon lines are logarithmically divergent individually, but it turns out that the sum over the 4! possible assignments of external momenta gives a convergent result. Therefore, there are only three types of primitively divergent graphs: electron self-energy (SE), photon SE, and vertex.

Assume that we know how to renormalize the primitive divergences. Then consider a connected nonvacuum graph. We define its *skeleton graph* as the graph obtained after the removal of all SE and vertex insertions. The skeleton may be convergent or divergent. If convergent, we reinsert the renormalized SE and vertex parts. If divergent, it must be primitively divergent. To prove the last statement, assume the contrary. Then, by cutting a sufficient number of internal lines, the graph can be reduced to a number of components (possibly disconnected), one of which is primitively divergent. The latter must be either an SE or vertex graph; but these have been removed by assumption. Therefore the divergent skeleton graph must be primitively divergent. The procedure to renormalize a connected nonvacuum graph therefore reduces to that for the primitive divergences.

13.2 ELECTRON SELF-ENERGY

The full electron propagator is defined as

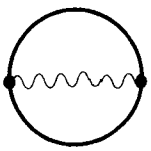
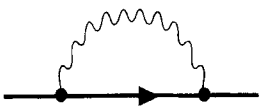
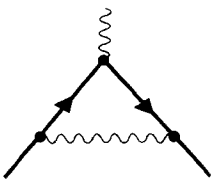
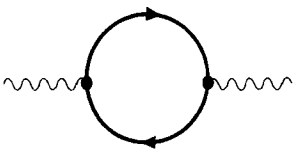
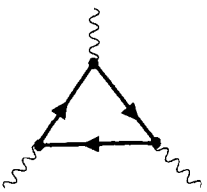
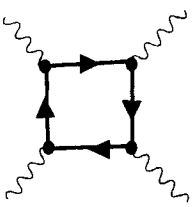
$$\begin{aligned} i[S'(x)]_{\alpha\beta} &= \langle 0|\mathrm{T}\psi^{\mathrm{h}}_{\alpha}(x)\bar{\psi}^{\mathrm{h}}_{\beta}(0)|0\rangle \\ &= \langle 0|\mathrm{T}\psi_{\alpha}(x)\bar{\psi}_{\beta}(0)e^{-i\int dy\mathcal{H}'(y)}|0\rangle \end{aligned} \tag{13.3}$$

where ψ^{h} is a Heisenberg operator and ψ is an interaction-picture operator. Its expansion in terms of Feynman graphs is shown in Fig. 13.1, where $\Sigma(p)$ denotes the proper self-energy part, which is the sum of all one-particle irreducible graphs. Without taking into account the mass counterterm, we have

$$\begin{aligned} iS'(p) &= iS_{\mathrm{F}} + iS_{\mathrm{F}}i\Sigma iS_{\mathrm{F}} + iS_{\mathrm{F}}i\Sigma iS_{\mathrm{F}}i\Sigma iS_{\mathrm{F}} + \cdots \\ &= iS_{\mathrm{F}}[1 + (i\Sigma iS_{\mathrm{F}}) + (i\Sigma iS_{\mathrm{F}})^2 + \cdots] \\ &= iS_{\mathrm{F}}(1 - i\Sigma iS_{\mathrm{F}})^{-1} \end{aligned} \tag{13.4}$$

where we denote the electron propagator in momentum space as S_{F} (without the tilde). Taking the inverse of both sides, we obtain

TABLE 13.1 Primitive Divergences in QED

P_e	E_e	$\mathcal{D}$	*Graph*	*Example*	*Remark*
0	0	4	Vacuum		Can be ignored
0	2	1	Electron SE		Actually logarithmically divergent
1	2	0	Vertex		Logarithmically divergent
2	0	2	Photon SE		Logarithmically divergent by gauge invariance
3	0	1	3-photon		Canceled (Furry's theorem)
4	0	0	Light–light scattering		Sum of 4! graphs convergent

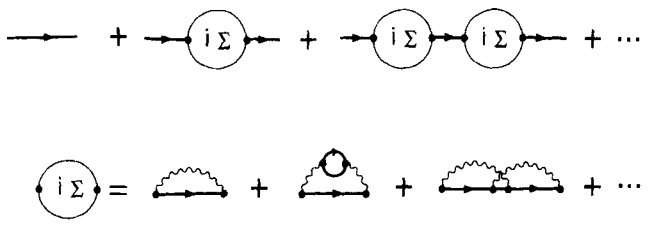


Figure 13.1 Graphs for full electron propagator and the proper self-energy part Σ.

$$(iS')^{-1} = (1 - i\Sigma iS_F)(iS_F)^{-1} = [(iS_F)^{-1} - i\Sigma]$$
$$= -i[\not{p} - m + \Sigma(p)] \tag{13.5}$$

where we have left the $i\eta$ term in the denominator understood.

To include the mass counterterm, we make the replacement $iS_F \rightarrow iS_F + iS_F i\delta m i S_F$ in all the Feynman graphs. This means that $\Sigma(p)$ is replaced by $\Sigma(p) + \delta m$. Consequently we have $(S')^{-1} = \not{p} - m + + \delta m + \Sigma(p)$, or

$$S'(p) = \frac{1}{\not{p} - m + \delta m + \Sigma(p)} \tag{13.6}$$

By Lorentz invariance, we can write $\Sigma(p)$ in the form

$$\Sigma(p) = A(p^2) + (\not{p} - m)B(p^2) \tag{13.7}$$

Thus

$$S'(p) = \frac{1}{[1 + B(p^2)](\not{p} - m) + A(p^2) + \delta m} \tag{13.8}$$

Mass renormalization consists of the assertion that the pole of the full propagator is located at m. Therefore

$$A(m^2) + \delta m = 0 \tag{13.9}$$

Since $A(m^2)$ is divergent, this shows that δm must also be divergent

Another divergent quantity is $B(m^2)$, and it is customary to put

$$1 + B(m^2) \equiv \frac{1}{Z_2} \tag{13.10}$$

so that $Z_2 \rightarrow 0$. However, we regard it as an infinite series in e_0^2, with divergent coefficients. At the mass pole we have

$$S'(p)\underset{p^2\to m^2}{\longrightarrow}\frac{Z_2}{\not p - m} = \frac{Z_2}{p^2 - m^2}\sum_s u(\mathbf{p}, s)\bar{u}(\mathbf{p}, s) \tag{13.11}$$

This shows that $\sqrt{Z_2}$ is the renormalization constant for electron wave functions. To verify this explicitly, we examine the self-energy insertions on an external electron line. The Feynman graphs in this case are the same as those in Fig. 13.1, except that the leftmost line is taken to be an external line. The sum of the graphs give

$$u'(\mathbf{p}) = \{1 + iS_F i(\Sigma(p) + \delta m) + [iS_F i(\Sigma(p) + \delta m)]^2 + \cdots\}u(\mathbf{p}) \tag{13.12}$$

The operator in brackets gives

$$[S_F^{-1}(p) + \Sigma(p) + \delta m]^{-1}S_F(p) = S'(p)(\not p - m) \underset{p^2\to m^2}{\longrightarrow} Z_2 \tag{13.13}$$

Therefore

$$u'(\mathbf{p}) = Z_2 u(\mathbf{p}) \tag{13.14}$$

We have shown that the wave function renormalization constant is $\sqrt{Z_2}$. Therefore, a factor $\sqrt{Z_2}$ goes toward the renormalization of the vertex that absorbs this particle.

We have yet to show that the electron propagator is finite after mass and wave function renormalization. To do this, we first analyze the skeleton self-energy graph. Later, we shall discuss how to make insertions. The skeleton is just the second-order proper self-energy graph:

$$\begin{aligned} i\Sigma(p) &= (-ie_0)^2\int \frac{d^4k}{(2\pi)^4}\gamma^\mu iS_F(p-k)\gamma_\mu iD_F(k) \\ &= -ie_0^2\int \frac{d^4k}{(2\pi)^4}\frac{\gamma^\mu(4m - 2(\not p - \not k + m))\gamma_\mu}{(k^2 + i\eta)[(p-k)^2 - m^2 + i\eta]} \\ &= -ie_0^2\int \frac{d^4k}{(2\pi)^4}\frac{-2(\not p - \not k) + 4m}{(k^2 + i\eta)[(p-k)^2 - m^2 + i\eta]} \end{aligned} \tag{13.15}$$

This integral is logarithmically divergent for large k, and must be cut off. Using the Feynman formula (12.53) to combine the factors in the denominator, we can rewrite

$$i\Sigma(p) = e_0^2\int \frac{d^4k}{(2\pi)^4}\int_0^1 dx \frac{2(\not p - \not k) - 4m}{\{(1-x)k^2 + x[(p-k)^2 - m^2] + i\eta\}^2} \tag{13.16}$$

The denominator has the form

$$(k - xp)^2 + x[(1-x)p^2 + m^2]$$

Changing the variable of integration to $q = k - xp$, and omitting in the numerator a term $\not{q}$, which integrates to 0, we have

$$i\Sigma(p) = \frac{2e_0^2}{(2\pi)^4}\int_0^1\int d^4q\,\frac{(1-x)\not{p} - 2m}{(q^2 - C + i\eta)^2}$$

$$C = x[m^2 - (1-x)p^2] \tag{13.17}$$

We now have to evaluate an integral of the type

$$I = \int d^4q\,\frac{1}{[q^2 - C + i\eta]^2} = \int d^3q\int_{-\infty}^{\infty} dq_0\,\frac{1}{[q_0 - (\mathbf{q}^2 + C) + i\eta]^2} \tag{13.18}$$

For simplicity, we cut off the spatial integral at $|\mathbf{q}| = \Lambda$. The q_0 integration can be performed through contour integration, which yields

$$\int_{-\infty}^{\infty}\frac{dq_0}{[q_0 - (\mathbf{q}^2 + C) + i\eta]^2} = \frac{i\pi}{2}\,\frac{1}{(|\mathbf{q}|^2 - C)^{3/2}} \tag{13.19}$$

Thus we have

$$I = 2i\pi^2\int_0^{\Lambda} dq\,\frac{q^2}{(q^2 + C)^{3/2}} \xrightarrow[\Lambda\to\infty]{} 2i\pi^2 \ln\frac{\Lambda}{m} + (\text{finite terms}) \tag{13.20}$$

Using this result, we obtain

$$\Sigma(p) = \frac{ie_0^2}{4\pi^2}\int_0^1 dx[(1-x)\not{p} - 2m]\ln\frac{\Lambda}{m} + (\text{finite terms}) \tag{13.21}$$

which gives

$$A(p^2) = -\frac{3e_0^2 m}{8\pi^2}\ln\frac{\Lambda}{m} + A_1(p^2)$$

$$B(p^2) = \frac{e_0^2}{8\pi^2}\ln\frac{\Lambda}{m} + B_1(p^2) \tag{13.22}$$

where A_1 and B_1 are finite, and vanish on mass shell. Thus

$$\frac{\delta m}{m} = \frac{3e_0^2}{8\pi^2}\ln\frac{\Lambda}{m} + O(e_0^4) \tag{13.23}$$

$$Z_2 = 1 - \frac{e_0^2}{8\pi^2}\ln\frac{\Lambda}{m} + O(e_0^4) \tag{13.24}$$

The result of mass and wave function renormalization to second order can be summarized by the statement

$$S'(p) = \frac{Z_2}{(\not{p} - m) + \Sigma_c(p)} \tag{13.25}$$

where

$$\Sigma_c(p) = \Sigma(p) - \Sigma(p_0) \tag{13.26}$$

where p_0 is a 4-momentum on mass shell. It remains to be shown that this form is correct to all orders.

Mass renormalization first made its appearance in Lorentz' calculation of the self-force of the classical electron. Lorentz modeled the electron as a spherically symmetric uniform charge distribution of radius a. The self-force $\mathbf{F}_s$ is the sum of all the forces between charge elements, with retardation taken into account. In the limit $a \to 0$, the result is [1]

$$\mathbf{F}_s = -\frac{4}{3c}\dot{\mathbf{v}} \int \frac{de\, de'}{4\pi r} + \frac{2e^2}{3c^3}\ddot{\mathbf{v}} + O(a) \tag{13.27}$$

The first term, which arises from the Coulomb self-energy, is divergent. The second term is independent of a, and gives the famous radiation damping. The other terms vanish when $a \to 0$. The equation of motion for an electron of "mechanical mass" m_0 is

$$m_0\dot{\mathbf{v}} = \mathbf{F}_{\text{ext}} + \mathbf{F}_s \tag{13.28}$$

where $\mathbf{F}_{\text{ext}}$ is the external force. When $a \to 0$, this has the form

$$m\dot{\mathbf{v}} = \mathbf{F}_{\text{ext}} + \frac{2e^2}{3c^3}\ddot{\mathbf{v}} \tag{13.29}$$

where

$$m = m_0 + \frac{4}{3c}\int \frac{de\, de'}{4\pi r} \tag{13.30}$$

is the renormalized mass, to be taken from experiments. The self-energy diverges like a^{-1} classically, but only logarithmically in QED. The difference can be attributed to the presence of the Dirac sea of negative-energy electrons [2].

13.3 VACUUM POLARIZATION

We can take the free photon propagator to be

$$D^{\mu\nu}(k) = -\frac{g^{\mu\nu}}{k^2} \tag{13.31}$$

where, for simplicity, we have dropped the subscript "F" and the $i\eta$ term in the denominator. The full photon propagator is given by the series of graphs shown in Fig. 13.2, which can be reduced to one-particle irreducible components. The sum of all one-particle irreducible graphs, with external photon lines omitted, is called the *vacuum polarization tensor* $\Pi^{\mu\nu}(k)$. It describes virtual electron–positron pairs produced by a photon propagating in the vacuum. In terms of this tensor, we can write the full photon propagator in the form

$$iD'^{\mu\nu}(k) = iD^{\mu\nu}(k) + iD^{\mu\alpha}(k)i\Pi_{\alpha\beta}(k)iD^{\beta\nu}(k) + \cdots \tag{13.32}$$

Because of current conservation, or gauge invariance, we should have

$$k^{\mu}\Pi_{\mu\nu}(k) = \Pi_{\nu\mu}k^{\mu} = 0 \tag{13.33}$$

We can therefore put

$$\Pi^{\mu\nu}(k) = (g^{\mu\nu}k^2 - k^{\mu}k^{\nu})e_0^2\Pi(k^2) \tag{13.34}$$

This form guarantees that the photon has zero mass, unless $\Pi(k^2)$ develops a pole at $k^2 = 0$, which does not happen in perturbation theory. Note that $\Pi(k^2)$ is defined with e_0^2 factored out.

To avoid a profusion of indices, we shall regard $D'^{\mu\nu}$ as a 4×4 matrix $\mathcal{D}'$, with matrix multiplication defined such that an upper index can be contracted only with a lower index. To be able to invert these matrices, we define the transverse and longitudinal projection operators $\mathcal{P}_T(k)$ and $\mathcal{P}_L(k)$ as follows:

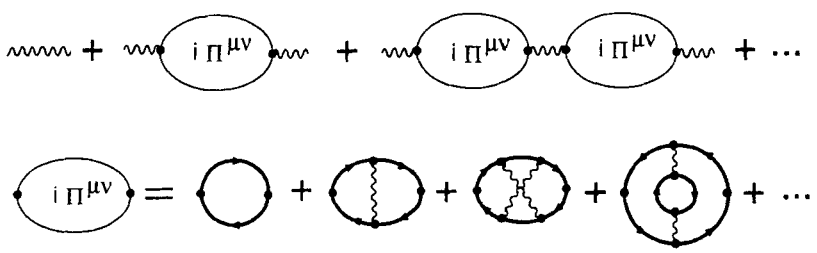


Figure 13.2 Full photon propagator and the vacuum polarization tensor.

$$\mathcal{P}_{\mathrm{T}}^{\mu\nu}(k) = g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}$$

$$\mathcal{P}_{\mathrm{L}}^{\mu\nu}(k) = \frac{k^\mu k^\nu}{k^2} \tag{13.35}$$

In matrix notation, they have the properties

$$\mathcal{P}_{\mathrm{T}}^2 = \mathcal{P}_{\mathrm{T}}$$

$$\mathcal{P}_{\mathrm{L}}^2 = \mathcal{P}_{\mathrm{L}}$$

$$\mathcal{P}_{\mathrm{T}} + \mathcal{P}_{\mathrm{L}} = 1 \tag{13.36}$$

The free propagator $\mathcal{D}$ is proportional to the unit matrix, while the vacuum polarization tensor takes the form

$$\mathcal{T}(k^2) = k^2 e_0^2 \Pi(k^2) \mathcal{P}_{\mathrm{T}} \tag{13.37}$$

In this notation, we have

$$\begin{aligned} \mathcal{D}'(k) &= \mathcal{D}(k) + i\mathcal{D}(k) i\mathcal{T}(k^2) i\mathcal{D}(k) + \cdots \\ &= \mathcal{D}[1 - i\mathcal{T}(k^2) i\mathcal{D}(k)]^{-1} \\ &= -\frac{1}{k^2}[1 - e_0^2 \Pi(k^2) \mathcal{P}_{\mathrm{T}}]^{-1} \end{aligned} \tag{13.38}$$

Now note that

$$\begin{aligned} (1 - \lambda\, \mathcal{P}_{\mathrm{T}})^{-1} &= 1 + \lambda\, \mathcal{P}_{\mathrm{T}} + \lambda^2 \mathcal{P}_{\mathrm{T}}^2 + \cdots \\ &= 1 + (\lambda + \lambda^2 + \cdot) \mathcal{P}_{\mathrm{T}} = \frac{1}{1 - \lambda} \mathcal{P}_{\mathrm{T}} + \mathcal{P}_{\mathrm{L}} \end{aligned} \tag{13.39}$$

Thus

$$\mathcal{D}'(k) = -\frac{\mathcal{P}_{\mathrm{T}}}{k^2[1 - e_0^2 \Pi(k^2)]} - \frac{\mathcal{P}_{\mathrm{L}}}{k^2} \tag{13.40}$$

or

$$D'^{\mu\nu}(k) = -\frac{1}{k^2[1 - e_0^2 \Pi(k^2)]} \left(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2} \right) - \frac{k^\mu k^\nu}{k^4} \tag{13.41}$$

The terms proportional to k^μ can be dropped, because they do not contribute in Feynman graphs owing to current conservation. Hence we take

$$D'^{\mu\nu}(k) = g^{\mu\nu}D'(k) \tag{13.42}$$

where

$$D'(k) = \frac{-1}{k^2[1 - e_0^2\Pi(k^2)]} \tag{13.43}$$

The divergences are now isolated in the function $\Pi(k^2)$, whose skeleton can be obtained from that of the vacuum polarization tensor, given by its second-order Feynman graph in Fig. 13.2:

$$\Pi^{\mu\nu}(k) = ie_0^2 \int \frac{d^4p}{(2\pi)^4} \mathrm{Tr}\left(\gamma^\mu \frac{1}{\not p - m} \gamma^\nu \frac{1}{\not p - \not k - m}\right) \tag{13.44}$$

which is quadratically divergent. By (13.34), the degree of divergence of $\Pi(k^2)$ should be reduced by 2, and we expect the latter to be logarithmically divergent. But (13.44) violates gauge invariance, since $k_\mu \Pi^{\mu\nu}(k) \neq 0$.

The violation of gauge invariance originates in the singular nature of the current, in which the electron field and its canonical conjugate are coupled at the same point, in a product of the form $\overline{\psi}_\alpha(x)\psi_\beta(x)$. By reversing the order of $\overline{\psi}$ and ψ, one would generate a meaningless term involving $\delta^3(0)$. To avoid the singularity, we could replace the factor by $\overline{\psi}_\alpha(x + \epsilon)\psi_\beta(x - \epsilon)$, and take the limit $\epsilon \to 0$ eventually; but to make the product gauge-invariant, we have to amend it further and take the current to be

$$j^\mu(x) = e_0 \overline{\psi}(x + \epsilon)\gamma^\mu \psi(x - \epsilon) e^{ie_0 \int_{x-\epsilon}^{x+\epsilon} dy_\nu A^\nu(y)} \tag{13.45}$$

with $\epsilon \to 0$. This procedure is called the "point-splitting method, " and will yield a gauge-invariant vacuum polarization tensor [3].

We use the following shortcut, which yields the correct answer. By (13.34), we can calculate $\Pi(k)$ through the relation

$$e_0^2\Pi(k^2) = \frac{1}{3k^2}\Pi_\mu^\mu(k) \tag{13.46}$$

To ensure $\Pi(0) = 0$, we replace this by

$$e_0^2\Pi(k^2) = \frac{1}{3k^2}[\Pi_\mu^\mu(k) - \Pi_\mu^\mu(0)] \tag{13.47}$$

The subtraction reduces the quadratic divergence to a logarithmic one. Explicitly, we have

$$\Pi(k^2) = -\frac{1}{12\pi^2} \ln \frac{\Lambda}{m} + R(k^2) + O(e_0^2) \tag{13.48}$$

where

$$R(k^2) = \frac{1}{2\pi^2}\int_0^1 dx\, x(1-x)\ln\left[1 - x(1-x)\frac{k^2}{m^2}\right] \tag{13.49}$$

We now define the renormalized charge. The convergent part of $\Pi(k^2)$ is defined by making one subtraction:

$$\Pi_c(k^2) \equiv \Pi(k^2) - \Pi(\mu^2) \tag{13.50}$$

where μ^2 is an arbitrary scale parameter. To second order, we have

$$D'(k) = -\frac{d(k^2/\mu^2)}{k^2} \tag{13.51}$$

where

$$d\left(\frac{k^2}{\mu^2}\right) = \frac{1}{[1 - e_0^2\Pi(\mu^2)] - e_0^2\Pi_c(k^2)} = \frac{Z(\mu^2)}{1 - e_0^2 Z(\mu^2)\Pi_c(k^2)} \tag{13.52}$$

where

$$Z(\mu^2) \equiv \frac{1}{1 - e_0^2\Pi(\mu^2)} \tag{13.53}$$

with normalization

$$d(1) = 1 \tag{13.54}$$

Defining a scale-dependent charge $e(\mu^2)$ by

$$e^2(\mu^2) \equiv e_0^2 Z(\mu^2) \tag{13.55}$$

we can write

$$e_0^2 d\left(\frac{k^2}{\mu^2}\right) = \frac{e^2(\mu^2)}{1 - e^2(\mu^2)\Pi_c(k^2)} \tag{13.56}$$

The conventional electronic charge is defined at the value $\mu^2 = 0$, which corresponds to a static interaction with zero 4-momentum transfer. The fine-structure constant is therefore related to $e^2(0) = e_0{}^2 Z(0)$. In the conventional notation

$$Z_3 \equiv Z(0) \tag{13.57}$$

we have

$$\alpha \equiv \frac{e_0^2 Z_3}{4\pi} \approx \frac{1}{137} \tag{13.58}$$

Defining

$$e^2(0)d_c(k^2) = \lim_{\mu^2 \to 0} e^2(\mu^2) d \frac{k^2}{\mu^2}$$

we can write the photon propagator in the form

$$\frac{e_0^2}{4\pi} D'(k) = -\frac{\alpha d_c(k^2)}{k^2} \tag{13.59}$$

where the right side involves only finite observable quantities.

13.4 RUNNING COUPLING CONSTANT

Using the momentum transfer as the renormalization scale, we have

$$e_0^2 D'(k) = -\frac{e^2(k^2)}{k^2} \tag{13.60}$$

That is, to the order considered, the full propagator describes a free photon coupled through the scale-dependent charge $e(k^2)$, also called the "running coupling constant" for this reason. We can define a running fine-structure constant

$$\alpha(k^2) = \frac{e^2(k^2)}{4\pi}$$

To relate it to the value at $k^2 = 0$, use the relations

$$\begin{aligned} \frac{1}{e^2(\mu^2)} &= \frac{1}{e_0^2} + \frac{1}{12\pi^2} \ln \frac{\Lambda}{m} - R(k^2) \\ \frac{1}{e^2(0)} &= \frac{1}{e_0^2} + \frac{1}{12\pi^2} \ln \frac{\Lambda}{m} \end{aligned} \tag{13.61}$$

where we have used the fact $R(0) = 0$. These are, of course, correct only to order e_0^2. Subtracting one equation from the other, we obtain the following after some rearrangement:

$$\frac{\alpha(k^2)}{\alpha} = 1 + 4\pi\alpha R(k^2) + O(e_0^4) \tag{13.62}$$

The function $R(k^2)$ is real for $k^2 < 4m^2$, below the pair-production threshold. For k^2/m^2 negative and large

$$\frac{\alpha(k^2)}{\alpha} = 1 + \frac{\alpha}{3\pi} \ln \frac{|k^2|}{m^2} + O(e_0^4) \tag{13.63}$$

13.5 FULL VERTEX

We represent the proper full vertex in the form

$$\Gamma_\mu(p_2, p_1) = \gamma_\mu + \Lambda_\mu(p_2, p_1) \tag{13.64}$$

with graphical expansion as shown in Fig. 13.3. There are an infinite number of skeleton graphs, whose sum is denoted by Λ_μ^*. Current conservation implies

$$(p_2 - p_1)^\mu \bar{u}(\mathbf{p}_2)\Gamma_\mu(p_2, p_1)u(\mathbf{p}_1) = 0 \tag{13.65}$$

The second-order skeleton graph gives

$$\Lambda^\mu(p_2, p_1) = -ie_0^2 \int \frac{d^4k}{(2\pi)^4} \frac{\gamma^\lambda(\not{p}_1 - \not{k} + m)\gamma^\mu(\not{p}_2 - \not{k} + m)\gamma_\lambda}{k^2[(p_1 - k)^2 - m^2][(p_2 - k)^2 - m^2]} \tag{13.66}$$

where the $i\eta$ devices have been left understood. (To this order there is no difference between Λ and Λ^*.) This integral is both ultraviolet- and infrared-divergent, and we

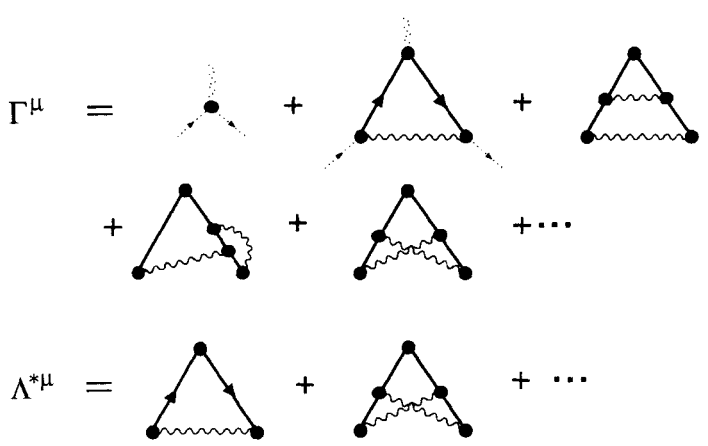


Figure 13.3 The proper full vertex Γ^μ and the skeleton $\Lambda^{*\mu}$. Note that external lines are omitted, by definition.

cut it off and high and low momenta. The ultraviolet divergence is logarithmic, and can be isolated by subtracting the integral at $p_1 = p_2 = p_0$, where p_0 is a 4-momentum on mass shell:

$$\Lambda^\mu(p_0, p_0) = L\gamma^\mu$$
$$L = \frac{e_0^2}{8\pi^2} \ln \frac{\Lambda}{m} \tag{13.67}$$

Thus

$$\Lambda^\mu(p_2, p_1) = L\gamma^\mu + \Lambda_c^\mu(p_2, p_1) \tag{13.68}$$

where Λ_c^μ is ultraviolet convergent. We will not deal with the "infrared catastrophe." (See the remarks in Section 12.4.) The proper full vertex now takes the form

$$\Gamma^\mu(p_2, p_1) = \frac{1}{Z_1}\gamma^\mu + \Lambda_c^\mu(p_2, p_1) + O(e_0^4) \tag{13.69}$$

where

$$Z_1 = \frac{1}{1 + L} \tag{13.70}$$

is the vertex renormalization constant. Like Z_2, it must be considered a power series in e_0^2 with divergent coefficients, even though formally it approaches 0 when $\Lambda \to \infty$. To second order, we can rewrite our results in a suggestive manner:

$$\Gamma^\mu(p_2, p_1) = \frac{1}{Z_1}\Gamma_c^\mu(p_2, p_1) \tag{13.71}$$

where the convergent part is given by

$$\Gamma_c^\mu(p_2, p_1) = \gamma^\mu + \Lambda_c^\mu(p_2, p_1) \tag{13.72}$$

13.6 WARD IDENTITY

Equivalent forms of the *Ward identity* are

$$\Gamma^\mu(p, p) = \frac{\partial}{\partial p_\mu}[S'(p)]^{-1}$$

$$\Lambda^{\mu}(p, p) = \frac{\partial}{\partial p_{\mu}} \Sigma(p) \tag{13.73}$$

To prove it, we can use the identity

$$\frac{1}{a+b} = \frac{1}{a} + \frac{1}{a} b \frac{1}{a} + \frac{1}{a} b \frac{1}{a} b \frac{1}{a} + \cdots \tag{13.74}$$

to show

$$\begin{aligned} \frac{\partial}{\partial p_{\mu}} \frac{1}{\not{p} - m} &= \lim_{\Delta p_{\mu} \to 0} \frac{1}{\Delta p_{\mu}} \left[\frac{1}{\not{p} + \Delta \not{p}_{\mu} - m} - \frac{1}{\not{p} - m} \right] \\ &= \frac{1}{\not{p} - m} \gamma^{\mu} \frac{1}{\not{p} - m} \end{aligned} \tag{13.75}$$

This states that differentiating a free electron propagator is equivalent to the insertion of a vertex that emits a photon of 4-momentum $k = 0$. This interpretation comes from the form of the current $j^{\mu} = \bar{\psi} \gamma^{\mu} \psi$, which embodies current conservation. The proof can then be stated graphically, as shown in Fig. 13.4. The more general *Ward–Takahashi identity* states, in equivalent forms (see Problem 13-1),

$$\begin{aligned} (p_2 - p_1)^{\mu} \Gamma_{\mu}(p_2, p_1) &= [S'(p_2)]^{-1} - [S'(p_1)]^{-1} \\ (p_2 - p_1)^{\mu} \Lambda_{\mu}(p_2, p_1) &= \Sigma(p_2) - \Sigma(p_1) \end{aligned} \tag{13.76}$$

From (13.25) and (13.71), we have

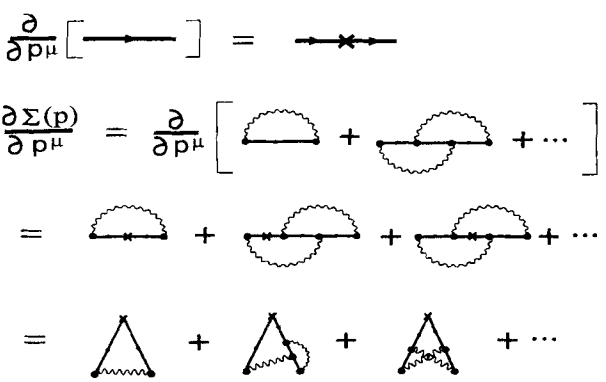


Figure 13.4 Graphical proof of the Ward identity. A cross indicates the insertion of a photon of zero 4-momentum.

$$\frac{\partial}{\partial p_\mu}[S'(p)]^{-1} = \frac{\gamma^\mu}{Z_2}$$

$$\Gamma^\mu(p, p) = \frac{\gamma^\mu}{Z_1} \tag{13.77}$$

Thus, the Ward identity states

$$Z_1 = Z_2 \tag{13.78}$$

This is verified in our second-order calculations.

13.7 RENORMALIZATION TO SECOND ORDER

We can summarize our results so far as follows:

$$S'(p) = \frac{Z_2}{\not p - m + \Sigma_c(p)} \equiv Z_2 S_c(p)$$

$$D'(k) = -\frac{Z_3 d_c(k^2)}{k^2} \equiv Z_3 D_c(k)$$

$$\Gamma^\mu(p_2, p_1) = \frac{1}{Z_1}[\gamma^\mu + \Lambda_c^\mu(p_2, p_1)] \equiv \frac{1}{Z_1}\Gamma_c^\mu(p_2, p_1) \tag{13.79}$$

where the quantities with a subscript "c" are finite. We have proved (13.79) to second order, but, as we shall show, they are true to all orders.

Consider the full vertex Υ^μ, including improper (one-particle reducible) graphs, as shown by the second-order graphical expansion in Fig. 13.5. Omitting momentum arguments for simplicity, we can write it in the form

$$\Upsilon^\mu = -ie_0 S'\Gamma^\mu S' D' \tag{13.80}$$

Rewriting this in terms of finite functions, we have

$$\Upsilon^\mu = -ie_0 \frac{Z_2 Z_2 Z_3}{Z_1} S_c \Gamma_c^\mu S_c D_c = -iZ_2\sqrt{Z_3}\left[\frac{Z_2}{Z_1}\sqrt{Z_3}\, e_0\right] S_c \Gamma_c^\mu S_c D_c \tag{13.81}$$

The factor $Z_2\sqrt{Z_3}$ goes toward the renormalization of other vertices and/or external lines. The factor in brackets is the renormalized charge:

$$e \equiv \frac{Z_2}{Z_1}\sqrt{Z_3} e_0 = \sqrt{Z_3} e_0 \tag{13.82}$$

It is important to note that the renormalized charge e depends only on the pho-

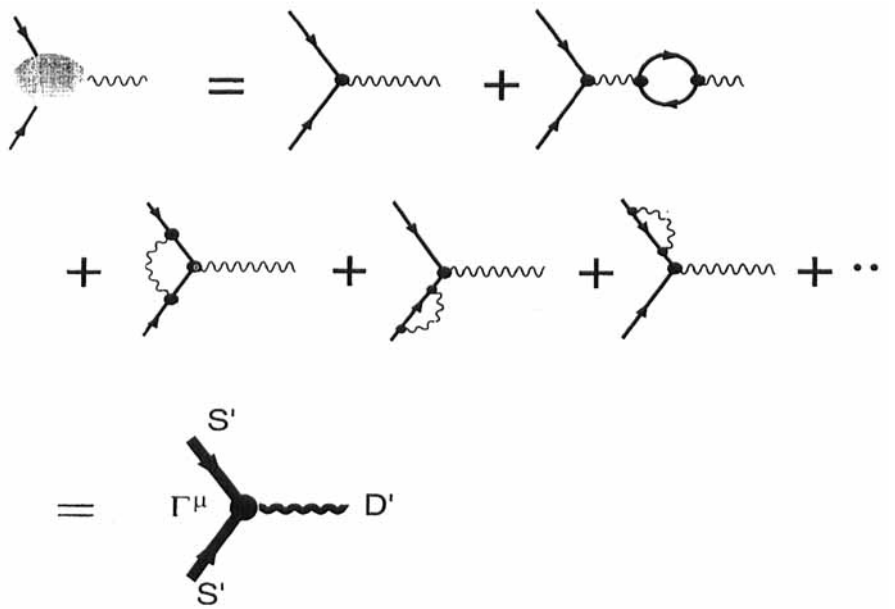


Figure 13.5 The full vertex can be expressed as a product of full proper vertex and full propagators.

ton renormalization constant Z_3, not on the renormalization of the electron propagator or the vertex. This is because $Z_1 = Z_2$, as guaranteed by the Ward identity. Because of this fact, Z_3 is universal. If we bring in other charged fields into the system, such as nucleons or π mesons, then Z_3 will acquire contribution from all fields through vacuum polarization, but it will be the common charge renormalization constant for all fields.

13.8 RENORMALIZATION TO ALL ORDERS

We now describe the renormalization program of Dyson and Ward [4], following a version due to Gell-Mann and Low [5]. The first item of business is to specify how SE and vertex parts are to be reinserted into a skeleton graph. Consider first the vertex part. We regard the skeleton Λ_μ^* as a functional of the free electron propagator S, the free photon propagator D, and the free vertex γ:

$$\text{Sum of skeleton vertex graphs} \equiv \Lambda_\mu^*[S, D, \gamma; e_0, p, p'] \tag{13.83}$$

It also depends on the bare charge e_0 and external momenta p, p' as parameters. With this notation, we can write

$$\Lambda_\mu(p, p') = \Lambda_\mu^*[S', D', \Gamma; e_0, p, p'] \tag{13.84}$$

That is, Λ_μ is obtainable from Λ_μ^* by inserting the full propagators S', D' in place of the free propagators S,D and full vertex Γ in place of the free vertex γ. The skeleton $\Lambda_\mu^*[S, D, \gamma; e_0, p, p']$ is only logarithmically divergent, but Λ_μ is much more divergent, due to the insertions.

For the electron self-energy, there is an ambiguity known as an "overlapping divergence," as illustrated in Fig. 13.6. It is not clear whether A should be regarded as an insertion into B, or vice versa. We avoid this problem by using the Ward–Takahashi identity to expresses S' in terms of the vertex:

$$[S'(p)]^{-1} = [S'(p_0)]^{-1} + (p - p_0)^\mu \Gamma_\mu(p, p_0) \tag{13.85}$$

The right-hand side is actually independent of p_0, but for definiteness we take p_0 to be a 4-momentum on mass shell. Mass renormalization consists of the statement

$$[S'(p_0)]^{-1} = \frac{1}{Z_2}(\not{p}_0 - m) \tag{13.86}$$

The photon self-energy also contains overlapping divergences. To circumvent them, we mimic the Ward identity by defining an auxiliary function $W\mu(k)$ by differentiating the inverse photon propagator:

$$W_\mu(k) \equiv \frac{\partial}{\partial k^\mu}[iD'(k)]^{-1} \tag{13.87}$$

Using (13.34), we can write

$$W_\mu(k) = 2ik_\mu - ik_\mu T(k) \tag{13.88}$$

where

$$T(k) = \frac{k^\mu}{k^2}\frac{\partial}{\partial k^\mu}[k^2 e_0^2 \Pi(k)] \tag{13.89}$$

We then define the skeleton of T, denoted by T^*, by removing all SE, vertex, as well as W_μ insertions. The skeleton is logarithmically divergent, and we can again isolate the divergence through one subtraction. To recover D' from W_μ, use the formula

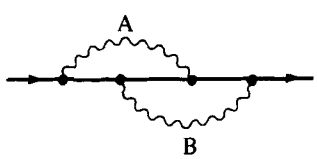


Figure 13.6 An overlapping divergence.

$$[iD'(k)]^{-1} = \int_0^1 dx\, k^\mu W_\mu(xk) \tag{13.90}$$

An ambiguity occurs in the definition of the skeleton of T, because the external momentum k can be routed through a graph in more than one way. The difficulty occurs in graphs of W_μ containing at least three closed electron loops, and are therefore at least of order e_0^{14}, but can be overcome by the proper convention for momentum routing [6].

Divergences in perturbation theory can be absorbed into renormalized coupling constants because the effective coupling constant at one energy scale can be related to that at any other energy scale. This property depends on the scaling properties the insertions, which lie at the heart of renormalizability.

The graphs in Λ^* are all of even order, and a graph of order $2n$ contains factors of e_0, S, D, and γ to various powers, as indicated schematically below:

$$\Lambda^*_{2n} \sim e_0^{2n} S^{2n} D^n \gamma^{2n+1} \tag{13.91}$$

Under the scale transformation

$$\begin{aligned} \gamma &\to a\gamma \\ D &\to bD \\ S &\to a^{-1}S \end{aligned} \tag{13.92}$$

where a and b are arbitrary number, we have

$$\Lambda^*_{2n} \to ab^n \Lambda^*_{2n} \tag{13.93}$$

Therefore, we have the scaling law

$$a\Lambda^*_\mu[S, D, \gamma; e_0, p, p'] = \Lambda^*_\mu[a^{-1}S, bD, a\gamma; b^{-1}e_0, p, p'] \tag{13.94}$$

For the functional T^*, a graph of order $2n$ has the structure

$$T^*_{2n} \sim e_0^{2n} S^{2n+\sigma} \gamma^{2n+\sigma} (2ik_\mu)^{1-\sigma} \tag{13.95}$$

where σ is an integer that receives an additive contribution 1 from each differentiation of an electron line, and 0 from that of a photon line. Under the scale transformation (13.92) supplemented by $2ik_\mu \to 2ik_\mu/b$, we obtain

$$T^*_{2n} \to b^{n-1} T_{2n}{}^* \tag{13.96}$$

Therefore a second scaling law is

$$b^{-1}T^*[S, D, \gamma, 2ik_\mu; e_0, k] = T^*[a^{-1}S, bD, a\gamma, b^{-1}2ik_\mu; b^{-1}e_0, k] \tag{13.97}$$

The functions S', D', and Γ satisfy the following coupled functional equations:

$$\begin{aligned}
\Gamma_\nu(p, p') &= \gamma_\nu + \Lambda_\nu^*[S', D', \Gamma; e_0, p, p'] \\
W_\nu &= 2ik_\nu + ik_\nu T^*[S', D', \Gamma, W; e_0, k] \\
[S'(p)]^{-1} &= [S'(p_0)]^{-1} + (p - p_0)^\nu \Gamma_\nu(p, p_0) \\
[D'(k)]^{-1} &= \int_0^1 dx\, k^\nu W_\nu(xk)
\end{aligned} \tag{13.98}$$

The solutions to these equations are divergent, because the functionals Λ^* and T^* have skeletal divergences. However, they are only logarithmically divergent, and can be made finite through one subtraction. Using the abbreviations $\Lambda^*(p, p')$ and $T^*(k)$ for the functionals, we define two finite functionals by[1]

$$\begin{aligned}
\tilde{\Lambda}_\nu^*(p, p') &\equiv \Lambda_\nu^*(p, p') - [\Lambda_\nu^*(p_0, p_0')]_{\not{p}_0 = m} \\
\tilde{T}(k) &\equiv T^*(k) - T^*(\mu)
\end{aligned} \tag{13.99}$$

where μ is an arbitrary reference momentum and p_0 is the momentum of an electron on mass shell, with $p_0^2 = m^2$. The subscript $\not{p}_0 = m$ instructs us to commute $\not{p}_0$ all the way to the right, and then replace it by m. Thus

$$[\Lambda_\nu^*(p_0, p_0)]_{\not{p}_0 = m} = L\gamma_\nu \tag{13.100}$$

where L is a power series in e_0^2 with logarithmically divergent coefficients. The same is true of $T^*(\mu)$.

By replacing Λ^*, T^* respectively by $\tilde{\Lambda}^*, \tilde{T}^*$, we have, from (13.98), a set of finite functional equations. The renormalized functions $\tilde{S}, \tilde{D}, \tilde{\Gamma}, \tilde{W}$ are defined as solutions to these finite equations, with e_0 replaced by an appropriately defined renormalized charge $e(\mu)$:

$$\begin{aligned}
\tilde{\Gamma}_\nu(p, p') &= \gamma_\nu + \tilde{\Lambda}\%_\nu[\tilde{S}', \tilde{D}', \tilde{\Gamma}; e, p, p'] \\
\tilde{W}_\nu &= 2ik_\nu + ik_\nu \tilde{T}^*[\tilde{S}', \tilde{D}', \tilde{\Gamma}, \tilde{W}; e, k] \\
[\tilde{S}'(p)]^{-1} &= [\tilde{S}'(p_0)]^{-1} + (p - p_0)^\nu \tilde{\Gamma}_\nu(p, p_0) \\
[\tilde{D}'(k)]^{-1} &= \int_0^1 dx\, k^\nu \tilde{W}_\nu(xk)
\end{aligned} \tag{13.101}$$

We fix the normalization of $\tilde{S}'$ by the condition

$$[\tilde{S}'(p_0)]^{-1} = \not{p}_0 - m \tag{13.102}$$

[1]For simplicity, we have chosen to subtract Γ_μ at a mass-shell momentum p_0. Actually the subtraction can be made at any momentum, whose invariant mass would then serve as an extra renormalization parameter in addition to μ.

Then we have

$$[\tilde{S}'(p)]^{-1} = \not{p} - m + (p - p_0)^{\nu}\tilde{\Lambda}_{\nu}(p, p_0) \tag{13.103}$$

with the property

$$[(\not{p} - m)\tilde{S}'(p)]_{p=p0} = 1 \tag{13.104}$$

The normalization of $\tilde{D}$ is such that

$$[ik^2\tilde{D}(k)]_{k=\mu} = 1 \tag{13.105}$$

To complete the renormalization scheme, we show that the renormalized quantities are proportional to the unrenormalized ones. Note that $\tilde{\Gamma}_{\mu}$ can be rewritten as follows:

$$\begin{aligned}\tilde{\Gamma}_{\nu} &= \gamma_{\nu} + \Lambda_{\nu}^{*} - L\gamma_{\nu} = (1 - L)\left(\gamma_{\nu} + \frac{1}{1 - L}\Lambda_{\nu}^{*}\right) \\ &= Z'\left\{\gamma_{\nu} + \frac{1}{Z'}\Lambda_{\nu}^{*}[\tilde{S}, \tilde{D}, \tilde{\Gamma}; e, p, p']\right\}\end{aligned} \tag{13.106}$$

where

$$Z' = 1 - L \tag{13.107}$$

This shows that the subtraction is equivalent to rescaling. Similarly, we can write

$$\begin{aligned}\tilde{W}_{\nu} &= 2ik_{\nu} + ik_{\nu}[T^{*}(k) - T^{*}(\mu)] \\ &= Z(\mu^2)\left\{2ik_{\nu} + \frac{1}{Z(\mu^2)}ik_{\nu}T^{*}[\tilde{S}, \tilde{D}, \tilde{\Gamma}, \tilde{W}; e, k]\right\}\end{aligned} \tag{13.108}$$

where

$$Z(\mu^2) = 1 - \tfrac{1}{2}T^{*}(\mu) \tag{13.109}$$

Using the scaling properties (13.94) and (13.95), we obtain

$$\begin{aligned}\frac{1}{Z'}\tilde{\Gamma}_{\nu} &= \gamma_{\nu} + \Lambda_{\nu}^{*}\left[Z'\tilde{S}; Z\tilde{D}, \frac{\tilde{\Gamma}}{Z'}; Z^{-1/2}e, p, p'\right] \\ \frac{1}{Z(\mu^2)}\tilde{W}_{\nu} &= ik_{\nu}T^{*}\left[Z'\tilde{S}, Z\tilde{D}, \frac{\tilde{\Gamma}}{Z'}, \frac{\tilde{W}}{Z}; Z^{-1/2}e, k\right]\end{aligned} \tag{13.110}$$

Thus, the system of equations (13.101) can be reduced to (13.98) by putting

$$\Gamma = \frac{1}{Z'}\tilde{\Gamma}$$

$$W_\mu = \frac{1}{Z(\mu^2)}\tilde{W}_\mu$$

$$S' = Z'\tilde{S}$$

$$D' = Z(\mu^2)\tilde{D}$$

$$e_0^2 = \frac{e^2(\mu^2)}{Z(\mu^2)} \tag{13.111}$$

This explicitly shows the connection between the unrenormalized and renormalized quantities.

The system of functional equations is a formal property of perturbation expansions. There is no guarantee that the expansions converge, or that the functionals actually exist mathematically. Our best indication that the process makes some sense is still the good agreement between perturbation theory and experiments.

13.9 CALLAN–SYMANZIK EQUATION

Under a change of scale, the running coupling constant changes according to

$$e^2(\mu^2) = e_0^2 Z(\mu^2) \tag{13.112}$$

where the renormalization constant $Z(\mu^2)$ depends on the cutoff Λ, the electron mass m, and the bare coupling $e_0^2 = e^2(\Lambda^2)$. It can depend only on the ratios $\mu^2/\Lambda^2$and m^2/Λ^2, because it is dimensionless. In a more general discussion, we would treat the mass as another running coupling constant, but here we shall assume $\mu^2 \gg m^2$ and set $m = 0$. Thus we rewrite (13.112) in the form

$$\alpha(\mu^2) = \alpha(\Lambda^2) Z\left(\frac{\mu^2}{\Lambda^2}, \alpha(\Lambda^2)\right) \tag{13.113}$$

or, putting $x = \mu^2$ and $y = \Lambda^2$,

$$\alpha(x) = \alpha(y)\, Z\left(\frac{x}{y}, \alpha(y)\right) \tag{13.114}$$

Differentiating both sides with respect to x at fixed y, and then setting $y = x$, we obtain

$$x\frac{d\alpha(x)}{dx} = \beta(\alpha(x)) \tag{13.115}$$

where

$$\beta(\alpha) = \alpha\frac{\partial Z(x, \alpha)}{\partial x}\bigg|_{x=1} \tag{13.116}$$

The function $\beta(\alpha)$ is the "β function," also called the Gell-Mann–Low function [7]. It generates scale transformations called *renormalization-group transformations.* The β function of QED to lowest order in perturbation theory can be obtained from (13.63):

$$\beta(\alpha) = \frac{\alpha^2}{3\pi} \tag{13.117}$$

Since this is positive, the coupling grows at high energies, and consequently we cannot investigate the high-energy behavior of QED using perturbation theory. In quantum chromodynamics, on the other hand, the β function is negative at small coupling, and thus the coupling vanishes at high energies—a behavior known as *asymptotic freedom.*

An important property of the β function is that it depends only on α. We can obtain the running coupling constant $\alpha(x)$ by integrating (13.115) to obtain

$$\int_{\alpha_1}^{\alpha_2} \frac{d\alpha}{\beta(\alpha)} = \ln\frac{x_2}{x_1} \tag{13.118}$$

where $\alpha_2 = \alpha(x_2)$, $\alpha_1 = \alpha(x_1)$. This equation expresses the essence of renormalization: *A dimensionless coupling constant at one energy scale can be related to that at another energy scale, without reference to any intrinsic scale.* This means that there is nothing special about the cutoff; it is just a scale like any other. Where $\beta(\alpha) = 0$, the value of α is at a fixed point, where it remains invariant under scale changes. These fixed points define possible quantum field theories, and are therefore physically interesting. We shall take up this subject in Chapter 16.

We can abstract from QED a general property of renormalizable field theories. Let $G'(p; \Lambda, g_0)$ be a renormalized correlation function, which is generally divergent, and depends on a high-momentum cutoff Λ. The external momenta are collectively denoted by p, and g_0 is a dimensionless coupling constant at the energy scale Λ. "Renormalizability" means that

$$\begin{aligned} G'(p; \Lambda, g_0) &= Z\left(\frac{\Lambda}{\mu}, g_0\right)G(p; \mu, g) \\ g &= g\left(\frac{\Lambda}{\mu}, g_0\right) \end{aligned} \tag{13.119}$$

where g is the renormalized coupling at the energy scale μ. The renormalized correlation function G is assumed to be a finite function of its arguments. The cutoff dependence is isolated in the dimensionless renormalization constant Z. We assume that μ and p are much greater than the particle masses, which have been neglected. Since the left side of the equation is independent of μ, so must be the right side. Thus, we can write

$$\frac{d}{d\mu}\left[Z\left(\frac{\Lambda}{\mu}, g_0\right)G(p; \mu, g)\right] = 0 \tag{13.120}$$

Carrying out the differentiation, we obtain

$$\frac{\partial G}{\partial \mu} + \frac{\partial \ln Z}{\partial \mu} G + \frac{\partial g}{\partial \mu}\frac{\partial G}{\partial g} = 0 \tag{13.121}$$

where partial derivatives are carried out with all other arguments held fixed. We rewrite this in the following form, which is called the *Callan–Symanzik equation* [8]:

$$\left[\mu\frac{\partial}{\partial \mu} + \eta(\mu) + \beta(g)\frac{\partial}{\partial g}\right]G(p; \mu, g) = 0 \tag{13.122}$$

where

$$\begin{aligned}\beta(g) &\equiv \mu\frac{\partial}{\partial \mu} g\left(\frac{\Lambda}{\mu}, g_0\right)\\ \eta(\mu) &\equiv \mu\frac{\partial}{\partial \mu} \ln Z\left(\frac{\Lambda}{\mu}, g_0\right)\end{aligned} \tag{13.123}$$

The first equation gives the β function, while the second is called "anomalous dimension."

13.10 TRIVIALITY

Landau [9] noted that, by calculating the renormalized charge using a plausible partial summation of Feynman graphs, one arrives at the conclusion that it vanishes in the limit of infinite cutoff. He concluded that this absurd result invalidates quantum field theory, which should be "buried with due honors."

We can derive Landau's result as follows. Substituting (13.117) into (13.118), with $\alpha_2 = \alpha(\Lambda^2)$ and $\alpha_1 = \alpha(k^2)$, we obtain

$$\frac{1}{\alpha(\mu^2)} = \frac{1}{\alpha(\Lambda^2)} + \frac{1}{24\pi^2}\ln\frac{\Lambda^2}{\mu^2} \tag{13.124}$$

This is essentially (13.61), though it pretends to be more accurate. Taken seriously, this implies that, if $\alpha(\Lambda^2) \geq 0$, then

$$\alpha(\mu^2) \underset{\Lambda\to\infty}{\longrightarrow} 0 \tag{13.125}$$

This is referred to as "triviality."

Similar perturbative calculations also point to triviality for scalar ϕ^4 theory. (See Problem 13.4.) The conclusion for ϕ^4 theory is supported by numerical calculations, and can be established more rigorously, as we shall show in Chapter 17. The result for QED is plausible although not proven; but even if proven, it hardly invalidates quantum field theory.

To view this alarming result in proper light, we have to understand what renormalization is really about, and we shall explain that in Chapter 17. In the meantime, we offer the following observations:

- Triviality has no impact on practical applications, because the renormalized charge is insensitive to Λ, as it goes like $(\ln \Lambda)^{-1}$. One can therefore choose a finite value of Λ to fit experiments.
- By accepting a finite Λ, however, one seems to negate renormalization, which says that we can hold the renormalized charge fixed at an arbitrary value, and send the cutoff to infinity. We shall see in Chapter 16 that this is possible only if the theory is based on an "ultraviolet fixed point," as in one with asymptotic freedom. It is not correct for a theory governed by an "infrared fixed-point," as is the case for ϕ^4 theory. In the latter case, the renormalized charge assumes the fixed-point value.
- Although the mathematical QED appears to be an "infrared"-type theory; the physical QED may not be, for it is embedded in the standard model, which is, in turn, embedded in some grand, yet unknown, unified theory. We offer a scenario for this in Section 16.9.

PROBLEMS

13.1 Ward–Takahashi Identity

(a) Show that the Heisenberg operators ψ and $\overline{\psi}$ satisfy the equations of motion

$$(i\gamma^\mu \partial_\mu - m_0)\psi = e_0 \gamma^\mu A_\mu \psi$$
$$\overline{\psi}(-i\gamma^\mu \overleftarrow{\partial}_\mu - m_0) = e_0 \overline{\psi}\gamma^\mu A_\mu$$

and from these show that the current $j^\mu = \overline{\psi}\gamma^\mu\psi$, as a Heisenberg operator, is formally conserved: $\partial_\mu j^\mu = 0$.

(b) Consider the quantity

$$V^\mu(x_1, x_2, y) \equiv \langle 0|\mathrm{T}\ \psi(x_1)\overline{\psi}(x_2)j^\mu(y)|0\rangle$$

where the operators are in the Heisenberg picture. Because of translational invariance, the Fourier transform of V^μ can be written in the form

$$\int d^4x_1\, d^4x_1\, d^4y\, e^{-i(p_1\cdot x_1 - p_2\cdot x_2 + k\cdot y)} V^\mu(x_1, x_2, y) = (2\pi)^4 \delta^4(p_1 - p_2 + k)\tilde{V}^\mu(p_1, p_2)$$

Obtain $\tilde{V}^\mu$ as an expansion in Feynman graphs, and show that

$$\tilde{V}^\mu(p_1, p_2) = S'_F(p_2)\Gamma^\mu(p_1, p_2)S'_F(p_1)$$

where Γ^μ is the full vertex.

(c) Using current conservation, show that

$$\frac{\partial}{\partial z^\mu} V^\mu(x_1, x_2, y) = [\delta^4(x_1 - y) - \delta^4(x_2 - y)]iS'_F(x_1 - x_2)$$

This is the spatial form of the Ward–Takahashi identity. Take the Fourier transform to obtain the form in momentum space:

$$(p - p_0)^\mu \Gamma_\mu(p, p_0) = [S'(p_0)]^{-1} - [S'(p)]^{-1}$$

13.2 Corrections to Coulomb's Law Consider the scattering of an electron from an infinitely heavy point charged, with 4-momentum transfer $k^\mu = (0, \mathbf{k})$. The electrostatic potential $V(r)$ is the Fourier transform of the scattering amplitude with respect to $\mathbf{k}$.

(a) Show that the potential is related to the running coupling constant $\alpha(k^2)$:

$$V(r) = e_0^2 \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} D'(k)|_{k0=0} = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} \frac{4\pi\alpha(-\mathbf{k}^2)}{\mathbf{k}^2}$$

(b) The deviation from Coulomb's law may be attributed to the fact that the electron is a charge distribution. Show that the charge density is given by

$$\rho(\mathbf{r}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} 4\pi\alpha(-\mathbf{k}^2)$$

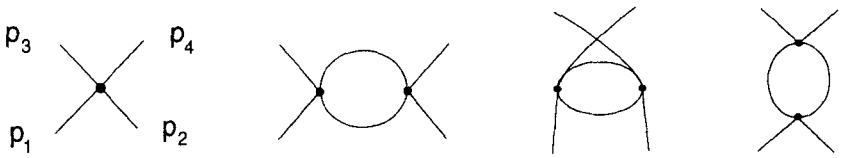
(c) Evaluate $\rho(\mathbf{r})$ for small r, and compare the result with the Coulomb potential.

13.3 Landau Ghost According to (13.60), the full photon propagator is related to the running coupling constant through $e_0^2 D'(k) = -4\pi\alpha(k^2)/k^2$. Show that the Landau formula (13.124) leads to

$$D'(k) = -k^{-2}\left(1 - \frac{\alpha}{24\pi^2} \ln \frac{k^2}{\Lambda^2}\right)^{-1}$$

There is a pole other than the photon pole at $k^2 = 0$. Show that it cannot correspond to a physical particle because the residue has the wrong sign. Estimate its location in physical terms. This is called the "Landau ghost," a curiosity that belongs to the same category as tachyons and Maxwell demons.

13.4 Coupling-Constant Renormalization Consider scalar ϕ^4 theory. Ignore mass renormalization and set the bare mass to zero. The proper vertex G is given to second order by the following Feynman graphs:



(a) Show that

$$G(p_i;\Lambda, \lambda_0) = \lambda_0 - \tfrac{1}{2}\lambda_0^2[I(p_1+p_2) + I(p_2+p_3) + I(p_3+p_1)]$$

where Λ is the cutoff and

$$I(p) = -i\int \frac{d^4k}{(2\pi)^4}\frac{1}{(k^2+i\eta)[(k+p)^2+i\eta]} = \frac{1}{16\pi^2}\ln\frac{\Lambda^2}{-p^2+i\eta}$$

(b) Define the renormalized coupling constant at scale μ by $\lambda(\Lambda/\mu, \lambda_0) = G(p_i; \Lambda, \lambda_0)$ with all $p_i^2 = -\mu^2$. Show that

$$\lambda\left(\frac{\Lambda}{\mu}, \lambda_0\right) = \lambda_0 - \frac{3\lambda_0^2}{16\pi^2}\ln\frac{\Lambda}{\mu}$$

Use the notation $\lambda(\mu) = \lambda(\Lambda/\mu, \lambda_0)$, $\lambda_0 = \lambda(\Lambda)$, rewrite this as

$$\frac{1}{\lambda(\mu)} = \frac{1}{\lambda(\Lambda)} + \frac{3}{16\pi^2}\ln\frac{\Lambda}{\mu}$$

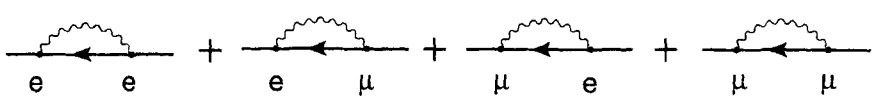
which is valid to second order. This exhibits triviality; for the theory to be physical, $\lambda(\Lambda) \geq 0$. Hence $\lambda(\mu) \to 0$ when $\Lambda \to \infty$.

13.5 Neutron–Proton Mass Difference [10] Since the proton has charge whereas the neutron is neutral, one might expect the proton to be heavier because of the electromagnetic self-energy; but in fact, the opposite is true. To understand this, one must include the magnetic self-energy due to the anomalous magnetic moment. This can be done phenomenologically using the Pauli term described in Problem 12.3:

$$\mathcal{H}_{\text{Pauli}}(x) = \frac{\kappa\mu_0}{2}\sigma^{\mu\nu}F_{\mu\nu}(x)$$

$$\kappa_P = 1.79 \qquad \kappa_N = -1.91$$

where $\mu_0 = |e|/2M$ is the nuclear magneton, with M the nucleon mass. The self-energy now is the sum of the graphs shown in the accompanying figure.



(a) Show that, for large momentum cutoff Λ, the self-mass δM is given by

$$\frac{\delta M}{M} = \frac{3\alpha}{2\pi} \ln \frac{\Lambda}{M} - \frac{\alpha}{2\pi}\left(3\kappa + \frac{5}{4}\kappa^2\right)\left[\left(\frac{\Lambda}{M}\right)^2 - \ln \frac{\Lambda}{M}\right]$$

where $\alpha \leqslant \frac{1}{137}$ is the fine-structure constant.

(b) The experimental value of the neutron–proton mass difference is approximately 1.26 MeV, or

$$\frac{\Delta M}{M} \approx 1.23 \frac{\alpha}{2\pi}$$

Show that this can be fit with a value $\Lambda/M \approx 1$, which is physically reasonable.

REFERENCES

1. H. A. Lorentz, *The Theory of the Electron*, 2nd ed., Dover Publications, New York, 1952, Note 18, p. 252.
2. V. F. Weisskopf, *Phys. Rev.* **56,** 72 (1939); K. Huang, *Phys. Rev.* **101,** 1173 (1956).
3. J. Schwinger, *Phys. Rev.* **82,** 664 (1981); K. Johnson, in *Particles and Field Theory,* S. Deser and K. W. Ford, Eds., Prentice-Hall, Englewood Cliffs, NJ, 1965.
4. F. J. Dyson, *Phys. Rev.* **75,** 486, 1736 (1949); J. C. Ward, *Proc. Phys. Soc.* (London) **A64,** 54 (1951).
5. M. Gell-Mann and F. E. Low, *Phys. Rev.* **95,** 1300 (1954).
6. T. T. Wu, *Phys. Rev.* **125,** 1436 (1962).
7. M. Gell-Mann and F. E. Low, *Phys. Rev.* **95,** 1300 (1954).
8. C. G. Callan, *Phys. Rev.* **D12,** 1541 (1970); K. Symanzik, *Commun. Math. Phys.* **18,** 227 (1970).
9. L. D. Landau, in *Niels Bohr and the Development of Physics*, W. Pauli, Ed., McGraw-Hill, New York, 1955).
10. R. P. Feynman and G. Speisman, *Phys. Rev.* **94,** 500 (1954); K. Huang, *Phys. Rev.* **101,** 1173 (1956).

CHAPTER FOURTEEN

Path Integrals

14.1 PATH INTEGRALS IN QUANTUM MECHANICS

Quantum mechanics is based on the commutation relation between the momentum operator p_{op} and the coordinate operator q_{op} at equal times, say, $t = 0$:

$$[p_{op}, q_{op}] = -i \tag{14.1}$$

In the Heisenberg picture, the coordinate operator at time t is given by

$$q_{op}(t) = e^{iHt} q_{op} e^{-iHt} \tag{14.2}$$

where H is the Hamiltonian operator. If $|q'\rangle$ denotes the eigenstate of q_{op} with eigenvalue q', then the eigenstate of $q_{op}(t')$ with eigenvalue q' is given by

$$|q', t'\rangle = e^{iHt'}|q'\rangle \tag{14.3}$$

The transition amplitude defined by

$$\langle q'', t''|q', t'\rangle \equiv \langle q'|e^{-iH(t''-t')}|q\rangle \tag{14.4}$$

contains a complete description of the quantum-mechanical system. Through the Feynman path integral [1], we can express this amplitude in terms of the classical Lagrangian, and thus obtain an alternative formulation of quantum mechanics that makes no reference to a Hilbert space.

To derive the path-integral representation, we break up the time interval $t'' - t'$ into N equal pieces, and use the identity

$$e^{-iH(t''-t')} = \lim_{N\to\infty}\left[1 - \frac{iH(t''-t')}{N}\right]^N \tag{14.5}$$

To simplify the notation, let

$$T \equiv 1 - i\mathrm{H}\Delta$$
$$\Delta = \frac{t'' - t'}{N} \tag{14.6}$$

Then we can write

$$e^{-iH(t''-t')} = T^N \tag{14.7}$$

with the understanding that $N \to \infty$. Now insert complete sets of coordinate eigenstates between factors of T to write

$$\langle q'', t''|q', t'\rangle = \int_q \langle q_N|T|q_{N-1}\rangle \cdots \langle q_2|T|q_1\rangle \langle q_1|T|q_0\rangle \tag{14.8}$$

where

$$q_0 \equiv q'$$
$$q_N \equiv q'' \tag{14.9}$$

and

$$\int_q \equiv \int dq_1 \cdots dq_{N-1} \tag{14.10}$$

Next replace the matrix element $\langle \mathrm{q}'|T|q\rangle$ by the mixed matrix element $\langle p|T|q\rangle$, by inserting complete sets of momentum eigenstates in appropriate places:

$$\langle q'', t''|q', t'\rangle = \int_q \int_p \langle q_N|p_{N-1}\rangle\langle p_{N-1}|T|q_{N-1}\rangle \cdots \langle p_1|T|q_1\rangle\langle q_1|p_0\rangle\langle p_0|T|q_0\rangle \tag{14.11}$$

where

$$\int_p \equiv \int \frac{dp_0}{2\pi} \cdots \frac{dp_{N-1}}{2\pi} \tag{14.12}$$

A classical Hamiltonian $H(p, q)$ can be defined through the relation

$$\langle p|H|q\rangle \equiv \langle p|q\rangle H(p, q) \tag{14.13}$$

To obtain $H(p, q)$, we pull all the p_{op} occurring in H to the far left, commuting across the q_{op} if necessary, and then replace them by the number p. Writing

$$T(p, q) = 1 - iH(p, q)\Delta \tag{14.14}$$

we have

$$\langle q'', t''|q', t'\rangle = \int_q \int_p [\langle q_N|p_{N-1}\rangle \cdots \langle q_1|p_0\rangle][T(p_{N-1}, q_{N-1}) \cdots T(p_0, q_0)] \tag{14.15}$$

The time evolution over the interval $\Delta t = \Delta$ is now effected by $T(p, q)$, a c-number. The crucial step toward obtaining the final result is to make the replacement

$$T(p, q) \to e^{-iH(p,q)\Delta} \tag{14.16}$$

which is exact in the limit $\Delta \to 0$. The purpose is to make $T(p, q)$ a unitary transformation over the finite time interval Δ. Putting $\langle p|q\rangle = \exp(ipq)$, we obtain

$$\langle q'', t''|q', t'\rangle = \int_q \int_p \exp\left\{i \sum_{j=1}^{N} [p_j(q_{j-1} - q_j) - H(p_j, q_j)\Delta]\right\} \tag{14.17}$$

We can think of $\{p_j, q_j\}$ as successive samplings of a path in phase space at equal time intervals Δ, and the whole path $\{p(t), q(t)\}$ is covered when $\Delta \to 0$. In that limit $(q_{i-1} - q_i)/\Delta$ becomes the velocity $\dot{q}(t)$, and the integrals over q_i and p_j become functional integrals:

$$\langle q'', t''|q', t'\rangle = \int_{q'}^{q''} Dq\, Dp \exp\left\{i \int_{t'}^{t''} dt[p(t)\dot{q}(t) - H(p(t), q(t))]\right\} \tag{14.18}$$

where the limits of integration denote the endpoint

$$\begin{aligned} q(t') &= q' \\ q(t'') &= q'' \end{aligned} \tag{14.19}$$

The measures of the functional integration are given by

$$\begin{aligned} Dq &= \prod_{i=1}^{N-1} dq_i \\ Dp &= \prod_{i=1}^{N} \int \frac{dp_j}{2\pi} \end{aligned} \tag{14.20}$$

which, however, do not have well-defined limits when $\Delta \to 0$. For this reason, we retain the discrete time steps for all intermediate computations, and take the desired limit only in the final answer.

Assume that the classical Hamiltonian has the form

$$H(p, q) = \frac{p^2}{2m} + V(q) \tag{14.21}$$

We can then perform the momentum integrations in (14.18) to obtain the final result

$$\langle q'', t''|q', t'\rangle = \mathcal{N} \int_{q'}^{q''} Dq\, e^{iS[q;t'',t']} \tag{14.22}$$

where $\mathcal{N}$ is a normalization constant, S is the classical action between the times t' and t'':

$$S[q; t'', t'] \equiv \int_{t'}^{t''} dt\, L(q(t), \dot{q}(t)) \tag{14.23}$$

and $L(q, \dot{q})$ is the classical Lagrangian:

$$L(q, \dot{q}) = \tfrac{1}{2} m\dot{q}^2 - V(q) \tag{14.24}$$

The representation (14.22) is the *Feynman path integral.* It gives the transition amplitude as a "sum over histories"—a sum over all possible paths leading from the initial state to the final state, weighted by i times the classical action of the path. An illustration of such a path is shown in Fig. 14.1. In the limit $\Delta \to 0$, the limiting paths are generally discontinuous.

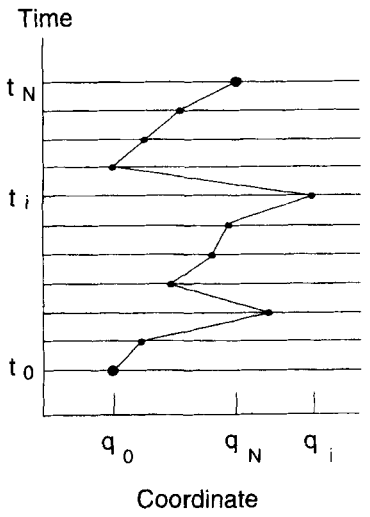


Figure 14.1 The coordinates $q_i (i = 1, \ldots, N-1)$ that make up the path are to be integrated independently over all possible values.

Consider the following matrix element of coordinate operators:

$$\langle q'', t''|q_{\rm op}(t_2)q_{\rm op}(t_1)|q', t'\rangle = \langle q''|e^{-iH(t''-t_2)}q_{\rm op}e^{-iH(t_2-t_1)}q_{\rm op}e^{-iH(t_1-t')}|q'\rangle \quad (14.25)$$

If $t_2 > t_1$, we can proceed with the same treatment as before. The only difference is that, when complete sets of coordinate eigenstates are inserted, we will pick up a factor of the eigenvalue $q(t_2)q(t_1)$. If $t_2 < t_1$, on the other hand, the procedure fails to go through. We can state the following general result:

$$\langle q'', t''|\mathrm{T}\, q_{\rm op}(t_1)\cdots q_{\rm op}(t_n)|q', t'\rangle = \mathcal{N}\int_{q'}^{q''} Dq\, q(t_1)\cdots q(t_n)e^{iS[q;t'',t']} \quad (14.26)$$

where T is the time-ordering operator. If we try to express a matrix element of field operators that are not chronologically ordered, we would be faced with multivalued paths, and new methods would be needed to render the paths well-defined.

14.2 IMAGINARY TIME

The time in the transition amplitudes can be analytically continued to pure imaginary values. In this domain the integrand of the path integral is real instead of pure imaginary, and this makes it convenient for some computations. The pure-imaginary time is called "Euclidean time," because it converts Minkowskian space–time to a 4-dimensional Euclidean space.

Let us go back to (14.4) at the beginning, and insert a complete set of energy eigenstates:

$$\begin{aligned} f(t) &\equiv \langle q'', t''|q', t'\rangle = \langle q''|e^{iHt}|q'\rangle \\ &= \sum_n \langle q''|n\rangle\langle n|q'\rangle e^{-iE_n t} \end{aligned} \quad (14.27)$$

where $t = t'' - t'$, and the nth eigenstate defined by

$$\begin{aligned} H|n\rangle &= E_n|n\rangle \\ E_0 &= 0 \end{aligned} \quad (14.28)$$

The energy spectrum is bounded from below by assumption. We can therefore analytically continue $f(t)$ to negative imaginary time

$$t = -i\tau \qquad (\tau > 0) \quad (14.29)$$

The result may be written

$$f(-i\tau) = \sum_n \rho_n e^{-E_n\tau} \tag{14.30}$$

where

$$\rho_n \equiv \langle q''|n\rangle\langle n|q'\rangle = \psi_n(q'')\psi_n^*(q') \tag{14.31}$$

where $\psi_n(q)$ is the wave function of the nth eigenstate. The oscillatory terms are now replaced by damped exponentials. For example, we can extract the energy eigenvalues one by one from the asymptotic behavior

$$f(-i\tau) \underset{\tau\to\infty}{\to} \rho_0 + \rho_1 e^{-E_1\tau} + \cdots \tag{14.32}$$

The Feynman path integral for imaginary times can be obtained by repeating the derivation of (14.22) using imaginary time intervals, with the result

$$\langle q'', -i\tau''|q', -i\tau'\rangle = \mathcal{N}\int_{q'}^{q''} Dq\, e^{-S_E[q;\tau'',\tau']} \tag{14.33}$$

This is an integral over paths $q(\tau)$ in imaginary time, with given endpoints q'',q'. The "Euclidean action" S_E is defined as

$$S_E[q;\tau'',\tau'] = \int_{\tau'}^{\tau''} d\tau\left[\frac{m}{2}\left(\frac{dq}{d\tau}\right)^2 + V(q)\right] \tag{14.34}$$

Note that the Hamiltonian, rather than the Lagrangian, appears in the integrand. The ground-state wave function can be related to the path integral as follows:

$$\psi_0(q'')\psi_0^*(q') = \lim_{\tau\to\infty} f(-i\tau) = \mathcal{N}\int_{q'}^{q''} Dq\, e^{-S_E[q]} \tag{14.35}$$

where

$$S_E[q] = \int_{-\infty}^{\infty} d\tau\left[\frac{m}{2}\left(\frac{dq}{d\tau}\right)^2 + V(q)\right] \tag{14.36}$$

14.3 PATH INTEGRALS IN QUANTUM FIELD THEORY

Consider a scalar field $\phi(\mathbf{x})$ in D spatial dimensions in the Schrödinger picture, with Hamiltonian density

$$\mathcal{H}(\pi(\mathbf{x}), \phi(\mathbf{x})) = \tfrac{1}{2}\pi^2(\mathbf{x}) + \tfrac{1}{2}|\nabla\phi(\mathbf{x})|^2 + V(\phi(\mathbf{x})) \tag{14.37}$$

We enclose the system in a large periodic box, and introduce a high-momentum cutoff Λ, so that the number of degrees of freedom is finite. The path-integral formulation can be taken over from that in quantum mechanics. We can view $\phi(\mathbf{x})$ as a coordinate labeled by $\mathbf{x}$, and denote the eigenstate of $\phi(\mathbf{x})$ by $|\phi'\rangle$. The eigenstate of the Heisenberg operator $\phi(\mathbf{x}, t)$ is then

$$|\phi', t'\rangle = e^{iHt'}|\phi'\rangle \tag{14.38}$$

The transition amplitude between two such states is

$$\langle\phi'', t''|\phi', t'\rangle = \langle\phi''|e^{-iH(t''-t')}|\phi'\rangle \tag{14.39}$$

From our earlier discussion, we can write down the path-integral representation

$$\langle\phi'', t''|\phi', t'\rangle = \mathcal{N}\int_{\phi'}^{\phi''} D\phi \int D\pi \exp\left\{i\int_{t'}^{t''} dt \int d^D\mathbf{x}[\pi(\mathbf{x})\dot{\phi}(\mathbf{x}) - \mathcal{H}(\pi(\mathbf{x}), \phi(\mathbf{x}))]\right\} \tag{14.40}$$

where the limits on the $D\phi$ integration refer to the endpoint condition

$$\phi(\mathbf{x}, t'') = \phi''(\mathbf{x})$$
$$\phi'(\mathbf{x}, t') = \phi'(\mathbf{x})$$

The Feynman path integral is obtained by carrying out the $D\pi$ integration:

$$\langle\phi'', t''|\phi', t'\rangle = \mathcal{N}\int_{\phi'}^{\phi''} D\phi\, e^{iS[\phi,t'',t']} \tag{14.41}$$

where S is the classical action:

$$S[\phi; t'', t'] = \int_{t'}^{t''} dt \int d^D\mathbf{x}\mathcal{L}(\mathbf{x}, t) \tag{14.42}$$

and $\mathcal{L}$ is the classical Lagrangian density:

$$\mathcal{L}(\mathbf{x}, t) = \tfrac{1}{2}\partial^\mu\phi(\mathbf{x}, t)\partial_\mu\phi(\mathbf{x}, t) - V(\phi(\mathbf{x}, t)) \tag{14.43}$$

It is straightforward to generalize (14.41) to

$$\langle\phi'', t''|T\phi_{\text{op}}(\mathbf{x}_1, t_1)\cdots\phi_{\text{op}}(\mathbf{x}_n, t_n)|\phi', t'\rangle$$
$$= \mathcal{N}\int_{\phi'}^{\phi''} D\phi\, \phi(\mathbf{x}_1, t_1)\cdots\phi(\mathbf{x}_n, t_n)\, e^{iS[\phi,t'',t']} \tag{14.44}$$

14.4 EUCLIDEAN SPACE–TIME

As in quantum mechanics, we can use imaginary time $\tau = it$. Together with the spatial coordinates $\mathbf{x}$, we have Euclidean space–time coordinates $x_E = (\mathbf{x}, \tau)$ of dimension $d = D + 1$. The generalized transition amplitude becomes

$$\begin{aligned}&\langle \phi'', -i\tau'' | T\phi_{op}(\mathbf{x}_1, -i\tau_1) \cdots \phi_{op}(\mathbf{x}_n, -i\tau_n) | \phi', -i\tau' \rangle \\ &= \mathcal{N} \int_{\phi'}^{\phi''} D\phi\ \phi(\mathbf{x}_1, \tau_1) \cdots \phi(\mathbf{x}_n, \tau_n) e^{-S_E[\phi, \tau'', \tau']}\end{aligned} \tag{14.45}$$

where S_E denotes the Euclidean action:

$$S_E[\phi; \tau'', \tau'] = \int_{\tau'}^{\tau''} d\tau \int d^D\mathbf{x}\ [\tfrac{1}{2} \partial^i \phi(\mathbf{x}, \tau) \partial^i \phi(\mathbf{x}, \tau) + V(\phi(\mathbf{x}, \tau))] \tag{14.46}$$

The path integral extends over all such fields with the specified endpoints. Note that there is no longer a distinction between upper and lower indices on ∂^i. For infinite time interval, $\tau'' - \tau' \to \infty$, we write

$$S_E[\phi] = \int d^d x_E [\tfrac{1}{2} \partial^i \phi(x_E) \partial^i \phi(x_E) + V(\phi(x_E))] \tag{14.47}$$

It is assumed that the initial and final constraints become irrelevant in this limit.

For future reference, we summarize the relation between Minkowskian space–time $x = (\mathbf{x}, x_0)$ and Euclidean space–time $x_E = (\mathbf{x}, x_d)$:

$$\begin{aligned}x_0 &= -ix_d \\ d^d x &= -id^d x_E \\ x_E^2 &= \sum_{i=1}^{d} x_i^2 = -x^2\end{aligned} \tag{14.48}$$

The relation between Minkowskian momentum k and Euclidean momentum k_E is defined such that such that $k_d x_d = k_0 x_0$:

$$\begin{aligned}k_0 &= ik_d \\ d^d k &= id^d k_E \\ k_E^2 &= \sum_{i=1}^{d} k_i^2 = -k^2\end{aligned} \tag{14.49}$$

We will omit the subscript "E" on Euclidean quantities when the context makes it unnecessary.

14.5 VACUUM AMPLITUDES

We have obtained path integrals for transition amplitudes between field eigenstates. For practical calculations, however, it is more convenient to work with amplitudes taken between vacuum states. To derive path integrals for the latter, we first couple the field to an external source $J(\mathbf{x}, t)$, which is switched on and off adiabatically:

$$J(\mathbf{x}, t) \xrightarrow[|t|\to\infty]{} 0 \tag{14.50}$$

The classical Lagrangian density becomes

$$\mathcal{L}_J(\mathbf{x}, t) = \mathcal{L}(\mathbf{x}, t) - J(\mathbf{x}, t)\phi(\mathbf{x}, t) \tag{14.51}$$

The transition amplitude in the presence of external field is denoted by

$$\langle \phi'', t''|\phi', t'\rangle_J = \mathcal{N}\int_{\phi'}^{\phi''} D\phi \exp\left\{ i\int_{t'}^{t''} dt \int d^D\mathbf{x}[\mathcal{L}(x) - J(x)\phi(x)]\right\} \tag{14.52}$$

For simplicity we have written $x = (\mathbf{x}, t)$. By taking the functional derivative of this amplitude with respect to $J(x)$, we insert a factor $-i\phi(x)$ in the integrand. Thus

$$\langle \phi'', t''|T\phi_{\text{op}}(x_1)\cdots\phi_{\text{op}}(x_n)|\phi', t'\rangle = i^n \frac{\delta\langle \phi'', t''|\phi', t'\rangle_J}{\delta J(x_1)\cdots\delta J(x_n)}\bigg|_{J=0} \tag{14.53}$$

We now show that the vacuum correlation function can be obtained by letting $t' \to -\infty$ and $t'' \to \infty$.

Assume for simplicity that the external source is present only during a finite but large time interval T:

$$J(x) = 0 \quad \text{for} \quad |t| > T \tag{14.54}$$

Eventually, we shall take the limit $T \to \infty$. Consider the transition amplitude $\langle \phi_2, t_2|\phi_1, t_1\rangle_J$ for t_1 before the source was turned on, and t_2 after the source was turned off. Assuming that the eigenstates of the field form a complete set at any time t, we have

$$\int D\phi|\phi, t\rangle\langle\phi, t| = 1$$

Thus, for $t_1 < -T$, $t_2 > T$, we can write

$$\langle \phi_2, t_2|\phi_1, t_1\rangle_J = \int D\phi\, D\phi'\, \langle \phi_2, t_2|\phi, T\rangle\langle \phi, T|\phi', -T\rangle_J\langle \phi', -T||\phi_1, t_1\rangle \tag{14.55}$$

The amplitudes $\langle \phi_2, t_2|\phi, T\rangle$ and $\langle \phi', -T||\phi_1, t_1\rangle$ pertain to source-free time intervals, and may be calculated as in the following example:

$$\langle \phi_2, t_2|\phi, T\rangle = \langle \phi_2|e^{-iH(t_2-T)}|\phi\rangle = \sum_n \langle \phi_2|n\rangle \langle n|\phi\rangle e^{-iE_n(t_2-T)} \tag{14.56}$$

where we have inserted a complete set of energy eigenstates, with vacuum state satisfying $H|0\rangle = 0$. To pick it out, we go to the limit of large imaginary $t_2 - T$, so that contributions from all other states are damped out:

$$\langle \phi_2, t_2|\phi, T\rangle \xrightarrow[t_2-T\to -i\infty]{} \langle \phi_2|0\rangle \langle 0|\phi\rangle \tag{14.57}$$

In the original amplitude, then, let us make $t_1 \to i\infty$, $t_2 \to -i\infty$, $T \to i\infty$:

$$\langle \phi_2, t_2|\phi_1, t_1\rangle_J \to \langle \phi_2|0\rangle \langle 0|\phi_1\rangle \int D\phi\, D\phi'\, \langle 0|\phi\rangle \langle \phi, T|\phi', -T\rangle_J \langle \phi'|0\rangle \tag{14.58}$$

The integral on the right side can be written as

$$\begin{aligned}\int D\phi D\phi'\, \langle 0|\phi\rangle \langle \phi, T|\phi', -T\rangle_J \langle \phi'|0\rangle &= \int D\phi D\phi'\, \langle 0|\phi\rangle \langle \phi|e^{-2iH_J T}|\phi'\rangle \langle \phi'|0\rangle \\ &= \langle 0|e^{-2iH_J T}|0\rangle = \langle 0^+|0^-\rangle_J \end{aligned} \tag{14.59}$$

where H_J is the total Hamiltonian in the presence of source and $\langle 0^+|0^-\rangle_J$ is the vacuum to vacuum amplitude in the presence of source discussed in Section 10.4. We can now write the following for this amplitude:

$$\langle 0^+|0^-\rangle_J = \lim_{\substack{t_1\to i\infty \\ t_2\to -i\infty}} \frac{\langle \phi_2, t_2|\phi_1, t_1\rangle_J}{\langle \phi_2|0\rangle\langle 0|\phi_1\rangle} \tag{14.60}$$

In taking the limit of large imaginary time, the real part of the time is kept arbitrary. All this does is make sure that oscillating phases damp out. Using (14.52) to rewrite the right side above in terms of path integrals, we have

$$\langle 0^+|0^-\rangle_J = \frac{\int D\phi e^{iS[\phi]-i(J,\phi)}}{\int D\phi e^{iS[\phi]}} \tag{14.61}$$

where $(J, \phi) = \int d^dx J(x)\phi(x)$. Taking functional derivatives of both sides, we obtain

$$\left.\frac{\delta\langle 0^+|0^-\rangle_J}{\delta J(x_1)\cdots\delta J(x_n)}\right|_{J=0} = (-i)^n \frac{\int D\phi\phi(x_1)\cdots\phi(x_n)\, e^{iS[\phi]}}{\int D\phi e^{iS[\phi]}} \tag{14.62}$$

On the other hand, we have seen in Section 10.4 that

$$\left.\frac{\delta\langle 0^+|0^-\rangle_J}{\delta J(x_1)\cdots\delta J(x_n)}\right|_{J=0} = (-i)^n\langle 0|T\phi_{\text{op}}(x_1)\cdots\phi_{\text{op}}(x_n)|0\rangle \tag{14.63}$$

Therefore

$$\langle 0|T\phi_{\rm op}(x_1)\cdots\phi_{\rm op}(x_n)|0\rangle = \frac{\int D\phi\phi(x_1)\cdots\phi(x_n)e^{iS[\phi]}}{\int D\phi e^{iS[\phi]}} \tag{14.64}$$

According to the linked-cluster theorem (10.45), $iW[J] = \ln\langle 0^+|0^-\rangle_J$ is the generating functional of connected correlation functions. We now have the representation

$$iW[J] = \ln\int D\phi e^{iS[\phi]-i(J,\phi)} - \ln\int D\phi e^{iS[\phi]} \tag{14.65}$$

The last term is irrelevant, since it is independent of $J(x)$.

14.6 STATISTICAL MECHANICS

As (14.64) shows, there is a similarity between a vacuum correlation function in path-integral form and an ensemble average in statistical mechanics. In fact, in Euclidean space–time the generating functional (14.65) is just the logarithm of the partition function in the presence of an external field:

$$Z[J] = \int D\phi\, e^{-S[\phi]-(J,\phi)} \tag{14.66}$$

where $S[\phi]$ is the Euclidean action (14.47), with subscript "E" suppressed, and

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

is an integral over Euclidean space. In statistical mechanics, we would make the identification

$$S[\phi] = \beta E[\phi] \tag{14.67}$$

where $E[\phi]$ is the energy functional of a classical field $\phi(x)$ and β is the inverse temperature. The thermodynamic free-energy density is proportional to the generating functional:

$$f[J] = -\frac{1}{\beta}\lim_{\Omega\to\infty}\frac{\ln Z[J]}{\Omega} \tag{14.68}$$

where Ω is the volume of the d–dimensional space. Like the generating functional, it is determined only up to an additive constant.

We see that (14.66) gives on one hand the generating functional for a quantum field theory in d space–time dimensions, and on the other hand the partition func-

tion for a classical field in *d spatial* dimensions. This is a special case of an equivalence between quantum field theory and quantum statistical mechanics, based on a correspondence between imaginary time and inverse temperature:

$$\begin{array}{ccc} e^{-itH} & \longleftrightarrow & e^{-\beta H} \\ \text{translation operator} & & \text{density matrix} \end{array} \tag{14.69}$$

On the left, the Hamiltonian H appears as the generator of the Lie group of time translations. On the right, the same operator appears in the Boltzmann weight in a statistical ensemble. The deeper reason for this correspondence remains one of the great mysteries in physics.

For practical use of the correspondence, consider a quantum field at absolute temperature β^{-1}, in $d-1$ spatial dimensions. Denote the field by $\phi(\mathbf{x}, \tau)$, where $(\mathbf{x}, \tau)$ denotes Euclidean space–time coordinates. We use field eigenstates $|\psi\rangle$ as a basis to calculate the partition function (with no external field):

$$Z = \operatorname{Tr} e^{-\beta H} = \int D\psi \langle \psi | e^{-\beta H} | \psi \rangle \tag{14.70}$$

where $\psi(\mathbf{x})$ denotes a time-independent field. The matrix element in the integrand above is a transition amplitude with the same initial and final states. We can write it as a path integral in Euclidean time:

$$\langle \psi | e^{-\beta H} | \psi \rangle = \int_{\substack{\phi(\mathbf{x},\beta)=\psi(\mathbf{x}) \\ \phi(\mathbf{x},0)=\psi(\mathbf{x})}} D\phi \exp\left\{ \int_0^\beta d\tau \int_\Omega d^{d-1}x \mathcal{H}(\phi(\mathbf{x}, \tau)) \right\} \tag{14.71}$$

where $\mathcal{H}(\phi(\mathbf{x}, \tau))$ is the Hamiltonian density and Ω is the volume of the $(d-1)$-dimensional spatial box. To obtain the partition function, we integrate over the initial field $\psi(\mathbf{x})$. This removes the restrictions to particular initial and final fields, but the initial field is still constraint to be equal to the final field. Thus

$$Z = \int_{\phi(\mathbf{x},\beta)=\phi(\mathbf{x},0)} D\phi \exp\left\{ \int_0^\beta d\tau \int_\Omega d^{d-1}x \mathcal{H}(\phi(\mathbf{x}, \tau)) \right\} \tag{14.72}$$

In the limit $\Omega \to \infty$, we recognize this as the generating functional of a Euclidean quantum field theory (at absolute zero) in a flat box of thickness β, with periodic boundary conditions required along the short edge. Along the long edges, which eventually tend to infinite length, we normally impose periodic boundary conditions anyway. The box is depicted in Fig. 14.2. In the limit of infinite temperatures ($\beta \to 0$) the volume flattens to a box in $d-1$ dimensions, and we recover the classical system mentioned earlier. In the limit of absolute zero ($\beta \to \infty$) we have a quantum field theory in d Euclidean dimensions.

In statistical mechanics, there is the *Ginsburg-Landau theory,* in which the partition function has the form

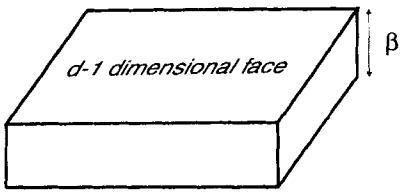


Figure 14.2 A flat box in d dimensions, of thickness β, contains a quantum field with periodic boundary conditions. The partition function in path-integral form describes either a system of space–time dimension $(d-1) \times \beta$ at absolute zero, or a system of spatial dimension $(d-1)$ at temperature β^{-1}.

$$Z = \int D\phi e^{-\beta E[\phi]} \tag{14.73}$$

where $\phi(x)$ is a classical field enclosed in a d-dimensional volume. One imagines that this was obtained by summing over all variables in a microscopic partition sum, while keeping a certain "order parameter" $\phi(x)$ fixed. This order parameter characterizes phase transitions. It vanishes at the critical point of a second-order phase transition, and in the neighborhood of such a point one can expand the energy functional in powers of the order parameter:

$$E[\phi] = \int d^d x \left[\frac{\sigma}{2}(\partial\phi)^2 + \frac{r}{2}\phi + u\phi^4 + \cdots \right] \tag{14.74}$$

where the parameter σ, r, u, $\cdots$ are phenomenological quantities that may depend on the temperature. The functional $E[\phi]$ is called the "Ginsburg-Landau free energy." In this fashion, the theory gives a macroscopic description of phase transitions, yielding phase diagrams and critical exponents. We see that the Ginsburg-Landau theory also gives a quantum field theory. The significance of the critical point will be discuss in more detail in Chapter 16.

14.7 GAUSSIAN INTEGRALS

When it comes to actually performing path integrations, there is only one path integral we can do, and that is the Gaussian integral. It is based on the elementary formula

$$\int_{-\infty}^{\infty} du\, e^{-\lambda u^2} = \sqrt{\frac{\pi}{\lambda}} \tag{14.75}$$

which can be generalized to

$$\int_{-\infty}^{\infty} \frac{du}{\sqrt{2\pi}}\, e^{-(1/2)Ku^2 - Ju} = \frac{e^{J^2/2K}}{\sqrt{K}} \tag{14.76}$$

Now consider a product of n such integrals, and regard the integration variables $u_1, \cdots u_n$ as components of a vector. Then we can write

$$\int Du\, e^{-(1/2)(u,Ku)-(J,u)} = \frac{e^{(1/2)(J,K^{-1}J)}}{\sqrt{\det K}} \tag{14.77}$$

where K is a real symmetric matrix, and J a vector with real components:

$$\begin{aligned}(u, Ku) &= \sum_{\alpha\beta} u_\alpha K_{\alpha\beta} u_\beta \\ (J, u) &= \sum_{\alpha} J_\alpha u_\alpha \end{aligned} \tag{14.78}$$

The measure Du is defined by

$$\int Du = \int_{-\infty}^{\infty} \frac{du_1}{\sqrt{2\pi}} \cdots \int_{-\infty}^{\infty} \frac{du_n}{\sqrt{2\pi}} \tag{14.79}$$

Consider now an average with respect to a Gaussian weight:

$$\langle f \rangle \equiv \frac{\int Du\, e^{-(1/2)(u,Ku)} f(u)}{\int Du\, e^{-(1/2)(u,Ku)}} \tag{14.80}$$

If f can be expanded in powers of its arguments, all we need is the average of a product of the u values. To calculate the latter, introduce an "external field" J, and consider

$$\begin{aligned}\langle u_{\alpha_1} \cdots u_{\alpha_m} \rangle_J &\equiv \frac{\int Du\, e^{-(1/2)(u,Ku)-(J,u)} u_{\alpha 1} \cdots u_{\alpha m}}{\int Du\, e^{-(1/2)(u,Ku)-(J,u)}} \\ &= \int Du\, e^{-(1/2)(u,Ku)-(J,u)} \left[\frac{\partial}{\partial J_{\alpha_1}} \cdots \frac{\partial}{\partial J_{\alpha_m}}\right] \int Du\, e^{-(1/2)(u,Ku)-(J,u)} \\ &= (-1)^m e^{-(1/2)(J,K^{-1}J)} \left[\frac{\partial}{\partial J_{\alpha_1}} \cdots \frac{\partial}{\partial J_{\alpha_m}}\right] e^{(1/2)(J,K^{-1}J)} \end{aligned} \tag{14.81}$$

The average is obtained by setting $J_\alpha = 0$. Then, we get a nonzero answer only if m is even:

$$\langle u_{\alpha 1} \cdots u_{\alpha m} \rangle = \frac{\partial}{\partial J_{\alpha_1}} \cdots \frac{\partial}{\partial J_{\alpha_m}} e^{(1/2)(J,K^{-1}J)} \Big|_{J=0} \tag{14.82}$$

On carrying out the differentiations, we get a sum of terms, but only those free of J terms can survive. Thus, the final result will be the sum of all possible matrix elements $\frac{1}{2}(K^{-1})_{\beta\gamma}$, in which $\{\beta, \gamma\}$ is a distinct pairing of indices among the set $\{\alpha_1, \cdots, \alpha_m\}$. For a symmetric matrix K, we can omit the factor $\frac{1}{2}$ in $\frac{1}{2}(K^{-1})_{\beta\gamma}$ and count the pairs $\{\beta, \gamma\}$ and $\{\gamma, \beta\}$ as the same. Defining a contraction between $u_{\alpha 1}, u_{\alpha 2}$ as

$$\underbrace{u_{\alpha_1} u_{\alpha 2}} \equiv (K^{-1})_{\alpha_1 \alpha_2} \tag{14.83}$$

we have, once again, Wick's theorem for averages:

$$\langle u_{\alpha_1} \cdots u_{\alpha m} \rangle = (\underbrace{u_{\alpha_1} u_{\alpha_2}} u_{\alpha_3} u_{\alpha_4} \cdots) + \cdots$$
$$\text{(sum of all possible pairings)} \tag{14.84}$$

To generalize the preceding calculations to functional integrals, we need to go to the limit where the vector components u_α become a field variable $\phi(x)$, with the discrete label α replaced by the continuous coordinate x. Correspondingly, the matrix element $K_{\alpha\beta}$ becomes a continuous function $K(x, y)$. We can use Dirac notation by regarding $\phi(x)$ as the coordinate representative of a vector in a Hilbert space, and $K(x, y)$ as the coordinate representative of an operator on that space. The basis vectors $|x\rangle$ satisfy

$$\int d^d x |x\rangle \langle x| = 1 \tag{14.85}$$

We write

$$\phi(x) = \langle x|\phi\rangle$$
$$K(x, y) = \langle x|K|y\rangle \tag{14.86}$$

and

$$(J, \phi) \equiv \langle J|\phi\rangle = \int d^d x J(x)\phi(x)$$

$$(\phi, K\eta) \equiv \langle \phi|K|\eta\rangle = \int d^d x\, d^d y \phi(x) K(x, y)\eta(y) \tag{14.87}$$

The Gaussian functional integral is

$$\int D\phi\, e^{-(1/2)(\phi, K\phi)+(J,\phi)} = \frac{e^{1/2(J,K^{-1}J)}}{\sqrt{\det K}} \tag{14.88}$$

where the measure $D\phi$ is defined only up to a multiplicative constant. The determinant det K can be calculated through the relation

$$\ln \det K = \mathrm{Tr} \ln K \tag{14.89}$$

which is easily proved for a finite matrix by going the representation in which K is diagonal.

As an example, consider the kinetic operator of a free field in Euclidean space

$$K = -\partial^2 + m^2 \tag{14.90}$$

where $\partial^2 = \Sigma^d_{i=1} \partial^2/\partial x_i^2$. The matrix element is given by

$$\langle y|K|x\rangle = \delta^d(y-x)\left[-\sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} + m^2\right] \tag{14.91}$$

We introduce momentum basis $|k\rangle$ through

$$\langle x|k\rangle = e^{ik\cdot x} \tag{14.92}$$

with

$$\int \frac{d^d k}{(2\pi)^d} |k\rangle \langle k| = 1 \tag{14.93}$$

The matrix is diagonal in the momentum representation:

$$\begin{aligned}\langle p|K|k\rangle &= \int d^d y\, d^d x \delta^d(y-x)\, e^{-ip\cdot y}\left[-\sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} + m^2\right] e^{ik\cdot x} \\ &= (2\pi)^d \delta^d(p-k)(k^2+m^2)\end{aligned} \tag{14.94}$$

Thus

$$\langle p|K^{-1}|k\rangle = (2\pi)^d \delta^d(p-k)\frac{1}{(k^2+m^2)}$$

$$\langle y|K^{-1}|x\rangle = \int \frac{d^d k}{(2\pi)^d} \frac{e^{i(y-x)\cdot k}}{k^2+m^2} = \Delta_{\mathrm{E}}(y-x) \tag{14.95}$$

The determinant is given through

$$\ln \det K = \mathrm{Tr} \ln K = \int \frac{d^d k}{(2\pi)^d} \ln(k^2+m^2) \tag{14.96}$$

This is divergent, and requires an upper cutoff.

14.8 PERTURBATION THEORY

Consider a scalar field $\phi(x)$ in d–dimensional Euclidean space–time, with Euclidean action

$$S[\phi] = S_0[\phi] + S_1[\phi] \tag{14.97}$$

which consists of a free and an interaction term:

$$S_0[\phi] = \tfrac{1}{2}\int d^dx[\partial^\mu\phi(x)\partial^\mu\phi(x) + m^2\phi^2(x)] \tag{14.98}$$

$$S_1[\phi] = \int d^dx\mathcal{H}_1(x) \tag{14.99}$$

We can rewrite, in a matrix notation,

$$S_0[\phi] = \tfrac{1}{2}(\phi, K\phi) \tag{14.100}$$

where $K = -\partial^2 + m^2$. In the Fourier representation

$$\phi(x) = \int \frac{d^dk}{(2\pi)^4} e^{ik\cdot x}\tilde{\phi}(k) \tag{14.101}$$

we have

$$(\phi, K\phi) = \int \frac{d^dp}{(2\pi)^4}(p^2 + m^2)\tilde{\phi}^*(p)\tilde{\phi}(p) \tag{14.102}$$

The partition function is[1]

$$Z[J] = \int D\phi\, e^{-(1/\hbar)S[\phi]-(1/\hbar)(J,\phi)} \tag{14.103}$$

The generating functional for connected correlation functions is

$$W[J] = -\hbar \ln Z[J] \tag{14.104}$$

The unperturbed system is described by the free partition function

[1]We temporarily restore $\hbar$ for scaling purposes, and as a convenient device to show the classical limit.

$$Z^{(0)}[J] = \int D\phi\, e^{-(1/\hbar)(\phi, K\phi) - (1/\hbar)(J,\phi)} \tag{14.105}$$

Using (14.77), we obtain

$$Z^{(0)}[J] = \mathcal{N}[\det(\hbar\Delta_E)]^{1/2} e^{(J,\Delta_E J)/2\hbar} \tag{14.106}$$

where $\mathcal{N}$ is a normalization constant and Δ_E is the Euclidean propagator:

$$\Delta_E(x) = K^{-1}(x) = \int \frac{d^d p}{(2\pi)^4} \frac{e^{ip\cdot x}}{p^2 + m^2} \tag{14.107}$$

The correlation functions in the unperturbed system can be calculated using Wick's theorem:

$$\langle \phi(x_1) \cdots \phi(x_n) \rangle_0 = \text{sum of all possible pairings} \tag{14.108}$$

where the average $\langle\ \rangle_0$ is taken with respect to $S_0[\phi]$, and "pairing" means a grouping of all members of the set $\{\phi(x_1), \cdots, \phi(x_n)\}$ into distinct pairs, with each pair contributing a factor

$$\underbrace{\phi(x)\phi(y)} = \Delta_E(x - y) \tag{14.109}$$

We now outline different methods to get a perturbation expansion for the partition function. The most straightforward is an expansion in powers of $S_1[\phi]$:

$$\begin{aligned} Z[J] &= \int D\phi\, e^{-(1/\hbar)S_0[\phi] - (1/\hbar)(J,\phi)} \sum_{n=0}^{\infty} \frac{\{-S_1[\phi]/\hbar\}^n}{n!} \\ &= \int D\phi\, e^{-(1/\hbar)S_0[\phi] - (1/\hbar)(J,\phi)} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\hbar^n} \int d^d x_1 \cdots d^d x_n\, \mathcal{H}_1(x_1) \cdots \mathcal{H}_1(x_n) \\ &= Z^{(0)}[0] \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\hbar^n} \int d^d x_1 \cdots d^d x_n \langle e^{-(1/\hbar)(J,\phi)} \mathcal{H}_1(x_1) \cdots \mathcal{H}_1(x_n) \rangle_0 \end{aligned} \tag{14.110}$$

The quantity inside the average $\langle\ \rangle_0$ can be expanded as an infinite series in $\phi(x)$, and the average calculated using Wick's theorem. This gives an expansion in terms of Feynman graphs for $Z[J]$, and we get a sum of connected graphs when we calculate $\ln Z[J]$. The constant $Z^{(0)}[0]$ then becomes an irrelevant additive constant.

An alternative method is to use the "cumulant expansion." In its generic form, this gives an expansion of $\ln \langle e^x \rangle$ in terms of the moments $\langle x^n \rangle$, and works for any definition of the average $\langle\ \rangle$. It is straightforward to write down the expansion:

$$\ln \langle e^x \rangle = \ln \sum_{n=0}^{\infty} \frac{\langle x^n \rangle}{n!}$$

$$= \langle x \rangle + \frac{1}{2}[\langle x^2 \rangle - \langle x \rangle^2] + \frac{1}{6}[\langle x^3 \rangle - 3\langle x \rangle \langle x^2 \rangle + \langle x \rangle^3] + \cdots \tag{14.111}$$

To apply this to the partition function, we write

$$Z[J] = Z^{(0)}[0]\langle e^{-(1/\hbar)S_1[\phi]-(1/\hbar)(J,\phi)} \rangle_0 \tag{14.112}$$

Thus

$$\ln Z[J] = \ln Z^{(0)}[0] + \langle A \rangle + \tfrac{1}{2}[\langle A^2 \rangle_0 - \langle A \rangle_0^2] + \cdots \tag{14.113}$$

where $A = -(1/\hbar)\{S_1[\phi] + (J, \phi)\}$. This will generate connected Feynman graphs for correlation functions.

Another method is to regard the interaction Hamiltonian as a function of $\phi(x)$, and write

$$\mathcal{H}_1(x) = V(\phi(x)) \tag{14.114}$$

even though it may depend on $\partial\phi(x)$. The partition function may be expanded in powers of $S_1[\phi]$, and further developed as follows:

$$Z[J] = \int D\phi \, e^{-(1/\hbar)S_0[\phi]-(1/\hbar)(J,\phi)} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!\hbar^n} \int d^d x_1 \cdots d^d x_n \mathcal{H}_1(x_1) \cdots \mathcal{H}_1(x_n)$$

$$= \int D\phi \, e^{-(1/\hbar)S0[\phi]} \sum_{n=0}^{\infty} \frac{1}{n!\hbar^n}$$

$$\times \int d^d x_1 \cdots d^d x_n \left[V\left(-\hbar \frac{\delta}{\delta J(x_1)}\right) \cdots V\left(-\hbar \frac{\delta}{\delta J(x_n)}\right) \right] e^{(J,\phi)/\hbar} \tag{14.115}$$

or

$$Z[J] = \left[\exp \frac{1}{\hbar} \int d^d x V\left(-\hbar \frac{\delta}{\delta J(x)}\right) \right] Z^{(0)}[J] \tag{14.116}$$

This is an alternative starting point for the Feynman graph expansion.[2]

[2]For an explicit example, see Huang [2].

14.9 THE LOOP EXPANSION

Connected Feynman graphs can be classified according to the number of closed loops in the graph. For any connected Feynman graph, let

$$\begin{aligned} n &= \text{number of vertices} \\ I &= \text{number of internal lines} \\ E &= \text{number of external lines} \end{aligned} \tag{14.117}$$

Each internal line carries an internal momentum that is integrated over. Not all the internal momenta are independent, on account of momentum conservation, but each closed loop can be associated with a loop momentum, and thus the number of loops is equal to the number of independent internal momenta. To find this number, we note that each vertex imposes one condition of momentum conservation, and there is one condition of overall momentum conservation. Thus, the number of independent internal momentum is

$$\ell = I - n + 1 \tag{14.118}$$

which is equal to the number of loops.

Now consider the power of $\hbar$ multiplying a Feynman graph. We can find this number by examination of (14.116). Each vertex comes with a factor $\hbar^{-1}$, since it comes from a factor $\hbar^{-1}V$. Each internal line results from an application of $\hbar\delta/\delta J$ twice to $Z_0[J] = C\exp[(J, \Delta_E J)/2\hbar]$ to bring down a factor $\hbar^{-1}\Delta_E$. Thus it comes with a factor $\hbar$. Thus, when external lines are ignored, a connected graph is proportional to

$$\hbar^{I-n+1} = \hbar^{\ell} \tag{14.119}$$

where a term +1 is added to the exponent because the generating functional is $\hbar$ ln Z. In the classical limit $\hbar \to 0$, we have only *tree graphs*—graphs with no closed loops. The first quantum correction is given by one-loop graphs, and so forth.

As $\hbar \to 0$, we can calculate the partition function

$$Z[J] = \int D\phi\, e^{-(1/\hbar)\{S[\phi]+(J,\phi)\}} \tag{14.120}$$

through a saddle-point expansion. Suppose that the exponent in the integrand has an extremum at $\phi = \phi_0$:

$$\left.\frac{\delta S[\phi]}{\delta\phi(x)}\right|_{\phi=\phi_0} = -J(x)$$

$$\left.\frac{\delta^2 S[\phi]}{\delta\phi(x)\delta\phi(x)}\right|_{\phi=\phi_0} = Q(x, y) \tag{14.121}$$

This is a saddle point, on the assumption that $S[\phi]$ is analytic, and therefore cannot have an absolute maximum or minimum. To carry this out, it is convenient to redefine the variable of integration as $\phi + \phi_0$, and write

$$\begin{aligned} S[\phi + \phi_0] &= S[\phi_0] + \int d^d x \phi(x) \left[\frac{\delta S[\phi]}{\delta\phi(x)}\right]_{\phi=\phi_0} \\ &\quad + \frac{1}{2}\int d^d x d^d y \phi(x)\phi(y) \left[\frac{\delta^2 S[\phi]}{\delta\phi(x)\delta\phi(x)}\right]_{\phi=\phi_0} + \cdots \\ &= S[\phi_0] - (J, \phi) + \frac{1}{2}(\phi, Q\phi) + \cdots \end{aligned} \tag{14.122}$$

Then

$$\begin{aligned} Z[J] &= \int D\phi\, e^{-(1/\hbar)\{S[\phi+\phi_0]+(J,\phi+\phi_0)\}} \\ &= e^{-(1/\hbar)\{S[\phi_0]+(J,\phi_0)\}} \int D\phi\, e^{(1/\hbar)(\phi,Q\phi)+\cdots} \end{aligned} \tag{14.123}$$

with

$$\int D\phi\, e^{(1/\hbar)(\phi,Q\phi)} = \det{}^{-1/2}\frac{Q}{\hbar} \tag{14.124}$$

Thus we have an expansion in powers of $\hbar$, and therefore in the number of loops:

$$\ln Z[J] = -\frac{1}{\hbar}\{S[\phi_0] + (J, \phi_0)\} - \frac{1}{2}\ln\det Q + O(\hbar) \tag{14.125}$$

where we have dropped a J-independent constant. It should be noted that ϕ_0 and Q are functionals of J. In this expansion, the first term comes from tree graphs, the second term originates from a one-loop graph, and the $O(\hbar)$ term represents two or more loop contributions. Applications of this formula can be found in Section 15.7.

14.10 BOSON AND FERMION LOOPS

We have seen in Chapter 11, in particular (11.62), that a closed fermion loop carries an extra minus sign, because fermion operators anticommute. The point is that a closed fermion loop is the result of a contraction of the form

$$\overbrace{\psi_1^\dagger \psi_1 \psi_2^\dagger \psi_2} \tag{14.26}$$

To write this as a product of two fermion propagators, we reverse the order of $\psi_1^\dagger \psi_2$, thus getting a minus sign. Fermions are set apart from bosons just by a minus sign. In the wave function of a many-particle system, the minus sign occurs in the signature of a permutation of two particles. In Feynman graphs, the minus sign occurs in closed loops.

To illustrate this difference, let us compare free boson and fermion theories, coupled to external field in such a manner that all Feynman graphs are one-loop graphs. Consider the classical Lagrangian density

$$\mathcal{L}(x) = \partial^\mu \psi^* \partial_\mu \psi - m^2 \psi^* \psi + W \psi^* \psi \tag{14.127}$$

where $W(x)$ is an external field coupled to a pair of fields. Decomposing the field into real and imaginary parts

$$\psi = \frac{1}{\sqrt{2}}(\psi_1 + i\psi_2)$$

$$\psi^* = \frac{1}{i\sqrt{2}}(\psi_1 - i\psi_2) \tag{14.128}$$

we can write

$$\mathcal{L}(x) = \tfrac{1}{2}[\partial^\mu \psi_1 \partial_\mu \psi_1 + (W - m^2)\psi_1^2] + \tfrac{1}{2}[\partial^\mu \psi_2 \partial_\mu \psi_2 + (W - m^2)\psi_2^2] \tag{14.129}$$

The Euclidean action is given by

$$S[\psi_1, \psi_2, W] = (\psi_1, (K - W)\psi_1) + (\psi_2, (K - W)\psi_2) \tag{14.130}$$

where $K = -\partial^2 + m^2$. To quantize ψ as a boson field, we write the path-integral representation for the generating functional:

$$\begin{aligned} Z_{\text{boson}}[W] &= \int D\psi_1 \, D\psi_2 \, e^{-(\psi_1,(K-W)\psi_1)-(\psi_2,(K-W)\psi_2)} \\ &= \int D\psi^* \, D\psi e^{-(\psi^*,(K-W)\psi)} \\ &= \frac{1}{\det(K - W)} \end{aligned} \tag{14.131}$$

The connected Feynman graphs are generated by

$$\ln Z_{\text{boson}}[W] = -\ln \det(K - W) = -\text{Tr} \ln(K - W)$$

$$= \sum_{n=1}^{\infty} \frac{1}{n} \text{Tr}(K^{-1}W)^n + \text{constant} \tag{14.132}$$

The traces can be calculated in the coordinate representation:

$$\text{Tr}(K^{-1}W) = \int d^dx \, \langle x|K^{-1}W|x\rangle = \int d^dx \, \Delta_E(0)W(x)$$

$$\text{Tr}(K^{-1}W)^2 = \int d^dx d^dy \, \langle x|K^{-1}W|y\rangle \, \langle y|K^{-1}W|x\rangle = \int d^dx d^dy \, \Delta_E(x-y)W(y)\Delta_E(y-x)W(x)$$

$$\vdots$$

$$\text{Tr}(K^{-1}W)^n = \int d^dx_1 \cdots d^dx_n \, \Delta_E(x_1 - x_2)W(x_2)\Delta_E(x_2 - x_3)W(x_3) \cdots \Delta_E(x_n - x1)W(x_1) \tag{14.133}$$

These give the Feynman graphs shown in Fig. 14.3. The factor $1/n$ in (14.132) corresponds to the symmetry number n. If the field ψ is a fermion field, each loop will give a factor -1, and thus the generating functional becomes

$$\ln Z_{\text{fermion}}[W] = -\sum_{n=1}^{\infty} \frac{1}{n} \text{Tr}(K^{-1}W)^n + \text{constant} \tag{14.134}$$

This result would follow if we can redefine the functional integral to give

$$Z_{\text{fermion}}[W] = \det(K - W) \tag{14.135}$$

How to do this is discussed in the next section.

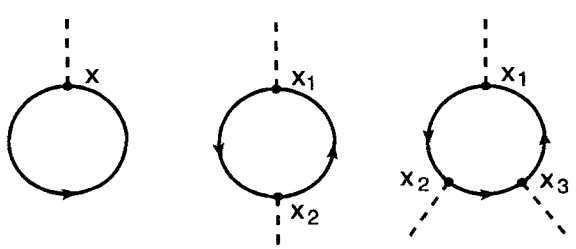


Figure 14.3 One-loop graphs generated by ln det K in a boson theory. For the corresponding fermion theory, each graph is replaced by its negative, and thus det K is replaced by $(\det K)^{-1}$.

14.11 GRASSMANN VARIABLES

We may regard a general Gaussian integral as a way to represent the reciprocal of a determinant. The question is how to represent a determinant instead of its reciprocal. Let us go back to an elementary Gaussian integral. We want to define variables η^* and η, and a sense of integration, such that we can write

$$\int d\eta^* d\eta \, e^{A\eta^*\eta} = A \tag{14.136}$$

Clearly this is impossible if η^* and η are ordinary numbers; but it is possible if they are Grassmann variables, which are anticommuting objects defined in the following fashion.

Consider first two Grassmann variables η_1 and η_2, defined by the relations

$$\begin{aligned} \eta_1^2 &= \eta_2^2 = 0 \\ \{\eta_1, \eta_2\} &= 0 \end{aligned} \tag{14.137}$$

This is different from fermion creation and annihilation operators because the anticommutator is always zero. The most general function of two Grassmann variables must be linear in each:

$$f(\eta_1, \eta_2) = C_0 + C_1\eta_1 + C_2\eta_2 + C_3\eta_1\eta_2 \tag{14.138}$$

where C_i is a complex number. We now define integration to mean

$$\begin{aligned} \int d\eta_1 = 0 \qquad & \int d\eta_2 = 0 \\ \int d\eta_1 \eta_1 = 1 \qquad & \int d\eta_2 \eta_2 = 1 \end{aligned} \tag{14.139}$$

We also define $d\eta_1$ and $d\eta_2$ such that they anticommute with each other, and with η_1 and η_2. Thus

$$\int d\eta_1 \eta_2 = -\eta_2 \int d\eta_1 = 0 \tag{14.140}$$

It follows that

$$\int d\eta_1 f(\eta_1 - \eta_2) = \int (d\eta + d\eta_2) f(\eta) = \int d\eta f(\eta) \tag{14.141}$$

Integrating $f(\eta_1, \eta_2)$ by these rules, we have

$$\int d\eta_1 f(\eta_1, \eta_2) = C_1 + C_3\eta_2$$

$$\int d\eta_2 \int d\eta_1 f(\eta_1, \eta_2) = C_3 \tag{14.142}$$

The exponential function reduces to a bilinear form:

$$e^{A\eta_1\eta_2} = 1 + A\eta_1\eta_2$$

Therefore

$$\int d\eta_2 \int d\eta_1 e^{A\eta_1\eta_2} = A \tag{14.143}$$

To get (14.136), we must therefore take η and η^* to be two independent Grassmann variables.

We can now write a representation for a determinant. Consider a set of Grassmann variables $\{\eta_i\}$, such that

$$\begin{aligned} \eta_i^2 &= 0 \\ \{\eta_i, \eta_j\} &= 0 \\ \int d\eta_i &= 0 \\ \int d\eta_i \eta_i &= 1 \end{aligned} \tag{14.144}$$

Suppose there are an even number $2N$ of variables. We divide them into two sets, labeling one set $\{\eta_\alpha\}$and the other $\{\eta_\alpha^*\}$. The asterisk here serves merely as a distinguishing label, and does not denote conjugation of any sort. We form the quadratic form

$$X = \sum_{\alpha,\beta} \eta_\alpha^* A_{\alpha\beta} \eta_\beta \equiv (\eta^* A \eta) \tag{14.145}$$

where $A_{\alpha\beta}$ is a real symmetric matrix. Through a linear transformation, we can diagonalize A, and obtain

$$X = \sum_{\alpha=1}^{N} A_\alpha \eta_\alpha^* \eta_\alpha \tag{14.146}$$

where A_α are the eigenvalues. Any power of X higher than X^N must vanish, because at least one of the Grassmann variables must appear twice. Therefore

$$e^X = 1 + X + \frac{1}{2}X^2 + \cdots + \frac{1}{N!}X^N \tag{14.147}$$

When we integrate over all the Grassmann variables, only the last term survives, because all other terms must have at least one integral of the form $\int d\eta$, which vanishes. Thus

$$\int D\eta^* D\eta e^X = \frac{1}{N!}\int D\eta^* D\eta\, X^N \tag{14.148}$$

where $D\eta = \prod_\alpha d\eta_\alpha$. The sign of this quantity depends on the order of the factors. Now consider

$$X^N = (A_1\eta_1^*\eta_1 + \cdots + A_N\eta_N^*\eta_N)\cdots(A_1\eta_1^*\eta_1 + \cdots + A_N\eta_N^*\eta_N) \tag{14.149}$$

where there are N factors. In the expansion of this quantity, no η may appear twice in a one term. Therefore a typical term in the expansion is obtained by choosing one term from each factor, which gives $\prod A_\alpha\eta_\alpha^*\eta_\alpha$. The total number of terms in the expansion is the number of ways to choose one term from each factor, or $N!$ Thus

$$X^N = N!\prod_{\alpha=1}^{N} A_\alpha\eta_\alpha^*\eta_\alpha \tag{14.150}$$

Integrating over the Grassmann variables, we obtain

$$\int D\eta^* D\eta e^{(\eta^* A\eta)} = \pm \det A \tag{14.151}$$

where the sign $\pm$ depends on the ordering of variables in $D\eta^* D\eta$. We have thus represented a determinant as a integral of Gaussian form. A more general form is

$$\int D\eta^* D\eta e^{(\eta^* A\eta)+(b^*\eta)+(b\eta^*)} = \pm e^{b^* A^{-1} b}\det A \tag{14.152}$$

where $\{b_a\}$ and $\{b_a^*\}$are sets of Grassmann variables.

The fermion analog of (14.131) is therefore

$$\begin{aligned} Z_{\text{fermion}}[W] &= \int D\psi^*\, D\psi e^{-(\psi^*(K-W)\psi)} \\ &= \det(K - W) \end{aligned} \tag{14.153}$$

where $\{\psi(x)\}$ and $\{\psi^*(x)\}$ are Grassmann fields.

PROBLEMS

14.1 Show that a correlation function in terms of path integration is automatically time-ordered;

$$\langle q'', t''|T[q(t_1)q(t_2)]|q', t'\rangle = \int_{\substack{q(t')=q' \\ q(t'')=q''}} (Dq)q(t_1)q(t_2)\exp\left[\frac{i}{\hbar}\int_{t'}^{t''} L(q(t), \dot{q}(t))\right]$$

14.2 Calculate the transition amplitude $\langle x_2, t_2|x_1, t_1\rangle$ for a nonrelativistic free particle, using the path-integral method.

14.3 Let the partition function be $Z[J] = \int D\phi\, e^{-S[\phi]-(J,\phi)}$.

(a) Illustrate the fact that $Z[J]$ generates correlation functions, by showing

$$\frac{1}{Z[J]}\frac{\delta}{\delta J(x)}\frac{\delta}{\delta J(y)}Z[J] = \langle\phi(x)\phi(y)\rangle$$

(b) Illustrate the fact that $\ln Z[J]$ generates connected correlation functions, by showing

$$\frac{\delta}{\delta J(x)}\frac{\delta}{\delta J(y)}\ln Z[J] = \langle\phi(x)\phi(y)\rangle - \langle\phi(x)\rangle\langle\phi(y)\rangle$$

14.4 Consider a free scalar field whose Euclidean action is represented in the form $S_0[\phi] = (\phi, K\phi)$, where K is the operator $K = -\partial^2 + m^2$.

(a) Calculate the partition function, and from it obtain the correlation function $\langle\phi(x)\phi(y)\rangle$. Show that it is given by

$$\langle\phi(x)\phi(y)\rangle = \int \frac{d^dk}{(2\pi)^4}\frac{e^{ik\cdot(x-y)}}{k^2 + m^2}$$

(b) Continue the result to Minkowski space–time and verify that it is the Feynman propagator $i\Delta_F(x)$.

14.5 Consider the correlation function in Problem 14.4 for $m = 0$.

(a) Show through a scaling argument that

$$\langle\phi(x)\phi(y)\rangle \propto |x|^{2-d}$$

(b) Show for $d = 3$

$$\langle\phi(x)\phi(0)\rangle = \frac{1}{4\pi|x|}$$

(c) Show for $d = 2$

$$\langle\phi(x)\phi(0)\rangle = \frac{1}{2\pi}\ln\frac{|x|}{a}$$

where a is an arbitrary constant.

14.6 For a set of variables $u_\alpha (\alpha = 1, \cdots, n)$, Wick's theorem says that $\langle u_{\alpha_1} \cdots u_{\alpha_m} \rangle$ is the sum of all possible pairings. The theorem is correct, even when some of the factors are equal. To illustrate this, consider the case when they are all equal:

$$\langle u^{2n} \rangle \equiv \frac{\int_{-\infty}^{\infty} du\, e^{-\lambda u 2} u^{2n}}{\int_{-\infty}^{\infty} du\, e^{-\lambda u 2}}$$

(a) By direct computation, show

$$\langle u^2 \rangle = \frac{1}{2}\lambda$$

$$\langle u^4 \rangle = 3\langle u^2 \rangle^2$$

$$\langle u^6 \rangle = 15\langle u^2 \rangle^3$$

$$\vdots$$

$$\langle u^{2n} \rangle = C_n \langle u^2 \rangle^n$$

where $C_n = (2n-1)!/[2^{n-1}(n-1)!]$.

(b) Calculate the same using Wick's theorem. First, $\langle u^2 \rangle$ defines the contraction. To calculate $\langle u^4 \rangle$, note that there are three ways to form pairs among *uuuu*, and each gives $\langle u^2 \rangle^2$. Similarly, there are 15 ways to form pairs among *uuuuuu*, with each giving $\langle u^2 \rangle^3$. Show that C_n is the number of way to form pairs among u^{2n}.

14.7 To illustrate how a simple field theory emerges as an approximation to a more complicated system, consider a two-dimensional flexible membrane in thermal equilibrium with its environment. We describe the thermal fluctuations in terms of the local height $\phi(x, y)$ of the membrane over some reference x–y plane. Ignoring the possibility that the membrane might fold over itself, we take $\phi(x, y)$ to be single-valued, and take the energy to be proportional to the area of the membrane:

$$E[\phi] = \sigma \int dx\, dy \sqrt{g}$$

where σ is the surface tension, and, with $\mathbf{r} \equiv \{x, y, \phi(x, y)\}$,

$$g = \begin{vmatrix} \dfrac{\partial \mathbf{r}}{\partial x} \cdot \dfrac{\partial \mathbf{r}}{\partial x} & \dfrac{\partial \mathbf{r}}{\partial x} \cdot \dfrac{\partial \mathbf{r}}{\partial y} \\ \dfrac{\partial \mathbf{r}}{\partial y} \cdot \dfrac{\partial \mathbf{r}}{\partial x} & \dfrac{\partial \mathbf{r}}{\partial y} \cdot \dfrac{\partial \mathbf{r}}{\partial y} \end{vmatrix}$$

For small ϕ we find

$$g \approx \begin{vmatrix} 1 + \left(\dfrac{\partial \phi}{\partial x}\right)^2 & \left(\dfrac{\partial \phi}{\partial x}\right)\left(\dfrac{\partial \phi}{\partial y}\right) \\ \left(\dfrac{\partial \phi}{\partial y}\right)\left(\dfrac{\partial \phi}{\partial x}\right) & 1 + \left(\dfrac{\partial \phi}{\partial y}\right)^2 \end{vmatrix} = 1 + \left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2$$

Therefore,

$$E[h] \approx \sigma \int dx\, dy \sqrt{1 + |\nabla \phi(x, y)|^2} = C + \frac{\sigma}{2} \int dx\, dy |\nabla \phi(x, y)|^2$$

where C is a constant, and the terms left out are higher-order terms in ϕ and its derivatives.

14.8 Grassmann Variables The Anderson localization problem is the study of the following single-particle Hamiltonian:

$$H = -\frac{\nabla^2}{2m} + V(r)$$

where $V(r)$ is a random potential. It is useful to define the Green function

$$G_{\pm}(r, r') = \sum_{\alpha} \frac{\phi_{\alpha}^*(r)\phi_{\alpha}(r')}{E - E_{\alpha} \pm i\eta}$$

where

$$H\phi_{\alpha}(x) = E_{\alpha}\phi_{\alpha}(x)$$

that is, ϕ_{α} and E_{α} are respectively the exact eigenstate and energy of the random Hamiltonian. To make analytic progress, it is necessary to perform averaging over the random potential, denoted by $\langle\ \rangle_{\text{im}}$ of $G_{\pm}$ and $G_{+}G_{-}$. To illustrate the difficulty of the problem and the technique developed to solve it, let us consider $\langle G_{\pm} \rangle_{\text{im}}$.

(a) Derive the following functional integral representation of $G_{\pm}$:

$$G_{\pm}(r, r') = \frac{\int D\psi D\psi^*\ \psi^*(r)\psi(r') e^{-i\int dy\ \psi^*(y)(E - H \pm i\eta)\psi(y)}}{\int D\psi D\psi^*\ e^{-i\int dy\ \psi^*(y)(E - H \pm i\eta)\psi(y)}}$$

where ψ is a complex scalar field.

(b) It is difficult to compute $\langle G_{\pm} \rangle_{\text{im}}$ because of the denominator. Show that the denominator is $\det\{-i(E - H \pm i\eta)^{-1}\}$ and use the fermion representation to show that

$$G_{\pm}(r, r') = \int D\psi D\psi^*\, D\chi D\chi^*\ \psi^*(r)\psi(r') e^{-\mathcal{L}}$$

where χ is a Grassmann field and

$$\mathcal{L} = i \int dy \{\psi^*(y)(E - H \pm i\eta)\psi(y) + \chi^*(y)(E - H \pm i\eta)\chi(y)\}$$

(c) Assume that $V(r)$ is a Gaussian random variable obeying

$$\langle V \rangle_{\text{im}} = 0$$

$$\langle V(r)V(r') \rangle_{\text{im}} = u(r - r')$$

Using the fact that

$$\langle e^{iA}\rangle_{\text{im}} = \exp[-\tfrac{1}{2}\langle A^2\rangle_{\text{im}}]$$

if A is a Gaussian random variable whose mean is 0, show that

$$\langle G_{\pm}(r, r')\rangle_{\text{im}} = \int D\psi D\psi^* \, D\chi \, D\chi^* \, e^{-L_{\text{eff}}}$$

where

$$L_{\text{eff}} = i\int dy\left\{\psi^*(y)\left(E - \frac{\nabla^2}{2m}\right)\psi(y) + \chi^*(y)\left(E - \frac{\nabla^2}{2m}\right)\chi(y)\right\}$$
$$+ \frac{1}{2}\int dx\, dy[\psi^*(x)\psi^*(y)u(x-y)\psi(y)\psi(x) + \chi^*(x)\chi^*(y)u(x-y)\chi(y)\chi(x)$$
$$+ \chi^*(x)\psi(y)u(x-y)\psi(y)\chi(x) + \psi^*(x)\chi(y)u(x-y)\chi(y)\psi(x)]$$

This Langrangian now describes an interacting fermion–boson system. The random variables have been removed and standard field theory methods can be applied. A similar procedure can be used to compute the more useful quantity $\langle G^+G^-\rangle_{\text{im}}$.

REFERENCES

1. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965).
2. K. Huang, *Quarks, Leptons, and Gauge Fields*, 2nd ed., World Scientific, Singapore, 1992, Section 7.7.

CHAPTER FIFTEEN

Broken Symmetry

15.1 WHY BROKEN SYMMETRY

When the Hamiltonian of a system is invariant under a symmetry operation, but the ground state is not, we say that the symmetry is "spontaneously broken." By applying the symmetry operation to the ground state, we transformed it to a different but equivalent ground state. Thus the ground state is degenerate, and in the case of a continuous symmetry, infinitely degenerate. The actual ground state of the system, of course, can only be one of these degenerate states.

Let the degenerate ground states be denoted $|\alpha\rangle$. In the case of broken continuous symmetry, α is a continuous label, and we can construct a state $|\alpha(x)\rangle$ that coincides with different ground states at different points in space in a continuous manner. This state is orthogonal to the true ground state, and, in the limit of infinitely slow variation of $\alpha(x)$, a state of vanishing excitation energy. This is an intuitive picture of the *Goldstone mode*, an excitation whose energy vanishes in the limit of zero momentum. In a relativistic system with no long-range interactions, one can prove that such a mode corresponds to a zero-mass particle called the *Goldstone boson* (see Problem 15.4). In a ferromagnet, the Goldstone mode corresponds to spin waves, as we shall show later.

Broken symmetry is ubiquitous in nature. For example, a solid has rigidity because the atoms occupy fixed positions, breaking translational invariance. When you kick a stone, no doubt remains in your mind about this quality[1]—as forcefully demonstrated by Dr. Johnson in his famous retort to Bishop Berkeley.[2] In the breaking of translational invariance, the Goldstone mode corresponds to acoustic

[1]We paraphrase Anderson [1].

[2]From *James Boswell's Life of Johnson* [2]:

> After we came out of the Church we stood talking for some time together of Bishop Berkeley's ingenious sophistry to prove the nonexistence of matter, and that every thing in the Universe is merely ideal. I observed that though we are satisfied his doctrine is not true it is impossible to refute it. I never shall forget the alacrity with which Johnson answered, striking his foot with mighty force against a large stone till he rebounded from it "I refute it *thus*."

phonons. Less obvious is the breaking of gauge symmetry, manifested in Bose–Einstein condensation. This appears to be the most prevalent form of symmetry breaking in nature, exhibited in such diverse phenomena as superfluidity, superconductivity, mass for elementary particles, and the inflation of the early universe. Modern pioneers in this subject include Anderson, Nambu and Jona-Lasinio, and Goldstone [3].

It may seem puzzling that broken symmetry can occur at all. Consider a ferromagnet, where rotational symmetry is allegedly broken, through the fact that all the atomic spins point along the same direction. However, rotational invariance implies that all directions of the total spin are equally probable, and its average over a statistical ensemble should therefore be zero. This puzzle is usually resolved by considering the spin density $S(B, \Omega)$ in a system of volume Ω, in the presence of a small external magnetic field B. The point is that the limits $\Omega \to \infty$ and $B \to 0$ do not commute. That the ensemble average yields zero is the statement

$$\lim_{\Omega\to\infty} \lim_{B\to 0} S(B, \Omega) = 0 \tag{15.1}$$

whereas spontaneous magnetization means

$$\lim_{B\to 0} \lim_{\Omega\to\infty} S(B, \Omega) \neq 0 \tag{15.2}$$

Although mathematically useful, this formulation masks the physics.

The physical reason for symmetry breaking is that, during its dynamical evolution, the system gets stuck for a long time in a certain pocket of states. In a ferromagnet, neighboring atoms prefer to have parallel spins, and if thermal agitation is reduced, all the spins would line up. The total spin can still freely rotate in space, but the rotation is very slow, because it requires the cooperative effort of a large number of atoms. When the number of atoms becomes macroscopically large, the total spin is forever stuck in a definite direction. What causes spontaneous symmetry breaking, then, is a breakdown of ergodicity.

A generic example of spontaneous symmetry breaking is the real ϕ^4 theory, with energy functional

$$E[\phi] = \int d^4x \left\{ \frac{1}{2}[\partial\phi(x)]^2 + V(\phi(x)) \right\}$$

$$V(\phi(x)) = \frac{r}{2}\phi^2(x) + \frac{\lambda}{4}\phi^4(x) \tag{15.3}$$

The potential $V(\phi)$ is shown in Fig. 15.1. There is only one minimum at $\phi = 0$ if $r > 0$, but there are two minima at $\phi = \pm\sqrt{-r/\lambda}$ if $r < 0$. From the point of view of statistical mechanics, the two minima $\phi = \pm\sqrt{-r/\lambda}$ are equally probable, and therefore the ensemble average of ϕ must be zero. The time average of ϕ, however, is not necessarily the same as the ensemble average. In the case $r < 0$, the system can go

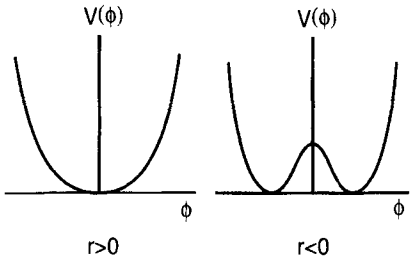


Figure 15.1 The potential has two minima. When the field $\phi(x)$ assumes the value at one of the minima at all x, it takes a long time before one sees a global transition to the other minimum, because fields at different x have to flip simultaneously. This is the underlying cause of spontaneous symmetry breaking.

back and forth between the two minima, but the transition takes a long time, because it requires a simultaneous transition at all x. On the basis of simple models in statistical mechanics, one can estimate this time to be of order e^{Ω}, where Ω is the total volume. Thus, ergodicity fails when $\Omega \to \infty$.

The failure of ergodicity can be demonstrated through a Monte Carlo simulation of the ϕ^4 model on a four-dimensional (4D) Euclidean lattice [4]. As the field configuration evolves in time by successive Monte Carlo updates, we record the instantaneous field $\langle\phi\rangle$, averaged over all lattice sites. This is shown in Fig. 15.2 for different values of r, with $\lambda = 1000$, for a lattice with 10^4 sites. We can see that there is a phase transition at a critical value $r_c \approx -165$. For $r > r_c$, the average field makes small fluctuations about zero. As we approach r_c, the fluctuations become more pronounced, with $\langle\phi\rangle$ making flip-flops between periods of positive and negative values. The time between flip-flops increases as $r \to r_c$, and when $r < r_c$ it becomes unobservably long, presumably of order e^{10^4}. The average field now makes small fluctuations about a nonzero value.

15.2 FERROMAGNETISM

Consider a nonrelativistic gas of magnetic atoms with spin $\frac{1}{2}$, described by the Hamiltonian

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i<j} V(\mathbf{r}_i - \mathbf{r}_j)\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \tag{15.4}$$

where $\boldsymbol{\sigma}_i$ is the 2×2 Pauli spin matrix of the ith atom. For simplicity we choose $V(\mathbf{r}_i - \mathbf{r}_j) = -J\delta^3(\mathbf{r}_i - \mathbf{r}_j)$, with $J > 0$. We thus specialize to short-ranged forces. Going over to a field representation, we take

$$H = \int d^3x \left[-\frac{1}{2m}\psi^\dagger \nabla^2 \psi - \frac{J}{2}(\psi^\dagger \boldsymbol{\sigma} \psi)^2 \right] \tag{15.5}$$

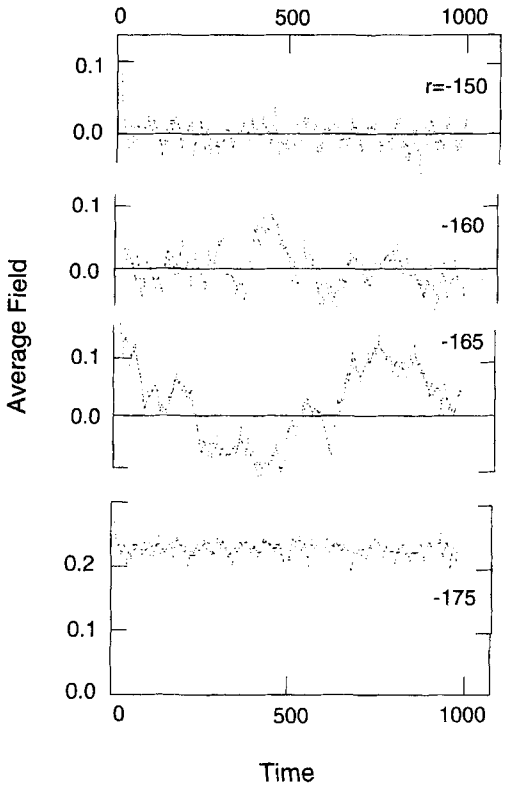


Figure 15.2 Time evolution of the spatial average of the instantaneous field in ϕ^4 theory, in a Monte Carlo simulation, for different values of the squared mass $r = m^2$. A symmetry-breaking phase transition occurs at $r = -165$. [After K. Huang, E. Manousakis, and J. Polonyi, *Phys. Rev.* **35**, 3187 (1987).]

dropping an irrelevant constant term. The use of a δ-function potential makes the ground-state energy divergent. Although this is easy to fix (see Huang [5]), we shall not be concerned with it here. The field operator satisfies fermion anticommutation relations

$$\{\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})\} = \delta_{\alpha\beta}\delta^3(\mathbf{x}-\mathbf{y}) \tag{15.6}$$

where α,β are spin indices. We enclose the system in a box of volume Ω, which eventually tends to infinity. The total number of particles $N = \int d^3x\, \psi^\dagger\psi$ also tends to infinity such that $n = N/\Omega$ is finite.

We use the mean-field approximation, through the replacement $\psi^\dagger\boldsymbol{\sigma}\psi \rightarrow \langle\psi^\dagger\boldsymbol{\sigma}\psi\rangle$, where $\langle\ \rangle$ denotes ground-state expectation. The mean-field Hamiltonian is

$$H_{\mathrm{MF}} = \int d^3x\left[-\frac{1}{2m}\psi^\dagger\nabla^2\psi - J(\psi^\dagger\boldsymbol{\sigma}\psi)\cdot\langle\psi^\dagger\boldsymbol{\sigma}\psi\rangle\right] \tag{15.7}$$

where the factor of $\frac{1}{2}$ in front of J has gone away, because there are two possible ways of making the replacement. Taking the direction of $\langle\psi^\dagger\boldsymbol{\sigma}\psi\rangle$ to be the z axis as, we write $\langle\psi^\dagger\boldsymbol{\sigma}\psi\rangle = \hat{\mathbf{z}}S$, where

$$S = \langle\psi^\dagger\sigma_z\psi\rangle \tag{15.8}$$

Thus

$$H_{\mathrm{MF}} = \int d^3x\ \psi^\dagger\left[-\frac{\nabla^2}{2m} - JS\sigma_z\right]\psi \tag{15.9}$$

The rotational symmetry of the system is spontaneously broken if $S \neq 0$.

The mean-field Hamiltonian describes two free Fermi gases, with up spin and down spin, respectively, for which the single-particle energies are $E_\mp = (p^2/2m) \pm JS$. In the ground state, the two gases must have the same Fermi energy E_{F}, for otherwise we can lower the total energy by transferring particles from one gas to the other. This is illustrated in Fig. 15.3. Denoting the number of particles in the two gases by $N_\pm$, and the total volume of the system by Ω, we have the conditions

$$\frac{N_+ + N_-}{\Omega} = n$$

$$\frac{N_+ - N_-}{\Omega} = S$$

$$\frac{p_+^2}{2m} - JS = \frac{p_-^2}{2m} + JS \equiv E_{\mathrm{F}} \tag{15.10}$$

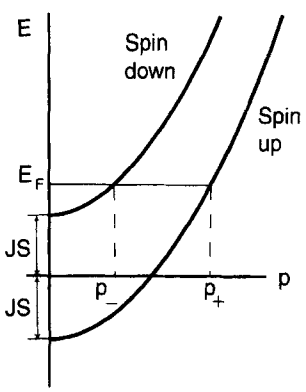


Figure 15.3 Energy-momentum relation for spin-up and spin-down fermions.

where $n = N/\Omega$ is the total density. The Fermi momenta of the two gases are defined by

$$\frac{4\pi}{3}\frac{p_\pm^3}{(2\pi)^3} = \frac{N_\pm}{\Omega} \tag{15.11}$$

in terms of which the conditions become

$$\begin{aligned} p_+^3 + p_+^3 &= 6\pi^2 n \\ p_+^3 - p_+^3 &= 6\pi^2 S \\ p_+^2 - p_-^2 &= 4mJS \end{aligned} \tag{15.12}$$

These equations determine the three unknowns p_+, p_-, S. Adding and subtracting the first two equations yield

$$p_\pm = 3\pi^2(n \pm S)^{1/3} \tag{15.13}$$

Substituting this into the third equation gives the condition for S:

$$(n+S)^{2/3} - (n-S)^{2/3} = \frac{4mJS}{9\pi^4} \tag{15.14}$$

Let

$$r \equiv \frac{S}{n} = \frac{\langle \psi^\dagger \sigma_z \psi \rangle}{\langle \psi^\dagger \psi \rangle} \tag{15.15}$$

Then the condition for r is

$$(1+r)^{2/3} - (1-r)^{2/3} = Jbr \tag{15.16}$$

where $b = (9\pi^4)^{-1} 4mn^{1/3}$. This may be solved graphically, as shown in Fig. 15.4. The symmetric case $r = 0$ is always a solution. Two symmetry-breaking solutions $r = \pm r_0$ arise when $J > J_c$, where

$$bJ_c = \tfrac{4}{3} \tag{15.17}$$

As discussed in Section 15.1, choosing one of these roots leads to spontaneous magnetization. From the form of the mean-field Hamiltonian, it is obvious that a symmetry-breaking solution has lower energy than the symmetric one.

The symmetry-breaking solution exists only if $b > 0$, which means $n > 0$. Thus it is possible in unbounded free space only in the limit $N \to \infty$.

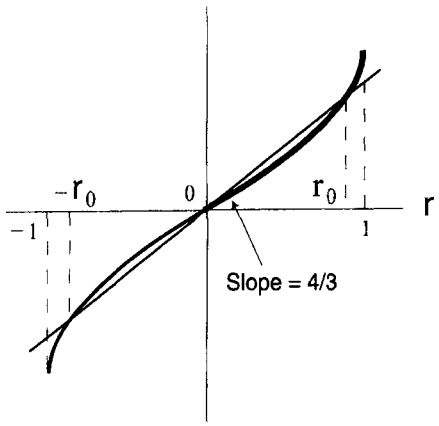


Figure 15.4 Graphical solution for the magnetization density. The horizontal axis is $r = \langle \sigma_z \rangle$. Spontaneous magnetization occurs when the slope of the straight line exceeds $\frac{4}{3}$, which happens when the spin–spin interaction is sufficiently strong.

15.3 SPIN WAVES

Spin waves are the Goldstone modes arising from the breaking of rotational symmetry. To demonstrate its existence, we perturb the system with a small external magnetic field $\delta\mathbf{B}(\mathbf{x}, t)$ transverse to the spontaneous magnetization, and calculate the linear response. The Hamiltonian becomes $H + \delta H(t)$, with

$$\delta H(t) = -\int d^3x[\psi^\dagger(x)\boldsymbol{\sigma}\psi(x)]\cdot\delta\mathbf{B}(x) \tag{15.18}$$

where $\hat{\mathbf{z}}\cdot\delta\mathbf{B} = 0$. The spin density can be represented in the form

$$\mathbf{S}(x) = \langle\psi_\alpha^\dagger(x)\boldsymbol{\sigma}_{\alpha\beta}\psi_\beta(x)\rangle = -i\lim_{y\to x}\mathrm{Tr}[\boldsymbol{\sigma}G(x, y)] \tag{15.19}$$

where $x = (\mathbf{x}, t)$ and G is the propagator:

$$G_{\alpha\beta}(x, y) = -i\langle \mathrm{T}\psi_\alpha(x)\psi_\beta^\dagger(y)\rangle \tag{15.20}$$

With the perturbing field, the propagator becomes $G + \delta G$, and the linear response in the spin density is given by

$$\delta\mathbf{S}(x) = -i\lim_{y\to x}\mathrm{Tr}[\boldsymbol{\sigma}\delta G(x, y)] \tag{15.21}$$

The mean-field Hamiltonian now becomes $H_{\mathrm{MF}} + \delta H_{\mathrm{MF}}(t)$, with

$$\delta H_{\mathrm{MF}}(t) = -\int d^3x\, \psi^\dagger(x)\boldsymbol{\sigma}\cdot[\delta\mathbf{B}(x) + J\,\delta\mathbf{S}(x)]\psi(x) \tag{15.22}$$

where the first term contains the external perturbation and the second terms comes from the induced magnetization. The propagator can be represented by the Feynman graphs

(15.23)

which gives

$$\delta G(x, y) = i\int d^4x'\; G^0(x - x')\boldsymbol{\sigma}\cdot[\delta\mathbf{B}(x') + J\,\delta\mathbf{S}(x')]G^0(x' - y) \tag{15.24}$$

where $G^0_{\alpha\beta}(x)$ is the unperturbed propagator. (See Problem 9.5.) Substituting this into (15.21), we obtain the equation

$$\delta S_i(x) = \int d^4x'\; \Pi_{ij}(x - x')[\delta B_j(x') + J\,\delta S_j(x')] \tag{15.25}$$

where $i = 1, 2$, and

$$\Pi_{ij}(x) = i\mathrm{Tr}[\sigma_i G^0(x)\sigma_j G^0(x)] \tag{15.26}$$

By symmetry in the xy plane, we can put

$$\Pi_{ij}(x) = \delta_{ij}\Pi(x) \tag{15.27}$$

where

$$\Pi(x) = i\mathrm{Tr}[\sigma_1 G^0(x)\sigma_1 G^0(x)] \tag{15.28}$$

Introducing the Fourier transforms

$$\begin{aligned}
\delta\tilde{S}_i(k) &= \int d^4x\; e^{ik\cdot x}\delta S_i(x) \\
\delta\tilde{B}_i(k) &= \int d^4x\; e^{ik\cdot x}\delta B_i(x) \\
\tilde{\Pi}(k) &= \int d^4x\; e^{ik\cdot x}\Pi(x)
\end{aligned} \tag{15.29}$$

where $k = (\mathbf{k}, \omega)$, we can write

$$\delta\tilde{S}_{\perp}(k) = \tilde{\Pi}(k)[\delta\tilde{B}_{\perp}(k) + J\delta\tilde{S}_{\perp}(k)] \tag{15.30}$$

where the subscript $\perp$ denotes a component in the x–y plane. The spin waves are solutions to the homogeneous equation

$$\delta\tilde{S}_{\perp}(k) = J\tilde{\Pi}(k)\delta\tilde{S}_{\perp}(k) \tag{15.31}$$

There is always a trivial solution. A nontrivial solution will indicate the existence of "excited states" in the field-free system. As we shall see below, it is easy to prove

$$\tilde{\Pi}(0) = \frac{1}{J} \tag{15.32}$$

which shows that there exist solutions at $k = 0$. However, these are states of constant density, in both space and time, and thus correspond to other ground states. For this reason, we might call this statement the "ground-state theorem." To show the existence of spin waves, we must prove the "spin-wave theorem"

$$\tilde{\Pi}(k) \xrightarrow[k\to 0]{} 1/J \tag{15.33}$$

This states that there exists a Goldstone mode, a non-uniform state that joins smoothly onto some ground state in the long-wavelength limit. To establish this, all we have to do is show that $\tilde{\Pi}(k)$ is regular at $k = 0$ (see Problem 15.1). An explicit calculation of $\tilde{\Pi}(k)$ yields more information:

$$\tilde{\Pi}(k) \xrightarrow[k\to 0]{} \frac{1}{J} + f(k) \tag{15.34}$$

The condition $f(k) = 0$ then determines a relation between $|\mathbf{k}|$ and ω, the dispersion law for spin waves.

To establish (15.32), turn on a spatially uniform external field $\mathbf{B}$. By rotational symmetry, the magnetization must have the form

$$\mathbf{S} = \frac{\mathbf{B}}{B} f(B^2) \tag{15.35}$$

where $B = |\mathbf{B}|$. Since we are dealing with a ferromagnet, $f(0) \neq 0$. The only nonvanishing Fourier component of $\mathbf{S}$ is that with $\mathbf{k} = 0$, $\omega = 0$:

$$\tilde{\mathbf{S}}(0) = \frac{\tilde{\mathbf{B}}(0)}{B} f(B^2) \tag{15.36}$$

Now rotate $\mathbf{B}$ slightly by letting $\mathbf{B} \to \mathbf{B} + \delta\mathbf{B}$, with $\mathbf{B}\cdot\delta\mathbf{B} = 0$. To first order we have

$$\delta\tilde{S}_{\perp}(0) = \frac{f(B^2)}{B}\delta\tilde{B}_{\perp}(0) \tag{15.37}$$

where the subscript $\perp$ denotes components transverse to the original direction. We can continue to use (15.30):

$$\delta\tilde{S}_{\perp}(0) = \tilde{\Pi}(0)[\delta\tilde{B}_{\perp}(0) + J\delta\tilde{S}_{\perp}(0)] \tag{15.38}$$

where $\tilde{\Pi}(0)$ now depends on B. This gives a second expression:

$$\delta\tilde{S}_{\perp}(0) = \frac{\tilde{\Pi}(0)}{1 - J\tilde{\Pi}(0)}\delta\tilde{B}_{\perp}(0) \tag{15.39}$$

Comparing the two equations for $\delta\tilde{S}_{\perp}(0)$, we obtain

$$\frac{1 - J\tilde{\Pi}(0)}{\tilde{\Pi}(0)} = \frac{B}{f(B^2)} \tag{15.40}$$

Thus, when $B \to 0$, we have $\tilde{\Pi}(0) = 1/J$. ∎

15.4 BREAKING GAUGE INVARIANCE

When Kamerlingh Onnes liquefied helium in 1908, and not long after observed strange behavior in liquid helium [6], and discovered superconductivity [7] (after trying unsuccessfully to fix what was thought to be a short circuit [8]), he observed for the first time broken gauge symmetry.

Both the strange behavior called "superfluidity" in liquid helium, and the superconductivity in certain metals, are manifestations of a Bose–Einstein condensate, formed by a macroscopic number of bosons in a single quantum state. The condensate wave function $\psi(\mathbf{x})$ is a complex number with a definite phase, and the existence of such a phase in the ground state breaks global gauge invariance—a symmetry associated with particle conservation. The ground state of the system is labeled by the phase, and hence infinitely degenerate.

Consider a nonrelativistic boson system described by a field operator $\hat{\Psi}(\mathbf{x})$, which is denoted with a caret to distinguish it from the c-number $\psi(\mathbf{x})$. Let $|N\rangle$ be the ground state with N particles, where N eventually approaches infinity. The macroscopic occupation of a single state is indicated by the fact that the amplitude to annihilate a particle at any point $\mathbf{x}$ is of order $N^{1/2}$:

$$\psi(\mathbf{x}) = \langle N-1|\hat{\Psi}(\mathbf{x})|N\rangle = O(N^{1/2}) \tag{15.41}$$

This amplitude defines the condensate wave function. Yang [9] calls this condition "off-diagonal long-range order" (ODLRO), and proved that it is possible only if

$\hat{\Psi}(\mathbf{x})$ is a boson field, or a product of an even number of fermion fields, but not for a product of an odd number of fermion fields. The distinction between $O(N)$ and $O(1)$ exists only in the limit $N \to \infty$, and this underscores the fact that spontaneous symmetry breaking is a macroscopic phenomenon.

Let us describe the system using states in the grand canonical ensemble, which are not eigenstates of particle number, but an average number N is determined by the chemical potential. Then the condition for ODLRO, or Bose–Einstein condensation, can be stated in the form

$$\psi(\mathbf{x}) = \langle \hat{\Psi}(\mathbf{x}) \rangle = O(N^{1/2}) \tag{15.42}$$

where $\langle\ \rangle$ denotes average with respect to the grand canonical ensemble. This condition was suggested much earlier by Penrose and Onsager [10], but we deviate from historical order in the interest of pedagogy. The condensate density is defined by

$$n_0 = \frac{1}{\Omega} \int d^3x |\psi(\mathbf{x})|^2 \tag{15.43}$$

where Ω is the volume of the system. This should remain finite in the thermodynamic limit $\Omega \to \infty$, $N \to \infty$, at fixed density $N/\Omega = n$.

To study the condensate in more detail, take the Hamiltonian to be

$$\hat{H} = \int d^3x \left[-\frac{1}{2m} \hat{\Psi}^\dagger \nabla^2 \hat{\Psi} + \frac{g}{2} (\hat{\Psi}^\dagger \hat{\Psi})^2 \right] \tag{15.44}$$

with

$$[\hat{\Psi}(\mathbf{x}), \hat{\Psi}^\dagger(\mathbf{y})] = \delta^3(\mathbf{x} - \mathbf{y}) \tag{15.45}$$

The particles interact through a δ-function potential, which reproduces low-energy scattering if

$$g = \frac{4\pi a}{m} \tag{15.46}$$

where a is the S-wave scattering length.[3] Clearly $\hat{H}$ is invariant under the global gauge transformation $\hat{\Psi}(\mathbf{x}) \to e^{i\alpha} \hat{\Psi}(\mathbf{x})$. The Heisenberg equation of motion reads

$$-\frac{1}{2m} \nabla^2 \hat{\Psi} + g \hat{\Psi}^\dagger \hat{\Psi} \hat{\Psi} = i \frac{\partial \hat{\Psi}}{\partial t} \tag{15.47}$$

Now put

[3]For derivation, see Huang [11].

$$\hat{\Psi}(\mathbf{x}, t) = \psi(\mathbf{x}, t) + \hat{\Phi}(\mathbf{x}, t) \tag{15.48}$$

where $\psi(\mathbf{x}, t) = \langle \hat{\Psi}(\mathbf{x}, t) \rangle$ is a c-number function and

$$\langle \hat{\Phi}(\mathbf{x}, t) \rangle = 0 \tag{15.49}$$

The operator $\hat{\Phi}(\mathbf{x}, t)$ annihilates a particle not in the condensate. Substituting (15.48) into the equation of motion, and taking the grand canonical average, we obtain

$$-\frac{1}{2m} \nabla^2 \psi + g\psi^* \psi^2 + g[2\psi \langle \hat{\Phi}^\dagger \hat{\Phi} \rangle + \psi^* \langle \hat{\Phi}^2 \rangle + \langle \hat{\Phi}^\dagger \hat{\Phi}^2 \rangle] = i \frac{\partial \psi}{\partial t} \tag{15.50}$$

Assuming that almost all particles are in the condensate, we neglect the terms in square brackets and obtain an equation for the condensate wave function:

$$-\frac{1}{2m} \nabla^2 \psi + g\psi^* \psi^2 = i \frac{\partial \psi}{\partial t} \tag{15.51}$$

This is called the *Gross–Pitaeveskii equation* [12]. The same equation occurs in optics and plasma physics, where the custom is to call it the "nonlinear Schrödinger equation [13]."

For a static solution, put

$$\psi(\mathbf{x}, t) = \varphi(\mathbf{x}) e^{-i\epsilon_0 t} \tag{15.52}$$

so that

$$-\frac{1}{2m} \nabla^2 \varphi - \epsilon_0 \varphi + g\varphi^* \varphi^2 = 0 \tag{15.53}$$

The spatially uniform solution is

$$\varphi = e^{i\alpha} \sqrt{\frac{\epsilon_0}{g}} \tag{15.54}$$

where α is arbitrary. The normalization condition (15.43) requires $|\varphi|^2 = n_0$, the condensate density. In this approximation $n_0 \approx n$, the total density. Thus, $\epsilon_0 = gn$. In physical units

$$\epsilon_0 = \frac{4\pi \hbar^2 n a}{m} \tag{15.55}$$

This gives the energy per particle in the condensate.

There is an uncertainty relation between the phase θ of a many-body wave function and the number of particles (see Problem 15.2):

$$\Delta\theta\,\Delta N \geq \tfrac{1}{2} \tag{15.56}$$

Thus an isolated system with a definite number of particles does not have a definite phase. If two such systems come into contact, and can exchange particles, then the relative phase becomes definite. Such a relative phase has been observed between two Bose condensates of sodium atoms, by Ketterle and his team at MIT [14], 89 years after Kamerlingh Onnes created a condensate in the laboratory. The interference fringes can be seen in the photograph in Fig. 15.5.

15.5 SUPERFLUIDITY

In quantum mechanics, the gradient of the phase of a wave function is a particle current density. Here, since the wave function is macroscopically occupied, the corresponding current density is a hydrodynamic quantity that describes superfluidity, the frictionless transport of particles. From the Gross–Pitaevskii equation, we have the conservation law

$$\nabla\cdot\mathbf{j} + \frac{\partial\rho}{\partial t} = 0 \tag{15.57}$$

where

$$\mathbf{j} = \frac{1}{2m} i[\psi^*\nabla\psi - \psi\nabla\psi^*]$$

$$\rho = \psi^*\psi \tag{15.58}$$

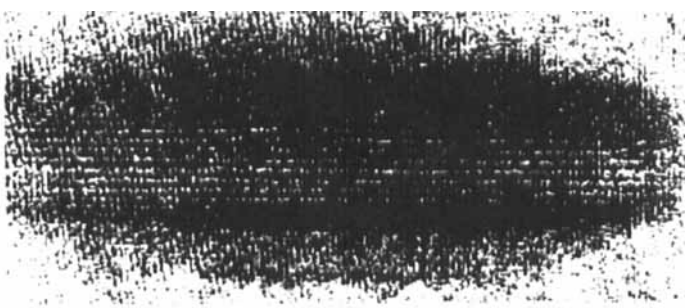


Figure 15.5 Interference fringes between two Bose–Einstein condensates of Sodium atoms observed by Ketterle and his MIT team. This shows the phase coherence of each condensate.. The fringe period is 15 μm, which is the de Broglie wavelength h/mv of the atoms, where v is the relative velocity of the two condensates as they approach each other. [Photograph courtesy W. Ketterle.]

In the ground state the entire system can be regarded as a superfluid at rest. When boundaries conditions are changed slowly, the adiabatic theorem tell us that the ground state changes slowly in response, and this leads to a frictionless superfluid flow described by $\mathbf{j}$ and ρ. Friction will occur as soon as the system can be excited from the ground state into a mode that lacks phase coherence. In the neighborhood of the ground state, there is the Goldstone mode, which corresponds to long-wavelength phonons [15], but they maintain the phase coherence. Vortex excitations, on the other hand, do disrupt phase coherence, and destroy superfluidity.

To investigate vortices, let us put

$$\psi(\mathbf{x}) = e^{i\theta(\mathbf{x})}\sqrt{\rho(\mathbf{x})} \tag{15.59}$$

Then the current can be expressed as

$$\mathbf{j} = \rho \mathbf{v}_s \tag{15.60}$$

where

$$\mathbf{v}_s = \frac{1}{m}\nabla\theta \tag{15.61}$$

is the "superfluid velocity." [In physical units, $\mathbf{v}_s = (\hbar/m)\nabla\theta$.] In order that ψ be single-valued, the phase angle θ can change only by a multiple of 2π over a closed path C in space. Therefore, the circulation of the superfluid velocity is quantized:

$$\oint_C d\mathbf{s}\cdot\mathbf{v}_s = \frac{2\pi n}{m} \tag{15.62}$$

where $n = 0, \pm 1, \ldots$. If $n \neq 0$, then C must encircle a line on which $\rho = 0$, for otherwise we could shrink C to a point, at which the phase angle θ becomes undefined. This line of zeros is the core of a vortex with quantized vorticity, and can terminate only on the boundary of the system, or terminate on itself by forming a closed curve.

To describe a vortex with core along the z axis, let us use cylindrical coordinates (r, φ, z), and put

$$\psi(r, \varphi) = \sqrt{\frac{\mu}{g}} f(r) e^{in\varphi} \tag{15.63}$$

with the boundary conditions

$$f(r) \xrightarrow[r\to\infty]{} 1 \qquad f(r) \xrightarrow[r\to 0]{} 0 \tag{15.64}$$

It is easily verified that this corresponds to a linear vortex with quantum number n.

We can rewrite the Gross–Pitaevskii equation is dimensionless form by introducing

$$s = \frac{r}{\xi}$$

$$\xi = \frac{1}{\sqrt{2m\mu}} \tag{15.65}$$

where ξ is the correlation length. Then f satisfies

$$s^2 \frac{d^2 f}{ds^2} + s\frac{df}{ds} + (s^2 - n^2)f - s^2 f^3 = 0 \tag{15.66}$$

which can be solved numerically [16], with the result for $n = 1$ shown in Fig. 15.6. The asymptotic behaviors are given by

$$f \approx 1 - \frac{n^2}{s^2} \qquad (s \gg 1)$$

$$f \approx C s^{-|n|} \qquad (s \ll 1) \tag{15.67}$$

where C is a constant. It is found that the energy is proportional to n^2. Thus vortices with $|n| > 1$ need not be considered, for they will break up into vortices with $|n| = 1$ to lower the energy. The linear vortex described has macroscopic energy, since it has a linear core whose dimension is that of the box containing the system. Finite-ener-

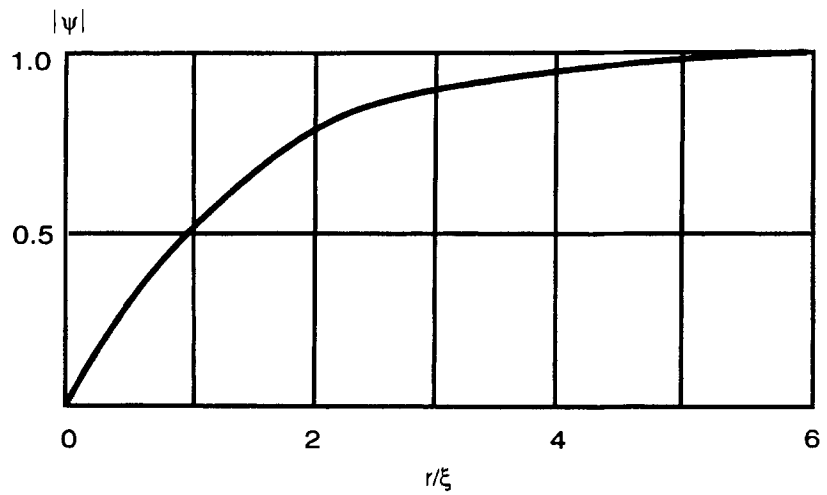


Figure 15.6 Vortex solution to Gross–Pitaevskii equation with one quantum of circulation. The correlation length is $\xi = \hbar/\sqrt{2m\mu}$, where μ is the chemical potential.

gy solutions are vortex rings—those in which the vortex core forms a closed curve. The solution above gives an approximate picture of the flow near the core.

The vortex core is devoid of superfluid and filled with particles not in the condensate (the ones whose field operator is $\hat{\Phi}$), whose effects were neglected in the Gross–Pitaevskii equation. The core renders the space nonsimply connected, and makes it possible to have $\oint_C d\mathbf{x}\cdot\mathbf{v}_s \neq 0$, even though $\mathbf{v}_s$ is the gradient of a function.

As a simplified model, we assume that the density of the superfluid is constant, except that it vanishes inside vortex cores, taken to be tubes of radius a, with energy per unit length υ. The system is then described by the energy functional

$$E[\mathbf{v}_s] = \frac{\rho_0}{2}\int d^3x\, \mathbf{v}_s^2 + \upsilon L \tag{15.68}$$

where L is the total length of vortex cores in the system. The integral $\int d^3x$ extends over the space outside vortex cores only. The superfluid velocity $\mathbf{v}_s$ is not irrotational, and may be decomposed it into longitudinal and transverse parts:

$$\begin{aligned}\mathbf{v}_s &= \mathbf{v}_L + \mathbf{v}_T\\ \mathbf{v}_L &= \nabla\omega\\ \mathbf{v}_T &= \nabla\times\mathbf{A}\end{aligned} \tag{15.69}$$

where $\nabla\omega$ describes irrotational flow, while $\nabla\times\mathbf{A}$ arises from vortices. Substituting this into (15.62) yields

$$\oint_C d\mathbf{s}\cdot\nabla\times\mathbf{A} = 2\pi n \tag{15.70}$$

(the mass is $m = 1$). Thus, $\mathbf{A}$ can be likened to the vector potential set up by steady currents flowing along the vortex cores. The energy functional can now be written as

$$E[\mathbf{v}_s] = \frac{\rho_0}{2}\int d^3x[|\nabla\omega|^2 + |\nabla\times\mathbf{A}|^2] + \upsilon L[\mathbf{A}] \tag{15.71}$$

At very low temperatures, vortices are suppressed because the cores cost energy. The low-energy excitations are longitudinal phonons, corresponding to the Goldstone mode.

The superfluid density ρ_s is a transport coefficient that describes the response of the system to an imposed motion. It measures that part of the system that does not respond to a shear force. This analysis is described in Appendix B, and we shall just make use of the result here:

$$\rho_s = \rho_0 - \frac{\beta\rho_0^2}{D-1}\int d^Dx\,\langle\mathbf{v}_T(x)\cdot\mathbf{v}_T(0)\rangle \tag{15.72}$$

where D is the dimension of space, and β the inverse temperature. The thermal average indicated by $\langle\rangle$ is weighted by the energy functional (15.71). This indicates that the superfluid density differs from ρ_0 as a result of vortex activity represented by $\mathbf{v}_T$.

Putting $D = 3$, and using as integration variable $\mathbf{w} = \mathbf{v}_T\sqrt{\beta\rho_0}$, we have

$$\frac{\rho_s}{\rho_0} = 1 - \frac{1}{2}\int d^3x \, \frac{\int D\mathbf{w}\mathbf{w}(\mathbf{x})\cdot\mathbf{w}(0)e^{-S[\mathbf{w}]}}{\int D\mathbf{w}\, e^{-S[\mathbf{w}]}} \tag{15.73}$$

where

$$S[\mathbf{w}] = \frac{1}{2}\int d^3y \, \mathbf{w}^2(\mathbf{y}) - \beta v L[\mathbf{w}] \tag{15.74}$$

where $L[\mathbf{w}]$ is the total length of vortex cores. The integration $\int D\mathbf{w}$ extends over the space of all possible vortex cores, which form arbitrary loops in space. This poses an insurmountable problem.

In contrast, the vortex cores are pointlike in 2D, and the superfluid density can be calculated *exactly*, for a dilute gas of vortices. As we will show in Chapter 18, this gives a complete understanding of the phase transition associated with 2D superfluidity.

15.6 GINSBURG–LANDAU THEORY

A generic model of gauge symmetry breaking is represented by the Ginsburg–Landau free energy

$$S[\psi, \psi^*] = \int d^dx[\partial^i\psi^* \, \partial^i\psi + u_2\psi^*\psi + u_4(\psi^*\psi)^2)] \tag{15.75}$$

where $\psi(x)$ is a complex order parameter. This is also the Euclidean action of a complex field. Writing

$$\psi = \frac{\phi_1 + i\phi_2}{\sqrt{2}} \qquad \psi^* = \frac{\phi_1 - i\phi_2}{\sqrt{2}} \tag{15.76}$$

we have

$$S[\phi_1, \phi_2] = \int d^dx[\tfrac{1}{2}(\partial\phi_1)^2 + \tfrac{1}{2}(\partial\phi_2)^2 + V(\phi(x))] \tag{15.77}$$

where

$$V(\psi) = \tfrac{1}{2}u_2R^2 + \tfrac{1}{4}u_4R^4 \tag{15.78}$$

with $R^2 = \Sigma_\alpha \phi_\alpha^2$. There is global gauge invariance associated with a constant phase change of ψ, or a rotation in ϕ_1–ϕ_2 space. This symmetry is spontaneously broken if $\langle\psi\rangle \neq 0$ in the infinite-volume limit.

In Fig. 15.7, the potential is plotted over the ϕ_1–ϕ_2 plane for $u_2 < 0$. It has a wine-bottle shape, which makes $|\psi| > 0$ in the ground state. In contrast to Fig. 15.1, where the potential breaks a discrete symmetry, the potential minimum here gives a continuous set of degenerate ground states.

The condition $|\psi| > 0$ is necessary but not sufficient for spontaneous breaking, which requires $\langle\psi\rangle \neq 0$. For the latter to be true, the fluctuations of the phase $\theta(x)$ of the field have to be small. Let us put

$$\psi(x) = \sqrt{\rho(x)}e^{i\theta(x)} \tag{15.79}$$

and rewrite the action in the form

$$S[\rho, \theta] = \int d^d x \left[\frac{1}{2}\rho(\nabla\theta)^2 + \frac{1}{8\rho}(\nabla\rho)^2 + V(\rho) \right] \tag{15.80}$$

As a indication of the fluctuations of θ, consider $\langle\theta(x)\theta(0)\rangle$ for fixed ρ. This gives the probable value of $\theta(x)$, when $\theta(0)$ is specified. The result can be obtained from Problem 14.5:

$$\langle\theta(x)\theta(0)\rangle \propto \begin{cases} |x|^{2-d} & (d \geq 3) \\ \ln|x| & (d = 2) \end{cases} \tag{15.81}$$

For $d \geq 3$, the fluctuations are bounded, and we expect $\langle\psi\rangle \neq 0$. For $d = 2$, they diverge at large x, indicating that the phase angle becomes random when taken modulo 2π. Hence we expect $\langle\psi\rangle = 0$. This case will be studied separately in Chapter 18.

The Ginsburg–Landau model can be used in different physical problems:

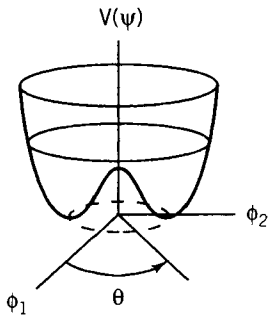


Figure 15.7 The "wine-bottle" potential makes $|\psi| > 0$ in the ground state. But gauge symmetry is broken only if $\langle\psi\rangle \neq 0$, which means that, in addition, the fluctuations of the phase θ must be small. This condition is fulfilled for $d \geq 3$, but not for $d = 2$.

- For $d = 3$, the model reduces to the Gross–Pitaevskii equation. We can obtain the vortex model of superfluidity by choosing u_2 and u_4 in such a manner as to trap the order parameter at a steep minimum. That is, we fix the value at $|\psi| = -u_2/u_4$, with $V' = 0$, $V'' \gg 1$. Then we introduce the vortex core as a cutoff, such that $\psi = 0$ inside.
- For $d = 4 - \epsilon$, we can model the λ transition in liquid helium at temperature T_c by choosing $u_2 = \alpha(T - T_c)$, where α is a constant and T is the absolute temperature. The critical exponents can be calculated as a expansion in powers of ϵ, and we extrapolate the results to $\epsilon = 1$. The reason we cannot do the calculations directly in $d = 3$ is that there are infrared divergences.
- For $d = 4$, we have a quantum field theory with broken symmetry. This is used in the standard model of particles, grand unified models, and the theory of the inflationary universe.

15.7 EFFECTIVE ACTION

We assume $d \geq 3$, with $\langle\psi\rangle \neq 0$. The value of $\langle\psi\rangle$ is not exactly at the minimum of the potential $V(\psi)$, because of fluctuations. We use real components ϕ, and define

$$e^{-W[J]} \equiv Z[J] = \int D\phi\, e^{-S[\phi]-(J,\phi)} \tag{15.82}$$

Then, in the presence of external source $J(x)$, the exact average field is given by

$$\eta_\alpha(x) \equiv \langle\phi_\alpha(x)\rangle = \frac{\delta W[J]}{\delta J_\alpha(x)} \tag{15.83}$$

Now we ask, "Given $\eta(x)$, what is the corresponding source function $J(x)$?" To address this question, we make the Legendre transformation

$$\Gamma[\eta] = W[J] - (J, \eta) \tag{15.84}$$

to use η as independent variable. The quantity $\Gamma[\eta]$ is called the *effective action.* Taking the functional derivative of (15.84), we have

$$\begin{aligned}
\frac{\delta\Gamma[\eta]}{\delta\eta_\alpha(x)} &= \frac{\delta W[J]}{\delta\eta_\alpha(x)} - \frac{\delta(J,\eta)}{\delta\eta_\alpha(x)} \\
&= \int d^dy\, \frac{\delta W[J]}{\delta J_\beta(y)} \frac{\delta J_\beta(y)}{\delta\eta_\alpha(x)} - \left[J_\alpha(x) + \int d^dy \frac{\delta J_\beta(y)}{\delta\eta_\alpha(x)}\eta_\beta(y)\right] \\
&= \int d^dy\, \eta_\beta(y)\frac{\delta J_\beta(y)}{\delta\eta_\alpha(x)} - \left[J_\alpha(x) + \int d^dy \frac{\delta J_\beta(y)}{\delta\eta_\alpha(x)}\eta_\beta(y)\right] = -J_\alpha(x)
\end{aligned} \tag{15.85}$$

Thus

$$\frac{\delta\Gamma[\eta]}{\delta\eta_\alpha(x)} = -J_\alpha(x) \tag{15.86}$$

The average field $\langle\phi_\alpha\rangle$ can be obtained by setting $J = 0$:

$$\left.\frac{\delta\Gamma[\eta]}{\delta\eta_\alpha(x)}\right|_{\eta_\alpha=\langle\phi_\alpha\rangle} = 0 \tag{15.87}$$

We can expand the effective action in terms of the average field η:

$$\Gamma[\eta] = \sum_{n=0}^{\infty} \frac{1}{n!} \int d^d x_1 \cdots d^d x_n \Gamma_n(x_1, \ldots, x_n)\eta(x_1) \cdots \eta(x_n) \tag{15.88}$$

The function $\Gamma_n(x_1, \ldots, x_n)$ is the sum of all one-particle irreducible Feynman graphs.

Proof. Consider the generating functional of a fictitious field theory whose action is $\Gamma[\eta]$:

$$e^{-U[J,a]} \equiv \int D\phi\, e^{-\{\Gamma[\phi]+(J,\phi)\}/a} \tag{15.89}$$

We put in the parameter a for mathematical purposes. As $a \to 0$, the integral is dominated by the saddle point of the exponent. The saddle-point condition is precisely (15.86), and thus

$$\lim_{a\to 0} aU[J, a] = \Gamma[\eta] + (J, \eta) \tag{15.90}$$

By (15.84), this states

$$\lim_{a\to 0} aU[J, a] = W[J] \tag{15.91}$$

The right side is the sum of all connected Feynman graphs of the original field theory. The left side is the classical limit of the generating functional of the fictitious field theory, which is given by the sum of all connected tree graphs of that theory. From (15.88), we see that the fictitious field theory is nonlocal, and the vertices in the tree graphs are the "blobs" $\Gamma_n(x_1, \ldots, x_n)$. We can represent the connected graphs in $W[J]$ as tree graphs with "blobs" for vertices, where each blob is a sum of all one-particle irreducible graphs with the appropriate number of external lines. Therefore, $\Gamma_n(x_1, \ldots, x_n)$ is the sum of all one-particle irreducible graphs of the original field theory. ∎

15.8 EFFECTIVE POTENTIAL

We choose a constant source function, so that the average field will be a constant $\eta(x) = v$. The effective action per unit volume will be a function of v, which we define as the *effective potential*:

$$V_{\text{eff}}(v) = \Omega^{-1}\Gamma[v] \tag{15.92}$$

where Ω is the volume of the d–dimensional Euclidean space. The minimum of V_{eff} gives the exact average field.

We now calculate the effective action to one-loop order. According to the method described in Section 14.9, we first make a saddle-point expansion of the partition function (15.82). At the saddle point $\phi(x) = \phi_0(x)$, we have

$$\begin{aligned} \frac{\delta S[\phi]}{\delta \phi_\alpha(x)}\bigg|_{\phi=\phi_0} &= -J_\alpha(x) \\ \frac{\delta^2 S[\phi]}{\delta\phi_\alpha(x)\delta\phi_\beta(y)}\bigg|_{\phi=\phi_0} &\equiv Q_{\alpha\beta}(x, y) \end{aligned} \tag{15.93}$$

Expanding the action about the saddle point to second order, we have, in an abbreviated notation,

$$S[\phi + \phi_0] = S[\phi_0] - (J, \phi) + \tfrac{1}{2}(\phi, Q\phi) \tag{15.94}$$

where Q is a functional of ϕ_0. Substituting this into (15.82) and performing the Gaussian integration, we obtain the partition function to one-loop order:

$$W[J] = S[\phi_0] + (J, \phi_0) + \tfrac{1}{2} \ln \det Q \tag{15.95}$$

The one-loop effective action is given by

$$\begin{aligned} \Gamma[\eta] &= W[J] - (J, \eta) \\ &= S[\phi_0] + (J, (\phi_0 - \eta)) + \tfrac{1}{2} \ln \det Q \end{aligned} \tag{15.96}$$

We still have to express the saddle point ϕ_0 in terms of the average field η. Treating $\phi_0 - \eta$ as a small quantity, we write

$$\begin{aligned} S[\phi_0] &= S[\eta] + \int d^d x[\phi_{0\alpha}(x) - \eta_\alpha(x)]\frac{\delta S[\psi]}{\delta\phi_\alpha(x)}\bigg|_{\phi=\phi_0} \\ &= S[\eta] - (J, (\phi_0 - \eta)) \end{aligned} \tag{15.97}$$

where we have used (15.93). Therefore, to one-loop order,

$$\Gamma[\eta] = S[\eta] + \tfrac{1}{2}\ln\det Q[\eta] \tag{15.98}$$

Now we use the explicit form of the action (18.14). With $\phi_\alpha(x) = v_\alpha + \zeta_\alpha(x)$, where v is constant, and $\zeta(x)$ small, the action to second order in ζ is

$$\begin{aligned} S[\phi] &= \int d^dx[\tfrac{1}{2}(\partial\zeta)^2 + V(v) + \zeta V'(v) + \tfrac{1}{2}\zeta^2 V''(v)] \\ &= \Omega V(v) + \int d^dx\{\zeta V'(v) + \tfrac{1}{2}\zeta[-\partial^2 + V''(v)]\zeta\} \end{aligned} \tag{15.99}$$

The saddle-point properties (15.93) are

$$\begin{aligned} V'(v) &= -J \\ Q_{\alpha\beta}(x, y) &= \delta_{\alpha\beta}[-\partial^2 + V''(v)]\delta^d(x - y) \end{aligned} \tag{15.100}$$

Therefore

$$\frac{1}{2}\ln\det Q = \frac{1}{2}\mathrm{Tr}\ln Q = \Omega\int\frac{d^dk}{(2\pi)^d}\ln[k^2 + V''(v)] \tag{15.101}$$

which gives[4]

$$V_{\text{eff}}(v) = V(v) + \int\frac{d^dk}{(2\pi)^d}\ln[k^2 + V''(v)] \tag{15.102}$$

PROBLEMS

15.1 Spin-Wave Theorem Consider the problem of ferromagnetism in the mean-field approximation.

(a) By examining the equation for $\tilde{\Pi}(k)$, show that it is regular at $k = 0$, and therefore $\tilde{\Pi}(k) \xrightarrow[k\to 0]{} \tilde{\Pi}(0)$.

(b) Verify $\tilde{\Pi}(0) = 1/J$ by direct calculation.

(c) Show that the dispersion law for spin waves is of the form $\omega \propto |\mathbf{k}|^2$.

15.2 Number-Phase Uncertainty Relation This problem addresses the question whether a Hermitian operator for the phase exists. For a boson field, write

$$\psi(\mathbf{x}) = \sqrt{\rho(\mathbf{x})}e^{i\theta(\mathbf{x})}$$

[4]This was first derived by Coleman and Weinberg [17] and Jackiw [18]. For an application of this formula in the Weinberg-Salam model, see Huang [19].

(a) Suppose that $\psi(\mathbf{x})$ is a classical field. Show that the following transformation preserves the measure for functional integration:

$$D\rho\, D\theta = D\psi\, D\psi^* \equiv D(\text{Re } \psi)D(\text{Im } \psi)$$

This result suggest that ρ and θ are canonically conjugate in quantum mechanics.

(b) Quantize $\psi(\mathbf{x})$ through the equal-time commutator $[\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \delta^3(\mathbf{x} - \mathbf{y})$. Assuming that $\rho(\mathbf{x})$ and $\theta(\mathbf{x})$ are Hermitian operators, show that they are canonically conjugate, by showing

$$[\theta(\mathbf{x}), \rho(\mathbf{y})] = -i\delta^3(\mathbf{x} - \mathbf{y})$$

(c) Let $\theta = \theta(\mathbf{x}_0)$ be the phase operator at some point $\mathbf{x}_0$, and let $N = \int d^3x\, \rho(\mathbf{x})$ be the number operator. Assuming the validity of the commutator derived in part (b), show the uncertainty relation

$$\Delta\theta\, \Delta N \geq \tfrac{1}{2}$$

(d) Let $N|n\rangle = n|n\rangle$, where $n = 0, 1, 2 \ldots$ (see Louisell [20]). Using the commutator derived in part (b), show that

$$\langle n|\theta|n'\rangle = \frac{\delta_{nn'}}{n - n'}$$

which is meaningless. Hence there does not exist a Hermitian phase operator θ.

(e) Show, however, that the matrix element above is meaningful when n and n' are large enough to be considered continuous. It then becomes the derivative of a delta function. The matrix element is analogous to $\langle p|x|p'\rangle$, where $[p, x] = -i$. This shows that the phase can be represented by a Hermitian operator if the system contains a large number of particles.

15.3 Effective Action The effective action can be expanded according to (15.88):

$$\Gamma[\eta] = \Gamma_0 + \int d^dx\, \Gamma_1(x)\eta(x) + \tfrac{1}{2}\int d^dx_1\, d^dx_2 \Gamma_2(x_1, x_2)\eta(x_1)\eta(x_2) + \cdots$$

(a) Find $\Gamma_1(x)$.

(b) Show $\int dz\, \Gamma_2(y - z)G_2(z - x) = -\delta(y - x)$, where $G_2(x)$ is the two-point correlation function.

(c) Take the Fourier transform to get $\tilde{\Gamma}_2(k)\, \tilde{G}_2(-k) = -1$. Find $\tilde{\Gamma}_2(k)$ from the known general form of $\tilde{G}_2(k)$.

15.4 Goldstone's Theorem [21] Consider a relativistic complex scalar field $\psi(x)$, whose Lagrangian density is invariant under a global gauge transformation. Corresponding to this invariance is a conserved Noether current $j_\mu(x)$, with

$$\partial^\mu j_\mu(x) = 0$$

The corresponding conserved charge is denoted by Q, with the property

$$[\psi(x), Q] = \psi(x), \qquad Q \equiv \int d^3x\, j_0(x)$$

Assume that the global gauge symmetry is spontaneously broken:

$$\langle\psi\rangle \equiv \langle 0|\psi(x)|0\rangle \neq 0$$

where $|0\rangle$ is the vacuum state. Goldstone's theorem states that a massless particle exists. This only relies on current conservation, and does not depend on the detailed form of the Lagrangian. Prove it following the steps outlined below. We write $\psi = \psi(0), j^\mu = j^\mu(0)$.

(a) Consider the quantity

$$\Gamma^\mu(x) \equiv \langle 0|[\psi, j^\mu(x)]|0\rangle$$

Write out the commutator, insert a complete set of states between the operators, and note

$$\sum_n \delta^4(p - p_n)\, \langle 0|\psi|n\rangle\, \langle n|j^\mu(x)|0\rangle = \theta(p_0)p^\mu F(p^2)$$

where p_n is the 4-momentum of the state $|n\rangle$ and $F(p^2)$ is a Lorentz-invariant function. With this, show that

$$\Gamma^\mu(x) = \int d^4p \int dm^2\, \delta(p^2 - m^2)p^\mu e^{-ip\cdot x}[F(m^2)\theta(p_0) + F^*(m^2)\theta(-p_0)]$$

(b) So far we have used Lorentz invariance and assumptions about the mass spectrum. Now use current conservation in the form $\partial^\mu\Gamma_\mu = 0$ to show that $F(m^2)$ is of the form

$$F(m^2) = C\delta(m^2)$$

Thus, a massless particle exists if and only if $C \neq 0$.

(c) Consider Γ^0, and use the fact $\langle\psi\rangle \neq 0$ to show that $C \neq 0$.

15.5 Higgs Mechanism [22] In the presence of long-range interactions, the Goldstone mode is transformed. For a system interacting with the electromagnetic field, the gauge symmetry is enlarged from a global to a local one. When this symmetry is broken, the Goldstone mode becomes the longitudinal component of the electromagnetic field, and the photon acquires mass. This is called the *Higgs mechanism*. In a superconductor, the symmetry breaking results from the presence of a condensate of bound electron pairs—the Cooper pairs.

Consider the relativistic classical Lagrangian density for a complex field ψ coupled to the electromagnetic field, in (3 + 1)-dimensional Minkowskian space–time:

$$\mathcal{L}(x) = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + (D^\mu\psi)^*(\mathrm{D}_\mu\psi) - V(\psi^*\psi)$$

where $D^\mu\psi = (\partial^\mu + iqA^\mu)\psi$, and $V(\psi^*\psi) = \lambda(\psi^*\psi - \rho_0)^2$ is the potential causing spontaneous symmetry breaking. For a superconductor $q = 2e$. The notation is that of Chapter 5.

(a) Show that the Hamiltonian is

$$H = \int d^3x [\tfrac{1}{2}(\mathbf{B}^2 + \mathbf{E}^2) + |\pi|^2 + |\mathbf{D}\psi|^2 + V]$$

and from this show that the lowest-energy solution is $A^\mu = 0$, $\psi = \sqrt{\rho_0}\, e^{i\alpha 0}$. Thus local gauge symmetry is broken.

(b) Choose the "unitary gauge, " in which $\psi(x)$ is real. Show that the classical equations of motion are

$$\partial_\mu F^{\mu\nu} = -2q^2\psi^2 A^\nu$$

$$D^\mu \mathrm{D}_\mu \psi = 2\lambda(\rho_0 - \psi^2)\psi$$

Since $\partial_\mu \partial_\nu F^{\mu\nu} \equiv 0$, we must have $\partial_\mu A^\mu = 0$ whereever $\psi \neq 0$.

(c) Put $\psi(x) = \rho_0 + \eta(x)$, and show that

$$(\Box^2 + 2q^2\rho_0)\, A^\mu = 0$$

$$(\Box^2 + 4\lambda\rho_0)\eta = 0$$

When quantized, the theory has a vector particle A^μ of mass $q\sqrt{2\rho_0}$, and a scalar particle η of mass $2\sqrt{\lambda\rho_0}$. There is no massless scalar particle.

REFERENCES

1. P. W. Anderson, *Basic Notions of Condensed Matter Physics*, Benjamin Cummings, Menlo Park, CA, 1984, Chapter 2.
2. M. Waingrow, Ed., *James Boswell's Life of Johnson,* Edinburgh Univ. Press, Edinburgh, 1944, p. 131 (entry for Aug. 6, 1763).
3. P. Anderson, *Phys. Rev.* **110,** 827 (1958); Y. Nambu and Jona-Lasinio, *Phys. Rev.* **122,** 345 (1961); J. Goldstone, *Nuovo Cimento* **19,** 154 (1961).
4. K. Huang, E. Manousakis, and J. Polonyi, *Phys, Rev.* **35,** 3187 (1987).
5. K. Huang, *Statistical Mechanics*, 2nd ed., Wiley, New York, 1987, Section 13.8.
6. H. Kamerlingh Onnes, *Proc. Roy. Acad. Amsterdam* **13,** 1903 (1911).
7. H. Kamerlingh Onnes, *Leiden Commun.* **122b, 124c** (1911).
8. J. de Nobel, "The Discovery of Superconductivity," *Physics Today* p. 40 (Sept. 1996).
9. C. N. Yang, *Rev. Mod. Phys.* **34,** 4 (1964).
10. O. Penrose and L. Onsager, *Phys. Rev.* **104,** 576 (1956).
11. K. Huang, *Statistical Mechanics*, 2nd ed., Wiley, New York, 1987, Section 10.5.
12. E. P. Gross, *Nuovo Cimento* **20,** 454 (1961); *J. Math. Phys.* **4,** 195 (1963); L. P. Pitaevskii, *Sov. Phys. JETP* **13,** 451 (1961).
13. G. B. Whitham, *Linear and Nonlinear Waves*, Wiley, New York, 1974, Chapter 16.
14. M. R. Andrews, C. G. Townsend, H.-J. Miesner, D. S. Durfee. D. M. Kurn, and W. Ketterle, *Science,* **275,** 637 (1997).
15. K. Huang, *Statistical Mechanics*, 2nd ed., Wiley, New York, 1987, Section 3.7.

16. V. L. Ginsburg and L. P. Pitaevskii, *Sov. Phys. JETP* **34,** 858 (1958).
17. S. Coleman and E. Weinberg, *Phys. Rev.* **D7,** 1888 (1973).
18. R. Jackiw, *Phys. Rev.* **D9,** 1686 (1974).
19. K. Huang, *Quarks, Leptons, and Gauge Fields*, 2nd ed., World Scientific, Singapore, 1992, Section 10.9.
20. W. H. Louisell, *Phys. Lett.* **7,** 60 (1963).
21. J. Goldstone, A. Salam, and S. Weinberg, *Phys. Rev.* **127,** 965 (1962).
22. P. W. Anderson, *Phys. Rev.* **130,** 48 (1963); P. W. Higgs, *Phys. Rev.* **Lett. 12,** 132 (1964).

CHAPTER SIXTEEN

Renormalization

16.1 THE CUTOFF AS SCALE PARAMETER

In perturbation theory, we encounter divergent Feynman graphs that must be cut off at high momentum. The cutoff marks the energy scale at which known physics ends and the unknown begins, and would be necessary even if there were no divergences. It is presumed that we observe the world at an energy scale far below that of the cutoff, but the Lagrangian of the system is specified at the cutoff scale, with parameters appropriate to that scale. It would be desirable to express them in terms of low-energy parameters, without explicit reference to the cutoff, and renormalization is the means for doing that.

To ensure that we remain in the low-energy domain, we would like to take the cutoff to be infinite, but this cannot be done by declaration. In the absence of external fields, the action of the system does not contain an intrinsic energy scale apart from the cutoff. Thus, the cutoff disappears from the action when we reduce all quantities to dimensionless form. The only way to tell whether it is finite or infinite is to calculate some physical quantity with dimension, such as the correlation length, from the theory. The cutoff is infinite when the correlation length diverges, in which case the system is said to be at a critical point. To approach the limit of infinite cutoff, therefore, we must adjust the parameters so as to make the system "go critical." We shall explain this in a concrete example.

Consider a scalar field $\phi(x)$ in d–dimensional Euclidean space–time, with classical action

$$S[\phi] = \int d^d x [\tfrac{1}{2}(\partial\phi)^2 + g_2\phi^2 + g_4\phi^4 + g_6\phi^6 + \cdots] \tag{16.1}$$

We place our system in a large hypercube of volume Ω, with periodic boundary conditions, and eventually let $\Omega \to \infty$. Fourier transforms are defined by

$$\phi(x) = \Omega^{-1/2} \sum_k e^{-ik\cdot x} \phi_k$$

$$\phi_k = \Omega^{-1/2} \int d^d x \, e^{ik\cdot x} \phi(x) \tag{16.2}$$

with $\phi_k^* = \phi_{-k}$. In the limit $\Omega \to \infty$, we write

$$\phi(x) = \int \frac{d^d q}{(2\pi)^d} e^{-iq\cdot x} \tilde{\phi}(q) \tag{16.3}$$

where

$$\tilde{\phi}(k) = \Omega^{1/2} \phi_k \tag{16.4}$$

The action can be rewritten in the form

$$S[\phi] = \frac{1}{2} \sum_{|p|<\Lambda} (p^2 + 2g_2) \phi_p \phi_{-p} + S_I[\phi] \tag{16.5}$$

The first term is called the "kinetic term," and $S_I[\phi]$ contains the interactions:

$$S_I[\phi] = \frac{g_4}{\Omega} \sum_{|p_i|<\Lambda} \delta(\Sigma p_i) \phi_{p_1} \phi_{p_2} \phi_{p_3} \phi_{p_4}$$

$$+ \frac{g_6}{\Omega^2} \sum_{|p_i|<\Lambda} \delta(\Sigma p_i) \phi_{p_1} \phi_{p_2} \phi_{p_3} \phi_{p_4} \phi_{p_5} \phi_{p_6} + \cdots \tag{16.6}$$

where δ denotes the Kronecker delta (δ). The partition function is given by

$$Z = \int D\phi \, e^{-S[\phi]} \tag{16.7}$$

where

$$\int D\phi = \mathcal{N} \prod_{|p|<\Lambda} \int d\phi_p d\phi_p^* \tag{16.8}$$

where $\mathcal{N}$ is a normalization constant.

With units such that $\hbar = c = 1$, the dimensionality of any quantity can be expressed as a power of length, or equivalently momentum. The dimensionality of a quantity X, denoted by $[X]$, can be deduced from the assertion that the action is dimensionless. Using the cutoff momentum Λ as unit, we have

$$[X] = \Lambda^{-p} \tag{16.9}$$

which means that X transforms like Λ^{-p} under a change of scale, or that $X\Lambda^p$ is dimensionless. It is straightforward to verify the following:

$$\begin{aligned}
[\phi(x)] &= \Lambda^{-1+d/2} \\
[\tilde{\phi}(p)] &= \Lambda^{-1-d/2} \\
[\phi_p] &= \Lambda^{-1} \\
[g_\alpha] &= \Lambda^{\alpha+d+-\alpha d/2}
\end{aligned} \tag{16.10}$$

The exponents above are the so-called canonical dimensions. It is convenient to use the following dimensionless quantities:

$$\begin{aligned}
\text{Momentum:} \quad & q = \frac{p}{\Lambda} \\
\text{Fourier component:} \quad & \varphi_q = \Lambda \phi_p \\
\text{Coupling constants:} \quad & u_\alpha = \Lambda^{-\alpha-d+\alpha d/2} g_\alpha
\end{aligned} \tag{16.11}$$

The partition function can then be written in the form

$$Z = \int D\varphi \, e^{-S[\varphi]} \tag{6.12}$$

The action is written

$$S[\varphi] = \tfrac{1}{2} \sum_{|q|<1} (q^2 + r)\varphi_q \varphi_{-q} + S_I[\varphi] \tag{6.13}$$

where

$$r \equiv 2u_2 \tag{6.14}$$

and the interaction term can be represented in the form

$$\begin{aligned}
S_I[\varphi] = {} & \frac{u_4}{\Omega} \sum_{|q_i|<1} \delta(\Sigma\, q_i)\varphi_{q_1}\varphi_{q_2}\varphi_{q_3}\varphi_{q_4} \\
& + \frac{u_6}{\Omega^2} \sum_{|q_i|<1} \delta(\Sigma\, q_i)\varphi_{q_1}\varphi_{q_2}\varphi_{q_3}\varphi_{q_4}\varphi_{q_5}\varphi_{q_6} + \cdots
\end{aligned} \tag{6.15}$$

We see that Λ has disappeared. It merely provides a scale to measure physical quantity with dimension, such as the correlation length $\tilde{\xi}$, defined through the asymptotic behavior of a correlation function:

$$\langle \phi(x)\phi(y)\rangle \xrightarrow[|x-y|\to\infty]{} C\, e^{-|x-y|/\tilde{\xi}} \tag{16.16}$$

Here, $\tilde{\xi}$ is measured in the same unit as x, such as meters. Using Λ^{-1} as unit for distance, we have

$$\frac{|x-y|}{\tilde{\xi}} = \frac{|x-y|\Lambda}{\xi} \tag{16.17}$$

where ξ is dimensionless:

$$\xi = \Lambda\tilde{\xi} \tag{16.18}$$

Ignoring the pathological case $\tilde{\xi} = 0$, we see that an infinite cutoff corresponds to the limit $\xi \to \infty$.

The theory at the cutoff scale, called the "bare theory," is specified by the coupling constants u_α. The value of the cutoff is reflected solely in these coupling constants. A renormalization-group (RG) transformation is a coarse-graining operation through which we lower the cutoff without changing the system. When the cutoff Λ is lowered to $\mu = \Lambda/b$, the "bare" couplings $u_\alpha(\Lambda)$ will change to the "renormalized" ones $u_\alpha(\mu)$. The system should remain unchanged, even though it appears to be changed, because the effective couplings are different. We shall define the RG transformation explicitly later. For the present, assume that such an operation has been defined.

Let us consider the parameter space spanned by all the u_α. This is a space of all possible Hamiltonians, if you like. Each point in this space specifies a system with a specific value of ξ, and therefore of Λ. Successive RG transformation generate an RG trajectory that flows in the coarse-graining direction. Since the cutoff is lowered in an RG transformation, the correlation length decreases along an RG trajectory.

There may exist "fixed points" that are invariant under RG transformations. Since ξ decreases under an RG transformation, we must have $\xi = 0$ or $\xi = \infty$ at a fixed point. We ignore the case $\xi = 0$, and concentrate on $\xi = \infty$. Since this corresponds to an infinite cutoff, we cannot place a bare system exactly on it, but only approach it as a limit. If the RG transformation is unique, which we assume, then two different trajectories cannot intersect except at a fixed point. This is why a trajectory can be continued backward, even though it is only defined for forward motion. This makes the RG a true group.

A fixed point is a source or sink of trajectories. To trajectories flowing away from it, the fixed point appears as an ultraviolet (UV) fixed point, since it represents the infinite-momentum limit. We call them *UV trajectories*. To trajectories flowing into it, the fixed point appears as an infrared (IR) fixed point, and we call them *IR trajectories*. Thus, whether a fixed point is UV or IR depends on the trajectory one chooses. On the other hand, the correlation length at the fixed point is an intrinsic property.

The correlation length along an IR trajectory must be infinite, since it must de-

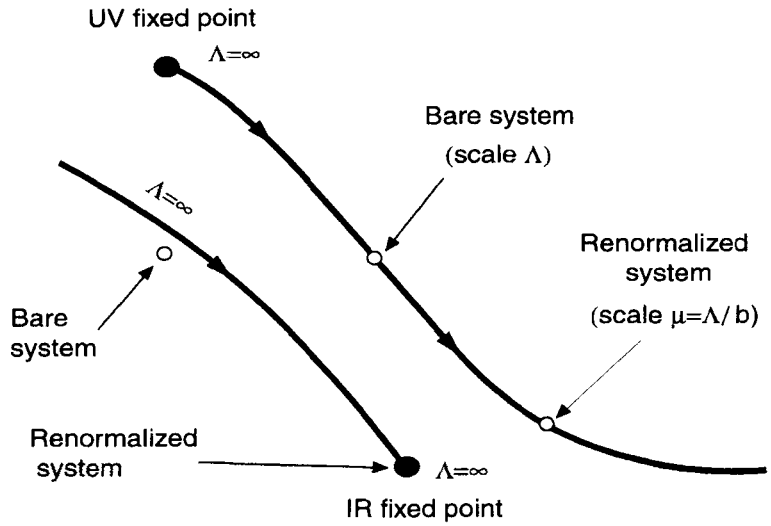


Figure 16.1 How to approach the infinite-cutoff limit depends on whether the theory is base on a UV or IR fixed point.

crease along the trajectory, and it is infinite at the endpoint. Thus, we cannot place a bare system on an IR trajectory, but must approach it as a limiting trajectory, and in that limit the system goes to the IR fixed point. A UV trajectory is different. Since the correlation length is finite, we can place our bare system on it. The difference between UV and IR trajectories is illustrated in Fig. 16.1.

The limit of infinite cutoff is also known as the *continuum limit,* because the unit of length approaches zero. As illustrated in Fig. 16.1, the continuum limit may be based on an UV fixed point, or an IR fixed point. They define different limiting theories. A theory of the UV kind conforms to our thinking in perturbative renormalization; namely, we can keep the renormalized parameters fixed at arbitrary values while letting the cutoff go to infinity.. A theory of the IR kind has no freedom. The renormalized parameters assume the fixed-point values.

An interesting possibility is that both UV and IR trajectories exist, and are near each other, as shown in Fig. 16.1. The system may appear to be following one trajectory at first, but make a switch later. As we speculate later, this may be the case for QED.

16.2 MOMENTUM SPACE RG

Wilson [1] proposes a RG transformation whereby the cutoff is reduced from 1 to $1/b$ by integrating out the modes in between. This gives a new action containing

fewer degrees of freedom. There are three steps, as illustrated in Fig. 16.2, and described in the following:

16.2.1 Designating Fast and Slow Modes

Decomposing the field into a "slow" part σ_q and "fast" part f_q:

$$\begin{aligned} \varphi_q &= \sigma_q + f_q \\ \sigma_k &= 0 \quad \text{unless } |k| < 1/b \\ f_k &= 0 \quad \text{unless } \frac{1}{b} \leq |k| \leq 1 \end{aligned} \tag{16.19}$$

where $b > 0$. The partition function can be rewritten in the form

$$Z = \int D\sigma \int Df\, e^{-S[\sigma+f]} \tag{16.20}$$

and we imagine that the $\int Df$ integration is carried out.

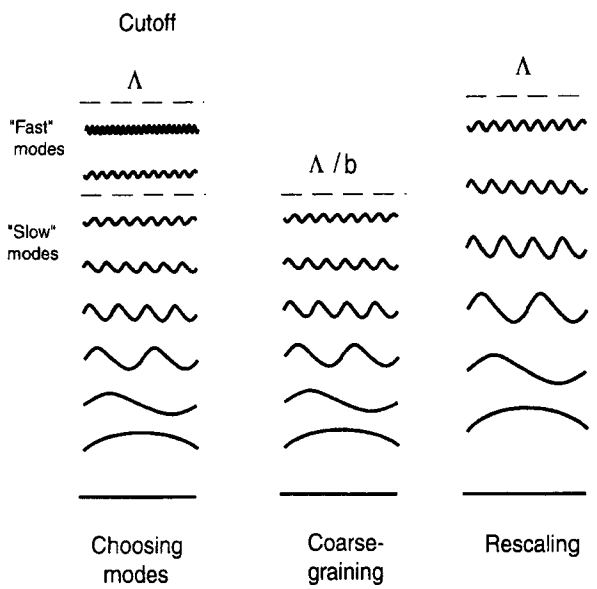


Figure 16.2 RG transformation in momentum space.

16.2.2 Coarse Graining

Define a new action $\tilde{S}[\sigma]$ dependent only on the slow fields through the relation

$$e^{-\tilde{S}[\sigma]} \equiv \mathcal{N} \int Df\, e^{-S[\sigma+f]} \tag{16.21}$$

where $\mathcal{N}$ is a constant. The new action is expanded in the form

$$\tilde{S}[\sigma] = \tfrac{1}{2} \sum_{|q|<1/b} [zq^2 + r_1]\sigma_q \sigma_{-q} + \tilde{S}_I[\sigma] \tag{16.22}$$

which defines z, r_1, and new parameters in $\tilde{S}_I[\sigma]$. The constant $\mathcal{N}$ is put in to absorb any constant terms generated, to make this expansion possible. We put

$$z = b^{-\eta} \tag{16.23}$$

which defines the exponent η, the *anomalous dimension*. (See Problem 16.1.) The partition function is now rewritten

$$Z = \mathcal{N} \int D\sigma\, e^{-\tilde{S}[\sigma]} \tag{16.24}$$

but its value is unchanged.

16.2.3 Rescaling

For comparison with the original action, rescale the cutoff to 1, and normalize the field such that the coefficient of the q^2 term is $\frac{1}{2}$. This is done by defining

$$\begin{aligned} q' &\equiv bq \\ \varphi'_{q'} &\equiv b^{-1-\eta/2}\sigma_{q'/b} \end{aligned} \tag{16.25}$$

The spatial volume changes to

$$\Omega' = b^{-d}\Omega \tag{16.26}$$

The action can be rewritten as

$$S'[\varphi'] \equiv \tilde{S}[\sigma] = \tfrac{1}{2} \sum_{|q|<1} (q^2 + r')\varphi'_q \varphi'_{-q} + S'_I[\varphi'] \tag{16.27}$$

where $r' = b^{2+\eta} r_1$. We use q in place of q', since it is just an summation variable.

We can expand the interaction term in the form

$$S_l'[\varphi'] = \frac{u_4'}{\Omega'} \sum_{|q_i|<1} \delta(\Sigma q_i)\varphi'_{q_1}\varphi'_{q_2}\varphi'_{q_3}\varphi'_{q_4} + \frac{u_6'}{\Omega'^2} \sum_{|q_i|<1} \delta(\Sigma q_i)\varphi'_{q_1}\varphi'_{q_2}\varphi'_{q_3}\varphi'_{q_4}\varphi'_{q_5}\varphi'_{q_6} + \cdots \tag{16.28}$$

and read off the renormalized parameters u_n'. The RG transformation is the operation

$$\{u_2, u_4, u_6, \cdots\} \to \{u_2', u_4', u_6', \cdots\} \tag{16.29}$$

Note that u_α' depends not only on the scale parameter b but also on the initial values $\{u_\alpha\}$ as well.

It is convenient to parametrize the trajectory by an additive parameter t defined by

$$b = e^{t-t_0} \tag{16.30}$$

with an arbitrary origin t_0. The coupling constants along the trajectory are then functions of t:

$$u_n' = u_n(t) \tag{16.31}$$

From (16.28), which gives $u(t)$ in terms of some initial value $u_0 = u(t_0)$, we can calculate the tangent vector along the trajectory:

$$\frac{\partial u_\alpha}{\partial t} = \beta_\alpha(u, u_0)$$

We can erase the memory of the initial state by taking the limit $u_0 \to u_*$, where u_* is a fixed point. If this can be done, we have the conventional β function of perturbative renormalization:

$$\beta_\alpha(u) \equiv \beta_\alpha(u, u_*) \tag{16.32}$$

16.3 REAL-SPACE RG

The coarse-graining process of the RG transformation can be performed in real space. This is Kadanoff's "block-spin" transformation [2], which historically precedes Wilson's momentum–space formulation. The system is defined on the discrete sites x of a d–dimensional hypercubic lattice, with lattice spacing $a = \Lambda^{-1}$. We consider again a scalar field whose value at site x is denoted by ϕ_x. We make the following correspondence with the continuum formulation:

$$\int d^d x \to a^d \sum_x$$
$$\partial\phi \to a^{-1}(\phi_x - \phi_y) \tag{16.33}$$

where x and y denote nearest-neighbor sites. The lattice action is accordingly

$$S[\phi] = \tfrac{1}{2} a^{d-2} \sum_{<xy>} (\phi_x - \phi_y)^2 + a^d \sum_x [g_2\phi_x^2 + g_4\phi_x^4 + g_6\phi_x^6 + \cdots] \tag{16.34}$$

where the sum extends over all distinct nearest-neighbor pairs $<xy>$. We introduce a dimensionless field φ_x and coupling constants u_n:

$$\varphi_x = a^{-1+d/2}\phi_x$$
$$u_\alpha = a^{-\alpha-d+\alpha d/2} g_\alpha \tag{16.35}$$

The action can then be written as

$$S[\varphi] = \tfrac{1}{2} \sum_{<xy>} (\varphi_x - \varphi_y)^2 + \sum_x [u_2\varphi_x^2 + u_4\varphi_x^4 + u_6\varphi_x^6 + \cdots] \tag{16.36}$$

where a does not appear explicitly. To approach the continuum limit, we must tune the parameters u_n such that the correlation length becomes infinite.

The partition function is

$$Z = \int \varphi\, e^{-S[\varphi]} \tag{16.37}$$

where

$$\int D\varphi = \prod_x \int d\varphi_x \tag{16.38}$$

The steps in an RG transformation are illustrated in Fig. 16.3, and described in the following pragraphs.

16.3.1 Making Blocks

We begin by grouping the sites into blocks, ℓ sites on a side, and denoting by X the position at the center of a block. (In Fig. 16.3 the blocks have $\ell = 2$.) The average field at average position X is defined as

$$\langle\varphi\rangle_X \equiv \ell^{-d} \sum_{x \in \text{block}} \varphi_x \tag{16.39}$$

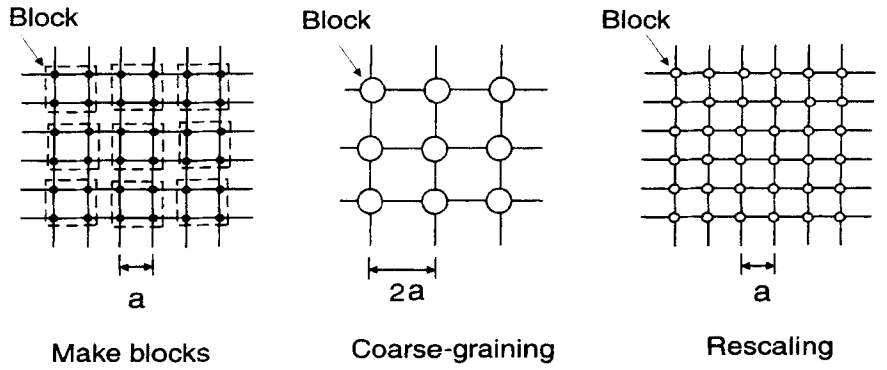


Figure 16.3 RG transformation in real space.

By introduce a block field $\tilde{\varphi}_X$, we can trivially rewrite the partition function in the form

$$Z = \int D\tilde{\varphi} \int D\varphi \prod_x \delta(\tilde{\varphi}_X - \langle\varphi\rangle_X) e^{-S[\varphi]} \tag{16.40}$$

where

$$\int D\tilde{\varphi} = \prod_x \int d\tilde{\varphi}_X \tag{16.41}$$

16.3.2 Coarse Graining

We integrate over the original field, holding the block field fixed, to define a new action $\tilde{S}[\tilde{\varphi}]$ dependent only on the block field:

$$e^{-\tilde{S}[\tilde{\varphi}]} \equiv \mathcal{N} \int D\varphi \prod_x \delta(\tilde{\varphi}_X - \langle\varphi\rangle_X)\, e^{-S[\varphi]} \tag{16.42}$$

The partition function now takes the form

$$Z = \int D\tilde{\varphi}\, e^{-\tilde{S}[\tilde{\varphi}]} \tag{16.43}$$

We can expand the new action in the form

$$\tilde{S}[\tilde{\varphi}] = \frac{z}{2} \sum_{<XY>} (\tilde{\varphi}_X - \tilde{\varphi}_Y)^2 + \sum_X [\tilde{u}_2 \tilde{\varphi}_X^2 + \tilde{u}_4 \tilde{\varphi}_X^4 + \tilde{u}_6 \tilde{\varphi}_X^6 + \cdots] \tag{16.44}$$

where $z = \ell^{-\eta}$.

16.3.3 Rescaling

The spacing between block sites X is ℓ times larger than that for the original sites x. We restore the original spacing by introducing site variables

$$x' = X/\ell \tag{16.45}$$

At the same time, the normalization of the kinetic term is restored to $\frac{1}{2}$ by transforming the block field to

$$\varphi'_{x'} \equiv \ell^{-\eta/2}\,\tilde{\varphi}_{\ell x'} \tag{16.46}$$

The action then takes the form

$$S'[\varphi'] \equiv \tilde{S}[\tilde{\varphi}] = \tfrac{1}{2}\sum_{<xy>}(\varphi'_x - \varphi'_y)^2 + \sum_x (u'_2\varphi'^2_x + u'_4\varphi'^4_x + u'_6\varphi'^6_x + \cdots) \tag{16.47}$$

where we use x instead of x' as summation index. The partition function can now be written as

$$Z = \mathcal{N}\int D\varphi'\; e^{-S'[\varphi']} \tag{16.48}$$

and the problem reduces to that in momentum space.

16.4 RENORMALIZATION OF CORRELATION FUNCTIONS

The n-point correlation function in momentum space is

$$G_n(p;\, g_0,\, \Lambda) \equiv \frac{\int D\phi[\phi_{p_1}\cdots\phi_{p_n}]e^{-S[\phi]}}{\int D\phi e^{-S[\phi]}} \tag{16.49}$$

where $|p_i| < \Lambda$, and $g_0(\Lambda)$ denotes the set of bare coupling constants. In terms of the dimensionless field $\varphi_{p/\Lambda} = \Lambda\phi_p$ and dimensionless bare coupling constants by u_0, we can write

$$G_n(p;\, g_0,\, \Lambda) = \Lambda^{-n}\mathcal{G}_n\left(\frac{p}{\Lambda};\, u_0\right) \tag{16.50}$$

where

$$\mathcal{G}_n(q;\, u_0) = \frac{\int D\varphi[\varphi_{q_1}\cdots\varphi_{q_n}]e^{-S[\varphi]}}{\int D\varphi e^{-S[\varphi]}} \tag{16.51}$$

We are interested only in the correlations among slow modes, with $|q_i| < 1/b$. Thus $\varphi_q = \sigma_q$, and we have

$$\mathcal{G}_n(q; u_0) = \frac{\int D\sigma \int Df[\sigma_{q_1} \cdots \sigma_{q_n}]e^{-S[\sigma+f]}}{\int D\sigma \int Df\, e^{-S[\sigma+f]}} = \frac{\int D\sigma[\sigma_{q_1} \cdots \sigma_{q_n}]e^{-\tilde{S}[\sigma]}}{\int D\sigma e^{-\tilde{S}[\sigma]}}$$

$$= z_0^{-n/2} b^n \frac{\int D\varphi'[\varphi'_{bq_1} \cdots \varphi'_{bq_n}]e^{-S'[\varphi']}}{\int D\varphi' e^{-S'[\varphi']}} \tag{16.52}$$

This gives the transformation law

$$\mathcal{G}_n(q; u_0) = z_0^{-n/2} b^n \mathcal{G}_n(bq; u') \tag{16.53}$$

To make contact with perturbative renormalization, choose

$$b = \frac{\Lambda}{\mu} \tag{16.54}$$

where μ is the renormalization scale, to be held fixed while we make $\Lambda \to \infty$ by making $u_0 \to u_*$. Multiplying both sides of (16.53) by $\Lambda^{-n} = (b\mu)^{-n}$, we obtain

$$G_n(p; g_0, \Lambda) = z_0^{-n/2} \mu^{-n} \mathcal{G}_n\left(\frac{p}{\mu}; u'\right) \tag{16.55}$$

In perturbative renormalization, this is usually written in the form

$$G_n(p; g_0, \Lambda) = \left[z_0\left(\frac{\Lambda}{\mu}, g_0\right)\right]^{-n/2} G'_n(p; g, \mu) \tag{16.56}$$

where g denotes the renormalized coupling constants and G_n' is the renormalized correlation function. We see that the cutoff dependence can be isolated in a factor, while the rest of the function depends on the renormalized couplings.

16.5 RELEVANT AND IRRELEVANT PARAMETERS

Consider a fixed point u_*. If there is a UV trajectory flowing out of it, then a point u on the trajectory goes away from the fixed point under the RG transformation. In this case u is said to be a *relevant* parameter, because $u - u_*$ grows. On the other hand, along an IR trajectory that flows into the fixed point, we have $(u - u_*) \to 0$, and u is said to be *irrelevant*. Relevancy and irrelevancy are properties associated with *directions* in parameter space with respect to a given fixed point.

In the neighborhood of a fixed point, we can neglect those couplings that vanish rapidly. Let us assume that we can limit the number of couplings to a finite number K, and represent them as components of a vector

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_K \end{pmatrix} \tag{16.57}$$

We parametrize the trajectory by t, so that the tangent vector to the trajectory is $d\vec{u}/dt$. In the neighborhood of u_*, to linear order in $\vec{u} - \vec{u}_*$, the RG is represented by a matrix equation

$$\frac{d}{dt}(\vec{u} - \vec{u}_*) = R(\vec{u} - \vec{u}_*) \tag{16.58}$$

where R is called the *RG matrix*. Let us diagonalize R:

$$R\vec{v} = \lambda\vec{v} \tag{16.59}$$

The eigenvectors $\vec{v}$ define the principal axes, along which we have

$$\vec{u}(t) - \vec{u}_* = \vec{c}e^{\lambda t} \tag{16.60}$$

where $\vec{c}$ is some constant vector. If $\lambda > 0$, the principal axes correspond to a relevant direction, for the trajectory flows away from the fixed point. If $\lambda < 0$, we have an irrelevant direction, and $\vec{u} - \vec{u}_* \to 0$. The eigenvalues of the RG matrix furnish information about the nature of the RG flow and consequently the geometric properties of parameter space in the neighborhood of the fixed point.

The linear analysis becomes inadequate if there are null eigenvalues $\lambda = 0$. The corresponding directions are said to be "marginal," and the flow patterns have to be investigated in a higher order.

16.6 THE FREE FIELD

We work out the RG for the free field, or Gaussian model, defined through the action

$$\begin{aligned} S_0[\phi] &= \tfrac{1}{2}\int d^dx[(\partial\phi)^2 + 2g_2\phi^2] \\ &= \tfrac{1}{2}\sum_{|q|<1}(q^2 + r)\varphi_q\varphi_{-q} \end{aligned} \tag{16.61}$$

where ϕ is a real scalar field, with dimensionless form φ, and

$$r \equiv 2u_2 \tag{16.62}$$

Since $\varphi_{-q} = \varphi_q{}^*$, we can write

$$S_0[\varphi] = \sum_{0<|q|<1} (q^2 + r)|\varphi_q|^2 \tag{16.63}$$

where the sum over q extends only over a hemisphere. We write the same action either as $S_0[\phi]$ or $S_0[\varphi]$. The partition function is

$$Z = \int D\varphi\, e^{-S_0[\varphi]} \tag{16.64}$$

where

$$\int D\varphi = \prod_{0<|q|<1} \int d\varphi_q^*\, d\varphi_q = \prod_{0<|q|<1} \int_{-\infty}^{\infty} d(\mathrm{Re}\ \varphi_q) \int_{-\infty}^{\infty} d(\mathrm{Im}\ \varphi_q) \tag{16.65}$$

The partition function can be calculated directly; but, for illustration, we shall carry out RG transformations "by the book."

Decomposing the field into fast and slow components, $\varphi = f + \sigma$, we have $f_q \sigma_{-q} = 0$, because their ranges are disjoint. Thus, the free action decomposes into separate sums with no cross-term:

$$S_0[f + \sigma] = S_0[f] + S_0[\sigma] \tag{16.66}$$

The partition function factorizes:

$$Z = \int D\sigma \int Df\, e^{-S_0[f+\sigma]} = \int D\sigma\, e^{-S_0[\sigma]} \int Df\, e^{-S_0[f]} \tag{16.67}$$

and integrating over the fast modes yields an overall factor:

$$\int Df\, e^{-S_0[f]} = \prod_{1/b<|q|<1} \frac{\pi}{q^2 + r} \equiv e^{-C_0(b,r)} \tag{16.68}$$

Thus, the slow modes undergo renormalization solely because of rescaling. With $q' = qb$, we have

$$S_0[\sigma] = \sum_{0<|q|<1/b} (q^2 + r)|\sigma_q|^2 = \sum_{0<|q'|<1} (b^{-2}q'^2 + r)|\sigma_{q'/b}|^2$$

$$= \sum_{0<|q'|<1} (q'^2 + 2b^2 r)|\varphi'_{q'}|^2 \tag{16.69}$$

where

$$\varphi'_{q'} = b^{-1}\sigma_{q'/b} \tag{16.70}$$

Thus the RG transformation gives

$$r' = b^2 r \tag{16.71}$$

We put $b = e^{t-t_0}$, and regard r as a function of t:

$$r' \equiv r(t) = r\, e^{2(t-t_0)} \tag{16.72}$$

Since r and t_0 are arbitrary, we have

$$r(t) = c_0 e^{2t} \tag{16.73}$$

where c_0 is a constant.

The RG trajectories are illustrated in Fig. 16.4. There is only one fixed point at $r = 0$, the Gaussian fixed point. It is an UV fixed point, corresponding to the high-momentum limit $t \to -\infty$. The parameter $r(t)$ exhibits asymptotic freedom, since it vanishes in the limit of infinite momentum scale. The negative r axis is unphysical, since it corresponds to systems whose energy spectrum is not bounded from below.

To verify that the correlation length diverges at the Gaussian fixed point, recall that the correlation function for $d > 2$ is given by

$$\langle \phi(x)\phi(y)\rangle = C\, e^{-\sqrt{2g_2}|x-y|} \tag{16.74}$$

The dimensionless correlation length is thus

$$\xi = r^{-1/2} \tag{16.75}$$

which diverges as $r \to 0$.

16.7 IR FIXED POINT AND PHASE TRANSITION

An IR fixed point is the endpoint of an IR trajectory that forms the dividing line between two different phases. We illustrate this with ϕ^4 theory in $d = 4$ dimensions.

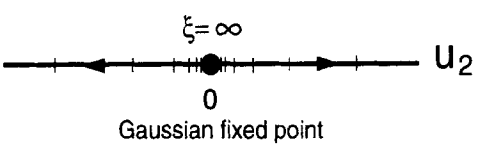


Figure 16.4 RG trajectories for u_2, the dimensionless squared mass. Arrows point along directions of coarse graining. There is an UV fixed point at $u_2 = 0$. Tick marks indicate equal increment of the dimensionless correlation length ξ, which diverges at the fixed point.

The RG trajectories are shown in Fig. 16.5 This parameter space is spanned by u_2 and u_4. Only the upper half-plane has physical relevance, since we must have $u_4 \geq 0$ for the energy spectrum to be bounded from below. The Gaussian fixed point occurs at the origin, and there is an IR trajectory flowing into it. The points P and Q represent cutoff theories, which approach continuum limits at any point on the IR trajectory. From any point in the neighborhood of the IR trajectory, coarse graining will tend to decrease u_4, bringing the system close to the origin. Thus, in the continuum limit, the renormalized system approaches the Gaussian fixed point, the free-field theory. This is the phenomenon of "triviality".

As illustrated in Fig. 16.5, the contours of constant ξ are parallel to the IR trajectory, along which $\xi = \infty$. The contours of equal separation in ξ become infinitely dense in the neighborhood of the IR trajectory, and the latter resembles a bottomless ravine. This ravine is in fact a phase transition line that divides the u_2–u_4 plane into the symmetry-broken phase to the left, and the symmetric phase to the right. The points P and Q approach continuum limits in the respective phases. In this limit we have $u_2 \to 0$ and $\Lambda \to \infty$, and thus the mass parameter is indeterminate:

$$m^2 \to 2u_2\Lambda^2 \tag{16.76}$$

Thus we can set it at an arbitrary value, and this is called "mass renormalization." In the symmetric phase, where m^2 is positive, it gives the squared mass of the field quanta. In the broken phase it is negative, and contributes to the vacuum field $\langle\phi\rangle = \sqrt{-m^2/4g_4}$.

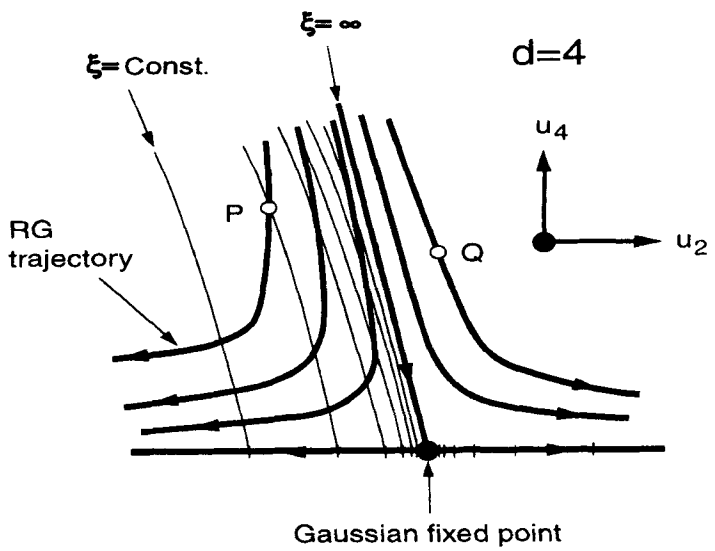


Figure 16.5 RG trajectories of ϕ^4 theory in four Euclidean dimensions. An IR trajectory flows into the Gaussian fixed point. It represents a phase-transition line, along which the correlation length is infinite.

We can interpret the ϕ^4 theory as a Ginsburg–Landau theory for an order parameter ϕ. In this case, u_2 is regarded as a temperature. The points P and Q in Fig. 16.5 correspond to thermodynamic states near a second-order phase transition, and the eigenvalues of the RG matrix determine the critical exponents (see, e.g., Huang [3]).

16.8 CROSSOVER

When a trajectory flows from the neighborhood of one fixed point to that of another, the qualitative nature of the system changes. This is called crossover behavior, and is illustrated in Fig. 16.6, for ϕ^4 theory in $d = 4 - \epsilon$ dimensions ($\epsilon \to 0^+$).

The Gaussian fixed point describes a massless free field, but the nontrivial fixed point describes something else. To illustrate the difference in an extreme limit, imagine that the nontrivial fixed point recedes to the far corner of the second quadrant. The potential will become a steep double well, with the field trapped in two possible values, and the fixed-point system will resemble an Ising model.

Consider the trajectory marked with point P in Fig. 16.6. At very short-distance scales, the system is near the Gaussian fixed point, and will remain there through orders of magnitude of coarse graining. Thus the system appears to be a massless free field through orders of magnitudes of scale change. When we continue to "zoom out," eventually the system leaves this neighborhood, and begins to move toward the nontrivial fixed point. Once it gets under way, the system moves quickly

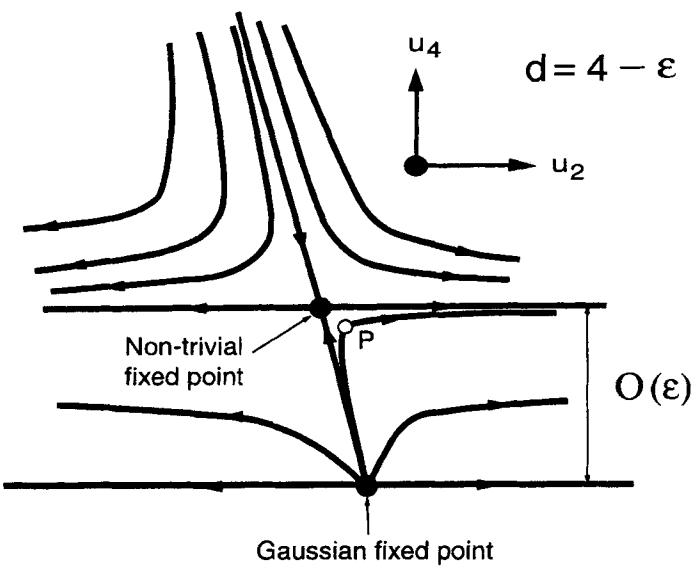


Figure 16.6 As P flows along the trajectory, the system crosses over from free-field behavior near the Gaussian fixed point to a different behavior near the nontrivial fixed point.

until it comes close to the nontrivial fixed point, like the point P referred to previously, where the system remains Isinglike through orders of magnitude of scale change.

If we imagine performing experiments on this system, we would first see the system at low energies, corresponding to a point far downstream on the trajectory. When the energy scale is increased, the system will quickly settle down to Ising behavior, and remain there until we enter the realm of "high-energy physics," when almost suddenly we see it as a free field whose mass exhibits "asymptotic freedom." The term "crossover" emphasizes the suddenness of the transition.

16.9 RELATION WITH PERTURBATIVE RENORMALIZATION

The Wilson RG supplies a physical basis for renormalization, but does not furnish tools for carrying it out. To implement it, the only analytic method we have is still perturbation theory. Perturbative renormalization, of course, predated and inspired the Wilson RG. It has been very successful in dealing with divergences in a practical way. It has even achieved some physical understanding when the theory happens to be asymptotically free, that is, based on an UV fixed point, but it fails to uncover the basic structure when the governing fixed point is IR. When guided by the Wilson RG, perturbation renormalization acquires a roadmap. It does not make calculations any easier, but at least one is more aware of the possible terrains.

What one does in perturbative renormalization may be summarized as follows. The immediate task is to calculate a correlation function by expanding it in terms of Feynman graphs. A graph is represented by a integral, which may be divergent. In that case, it is rendered finite by introducing a cutoff momentum Λ:

$$I(p,\Lambda) = \int_0^\Lambda dk\, f(k, p) \tag{16.77}$$

where p is an external momentum. Let us suppose, as in typical situations, that it diverges like $\ln \Lambda$. Then, one subtraction will render the integral finite. That is, we write

$$I(p,\Lambda) = I(p_0, \Lambda) + [I(p,\Lambda) - I(p_0, \Lambda)] \tag{16.78}$$

where p_0 is a momentum that sets the renormalization scale $|p_0|^2 = \mu^2$. The divergence is now isolated in the first term. One then shows that, when all the Feynman graphs are added, all such divergent terms can be absorbed into renormalized coupling constants. The result is summarized by formula (16.56). The miracle is that a subtraction made in a graph somehow turns into a multiplication of the coupling constant. The rederivation of (16.56) through the Wilson RG shows the equivalence of the result in the two approaches. The methodologies are also the same. The renormalization scale μ, established by the choice of subtraction point in perturbative

renormalization, corresponds in the Wilson RG to the effective cutoff $\mu = \Lambda/b$. The important difference is that, whereas in perturbative renormalization μ is seen as a subtraction point to get rid of divergences, in the Wilson RG it is the result of coarse graining, and thus has physical meaning. The Wilson RG has a better vantage point, for it works with the action as a whole, instead of individual Feynman graphs. One limitation of the present formulation of the Wilson method is the explicit use of a sharp momentum cutoff.[2]

Certain folk beliefs growing out of perturbative renormalization need to be revised. It is usually assumed that renormalized coupling parameters can be held fixed at arbitrary values, while we send the cutoff to infinity. This is not always correct. As we have discussed, renormalized parameters can be considered arbitrary only if the continuum limit is realized by placing the theory on a UV trajectory. If the continuum limit is realized by approaching an IR trajectory, then the renormalized coupling constants are fixed; they assume the values at the IR fixed point.

It is also a common belief that interactions are either "renormalizable" or "nonrenormalizable." In the former, Feynman graphs can be made finite through a finite number of subtractions, whereas in the latter category they would require an infinite number of subtractions. This is an artificial distinction based on the idea that the degree of divergence of a Feynman graph is determined solely by its topological structure, without reference to the scaling properties of the coupling constants. As illustrated in Problem 16.2, a proper counting of the degree of divergence must take into account the cutoff dependence of the coupling constants.

A better criterion for a "renormalizable" theory is that the Lagrangian contain the same interaction terms at all length scales. One can then make the subtractions in the Lagrangian itself, by introducing counterterms. As a general requirement, however, this is impossible. Although we can choose the Lagrangian at a particular energy scale, its form at lower energy scale is not under our control. The system evolves along an RG trajectory, and relevant interactions emerge, while irrelevant ones die out. The limitation to the same set of interaction terms merely means that we are sufficiently close to a fixed point that all irrelevant interactions can be ignored.

In discussing the relation between the Wilson RG and perturbative renormalization, the case of QED remains a puzzle. This is ironic, for perturbative renormalization scores its greatest triumph in QED, and yet the fixed-point structure is not clear. On one hand, the success of perturbative renormalization is based on the recipe of the UV type, that we can keep the renormalized parameters fixed at arbitrary values while letting the cutoff go to infinity. Yet, perturbation theory also shows that there is no UV fixed point—the theory is not asymptotically free. Instead, a partial summation of Feynman graphs suggests that there is an IR fixed point, which leads to triviality. What is going on? The importance of this theory impels us to offer a scenario.

We have to distinguish between the mathematical QED, the theory described in

[2]Polchinski [4] attempts a reformulation of the Wilson RG with a more general choice of cutoff function.

Chapter 11, and the physical QED, which is embedded in the larger standard model, the still larger grand unified model, and beyond. In the energy scale of current experiments, the mathematical QED is an excellent model. However, we must keep our distance from the IR trajectory, by keeping the cutoff large but finite. As the cutoff approaches infinity, the renormalized charge will tend to zero, but with logarithmic slowness, since all divergences in QED are logarithmic. From a phenomenological point of view, the cutoff dependence is so weak as to be undetectable; however, the effective charge can be set at an arbitrary value by adjusting the cutoff.

A possible scenario for the physical QED, which is embedded in a larger model of unified interactions, is the following. The true trajectory of the theory may well lie on an UV trajectory, similar to the upper curve in Fig. 16.1. On energy scales of our experiments, however, the true trajectory may be close to an IR trajectory, similar to the lower curve in Fig. 16.1. The mathematical QED is modeled after the IR case, because that yields a simplier description.

16.10 WHY CORRECT THEORIES ARE BEAUTIFUL

Physicists are always sure that they possess the correct picture of the world, because their theories are not only "true" but also "beautiful."

In the Newtonian view, the world was made up of particles ruled by the elegant canonical laws, and that encompassed everything. As Laplace said, given the positions and velocities of all the particles of the universe at any one time, one could determine the course of the universe for all times.

With the discovery of electromagnetism, the "luminiferous ether" joined "ponderable matter" as ingredients of the universe, and the picture was complete. Lord Kelvin was of the opinion that physics in the next century would be concerned only with "the next decimal place."

But then came relativity and quantum mechanics, more impressive and beautiful than ever. "Quantum mechanics," said Dirac, "has explained all of chemistry and most of physics."

We have since progressed from atoms to nuclei to quarks in one direction, and from galaxies to black holes in the other, both heading toward the Planck scale. As always, some believe that the end is in sight.

How is it that our effective theories at different scales are so compelling as to make physicists think they are gods? The answer is that, like Aesop's mouse, they walk in front of a lion, and the lion is renormalization.

Because of renormalization, we can understand the observed world on its own terms, without reference to detailed structures of a lower level. Thus the proper way to calculate satellite orbits is to use Newtonian mechanics, even in the postquantum era. And it is better not to mention quarks when we describe superconductivity. The fact is, a system in coarse grain can appear completely different from that in fine grain. Looking at a certain canvas through a microscope, one perceives only paint pigments. But when viewed with the naked eye, a Rembrandt might emerge. At different ranges of length scales, the world will appear to be governed by different

fixed points, because of the phenomenon of crossover. A fanciful interpretation is shown in Fig. 16.7.

The task of the physicist has been to find out where our RG trajectory came from. This is done by guessing the Hamiltonian, and working out its low-energy consequences to compare with experiments. This process is perhaps what Einstein had in mind when he said [5]

> The axiomatic basis of theoretical physics cannot be abstracted from experience, but must be freely invented. . . . Experience may suggest the appropriate mathematical concepts, but they most certainty cannot be deduced from it.

Through "free inventions," physicists have had remarkable success in the seemingly impossible task of reverse coarse graining. In this they have been guided by the faith that a "true" theory must be "beautiful." This mysterious unity prompted Wigner [6] to wonder about the "unreasonable effectiveness" of mathematics, and Dirac[3] to extol formalism. We can perhaps understand it by noting that what one

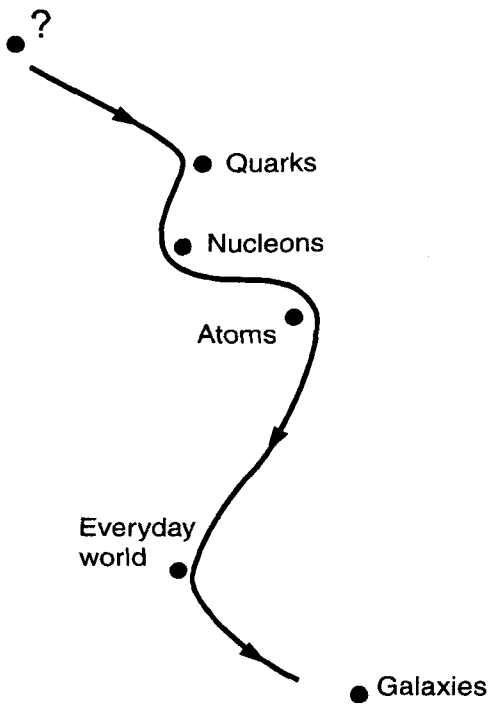


Figure 16.7 Under a change of length scale, the world's RG trajectory crosses over from the domain of one fixed point to another. This might explain why "true" theories are "beautiful."

[3]In a two-lecture series on creativity in physics at Harvard University in 1960, Dirac said in the first lecture that he was always guided by "correct formalism." But Heisenberg, in the second lecture, emphasized physical intuition.

really tries to do is to construct fixed points, which are purely mathematical objects endowed with a high degree of symmetry and universality.

PROBLEMS

16.1 **(a)** We have defined our field $\phi(x)$ with Fourier transform ϕ_k for finite volume, and $\tilde{\phi}(k)$ for infinite volume, such that the free action is dimensionless:

$$\frac{1}{2}\int d^dx\, \partial^\mu\phi(x)\partial^\mu\phi(x) = \frac{1}{2}\sum_k |\phi_k|^2 = \frac{1}{2}\int \frac{d^dk}{(2\pi)^d}|\tilde{\phi}(k)|^2$$

Verify the canonical dimensions for $\phi(x)$, ϕ_k, and $\phi(k)$.

(b) Show that, according to the canonical dimensions, the correlation function in the massless case should have the behavior

$$\langle\phi(x)\phi(0)\rangle \sim \frac{1}{|x|^{d-2}}$$

(c) A change of scale changes the cutoff from Λ to Λ/b without changing the physical content of the theory. In an interacting theory, this changes the free part of the action to $(b^{-\eta}/2)\int d^dx\, \partial^\mu\phi(x)\partial^\mu\phi(x)$, as indicated in (16.22). Show that

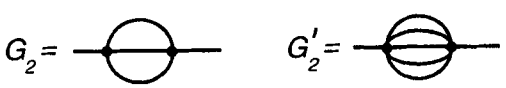
$$\langle\phi(x)\phi(0)\rangle \sim \frac{1}{|x|^{d-2+\eta}}$$

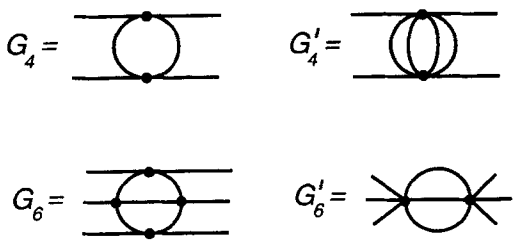
which gives rise to the term "anomalous dimension" for the exponent η.

16.2 The $p = 0$ component of ϕ_p does not appear in the action $\Sigma_p p^2|\tilde{\phi}(p)|^2$. Does this mean that ϕ_0 can be scaled differently from ϕ_p with $p \neq 0$? This would mean, in case of spontaneous symmetry breaking, that the vacuum field $\langle\phi\rangle$ scales differently from $\phi(x)$.

Solution. The answer is "no," for physical reasons. At finite x, $\phi(x)$ should be independent of the boundary conditions when the size of the box containing the system goes to infinity. Let boundary condition A be periodic, and B be that the field vanishes at the boundary. The vacuum field is constant in finite portions of space, but must drop to zero at the boundary. If a constant field scales differently from a nonconstant one, then we can tell the difference between A and B, even in finite portions of space. ■

16.3 Consider a scalar theory with interaction $g_4\phi^4 + g_6\phi^6$ in $d = 4$ space–time dimensions. Assuming that the only intrinsic scale is the cutoff momentum Λ, we have $g_4 = u_4$, $g_6 = u_6\Lambda^{-2}$, where u_α is dimensionless. In $d = 4$, g_4 is distinguished by the fact that it is dimensionless. As far as the S matrix is concerned, this theory is equivalent to a ϕ^4 theory with an effective coupling constant. To illustrate this fact, consider the various irreducible correlation functions through illustrative examples.





(a) Consider the graphs for the irreducible self-energy G_2, G_2' in the accompanying sketch. Show that they have the same degree of divergence Λ^2, as in pure ϕ^4 theory.

(b) Show that the irreducible vertex graphs G_4, G_4' diverge like $\ln \Lambda$, as in pure ϕ^4 theory.

(c) Show that the irreducible graph for the six-point function G_6 is convergent. The ones involving g_6, such as G_6', vanish when $\Lambda \to \infty$. This shows that the correlation function depends on g_6 only through the four-point function, as vertex insertions.

(d) Can we renormalize this theory like it is a ϕ^4 theory?

16.4 In a d-dimensional scalar theory with all powers ϕ^K in the interaction term, as represented in (16.1), there is a particular power M such that the coupling constant g_M is dimensionless. Show that, as far as the S matrix is concerned, this theory is equivalent to a ϕ^M theory with effective coupling constants dependent on all the g_K. To do this in a systematic way, repeat the analysis of Section 13.1 on the enumeration of primitive divergences. Show that the only primitively divergent graphs are those with M or fewer external lines.

REFERENCES

1. K. G. Wilson, *Rev. Mod. Phys.* **55,** 583 (1983). This is Wilson's Nobel lecture, which contains a historical account, and references to the original literature.
2. L. P. Kadanoff, *Physics,* **2,** 263 (1966).
3. K. Huang, *Statistical Mechanics*, 2nd ed., Wiley, New York, 1987, Section 18.7.
4. J. Polchinski, *Nucl. Phys.* **B231,** 269 (1984).
5. A. Einstein, *The World As I See It*, Covici Friede, New York, 1934, p. 34.
6. E. P. Wigner, *Commun. Pure Appl. Math.* **13**(1), 1 (1960).

CHAPTER SEVENTEEN

The Gaussian Fixed Point

17.1 STABILITY OF THE FREE FIELD

To study renormalization in greater depth, we investigate the space of Hamiltonians in the neighborhood of the one fixed point we know about—the Gaussian fixed point, corresponding to a massless free field with Hamiltonian density $\frac{1}{2}(\partial\phi)^2$.

We want to know, in particular, whether the fixed point is stable against perturbations. Were we to displace the system from the origin of parameter space along some direction, by adding extra terms to the action, would the system return to the origin under coarse graining, or would it go off on a tangent? *Stability* would mean that we have picked an "irrelevant" direction in parameter space, and *instability* would mean that the direction is "relevant." In the former case we would have "triviality," and in the latter, "asymptotic freedom." But since the parameter space was not there to begin with, we must invent it. The possibilities are vast, for the only condition is that the space contain the null element. In our investigation, the choice is dictated by what we can mathematically handle.

The simplest extension is to add a mass term $u_2\phi^2$. This is a relevant term, as we can see in Fig. 16.4. Under coarse graining, u_2 runs to infinity without reaching a fixed point. At smaller length scales, on the other hand, it approaches zero. This means that the mass is asymptotically free, which is hardly surprising, since it is intuitively clear that mass can be neglected at high energies.

The situation becomes more complicated when we introduce the simplest interaction $u_4\phi^4$. The parameter space is now spanned by u_2 and u_4, and the physical regions corresponds to the upper half plane $u_4 \geq 0$. The Gaussian fixed point is stable if $d \geq 4$, as indicated in Fig. 16.5; but it becomes unstable for $d = 4 - \epsilon\ (\epsilon \to 0^+)$, as indicated in Fig. 16.6. When d decreases from 4, what happens is that a nontrivial fixed point splits off from the Gaussian fixed point and moves to the upper half-plane. The Gaussian fixed point is now unstable, but the trajectories flowing from it are confined to an infinitesimal strip above the u_2 axis. When d changes in the opposite direction, the nontrivial fixed point moves toward the Gaussian fixed point,

merges with it at $d = 4$, and then moves off to the unphysical lower half-plane when $d > 4$.

In condensed-matter physics the interesting case is $d = 3$. In this case, the nontrivial fixed point is presumably located at a finite distance from the origin. The motivation for studying the case $d = 4 - \epsilon$ is the hope that the critical exponents at the nontrivial fixed point can be extrapolated to $\epsilon = 1$. For particle theory, however, the physical case is $d = 4$, and we are stuck with the fact that the ϕ^4 interaction is irrelevant. To avoid triviality, we must keep a finite cutoff Λ. Physical quantities depend only on $\ln \Lambda$, and are therefore not sensitive to changes in Λ. While this is acceptable from a phenomenological point of view, it does not seem very satisfactory. This provides an incentive to inquire whether there are relevant directions emanating from the Gaussian fixed point for $d \geq 4$.

17.2 GENERAL SCALAR FIELD

We shall study the real scalar field $\phi(x)$ in d Euclidean dimensions, whose action is given in (16.1). In terms of the dimensionless variables in (16.11), it has the form

$$S[\varphi] = \int d^d x [\tfrac{1}{2}(\partial\varphi)^2 + u_2\varphi^2 + u_4\varphi^4 + u_6\varphi^6 + \cdots] \tag{17.1}$$

In momentum space it has the form given in (16.13):

$$S[\varphi] = \frac{1}{2}\sum_{|q|<1}(q^2 + r)\varphi_q\varphi_{-q} + \frac{u_4}{\Omega}\sum_{|q_i|<1}\delta\Big(\sum q_i\Big)\varphi_{q_1}\varphi_{q_2}\varphi_{q_3}\varphi_{q_4}$$

$$+ \frac{u_6}{\Omega^2}\sum_{|q_i|<1}\delta\Big(\sum q_i\Big)\varphi_{q_1}\varphi_{q_2}\varphi_{q_3}\varphi_{q_4}\varphi_{q_5}\varphi_{q_6} + \cdots \tag{17.2}$$

where $r = 2u_2$. However, we shall start with a more general theory, to see whether the above form is closed under RG. In particular, we want to know whether RG generates derivative couplings and nonlocal interactions not included above. By *derivative coupling,* we mean terms containing derivatives of the field not of the form $(\partial\varphi)^2$, such as

$$\int d^d x(\partial^2\varphi)^2$$

A nonlocal term involves fields at different space–time points, such as

$$\int d^d x d^d y \varphi(x) K(x-y)\varphi(y)$$

Of course, in a cutoff field theory ostensibly "local" interactions are nonlocal within a spatial distance of order Λ^{-1}. By "nonlocal" terms, we specifically refer to

those for which the range of nonlocality is large compared to Λ^{-1}. We shall first consider a very general model containing arbitrary derivative and nonlocal couplings, in order to show that, in an infinitesimal RG, such interactions are not generated if they were not present originally. We shall then revert to the action of (16.1).

To generalize the action given in (16.13), we replace u_α by an arbitrary function of momenta

$$u_\alpha(q) \equiv u_\alpha(q_1, \ldots, q_\alpha) \tag{17.3}$$

Since this quantity multiplies the symmetric combination $\varphi_{q_1} \cdots \varphi_{q_\alpha}$, we may take it to be a symmetric function. The kinetic term $\frac{1}{2}\Sigma(q^2 + r)\varphi_q\varphi_{-q}$ is generalized by replacing r by the momentum-dependent quantity

$$\omega(q) = r + c_4 q^4 + c_6 q^6 + \cdots \tag{17.4}$$

Our starting point, then, is the action

$$\begin{aligned} S[\varphi] &= \sum_{\alpha=2}^{\infty} S_\alpha[\varphi] \\ S_2[\varphi] &= \tfrac{1}{2} \sum_{|q|<1} [q^2 + \omega(q)]\varphi_q\varphi_{-q} \\ S_\alpha[\varphi] &= \Omega^{1-\alpha/2} \sum_{|q_i|<1} \delta(q) u_\alpha(q) \varphi_{q_1} \cdots \varphi_{q_\alpha} \qquad (\alpha \geq 4) \end{aligned} \tag{17.5}$$

where $\delta(q)$ is an abbreviation for $\delta(q_1 + \cdots + q_\alpha)$. It will be understood that the sum over α extends over even integers only. This form is quite general. For example, in coordinate space the kinetic term has the form

$$S_2[\varphi] = \tfrac{1}{2} \int d^d x [(\partial\varphi)^2 + r\varphi^2 + c_4(\partial\varphi)^4 + c_6(\partial\varphi)^6 + \cdots] \tag{17.6}$$

where $(\partial\varphi)^{2n} = (\partial^j\varphi\partial^j\varphi)^n$. The other terms in the action contain similar derivative couplings; but in addition, nonlocal terms can arise. For example, $u_4(q)$ can contain a term of the form $\delta_{q_1,q_2}\delta_{q_3,q_4}s(q_1)\omega(q_3)$. This will give a nonlocal term with an infinite range of the nonlocality. But we are not actually going to use these interactions, and will not bother to impose physical constraints.

17.3 FEYNMAN GRAPHS

The RG has been described in Section 16.2. We introduce Feynman graphs in order to make certain arguments important for our later calculation. When the field is split into slow and fast parts with $\varphi = \sigma + f$, the kinetic part of the action is additive:

$$S_2[\sigma + f] = S_2[\sigma] + S_2[f] \tag{17.7}$$

because $\sigma_q f_{-q} = 0$, as their domains do not overlap. The partition function can be written

$$Z = \int D\varphi\, e^{-S[\varphi]} = \int D\sigma\, e^{-S_2[\sigma]} \int Df\, e^{-S_2[f] - S_I[\sigma+f]}$$

$$= \mathcal{N} \int D\sigma\, e^{-S_2[\sigma]} \langle e^{-S_I[\sigma+f]} \rangle_f \tag{17.8}$$

where $\mathcal{N}^{-1} = \int Df \exp\{-S_2[f]\}$ is a constant, and $\langle O \rangle_f$ denotes averaging over f with weight $\exp\{-S_2[f]\}$. The new action $\tilde{S}[\sigma]$ for the slow fields is given through

$$\langle e^{-S_I[\sigma+f]} \rangle_f \equiv e^{-\tilde{S}[\sigma]} \tag{17.9}$$

and after scaling transformations we extract the renormalized coupling constants.

To calculate $\tilde{S}[\sigma]$ in terms of Feynman graphs, we make the expansion

$$e^{-\tilde{S}[\sigma]} \equiv \langle e^{-S_I[\sigma+f]} \rangle_f = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \langle S_I^n[\sigma + f] \rangle_f \tag{17.10}$$

The interaction S_I is a sum of vertices illustrated in Fig. 17.1, where each line emanating from a vertex is a sum of one slow and one fast line. All the slow lines become external lines, and all the fast lines are internal lines to be integrated over. We substitute the sum of vertices into (17.10), and expand in powers of σ. The averaging with respect to f can be done using Wick's theorem:

$$\langle f_{q_1} \cdots f_{q_\alpha} \rangle_f = \text{sum of all contractions} \tag{17.11}$$

where a contraction is a pairing of the f values, with each pair contributing a factor

$$\langle f_p f_q \rangle_f = \frac{\delta_{pq}}{q^2 + \omega(q)} \tag{17.12}$$

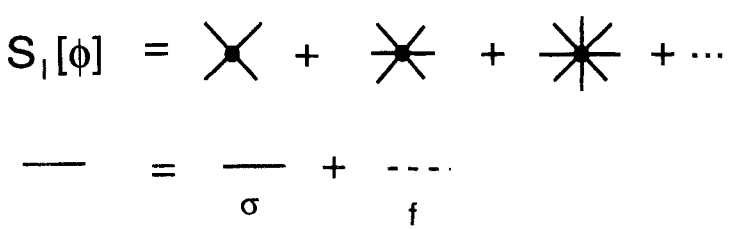


Figure 17.1 The interaction vertices. Each line is a sum of a slow line σ and a fast line f. The fast lines are internal lines to be integrated over, and the slow lines are external lines.

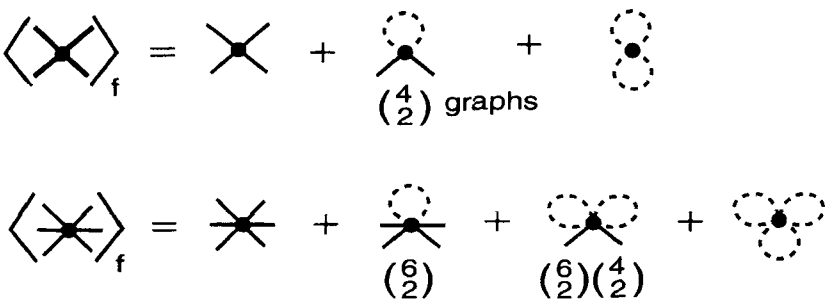


Figure 17.2 Graphs with n external lines contribute to the renormalized parameter u_n'. The vacuum graphs can be ignored.

The average of an odd number of f values is zero. This generates Feynman graphs in the usual way. For example, the first two interaction vertices give rise to the Feynman graphs shown in Fig. 17.2. Averaging the first vertex yields a four-line graph that contributes to u_4', and a two-line graph that contributes to u_2'. Similarly, averaging the second vertex yields contributions to u_6', u_4', u_2'. Vacuum graphs are irrelevant because they contribute only to the constant $\mathcal{N}$. With this expansion, we have the rules

$$\begin{aligned} e^{-\tilde{S}[\sigma]} &= \text{sum of all Feynman graphs} \\ -\tilde{S}[\sigma] &= \text{sum of all connected Feynman graphs} \end{aligned} \tag{17.13}$$

The second statement is the linked cluster theorem. We will not state the Feynman rules in detail, for we shall not need them for later calculations.

17.4 WEGNER–HOUGHTON FORMULA

Consider an infinitesimal RG transformation at the cutoff momentum. We put

$$b = e^t \qquad (t \to 0) \tag{17.14}$$

Wegner and Houghton [1] show that the renormalized action can be obtained exactly to order t, with the momentum-dependent interactions introduced earlier. What makes this possible is that all the internal momenta in Feynman graphs are confined to a thin shell ζ of thickness $O(t)$ in momentum space:

$$\zeta : 1 - t < |q| < 1 \tag{17.15}$$

Specifically, we have the following simplifications:

- To first order in t, only tree and one-loop graphs contribute. Some of the latter are shown in Fig. 17.3.
- A one-loop graph with two or more vertices is generally $O(t^2)$, except that it is $O(t)$ when all the internal lines carry exactly the same loop momentum. This requires that the total momentum of external lines at any one vertex be zero.

To show the second statement, consider graph c in Fig. 17.3, which is proportional to

$$\int_{q_1 \in \zeta} d^d q_1 \int_{q_2 \in \zeta} d^d q_2 \, \delta^d(p_1 + p_2 - q_1 - q_2) \delta^d(p_1' + p_2' - q_1 - q_2) \times \frac{u_4(p_1, p_2, q_1, q_2) u_4(q_1, q_2, p_1', p_2')}{v(q_1) v(q_2)}$$

where external momenta are denoted p, and internal momenta are denoted q. This integral is $O(t^2)$ in general, because there are two q integrations, and each ranges over a momentum shell of thickness $O(t)$. An exception occurs when $p_1 + p_2 = 0$. The integrations are then constrained by $\delta(q_1 + q_2)$, and the graph becomes $O(t)$. This argument applies to any vertex of a graph, even if it is a subgraph. ■

We have shown in Section 14.9 that the sum of tree and one-loop graphs is given by the first two terms in a saddle-point expansion of the partition function. In

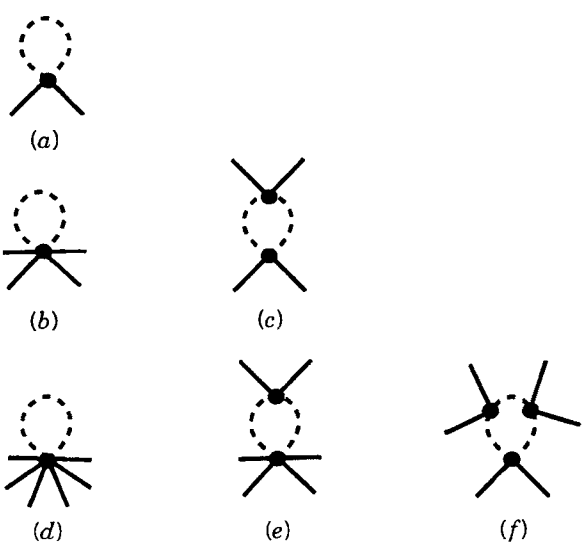


Figure 17.3 One-loop graphs.

this particular instance, the simplifications mentioned above make the calculation even easier. We expand the action in powers of f:

$$S[\sigma+f] = S[\sigma] + \sum_{q\in\zeta} P_q[\sigma] f_q + \tfrac{1}{2}\sum_{q\in\zeta} Q_q[\sigma] f_q f_{-q} + \cdots \tag{17.16}$$

with

$$P_q[\sigma] = \left.\frac{\partial S[\sigma+f]}{\partial f_q}\right|_{f=0}$$

$$Q_q[\sigma] = \left.\frac{\partial^2 S[\sigma+f]}{\partial f_q \partial f_{-q}}\right|_{f=0} \tag{17.17}$$

where $P_q[\sigma]$ is associated with tree graphs, while $Q_q[\sigma]$ is associated with one-loop graphs. The terms not displayed can be neglected.

Proof. Each f_q in the expansion corresponds to an internal line in a Feynman graph of momentum q. In the one-loop graphs we need, there emerge from every vertex two—and only two—internal lines with equal and opposite momenta (since the external lines must have total zero momentum). Thus we need not go beyond quadratic order in f_q, and of the quadratic term we only need to keep terms of the form $f_q f_{-q}$. ∎

Since $f_{-q} = f_q^*$, the modes q and $-q$ are not independent. We rewrite

$$\begin{aligned} S[\sigma+f] &= S[\sigma] + \sum_{q\in\zeta'} (P_q f_q + P_q^* f_q^*) + \sum_{q\in\zeta''} Q_q f_q^* f_q \\ &= S[\sigma] + \sum_{q\in\zeta'} \left[Q_q \left| f_q + \frac{P_q^*}{Q_q} \right|^2 - \frac{|P_q|^2}{Q_q} \right] \end{aligned} \tag{17.18}$$

where ζ' denotes a hemisphere of the thin shell ζ. We can now integrate over f:

$$\begin{aligned} Z &= \int D\sigma \int Df \exp\left\{ -S[\sigma] - \sum_{q\in\zeta'} \left[Q_q \left| f_q + \left(\frac{P_q^*}{Q_q} \right) \right|^2 - \frac{|P_q|^2}{Q_q} \right] \right\} \\ &= \int D\sigma\, e^{-S[\sigma]} \prod_{q\in\zeta'} e^{-|P_q|^2/Q_q} \frac{\pi}{Q_q} \equiv \mathcal{N} \int D\sigma\, e^{-\tilde{S}[\sigma]} \end{aligned} \tag{17.19}$$

The renormalized action to first order in t is given by the *Wegner–Houghton formula*

$$\begin{aligned} \tilde{S}[\sigma] &= S[\sigma] + tB[\sigma] \\ B[\sigma] &= \frac{1}{2t} \sum_{q\in\zeta} \left[\ln Q_q[\sigma] - \frac{|P_q[\sigma]|^2}{Q_q[\sigma]} \right] \end{aligned} \tag{17.20}$$

The quantity $B[\sigma]$ is $O(1)$ because the summation over the thin shell ζ is $O(t)$, and it is divided by t. This result is valid only to order t, but exact to all orders in the coupling constants $u_\alpha(q)$. With momentum-dependent couplings $u_\alpha(q)$, the theory is very general, but not all choices of $u_\alpha(q)$ are physical. What constraints are needed to make the theory physical is a problem that has not been investigated systematically.

17.5 RENORMALIZED COUPLINGS

We still have to transform to rescaled variables according to (16.26). To first order in t, we need only to transform $S[\sigma]$. From (17.5), we obtain

$$\begin{aligned} S_\alpha[\sigma] &= \Omega^{1-\alpha/2} \sum_{|q_i|<1/b} \delta(q) u_\alpha(q) \sigma_{q1} \cdots \sigma_{q\alpha} \\ &= \Omega^{1-\alpha/2} \sum_{|q_i'|<1} \delta(q') u_\alpha(q'/b) \sigma_{q_1'/b} \cdots \sigma_{q_\alpha'/b} \end{aligned} \tag{17.21}$$

Putting $b = 1 + t$ and keeping only first-order terms in t, we obtain

$$S_\alpha'[\varphi'] \equiv S_\alpha[\sigma] = \Omega^{1-\alpha/2} \sum_{|q_i'|<1} \delta(q') \varphi_{q_1'}' \cdots \varphi_{q_\alpha'}' \tag{17.22}$$

$$\times \left\{ 1 + t \left[d + \frac{\alpha}{2}(2 - d + \eta) - \sum_i q_i' \frac{\partial}{\partial q_i'} \right] \right\} u_\alpha(q') \tag{17.23}$$

where $\varphi_{q'}'$ is as defined in (16.25). The partition function can now be written as

$$Z = \mathcal{N} \int D\varphi' \, e^{-S'[\varphi']} \tag{17.24}$$

where

$$\begin{aligned} S'[\varphi'] &= S[\varphi'] + tB[\varphi'] + tC[\varphi'] \\ B[\varphi] &= \frac{1}{2t} \sum_{q \in \zeta} \left[\ln Q_q[\varphi] - \frac{|P_q[\varphi]|^2}{Q_q[\varphi]} \right] \\ C[\varphi] &= \sum_{\alpha=2}^{\infty} \Omega^{1-\alpha/2} \sum_{|q_i|<1} \delta(q) \varphi_{q_1} \cdots \varphi_{q_\alpha} \\ &\quad \times \left[d + \frac{\alpha}{2}(2 + \eta - d) - \sum_i q_i \frac{\partial}{\partial q_i} \right] u_\alpha(q) \end{aligned} \tag{17.25}$$

It is clear that if all $u_\alpha(q) = 0$, then the action is invariant under RG. This establishes the Gaussian fixed point

If the action does not contain odd powers of the field initially, then none will be generated to first order in t, because $Q_q[\varphi]$ is even in ω, as we can see from (17.17). If no derivative couplings were present initially, then to first order in t none will be induced. This can seen as follows. Derivatives are generated by momentum-dependent terms, and can occur only in $C[\varphi]$ in the term

$$\sum_i q_i \frac{\partial}{\partial q_i} u_\alpha(q)$$

If only nonderivative local couplings were present at the start, then the preceding vanishes except for $\alpha = 2$, for which it gives a term proportional to q^2. Therefore no derivative couplings are generated to first order in t. We have thus shown that the action (17.1) is closed under RG to first order in t.

To obtain the RG equations, we expand $B[\varphi]$ and $C[\varphi]$ in powers of φ to rewrite $S'[\varphi]$ in the form (17.1), and read off the new coupling functions $u'_\alpha(q)$. This is straightforward but messy, and we shall do it only to linear order in the couplings. But first, some general comments about the β function.

In terms of RG, the momentum q in $u_\alpha(q)$ is on the same footing as α, as it is a label for the type of interaction. To emphasize this, we write

$$u_{\alpha,q} \equiv u_\alpha(q) \tag{17.26}$$

The β function in our case is defined by

$$\beta_{\alpha,q}(u', u) \equiv \frac{1}{t}(u'_{\alpha,q} - u_{\alpha,q}) \tag{17.27}$$

where $u' - u = O(t)$ by construction. By considering a function $u_{\alpha,q}(t)$, we can write

$$\frac{du_{\alpha,q}(t)}{dt} = \beta_{\alpha,q}(u(t), u(0)) \tag{17.28}$$

where $u(t)$ stand for the set of all couplings at t and $t = 0$ corresponds to the cutoff scale. We have calculated β exactly to first order in t. The β function one uses in particle theory corresponds to

$$\beta_{\alpha,q}(u) = \beta_{\alpha,q}(u, u_*) \tag{17.29}$$

where u_* is a fixed point. To calculate this will require integrating the fast modes over a finite instead of an infinitesimal shell in momentum space. This seems difficult, but its equivalent has been achieved via perturbation expansions in special cases. Such calculations are the forte of perturbative renormalization, although in that context one is not aware of the role of u_*.

Our modest calculation has nontrivial consequences. By placing $u(0)$ at the Gaussian fixed point

$$u(0) = u_* = 0 \tag{17.30}$$

we obtain the *exact* β function at the Gaussian fixed point. This gives us the tangent vectors of all the trajectories at that point, and we can tell which directions are irrelevant or relevant.

17.6 THE RG MATRIX

We study the tangent vectors to RG trajectories at the Gaussian fixed point. For this, we need only the linearized RG equations. All tree graphs can be ignored, since they are at least of second order in the couplings. The only relevant graphs are the one-loop graphs with one vertex, the first graph in each row of Fig. 17.3 (the "octopuses"). It is easily verified that to this order $\eta = 0$.

We revert to the action (17.1), with momentum–space representation (17.1). From (17.2) we have

$$\begin{aligned} Q_q[\varphi] &= q^2 + r + \sum_{\alpha=4}^{\infty} \Omega^{1-\alpha/2} \sum_{|q_i|<1} \delta(q_1 + \cdots + q_\alpha) u_\alpha \\ &\quad \times \frac{\partial^2}{\partial f_q \, \partial f_{-q}} [(\varphi_{q_1} + f_{q_1}) \cdots (\varphi_{q_\alpha} + f_{q_\alpha})]|_{f=0} \\ &= q^2 + r + \tilde{Q} \end{aligned} \tag{17.31}$$

where

$$\tilde{Q} = \sum_{\alpha=2}^{\infty} (\alpha+1)(\alpha+2) \Omega^{-\alpha/2} u_{\alpha+2} \sum_{|q_i|<1} \delta(q_1 + \cdots + q_\alpha) \varphi_{q_1} \cdots \varphi_{q_\alpha} \tag{17.32}$$

For $B[\varphi]$ given in (17.25), we neglect the P_q term, and obtain

$$B[\varphi] = \frac{1}{2t} \sum_{q \in \zeta} \ln (q^2 + r + \tilde{Q}) \tag{17.33}$$

Since the shell ζ is thin, we set $q^2 = 1$, and replace the sum by the volume of ζ:

$$\sum_{q \in \zeta} = \frac{\Omega}{(2\pi)^d} t A_d \tag{17.34}$$

where A_d is the surface area of a unit d sphere, given by

$$S_d \equiv \frac{A_d}{(2\pi)^d} = \frac{2^{1-d}\pi^{-d/2}}{\Gamma(d/2)}$$

$$S_4 = (8\pi^2)^{-1} \tag{17.35}$$

Thus, to first order in u_α,

$$B[\varphi] = \frac{\Omega S_d}{2} \ln(1 + r + \tilde{Q}) = \frac{\Omega S_d}{2}[\ln(1 + r) + \tilde{Q}]$$

$$= \text{constant} + \frac{S_d}{2} \sum_{\alpha=2}^{\infty} (\alpha + 1)(\alpha + 2)\Omega^{1-\alpha/2} u_{\alpha+2} \sum_{|q_i|<1} \delta(q)\varphi_{q_1} \cdots \varphi_{q_\alpha} \tag{17.36}$$

The constant term can be ignored. From (17.25) we have, with $\eta = 0$,

$$C[\varphi] = \sum_{\alpha=2}^{\infty} \Omega^{1-\alpha/2} \sum_{|q_i|<1} \delta(q)\varphi_{q_1} \cdots \varphi_{q_\alpha} \left[d + \frac{\alpha}{2}(2 - d) \right] u_\alpha \tag{17.37}$$

which gives

$$B[\varphi] + C[\varphi] = \sum_{\alpha=2}^{\infty} \sum_{|q_i|<1} \Omega^{1-\alpha/2} \delta(q)\varphi_{q_1} \cdots \varphi_{q_\alpha}$$

$$\times \left\{ \frac{S_d}{2}(\alpha + 1)(\alpha + 2)u_{\alpha+2} + \left[d + \frac{\alpha}{2}(2 - d) \right] u_\alpha \right\} \tag{17.38}$$

Putting $\alpha = 2n$, we obtain the linearized RG equations [2]

$$\frac{du_{2n}}{dt} = (2n + d - nd)u_{2n} + S_d(n + 1)(2n + 1)u_{2n+2}$$

$$(n = 1, 2, \ldots, \infty) \tag{17.39}$$

Let ψ be the column matrix whose elements are $\psi_n = u_{2n}$:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix} \tag{17.40}$$

We can rewrite (17.39) in the form

$$\frac{d\psi}{dt} = M\psi \tag{17.41}$$

where M is the RG matrix:

$$M_{nm} = \delta_{nm}\left(n + d - 1 - \frac{nd}{2}\right) + \delta_{m,n+1}(n+1)\left(n + \frac{1}{2}\right)S_d \tag{17.42}$$

Now consider the eigenvalue problem

$$M\psi = \lambda\psi \tag{17.43}$$

The eigenvectors ψ correspond to "principal axes" in the parameter space, along which we have the behavior $d\psi/dt = \lambda\psi$, or

$$\psi(t) = \psi(t_0)e^{\lambda(t-t_0)} \tag{17.44}$$

The origin t_0 is arbitrary, except that it should be such that ψ is small; but it should not correspond to the Gaussian fixed point, where $\psi \equiv 0$.

The eigenvalue λ characterizes the trajectory tangent to the corresponding principal axis at the Gaussian fixed point:

1. If $\lambda < 0$, then $\psi \to 0$ as $t \to \infty$. The couplings constants are irrelevant. Under coarse graining, they tend to approach the Gaussian fixed point. On such a trajectory, the Gaussian fixed point is IR, and the theory is trivial.
2. If $\lambda > 0$, then ψ grows with t. The coupling constants are relevant. Under coarse graining, they tend to go away from the Gaussian fixed point. On such a trajectory the Gaussian fixed point is UV, and the theory is nontrivial. The trajectory is specified by some initial condition at an arbitrary point $t = t_0$, and it flows away from the Gaussian fixed point. The latter can be reached by letting $t \to -\infty$, in which limit the couplings vanish. This is asymptotic freedom.
3. The case $\lambda = 0$ corresponds to "marginal" coupling constants. In this case, we have to go beyond the linear approximation in order to determine the true behavior.

Using (17.39), we can put the eigenvalue equation (17.43) in the form

$$u_{2n+2} = \frac{n(d-2) - d + \lambda}{S_d(n+1)(2n+1)}u_{2n} \qquad (n = 1, 2, \cdots, \infty) \tag{17.45}$$

which is a recursion relation starting with $u_2 = r/2$. The case $d = 2$ will be treated separately. For $d > 2$, it is convenient to introduce a parameter a by writing the eigenvalue in the form

$$\lambda = 2 + (d-2)a \tag{17.46}$$

The recursion relation can then be put in the form

$$u_{2n+2} = \frac{(d-2)(a+n-1)}{2S_d(n+1)(n+\frac{1}{2})} u_{2n} \tag{17.47}$$

whose solution is

$$u_{2n} = \frac{r}{2}\left(\frac{d-2}{2S_d}\right)^{n-1} \frac{a(a+1)\cdots(a+n-2)}{n!(n-\frac{1}{2})(n-\frac{3}{2})\cdots\frac{3}{2}} \tag{17.48}$$

The potential with these coupling constants is referred to as the *eigenpotential.* Using the abbreviation

$$z = \frac{(d-2)\varphi^2(x)}{2S_d} \tag{17.49}$$

we have

$$U_a(\varphi^2(x)) \equiv \sum_{n=1}^{\infty} u_{2n}\varphi^{2n}(x) = \frac{rS_d}{2(a-1)(d-2)}[M(a-1, 1/2, z) - 1] \tag{17.50}$$

where $M(a, b, z)$ is the Kummer function [3]:

$$\begin{aligned} M(a, b, z) &= 1 + \frac{a}{b}\frac{z}{1!} + \frac{a(a+1)}{b(b+1)}\frac{z^2}{2!} + \cdots \\ &= \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_0^1 dt\, e^{zt}t^{a-1}(1-t)^{b-a-1} \end{aligned} \tag{17.51}$$

If a is a negative integer, the power series breaks off to become a polynomial of degree $|a|$. Otherwise, its asymptotic behavior for large z is given by

$$M(a, b, z) \approx \frac{\Gamma(b)z^{a-b}e^z}{\Gamma(a)}[1 + O(z^{-1})] \tag{17.52}$$

The eigenpotential $U_a(\phi^2)$ describes a field theory lying on a trajectory tangent to a particular principal axis with respect to the Gaussian fixed point. The principal axis is identified only through the eigenvalue parameter a.

For a polynomial potential of degree K in ϕ^2, then, we have $a = -K + 1$. The corresponding eigenvalues are

$$\lambda = 2[1 - (d-2)(K-1)] \qquad (K = 1, 2, \ldots) \tag{17.53}$$

The case $K = 1$ corresponds to a free field with squared mass r, which is relevant for all d. For $K \geq 2$, we have $\lambda < 0$ for $d \geq 4$. For $d = 3$ it is negative except for the marginal case of $K = 2$. This case corresponds to the ϕ^4 theory in $d = 4 - \epsilon$, with RG

flow as shown in Fig. 16.6. The Gaussian fixed point is, in fact, relevant in this case. The case $d = 2$ will be analyzed separately later.

In summary, all potentials that are polynomials in ϕ^2 lead to triviality for $d \geq 3$, except for the free field, and the ϕ^4 theory in $d = 3$.

17.7 NONTRIVIALITY AND ASYMPTOTIC FREEDOM

We investigate relevant directions for $d > 2$. They correspond to positive eigenvalues $\lambda > 0$, or

$$a > -\frac{2}{d-2} \tag{17.54}$$

The eigenpotentials have the following asymptotic behavior for large φ:

$$U(\varphi^2) \sim \exp\left[\frac{(d-2)\varphi^2}{2S_d}\right] \tag{17.55}$$

Nothing in canonical field theory rules out such a potential.

Sufficiently close to the Gaussian fixed point, the potential is proportional to r, which evolves in t according to

$$r(t) = r(t_0)e^{\lambda(t-t_0)} = Ce^{\lambda t} \tag{17.56}$$

with $C = r(t_0)\exp(-t_0)$. This is a running coupling constant, with an arbitrary renormalized value $r(t_0)$ at the reference point t_0. The theory is nontrivial, because the potential does not tend to zero in the low-momentum limit. Instead, we have asymptotic freedom, corresponding to the fact that the potential vanishes in the limit $t \to -\infty$, which corresponds to infinite momentum.

In order to have spontaneous symmetry breaking on the semiclassical level, the eigenpotential must have at least one minimum in ϕ. The power-series expansion for the eigenpotential reads

$$U_a(\varphi^2) = \frac{rS_d}{(d-2)}\left[z + \frac{az^2}{(3/2)2!} + \frac{a(a+1)z^3}{(3/2)(5/2)3!} + \cdots\right] \tag{17.57}$$

A sufficient condition is that $U'(0) < 0$, and $U > 0$ for large z. The first is satisfied by choosing $r < 0$. Asymptotically, U is proportional to $r[(a-1)\Gamma(a)]^{-1}$; the rest of the factors are positive. Thus we must have $(a-1)\Gamma(a) < 0$, which is equivalent to $\Gamma(a-1) < 0$. Using the formula $\Gamma(a)\Gamma(-a) = \pi/\sin(\pi a)$, and the fact that $\Gamma(a)$ is positive for $a > 0$, we find that a must be in one of the open intervals $(-1, -0)$, $(-3, -2)$, and so on. For a nontrivial theory, we have $\lambda > 0$, or $2 + (d-2)a > 0$. Combining these requirements, we obtain the sufficient condition

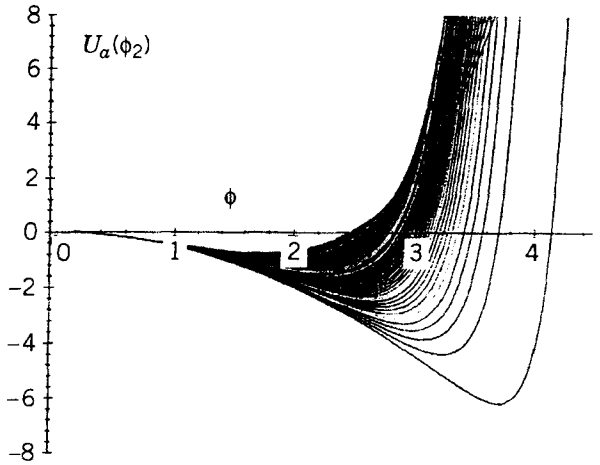


Figure 17.4 Eigenpotentials for $d = 4$. Theories with these potentials are covariant with respect to RG. The curves correspond, from top to bottom, to values of the eigenvalue parameter a uniformly spaced from –0.999 to –0.001. All potentials behave like exp ϕ^2 for large ϕ. The theories are all nontrivial and asymptotically free. The limiting case $a = -1$ is the ϕ^4 theory, which is trivial. (After Halpern and Huang [27].)

$$-\frac{2}{d-2} < a < 0 \tag{17.58}$$

A family of eigenpotentials for this range of a, and $d = 4$, is plotted in Fig. 17.4.

The eigenpotential U_a corresponds to a theory that lies on a trajectory tangent to a principal axis. Generally, we can consider a theory on an arbitrary trajectory, which is represented near the Gaussian fixed point by a linear superposition of the eigenpotentials. This gives us considerable freedom in choosing potentials.

17.8 THE CASE $d = 2$

We now calculate the eigenpotential for $d = 2$. Going back to (17.45), we have

$$U(\varphi^2) = \sum_{n=1}^{\infty} u_{2n}\varphi^{2n} \tag{17.59}$$

with the recursion relation

$$u_{2n+2} = \frac{2\pi(\lambda - 2)}{(n+1)(2n+1)} u_{2n} \tag{17.60}$$

where λ is the eigenvalue of the RG matrix. Write the recursion relation in the form

$$(2n^2 + 3n + 1)\, u_{2n+2} + \tfrac{1}{2}\gamma^2 u_{2n} = 0 \tag{17.61}$$

where

$$\gamma^2 \equiv 4\pi(2 - \lambda) \tag{17.62}$$

Multiplying both sides of the recursion equation by z^n, and summing over n from 1 to ∞, we obtain a differential equation for the eigenpotential:

$$2z\frac{d^2U}{dz^2} + \frac{dU}{dz} + \frac{1}{2}\gamma^2 U = \frac{r}{2} \tag{17.63}$$

where $z = \varphi^2$. We seek the solution that satisfies

$$U \xrightarrow[\varphi \to 0]{} \tfrac{1}{2} r\varphi^2 \tag{17.64}$$

Changing variables back to $\varphi = \sqrt{z}$, we have

$$\frac{d^2U}{d\varphi^2} + \gamma^2 U = r \tag{17.65}$$

The solution is

$$U(\varphi^2) = -\frac{r(t)}{\gamma^2}[\cos(\gamma\varphi) - 1]$$
$$r(t) = r(0)e^{2[1-(\gamma^2/8\pi)]t} \tag{17.66}$$

For $\gamma^2 > 0$, the eigenpotential is periodic, and thus ω is an angular variable. The Euclidean action

$$S[\varphi] = \int d^2x[\tfrac{1}{2}(\partial\varphi)^2 + U(\varphi^2(x))] \tag{17.67}$$

leads to the equation of motion

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\varphi - \frac{r(t)}{\gamma}\sin(\gamma\varphi) = 0 \tag{17.68}$$

which is called the "sine Gordon equation." For $0 < \gamma^2 < 8\pi$, the potential is relevant, and the theory is asymptotically free. For $\gamma^2 > 8\pi$, it is irrelevant, and the theory is trivial. These conclusions agree with those arrived at through perturbation renormalization [4].

As we shall see in the next chapter, the angular nature of the field dictates the

existence of vortices, which cannot be described within the linear approximation considered here.

PROBLEMS

17.1 Show that the anomalous dimension has the value $\eta = 0$ to one-loop order. (*Hint:* Obtain η from the Wegner–Houghton formula, which represents a sum of one-loop graphs.)

17.2 Consider a Ginsburg–Landau theory with free energy $\frac{1}{2}\int d^dx[(\partial\varphi)^2 + r\varphi^2]$. Show that $d^2 \ln Z/dr^2 = Cr^{d-4}$. (If we regard r as a temperature, this gives the heat capacity.) This shows that $d = 4$ is a "critical dimension" of the Gaussian fixed point, in that perturbation theory in powers of r breaks down for $d \leq 3$. This is why, in the theory of critical phenomena, one considers $d = 4 - \epsilon$, and uses a double perturbation expansion in r and ϵ. The $d = 2$ case in the last section is in a different category, for the $r\varphi^2$ term is replaced by $\cos\varphi$, making φ an angle. This is discussed in Chapter 18.

17.3 For $d = 2$, give the eigenpotential and the corresponding equations of motion for the field, when $\gamma = i\kappa$. Argue that this result is valid by analytic continuation from real γ.

REFERENCES

1. F. J. Wegner and A. Houghton, *Phys. Rev.* **A8,** 401 (1973).
2. K. Halpern and K. Huang, *Phys. Rev.* **53,** 3252 (1996).
3. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions,* National Bureau of Standards, Washington, DC, 1964, p. 503.
4. S. Coleman, *Phys. Rev.* **11,** 2088 (1975).

CHAPTER EIGHTEEN

In Two Dimensions

18.1 ABSENCE OF LONG-RANGE ORDER

We consider systems in two Euclidean dimensions, which may be statistical systems in two spatial dimensions, or quantum field theories in one spatial dimension. These interpretations are merely different aspects of the same partition function in the language of path integrals. For definiteness, we phrase our discussion in terms of statistical mechanics.

From a physical point of view, the essence of two-dimensionality is that long-range order cannot be maintained over arbitrarily large distances, due to large fluctuations of the Goldstone mode [1]. An equivalent statement is that spontaneous breaking of a continuous symmetry is impossible [2]. We shall illustrate this in a simple model.

Consider a collection of atoms in a plane, in which a crystal of size $L \times L$ begins to form. Let x denote an equilibrium lattice site and $\mathbf{u}(x)$, the deviation of an atom's instantaneous position from x. We can make the decomposition

$$\mathbf{u}(x) = \frac{1}{L^2} \sum_{\lambda,k} e^{ikx} \mathbf{q}_\lambda(k) \tag{18.1}$$

where $q_\lambda(k)$ is the amplitude of a normal mode of type λ, of wave vector k. In thermal equilibrium, the amount of energy residing in this mode is given by

$$E_\lambda(k) = \tfrac{1}{2}\, \omega_\lambda^2(k) |\mathbf{q}_\lambda(k)|^2 \tag{18.2}$$

where $\omega_\lambda(k)$ is the normal frequency. At absolute temperature T, this should be equal to T by equipartition, in units with Boltzmann's constant set to unity. Thus

$$|\mathbf{q}_\lambda(k)|^2 = \frac{2T}{\omega_\lambda^2(k)} \tag{18.3}$$

[1]Spontaneous breaking of a discrete symmetry is possible, as in the 2D Ising model.

For large L, the mean-square displacement is given by

$$\langle \mathbf{u}^2 \rangle = 2T \sum_{\lambda} \int \frac{d^2k}{(2\pi)^2} \frac{1}{\omega_\lambda^2(k)} \tag{18.4}$$

Since formation of a crystal means that translational invariance is broken, there exists a Goldstone mode whose frequency $\omega_0(k)$ approaches zero in the limit $k \to 0$. This corresponds to lattice phonons, with $\omega_0(k) = ck$, where c is the velocity of sound. Thus we have a lower bound:

$$\langle \mathbf{u}^2 \rangle > 2T \int \frac{d^2k}{(2\pi)^2} \frac{1}{\omega_0^2(k)} = \frac{T}{\pi c^2} \int_{2\pi/L}^{\Lambda} \frac{dk}{k} \tag{18.5}$$

where Λ is the inverse lattice spacing. As $L \to \infty$, the integral diverges like $\ln L$. This shows that crystalline order cannot be maintained over long distances, for it will be disrupted by long-wavelength density fluctuations.

18.2 TOPOLOGICAL ORDER

Although long-range correlations are eventually disrupted, they can extend over a considerable distance, because $\langle \mathbf{u}^2 \rangle$ diverges only logarithmically. In fact, there can exist patches of crystalline structure of macroscopic size. The boundaries of these patches are domain walls, which can be modeled by removing lines of atoms from the lattice, leaving chasms across which the atomic interactions are altered. Domain walls so created are illustrated in Fig. 18.1. The endpoint of a domain wall is a "dislocation center," or "center" for short. To find a dislocation center through a site-by-site inspection would be very tedious; however, there is an easier way. Let us go around a circuit made up of successive lattice steps, such that we make n steps in each of the four directions. For example, Fig. 18.1*a* shows a circuit with $n = 3$, starting at A. On a perfect lattice a circuit will return to the starting point, but it does not close if we go around a dislocation center. The difference between the starting point A and the endpoint B is called the *Burgers vector.* When we go around a single dislocation center, the Burgers vector is either +1 or −1, as illustrated in Figs. 18.1*a* and 18.1*b*. We call the +1 case a "center," and other an "anticenter." In general, the Burgers vector of a circuit is the sum of contributions from the centers enclosed. It is a topological property independent of the size and shape of the circuit, as long as it goes around the same set of centers. We now imagine that the imperfect crystal is very large, so we can draw a circuit C of macroscopic size, on whose scale the lattice appear to be almost a continuum. We associate with C a *topological order parameter* defined as

$$\Phi(C) = \langle N_+ - N_- \rangle \tag{18.6}$$

where $N_\pm$ is respectively the number of centers and anticenters enclosed by C, and

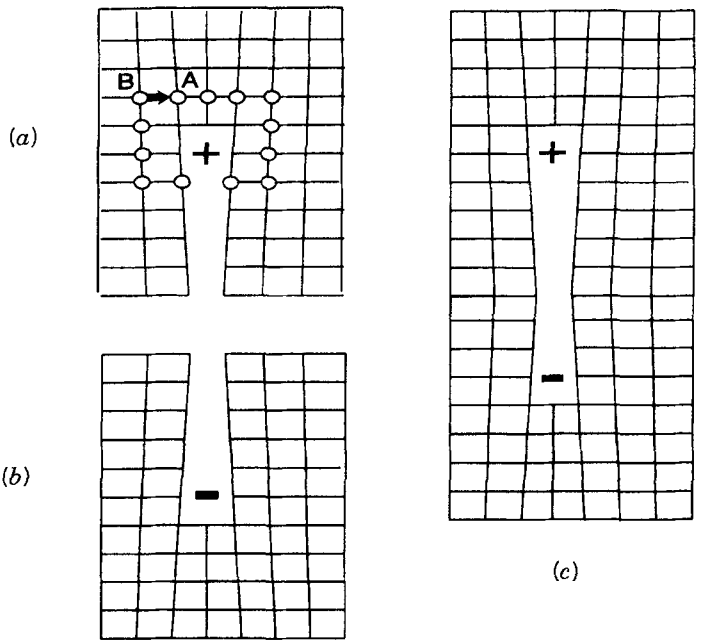


Figure 18.1 Dislocations in a 2D lattice. A circuit around a dislocation fails to close. The difference $A - B$ between the start and finish is called the *Burgers vector.*

$\langle\ \rangle$ denotes thermal average. This is a loop function, and quite different from a local order parameter.

A isolated center can cost a lot of energy, because a domain wall must extend from the center to the surface of the system, and there is a fixed energy per unit length. On the other hand, the domain wall that connects a center with an anticenter has energy proportional to their separation. Thus, in an infinitely large system, there must be an equal number of centers and anticenters, tied in pairs by domain walls. These pairs may form an "ionized" gas, or a collection of tightly bound "dipoles," as illustrated in Fig. 18.2. These configurations are characterized by different behaviors of the topological order parameter.

In the ionized state, the centers are uniformly distributed. The probability of finding a center or anticenter inside C is proportional to the area enclosed by C. The average value of $N_+ - N_-$ is therefore proportional to the statistical fluctuation $\sqrt{N_+}$, the square root of the area, or the perimeter $L(C)$ of the closed loop.

In the bound state, on the other hand, a bound pair makes no contribution to the order parameter, unless it is cut through by C. The number of pairs cut is proportional to $L(C)$, and a cut pair contributes ± 1 with equal probability. Therefore, the average contribution is proportional to the fluctuation $\sqrt{L(C)}$.

Thus, up to a proportionality constant, we obtain

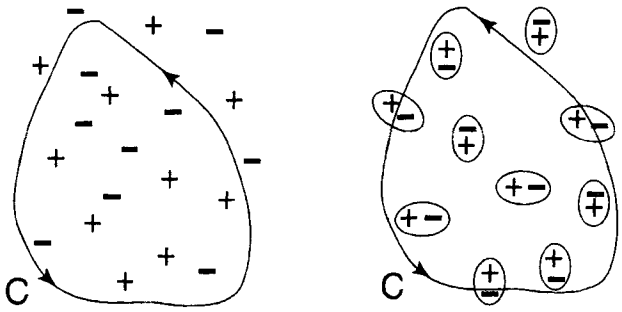


Figure 18.2 (*a*) Ionized gas of dislocation centers and anticenters; (*b*) bound pairs.

$$\Phi(C) = \begin{cases} L(C) & \text{(ionized state)} \\ \sqrt{L(C)} & \text{(bound state)} \end{cases} \tag{18.7}$$

where $L(C)$ is the perimeter of the loop C. The two cases become distinct for large loops, such that $L(C) \to \infty$. In a thermodynamic treatment of this system, one finds there is a phase transition between a low-temperature bound phase and a high-temperature ionized phase. In the bound phase the system responds to an external stress like an elastic solid, whereas the ionized phase cannot support any stress at all, but flows like a liquid. This model forms the basis of a theory of 2D melting [2]. We shall discuss the phase transition in the equivalent *XY* model.

18.3 *XY* MODEL

The XY model consists of a 2D square lattice of classical spins $\mathbf{s}_i$ of unit length, which are free to rotate in the x–y plane. The energy of the system is given by

$$E_{\mathrm{XY}}[\theta] = -J \sum_{\langle ij \rangle} \mathbf{s}_i \cdot \mathbf{s}_j = -J \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j) \tag{18.8}$$

where θ_i is the angle of $\mathbf{s}_i$ with respect to some fixed axis and $\langle ij \rangle$ denotes a nearest-neighbor pair. This is invariant under a simultaneous rotation of all the spins. The partition function is given by

$$Z_{XY} = \prod_i \int_0^{2\pi} d\theta_i \, e^{-\beta E_{\mathrm{XY}}[\theta]} \tag{18.9}$$

where β is the inverse temperature. The absence of spontaneous magnetization in 2D can be proved rigorously, and is known as the *Mermin–Wagner theorem* [3]. Our

physical discussion earlier shows that there can be large but finite patches of magnetization; however, long-range correlation is disrupted by defects, which in this case are vortices, as illustrated in Fig. 18.3.They have the property that the spin rotates through 2π around a vortex, and -2π around an antivortex. They bear a one-to-one correspondence with dislocation centers and anticenters, and we can take over the idea of topological order from the 2D crystal.

The XY model has been studied in great detail [4]; but we are interested only in the continuum limit, in which we can make the replacement

$$\cos(\theta_i - \theta_j) \approx 1 - \frac{1}{2}(\theta_i - \theta_j)^2 \rightarrow 1 - \frac{a^2}{2}(\nabla\theta)^2 \tag{18.10}$$

where a is the lattice spacing. This leads to an energy functional

$$E[\theta] = \frac{1}{2}\rho_0 \int d^2x(\nabla\theta)^2 \tag{18.11}$$

where ρ_0 is a constant. This looks just like a free field theory, and we seem to have lost the vortices. However, there is an important difference, namely, θ has physical meaning only modulo (2π). The system has a kind of gauge invariance, and, as pointed out by Berezinskii [5], this gives rise to vortices.

To explain the peculiarities of an angular field, consider the Ginsburg–Landau (GL) functional

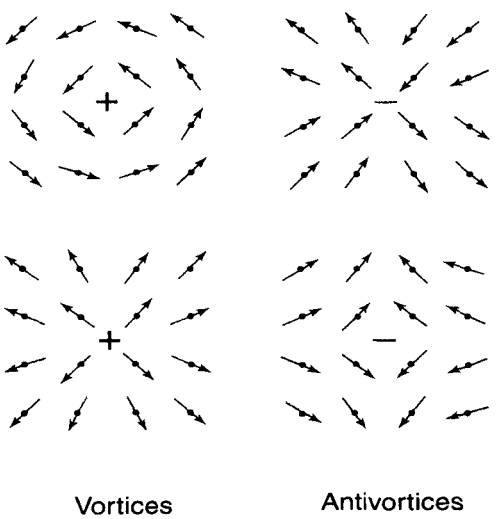


Figure 18.3 Vortices and antivortices in the XY model. Lower panel shows equivalent configurations after a global spin rotation through 90°.

$$E_{\mathrm{GL}}[\psi] = \int d^2x[\tfrac{1}{2}|\nabla\psi|^2 + u_2|\psi|^2 + u_4|\psi|^4] \tag{18.12}$$

where ψ is a complex field. By putting

$$\psi = \sqrt{\rho}e^{i\theta} \tag{18.13}$$

we can rewrite

$$E_{\mathrm{GL}}[\rho, \theta] = \int d^2x\left[\frac{1}{2}\rho(\nabla\theta)^2 + \frac{1}{8\rho}(\nabla\rho)^2 + V(\rho)\right]$$

$$V(\rho) = u_2\rho^2 + u_4\rho^4 \tag{18.14}$$

Now choose u_2 and u_4 in such a manner as to trap ρ at a steep minimum at $\rho_0 = \sqrt{-u_2/u_4}$, with

$$V'(\rho_0) = 0 \qquad V''(\rho_0) \gg 1 \tag{18.15}$$

With this, we can identify our continuum model with a limit of the GL model:

$$E[\theta] = E_{\mathrm{GL}}[\rho_0, \theta] + \text{Const.} \tag{18.16}$$

With this, it is convenient to take the spin variable as $\psi = \sqrt{\rho_0}e^{i\theta}$, as it is a complex representation of a planar spin. The spin–spin correlation function has been calculated by Berezinskii [5]:

$$\langle e^{i\theta(x)}e^{-i\theta(0)}\rangle = \frac{\int D\theta\, e^{-\beta E[\theta]}e^{i[\theta(x)-\theta(0)]}}{\int D\theta\, e^{-\beta E[\theta]}} = \left(\frac{|x|}{a}\right)^{-1/(2\pi\beta\rho_0)} \tag{18.17}$$

This is equal to $\langle\cos(\theta(x) - \theta(0))\rangle$, because $\langle\sin(\theta(x) - \theta(0))\rangle = 0$. This correlation function is "gauge-invariant," for it depends only on θ modulo (2π).

Consider the gauge-variant correlation function (see Problem 14.5):

$$\langle\theta(x)\theta(0)\rangle = \frac{\int D\theta\, e^{-\beta E[\theta]}\theta(x)\theta(0)}{\int D\theta\, e^{-\beta E[\theta]}} = \frac{1}{2\pi\beta\rho_0}\ln\frac{|x|}{a} \tag{18.18}$$

Formally this is just a correlation function of a free field in 2D. It says that the probable value of $\theta(x)$ increases with distance, when it is fixed at $x = 0$. If $\theta(x)$ were an ordinary scalar field, this would simply mean that the field amplitude gets larger and larger. But since $\theta(x)$ is an angle, this says that it makes an increasing number of revolutions as $|x|$ increases. Sufficiently far the origin, therefore, $\theta(x)$ mod(2π) will be randomized, and $e^{i\theta(x)}$ will average to zero. Therefore, in an infinite 2D system

$$\langle \psi \rangle = 0 \tag{18.19}$$

This statement is equivalent to the Mermin–Wagner theorem, and has been proven more rigorously by Hohenberg [6] and Coleman [7].

Consider now two paths P and Q with the same endpoints, as illustrated in Fig. 18.4. The phase change along P and Q may differ by a multiple of 2π, since only θ mod(2π) has physical meaning:

$$\left(\int_Q - \int_P\right) d\mathbf{s}\cdot\nabla\theta = 2\pi n \qquad (n = 0, \pm 1, \pm 2, \ldots) \tag{18.20}$$

We can regard the two paths as a closed path $C = Q - P$. When C is continuously deformed, the line integral cannot change continuously, but must jump by units of $\pm 2\pi$. These quanta are vortices or antivortices, defined by

$$\oint_C d\mathbf{s}\cdot\nabla\theta_\pm = \pm 2\pi \tag{18.21}$$

For counterclock-wise C, the sign $+1$ (-1) corresponds to a vortex (antivortex). We shall refer to the quanta generically as "vortices," and regard an antivortex as a vortex with vorticity -1.

The loop C cannot be shrunken to a point, because the angle becomes ill-defined. Therefore the function $\theta_\pm(x)$ has singularities at the location of the vortices, and we need a short-distance cutoff. This is introduced by decreeing that there is a "vortex core" of radius a, inside of which we set $\rho_0 = 0$. This renders the space effectively nonsimply connected.

An explicit solution for the vortex field is

$$\theta_\pm(r, \varphi) = \pm\varphi \tag{18.22}$$

where (r, φ) are cylindrical coordinates with respect to the vortex center. The reference axis $\varphi = 0$ is arbitrary, and changing this axis is a gauge transformation. That the phase angle increases by $\pm 2\pi$ around the vortex is the hallmark of a "gauge soliton": A transformation in spatial coordinates induces a "gauge transformation" of the internal coordinates. The velocity field of a vortex is gauge-invariant:

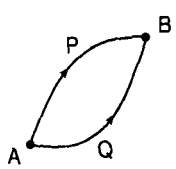


Figure 18.4 The phase change along P and Q may differ by an integer multiple of 2π. signifying the presence of vortices within the close path $Q - P$.

$$\nabla\theta_{\pm}(r, \varphi) = \pm\frac{\hat{\boldsymbol{\varphi}}}{r} \tag{18.23}$$

where $\hat{\boldsymbol{\varphi}}$ is the unit vector in the φ direction. The velocity field is tangent to circles about the vortex center. The velocity in this case corresponds to the difference between neighboring spins in Fig. 18.3.

The energy of a vortex is, up to an additive constant,

$$\varepsilon_{\pm} = \frac{\rho_0}{2}\int d^2x|\nabla\theta_{\pm}|^2 = \pi\rho_0 \ln\frac{R}{a} \tag{18.24}$$

where R is the radius of the whole system. The energy of an isolated vortex diverges as $R \to \infty$; but that of a dipole pair remains finite, because the corresponding velocity field is short-ranged. Thus, in an infinite system there must be equal numbers of vortices and antivortices. The vortices are just like dislocation centers in a 2D crystal, and we can take over the definition of the topological order parameter (18.6), with the property (18.7).

18.4 KOSTERLITZ–THOULESS TRANSITION

To see whether a system of vortices can undergo a phase transition, consider an ionized gas of $N/2$ vortices and $N/2$ antivortices, with uniform density. The free energy of the system at temperature T is given by

$$F = E - TS \tag{18.25}$$

where E is the total energy:

$$E = N\pi\rho_0 \ln\frac{R}{a} \tag{18.26}$$

and S is the total entropy, the logarithm of the number of ways to distribute the vortices and antivortices in space. The number of ways to place one vortex is $(R/a)^2$, the ratio of total area to the area of the core. Thus

$$S = N \ln\left(\frac{R}{a}\right)^2 \tag{18.27}$$

The free energy is thus

$$F = N(\pi\rho_0 - 2T)\ln\frac{R}{a} \tag{18.28}$$

This shows that there is a critical temperature

$$T_c = \frac{\pi\rho_0}{2} \tag{18.29}$$

For $T > T_c$ the free energy decreases with increasing t, favoring the creation of free vortices. If $T < T_c$, however, the free energy is minimum at $N = 0$. That means that the free vortices will form tightly bound pairs. This phase transition, which marks a change in topological order, is known as the *Kosterlitz–Thouless transition* [8].

18.5 VORTEX MODEL

We can now describe the continuum limit of the XY model taking vorticity into account, in a manner similar to that described in Section 15.5. Since the vortex core renders the space non–simply connected, $\nabla\theta$ has both longitudinal and transverse parts:

$$\nabla\theta = \nabla\omega + \nabla \times \mathbf{A} \tag{18.30}$$

where $\nabla\omega$ is a regular function, representing spin-wave contributions, while $\nabla \times \mathbf{A}$ is singular, coming from vortices. We can obtain $\mathbf{A}$ from the expression for $\nabla\theta_{\pm}$, but it is easier to find it from the defining condition

$$\oint_C d\mathbf{s}\cdot\nabla \times \mathbf{A} = 2\pi n \tag{18.31}$$

For pointlike vortices, this can be rewritten in differential form:

$$\nabla^2\mathbf{A} = -\hat{\mathbf{z}}n(\mathbf{x}) \tag{18.32}$$

where $\hat{\mathbf{z}}$ is the unit vector normal to the plane of the system and $n(\mathbf{x})$ is the "vortex charge density":

$$n(\mathbf{x}) = \sum_{i=1}^{N} 2\pi p_i \delta^2(\mathbf{x} - \mathbf{r}_i)$$

The vortices are centered at $\mathbf{r}_1 \cdots \mathbf{r}_N$, with "charges" $p_i = +1$ for vortex, -1 for antivortex. Using the 2D Green's function

$$G(\mathbf{x}) = -\frac{1}{2\pi} \ln \frac{|\mathbf{x}|}{a} \tag{18.33}$$

we obtain

$$\mathbf{A}(\mathbf{x}) = \hat{\mathbf{z}} \int d^2y\, G(\mathbf{x} - \mathbf{y}) n(\mathbf{y}) = -\hat{\mathbf{z}} \sum_{i=1}^{N} p_i \ln \frac{|\mathbf{x} - \mathbf{r}_i|}{a} \tag{18.34}$$

This holds outside vortex cores. Inside any vortex core, $\mathbf{A}(\mathbf{x}) = 0$.

In the presence of N vortices, the energy functional (18.11) should be generalized to

$$\begin{aligned} E_N[\theta] &= \frac{\rho_0}{2} \int d^2x |\nabla\theta|^2 + N\mu \\ &= \frac{\rho_0}{2} \int d^2x [|\nabla\omega|^2 + |\nabla \times \mathbf{A}|^2] + N\mu \end{aligned} \tag{18.35}$$

where μ is the chemical potential—the energy required to create a vortex core. We can calculate the vortex contribution more explicitly:

$$\begin{aligned} \frac{\rho_0}{2} \int d^2x (\nabla \times \mathbf{A})^2 &= \frac{\rho_0}{2} \int d^2x\, d^2y\, n(\mathbf{x}) G(\mathbf{x} - \mathbf{y}) n(\mathbf{y}) \\ &= -2\pi\rho_0 \sum_{i<j} p_i p_j \ln \frac{|\mathbf{r}_i - \mathbf{r}_j|}{a} \end{aligned} \tag{18.36}$$

Since the total vorticity should be zero for an infinite system,

$$\sum_{i=1}^{N} p_i = 0 \tag{18.37}$$

We see that the vortex system is equivalent to a neutral 2D Coulomb gas.

The energy functional can now be represented in the form given by Kosterlitz [9]:

$$\begin{aligned} E_N[\theta] &= E_{\text{wave}}[\omega] + E_{\text{vortex}}(\mathbf{r}_1 \cdots \mathbf{r}_N) \\ E_{\text{wave}}[\omega] &= \frac{\rho_0}{2} \int d^2x |\nabla\omega|^2 \\ E_{\text{vortex}}(\mathbf{r}_1 \cdots \mathbf{r}_N) &= -2\pi\rho_0 \sum_{i<j} p_i p_j \ln \frac{|\mathbf{r}_i - \mathbf{r}_j|}{a} + N\mu \end{aligned} \tag{18.38}$$

The partition function is given by

$$Z = Z_{\text{wave}} Z_{\text{vortex}}$$

$$Z_{\text{wave}} = \int D\omega \exp\left[-\frac{\beta\rho_0}{2} \int d^2x (\nabla\omega)^2 \right]$$

$$Z_{\text{vortex}} = \sum_{N=0}^{\infty} \frac{z^N}{[(N/2)!]^2} \sum_{\{p_i\}} \int \frac{d^2r_1}{a^2} \cdots \frac{d^2r_N}{a^2} \exp\left[2\pi\beta\rho_0 \sum_{i<j} p_i p_j \ln \frac{|\mathbf{r}_i - \mathbf{r}_j|}{a} \right] \tag{18.39}$$

where

$$z \equiv e^{\beta\mu} \tag{18.40}$$

is the vortex fugacity. The sum over N extends over even integers only, the sum over p_i is subject to $\Sigma p_i = 0$, and the r integrations are subject to $|\mathbf{r}_i - \mathbf{r}_j| > a$. As is usual in calculating grand partition functions, we keep the volume Ω large but finite, and take the limit $\Omega \to \infty$ in physical quantities, such as $\Omega^{-1} \ln Z_{\text{vortex}}$.

In the limit $z \to 0$ we have a dilute gas of vortices. To order z^2, we need only the trivial case $N = 0$, and the case $N = 2$, corresponding to one vortex–antivortex pair. Since $p_1 p_2 = -1$, we have to this order

$$Z_{\text{vortex}} = 1 + z^2 \int_{|\mathbf{r}_1 - \mathbf{r}_2| > a} \frac{d^2r_1\, d^2r_2}{a^4} \left[\frac{|\mathbf{r}_1 - \mathbf{r}_2|}{a} \right]^{-2\pi\beta\rho_0} + O(z^4) \tag{18.41}$$

The mean-square radius of a vortex-antivortex pair is

$$\begin{aligned} \langle r^2 \rangle &= \frac{1}{2\pi\Omega} \int_{|\mathbf{r}_1 - \mathbf{r}_2| > a} \frac{d^2r_1\, d^2r_2}{a^4} \left[\frac{|\mathbf{r}_1 - \mathbf{r}_2|}{a} \right]^{-2\pi\beta\rho_0} |\mathbf{r}_1 - \mathbf{r}_2|^2 \\ &= \int_1^{\infty} du\, u^{3-2\pi\beta\rho_0} = \frac{T}{4T_c} \left(1 - \frac{T}{T_c} \right)^{-1} \end{aligned} \tag{18.42}$$

which diverges at the critical temperature $T_c = \pi\rho_0/2$. This value agrees with (18.29) arrived at by an intuitive argument. The fugacity expansion is valid only when the vortices are bound in pairs, and fails when they become ionized.

18.6 2D SUPERFLUIDITY

We take the formula for the superfluid density from Appendix B:

$$\rho_s = \rho_0 - \frac{\beta\rho_0^2}{D-1} \int d^D x\, \langle \mathbf{v}_T(\mathbf{x}) \cdot \mathbf{v}_T(0) \rangle \tag{18.43}$$

where

$$\mathbf{v}_T = \nabla \times \mathbf{A} \tag{18.44}$$

Taking the Fourier transform of $\mathbf{v}_T$ in a large periodic box of volume Ω, we have

$$\mathbf{v}_T(\mathbf{x}) = \frac{1}{\Omega}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}\tilde{\mathbf{v}}_T(\mathbf{k})$$

$$\tilde{\mathbf{v}}_T(\mathbf{k}) = \int d^D x\, e^{-i\mathbf{k}\cdot\mathbf{x}}\,\mathbf{v}_T(\mathbf{x}) \tag{18.45}$$

Then

$$\rho_s = \rho_0 - \frac{\beta\rho_0^2}{(D-1)\Omega}\lim_{k\to 0}\langle\tilde{\mathbf{v}}_T(\mathbf{k})\cdot\tilde{\mathbf{v}}_T(-\mathbf{k})\rangle \tag{18.46}$$

Specializing to $D = 2$, we have

$$\tilde{\mathbf{v}}_T(\mathbf{k}) = \int d^2x\, e^{-i\mathbf{k}\cdot\mathbf{x}}\,\nabla\times\mathbf{A}(\mathbf{x}) = \int d^2x\, e^{-i\mathbf{k}\cdot\mathbf{x}}\,\nabla\times\left[\int d^2y\, G(\mathbf{x}-\mathbf{y})n(\mathbf{y})\right]$$

$$= \frac{i\mathbf{k}\times\hat{\mathbf{z}}}{k^2}n(\mathbf{k}) \tag{18.47}$$

where $n(\mathbf{k})$ is the Fourier transform of the vortex density:

$$n(\mathbf{k}) = \sum_{i=1}^{N} 2\pi p_i\, e^{-i\mathbf{k}\cdot\mathbf{r}_i} \tag{18.48}$$

Thus

$$\rho_s = \rho_0 - \beta\rho_0^2\lim_{k\to 0}\frac{\langle n(\mathbf{k})n(-\mathbf{k})\rangle}{\Omega k^2} \tag{18.49}$$

For small k, we write

$$n(\mathbf{k})n(-\mathbf{k}) = (2\pi)^2\sum_{i,j} p_i p_j e^{-i\mathbf{k}\cdot(\mathbf{r}i-\mathbf{r}j)}$$

$$= (2\pi)^2\sum_{i,j} p_i p_j\{1 - i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j) - \tfrac{1}{2}[\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)]^2 + \cdots\} \tag{18.50}$$

The first term vanishes because $\Sigma p_i = 0$. The second terms vanishes when integrated over $\mathbf{r}_i$. When the third term is integrated over the coordinates, we have

$$\int d^2r_1\, d^2r_2[\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)]^2 = k^\alpha k^\beta\int d^2r_1\, d^2r_2(\mathbf{r}_1-\mathbf{r}_2)^\alpha(\mathbf{r}_1-\mathbf{r}_2)^\beta$$

$$= \frac{k^2}{2}\int d^2r_1\, d^2r_2|\mathbf{r}_1-\mathbf{r}_2|^2 \tag{18.51}$$

Anticipating the integrations, we can take

$$n(\mathbf{k})n(-\mathbf{k}) = -\frac{(2\pi)^2 k^2}{4} \sum_{i,j} p_i p_j |\mathbf{r}_i - \mathbf{r}_j|^2 + O(k^4) \tag{18.52}$$

To order z^2 we need to average the preceding over configurations with only one vortex–antivortex pair:

$$\lim_{k\to 0} \frac{\langle n(\mathbf{k})n(-\mathbf{k})\rangle}{\Omega k^2} = z^2 2\pi^3 \langle r^2\rangle + O(z^4) \tag{18.53}$$

where $\langle r^2\rangle$ is given by (18.42). To this order, the denominator $Z_{\text{vortex}} = 1 + O(z^2)$ can be set to unity. Thus, the fugacity expansion of the superfluid density is given by

$$\rho_s = \rho_0 - 2\pi^3 z^2 \beta \rho_0^2 \langle r^2\rangle + O(z^4) \tag{18.54}$$

or

$$\frac{1}{K_s} = \frac{1}{K_0} + 2\pi^3 z^2 \int_1^\infty du\, u^{3-2\pi K_0} + O(z^4) \tag{18.55}$$

where

$$\begin{aligned} K_0 &\equiv \beta\rho_0 \\ K_s &\equiv \beta\rho_s \end{aligned} \tag{18.56}$$

As noted before, the fugacity expansion is good at low temperatures, but breaks down at the KT transition point.

18.7 RG TRAJECTORIES

We make an RG transformation through the scale change $a \to ae^t$, without altering the system. For infinitesimal t, we rewrite (18.55) by splitting the integral:

$$\int_1^\infty du\, u^{3-2\pi K_0} = \int_1^{1+t} du\, u^{3-2\pi K_0} + \int_{1+t}^\infty du\, u^{3-2\pi K_0} \tag{18.57}$$

The first term gives t. The lower limit in the second term can be restored to unity through rescaling. Thus we obtain

$$\frac{1}{K_s} = \frac{1}{K_0} + 2\pi^3 z^2 t + 2\pi^3 z^2 [1 + (4 - 2\pi K_0)t] \int_1^\infty du\, u^{3-2\pi K_0} + O(t^2) \tag{18.58}$$

Now defining the scale-dependent parameters

$$\frac{1}{K(t)} \equiv \frac{1}{K(0)} + 2\pi^3 t\, z^2(t) + O(t^2)$$

$$z^2(t) \equiv z^2(0)\{1 + [4 - 2\pi K(t)]t\} + O(t^2) \tag{18.59}$$

with $z(0) = z$ and $K(0) = K_0$. Then we can write

$$\frac{1}{K_s} = \frac{1}{K(t)} + 2\pi^3 z^2(t) \int_1^\infty du\, u^{3-2\pi K(t)} + O(t^2, z^4) \tag{18.60}$$

which has the same form as (18.55), except for a change of parameters. As we can see, K_s is scale-invariant.

The differential form of (18.59) gives the RG equations of the XY model [9]:

$$\frac{d}{dt}\frac{1}{K} = 2\pi^3 z^2 + O(z^4)$$

$$\frac{dz^2}{dt} = z^2(4 - 2\pi K) + O(z^4) \tag{18.61}$$

The line $z = 0$ is a continuous distribution of fixed points. An isolated fixed point occurs at $z = 0$, $K = 2/\pi$. This is the critical point of the KT transition, and we call it the *KT fixed point*. In terms of the temperature $T = \rho_0/K$, the critical point occurs at $T_c = \pi\rho_0/2$. Let

$$\tau \equiv \frac{T - T_c}{T_c} \tag{18.62}$$

Then, near the KT fixed point, to first order in τ and second order in z, we have

$$\frac{d\tau}{dt} = (2\pi)^2 z^2$$

$$\frac{dz}{dt} = 2z\tau \tag{18.63}$$

From the first equation, we see that $d\tau/dt > 0$. Thus trajectories always flow in the positive τ direction.

Dividing the first equation by the second, we obtain $d\tau/dz = 2\pi^2 z/\tau$, which leads to the orbit equation

$$\tau^2 - 2\pi^2 z^2 = C \tag{18.64}$$

where C is an integration constant. The trajectories are hyperbolas, as shown in Fig. 18.5. There are two asymptotes corresponding to $C = 0$. The left asymptotic is the

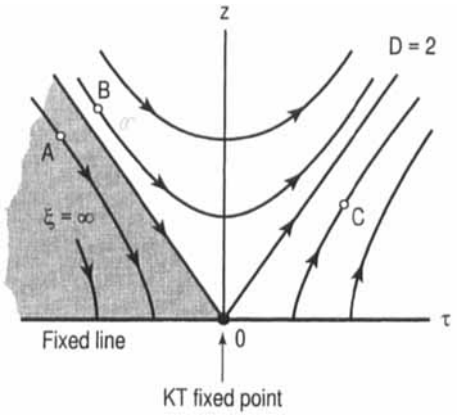


Figure 18.5 RG trajectories of the *XY* model in the neighborhood of the Kosterlitz–Thouless fixed point, in the parameter space spanned by a temperature τ and the vortex fugacity z. All points in the shaded region have infinite correlation length.

transition line of the *KT* transition. In the region to the left, all trajectories flow into the fixed line. This means that z is irrelevant, and there are no free vortices. The correlation length diverges in the entire region, and the correlation function has algebraic rather than exponential behavior.

The region to the right of the transition line represents the high-temperature phase, where all trajectories eventually tend toward large τ and z. Points A and B in Fig. 18.5 lie on opposite sides of the *KT* transition line, with A in the low-temperature phase and B in the high-temperature phase.

In the region between the right asymptote and the positive τ axis, all trajectories are UV, and the fixed line is unstable against vortex creation.

We have derived the RG equations in a fugacity expansion to order z^2, and this covers only an infinitesimal strip just above the negative τ axis. We can analytically continue the RG equations to positive τ, but the domain of validity remains an infinitesimal strip. Fortunately, this is sufficient to give a complete understanding of the *KT* transition, because the strip contains the fixed point.

The correlation function of this model is the same as (18.17), except that the phase angle should include contributions from vortices:

$$\theta(x) = \omega(x) + v(x)$$

where $\omega(x)$ represents the spin-wave part and

$$v(x) = \sum_i p_i \varphi_i \tag{18.65}$$

is a sum of vortex contributions (18.22). The correlation function is thus

$$G(x) = \langle e^{i\omega(x)} e^{-i\omega(0)} \rangle \langle e^{iv(x)} e^{-iv(0)} \rangle \tag{18.66}$$

The spin-wave factor is given by (18.17), and Kosterlitz [9] has calculated the vortex contribution in the low-temperature phase at the KT fixed point:

$$\langle e^{iv(x)} e^{-iv(0)} \rangle = \left(\frac{r}{a}\right)^{-1/4} \left[\ln\left(\frac{r}{a}\right)\right]^{1/8} \tag{18.67}$$

where $r = |x|$. Thus in the low-temperature phase

$$G(x) = \left[\ln\left(\frac{r}{a}\right)\right]^{1/8} \left(\frac{r}{a}\right)^{-(1/2\pi\beta\rho_0)-1/4} \tag{18.68}$$

which verifies that the correlation length is infinite.

In the high-temperature phase, the correlation length ξ is finite, but diverges at the critical point. Recall that $\xi \propto e^{-t}$ under a scale change. Let $\Delta\tau$ be the distance to the transition point at fixed z. To find how ξ depends on $\Delta\tau$, we shall find how $\Delta\tau$ depends on t.

In the high-temperature phase, since $C > 0$ in (18.64), we put $C = \alpha^2$. The transition line corresponds to $\alpha = 0$. A point such as B in Fig. 18.5 moves along the trajectory when t changes, but moves to a different trajectory when α changes. We want to find the correlation between these movements. The distance to the transition point at constant z is given by

$$\Delta\tau = c - \sqrt{c^2 - \alpha^2} \approx \frac{\alpha^2}{2c} \tag{18.69}$$

where $c = \sqrt{2}\pi z$. By integrating the RG equation for τ, we obtain its t dependence (with arbitrary origin):

$$-\frac{1}{\alpha}\cot^{-1}\frac{\tau}{\alpha} = 2t \tag{18.70}$$

which becomes $-\pi/2\alpha = t$ when $\alpha \to 0$. In terms of $\Delta\tau$, then, we have

$$t = -\frac{c'}{\sqrt{\Delta\tau}} \tag{18.71}$$

where c' is a constant. Therefore the correlation length diverges like

$$\xi \underset{\Delta\tau \to 0}{\to} \exp\frac{c'}{\sqrt{\Delta\tau}} \tag{18.72}$$

This exhibits an essential singularity, and the usual notion of critical exponent does not apply.

Fisher and Nelson [10] suggest the following generalization for the RG equations for $D = 2 + \epsilon$:

$$\frac{d}{dt}\frac{1}{K} = 2\pi^3 z^2 - \frac{\epsilon}{K}$$

$$\frac{dz^2}{dt} = z^2(4 - 2\pi K) \tag{18.73}$$

The extra term the first equation comes from the fact that K is no longer dimensionless. The second equation is unchanged because z remains dimensionless. We see that the KT fixed point is moved to a location above the z axis. But there is no longer a line of fixed points. As shown in Fig. 18.6, there is a line crossed by trajectories at normal incidence. It is this line that becomes the fixed line at $z = 0$ when $\epsilon \to 0$. The fact that the fixed point is off the z axis means that the fugacity expansion no longer yields exact answers.

18.8 UNIVERSAL JUMP OF SUPERFLUID DENSITY

Ordinarily, when a superfluid is heated, the superfluid density decreases to zero according to a power law. Nelson and Kosterlitz [11] pointed out that in 2D the superfluid density suddenly jumps to zero at the transition point, with a universal discontinuity. This can be seen by inspection of Fig. 18.5. We recall that K_s is scale invariant in 2D, and therefore has the same value along a RG trajectory. Thus, as point A approaches the transition line, $K_s \to 2/\pi$, the KT fixed point. At point B, on the other hand, $K_s = 0$, because it is equal to its value at high temperatures. There is thus a discontinuity across the phase transition line:

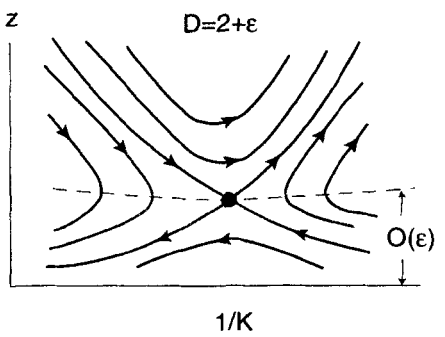


Figure 18.6 RG flow in $2 + \epsilon$ dimensions. The dotted line becomes the fixed line when $\epsilon \to 0$.

$$K_s = \begin{cases} \dfrac{2}{\pi} & \text{(low- temperature phase)} \\ 0 & \text{(high-temperature phase)} \end{cases} \tag{18.74}$$

The superfluid density is $\rho_s = \beta^{-1}K_s$, in some natural units, in which the energy functional is $\frac{1}{2}\rho_0 \int dx(\nabla\theta)^2$. In physical units the superfluid velocity is $(\hbar/m)\nabla\theta$, and all densities acquire the units $(\hbar/m)^2$. Thus the superfluidity density has a discontinuity given in physical units by

$$\Delta\rho_s = K_0 T_c$$
$$K_0 = \frac{2m^2 k_B}{\pi\hbar^2} = 3.491 \times 10^{-9}\ \text{g cm}^{-2}\ \text{K}^{-1} \tag{18.75}$$

where the numerical value refers to helium atoms. The jump is universal in that it depends only on atomic constants. It has been tested in experiments on helium films from different laboratories, using different methods to measure the superfluid density, with different film substrates and thicknesses.Very good agreement was obtained, as shown in Fig. 18.7. This indicates that only the temperature and vortex fugacity are relevant parameters. Other parameters that vary from experiment to experiment, such as film thickness and the nature of the substrate, are irrelevant.

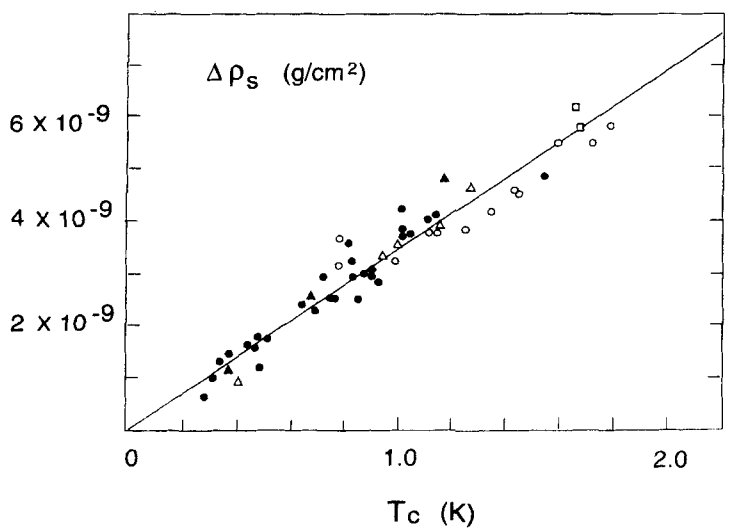


Figure 18.7 Compilation of data on the superfluid jump in thin helium film from different laboratories. The straight line is the prediction from the vortex model $\Delta\rho_s = K_0 T_c$, with $K_0 = 3.491 \times 10^{-9}$ g cm^{-2} K^{-1}. [After D. J. Bishop and J. Reppy, *Phys. Rev. Lett.* **40**, 1729 (1978).]

PROBLEMS

18.1 A general vortex is a configuration $\phi(x, y)$ satisfying $\oint_C dx^k\, \partial^k \phi(x, y) = 2\pi v_0$, where v_0 is the circulation. We have $v_0 = 1$ in (18.21). Clearly, the configuration has to be singular, like (18.22). A manifestation of the singularity is that the current defined by $j^k(x, y) = \epsilon^{kj} \partial^j \phi(x, y)$ fails to satisfy the expected identity $\partial^k j^k \equiv 0$.

(a) Verify that a solution to the vortex condition can be written

$$\phi(x, y) = v_0 \tan^{-1} \frac{x}{y}$$

(b) For this configuration show that

$$\partial^k j^k(x, y) = 2\pi v_0 \delta(x)\delta(y)$$

18.2 Obtain equations for the RG trajectories of the XY model by integrating the RG equations (18.63), which are valid for all τ, but only for small z. Use the relation (18.64) to express z in terms of τ. The constant C is negative for point B in Fig. 18.5, and positive for points A and C.

(a) For points C and B, put $C = b^2$ and show that

$$\tau(t) \approx b(1 + 2e^{4bt})$$
$$z(t) \approx \pi^{-1}\sqrt{2b^2}e^{2bt}$$

which are valid for $bt \to -\infty$, where $b < 0$ for A, and $b > 0$ for B. As a quantum field theory, the coupling constant z is trivial in the former case, and asymptotically free in the latter.

(b) For point B, put $C = -a^2$, and show that

$$\tau(t) \approx 2at$$
$$z(t) \approx \pi^{-1}\sqrt{\frac{a^2}{2}}(1 + 2t^2)$$

which are valid for $|a| \ll 1$, $|t| \ll 1$. There is no fixed point in this region.

REFERENCES

1. R. E. Peierls, *Helv. Phys. Acta* **7** (Suppl. II), 81 (1934).
2. D. R. Nelson and B. I. Halperin, *Phys. Rev.* **B19,** 2475 (1979).
3. N. D. Mermin and H. Wagner, *Phys. Rev. Lett.* **17,** 1133 (1966).
4. J. V. Jose, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson, *Phys. Rev.* **B19,** 2457 (1979); D. R. Nelson, in *Phase Transitions and Critical Phenomena,* Vol. 7, C. Domb and J. L Lebowitz, Eds., Academic Press, Orlando, FL 1983).
5. V. L. Berezinskii, *Sov. Phys. JETP* **32,** 493 (1970); **34,** 610 (1971).

6. P. C. Hohenberg, *Phys. Rev.* **158,** 383 (1967).
7. S. Coleman, *Commun. Math. Phys.* **31,** 259 (1973).
8. J. M. Kosterlitz and D. J. Thouless, *J. Phys.* **C6,** 1181 (1973).
9. J. M. Kosterlitz, *J. Phys.* **C7,** 1046 (1974).
10. D. S. Fisher and D. R. Nelson, *Phys. Rev.* **B16,** 4945 (1971).
11. D. R. Nelson and J. M. Kosterlitz, *Phys. Rev. Lett.* **39,** 1201 (1977).

CHAPTER NINETEEN

Topological Excitations

19.1 TOPOLOGICAL SOLITON

A soliton is a solution to a classical field theory that describes a localized disturbance with finite energy. A *topological soliton* is one stabilized by topology; that is, it cannot be continuously deformed to the vacuum. We introduce the subject through an simple example.

Consider the so-called sine–Gordon theory in 1 + 1 dimensional Minkowskian space–time, with Lagrangian density

$$\mathcal{L}(x, t) = \frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi}{\partial x}\right)^2 - W(1 - \cos\phi) \tag{19.1}$$

where the potential includes an appropriate constant to make it nonnegative. The action is

$$S[\phi] = \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} dx \left[\frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi}{\partial x}\right)^2 - W(1 - \cos\phi)\right] \tag{19.2}$$

which leads to the equation of motion

$$\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} + W\sin\phi = 0 \tag{19.3}$$

We encountered this theory in Chapter 17 in Euclidean space–time, as a quantum field theory with asymptotic freedom. It is also the XY model discussed in Chapter 18, in the presence of an external magnetic field W. We view it here as a classical field theory in one spatial dimension.

The topological essence of this model is as follows [1]. The field variable $\phi(x)$ is an angular variable defined on a circle S^1. We impose the boundary condition $e^{i\phi} = 1$ at $x = \pm\infty$, and this compactifies the manifold of x to S^1. Thus $\phi(x)$ represents

a map $S^1 \rightarrow S^1$, which falls into classes labeled by an integer winding number, which denotes the number of time the ϕ manifold is covered when the x manifold is covered once. The winding number is invariant under homotopic transformations—continuous deformations of the field configuration.

We now discuss properties of the solutions. The Hamiltonian is given by

$$H = \int_{-\infty}^{\infty} dx \left[\frac{1}{2} \left(\frac{\partial \phi}{\partial t} \right)^2 - \frac{1}{2} \left(\frac{\partial \phi}{\partial x} \right)^2 + W(1 - \cos \phi) \right] \tag{19.4}$$

The classical vacuum—the configuration of lowest energy—corresponds to $\phi(x, t) = 0 \bmod(2\pi)$. A topological soliton is a solution in which the field approaches different but equivalent values as $x \rightarrow \pm\infty$, such as

$$\phi(x, t) \rightarrow \begin{cases} 2\pi & (x \rightarrow \infty) \\ 0 & (x \rightarrow -\infty) \end{cases} \tag{19.5}$$

The solution can be obtained by minimizing the Hamiltonian. Consider first a static soliton. Since the potential is zero only when $\phi = 2\pi n$, to keep the potential energy small, ϕ would like to switch from 0 to 2π suddenly, but that would cost too much kinetic energy. A compromise is struck by having the transition occur in an interval of some width L. The kinetic energy is then of order $1/L$, and the potential energy is of order WL. Thus

$$\text{Energy} \sim \frac{1}{L} + WL \tag{19.6}$$

Minimization with respect to L gives

$$L \sim \frac{1}{\sqrt{W}}$$

$$\text{Energy} \sim \sqrt{W} \tag{19.7}$$

The soliton is a "kink" in the field centered at some location x_0, as illustrated in Fig. 19.1. The energy is independent of x_0.

We can construct a time-dependent soliton, which moves with a definite velocity without changing its shape. To do this, let

$$\xi = x - vt \tag{19.8}$$

The equation of motion can be written in the form

$$\phi'' = -\frac{\partial V(\phi)}{\partial \phi} \qquad V(\phi) = W \frac{\cos \phi}{1 - v^2} \tag{19.9}$$

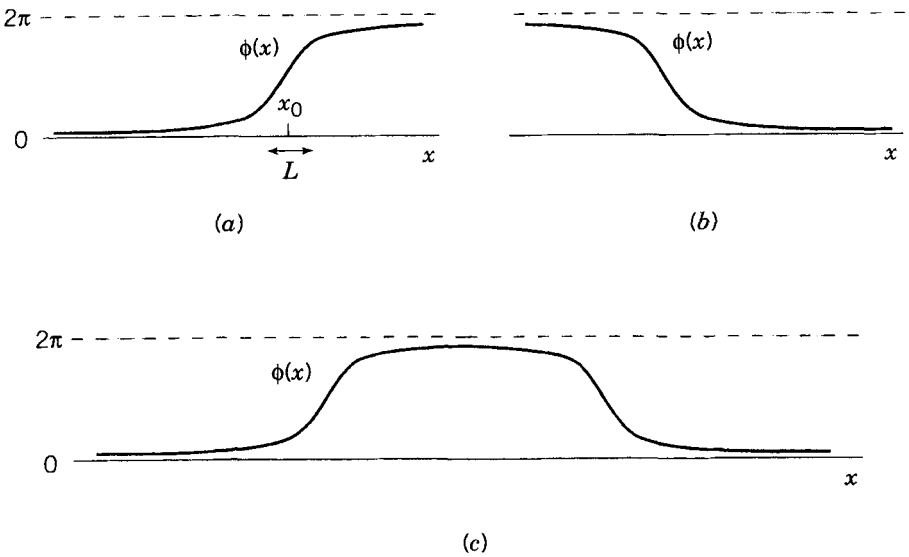


Figure 19.1 (*a*) Topological soliton stabilized by boundary condition; (*b*) antisoliton; (*c*) soliton–antisoliton pair.

where $\phi' = \partial\phi/\partial\xi$. This is like Newton's equation in mechanics, with conserved "energy" given by

$$C = \tfrac{1}{2}(\phi')^2 + V(\phi) \tag{19.10}$$

This can be integrated to yield

$$\int \frac{d\phi}{\sqrt{2[C - V(\phi)]}} = \xi \tag{19.11}$$

Choosing $C = W/(1 - v^2)$, we have

$$\frac{1}{2}\sqrt{\frac{1-v^2}{W}}\int \frac{d\phi}{\sin^2(\phi/2)} = \xi \tag{19.12}$$

which gives the solution

$$\phi(x, t) = 4\tan^{-1}(e^{\sqrt{W}(x-vt-x_0)}) \tag{19.13}$$

where x_0 is an arbitrary constant. The name soliton was derived from the "solitary" wave witnessed in 1834 by J. Scott Russell [2]:

> I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped—not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon.

Such solitary waves have been observed in plasmas and optical media. But our interest here is confined to the topological aspects of static solitons.

19.2 INSTANTON AND TUNNELING

The static soliton is obtained by putting $v = 0$:

$$\phi(x) = 4 \tan^{-1}(e^{\sqrt{W}(x-x_0)}) \tag{19.14}$$

Classically, this represents a time-independent field configuration in one dimension. However, in quantum theory, we can interpret x as Euclidean time. In this view, $\phi(x)$ is the coordinate of a particle, and the soliton (19.14) interpolates between the vacua $\phi(-\infty) = 0$ and $\phi(\infty) = 2\pi$ in Euclidean time. 'tHooft [3] named it the "instanton," for one could imagine that the vacuum "pops" at time x_0. We can show that the instanton give the transition amplitude for quantum tunneling between the two vacua.

Consider a quantum-mechanical particle of unit mass and coordinate ϕ, moving in a potential $V(\phi)$, which possesses two minima at ϕ_1 and ϕ_2, as shown in Fig. 19.2. The amplitude for a particle to tunnel from ϕ_1 to ϕ_2 at zero energy is given in the WKB (Wentzel–Kramers–Brillouin) approximation by

$$T_{\text{WKB}} = \exp\left[-\int_{\phi_1}^{\phi_2} d\phi \sqrt{2V(\phi)}\right] \tag{19.15}$$

In this semiclassical approximation, the particle is a wave packet of zero classical energy, whose motion is described by classical mechanics (in real time) except for the tunneling. It was located initially at ϕ_1, and makes a quantum jump with probability $|T_{\text{WKB}}|^2$, to appear at ϕ_2 with zero velocity. Thereafter, its motion is again governed by classical mechanics.

The quantum jump can be regarded as a process happening in imaginary time $\tau = it$, with transition probability amplitude

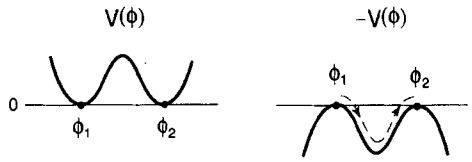


Figure 19.2 An instanton gives the dominant Feynman path for tunneling between the two minima of the potential V. It describes imaginary time development with inverted potential $-V$.

$$\langle \phi_2 | T | \phi_1 \rangle = \int D\phi(\tau)\, e^{-S[\phi]} \tag{19.16}$$

The Euclidean action is

$$S[\phi] = \int_{-\infty}^{\infty} d\tau \left[\frac{1}{2} \left(\frac{\partial \phi(\tau)}{\partial \tau} \right)^2 + V(\phi(\tau)) \right] \tag{19.17}$$

with the boundary conditions

$$\phi(\tau) \rightarrow \begin{cases} \phi_1 & (\tau \rightarrow \infty) \\ \phi_2 & (\tau \rightarrow -\infty) \end{cases} \tag{19.18}$$

In the semiclassical limit, the dominant path is a solution to the classical equation of motion

$$\frac{d^2\phi}{d\tau^2} - \frac{dV}{d\phi} = 0 \tag{19.19}$$

which describes the motion of a particle in the potential $-V(\phi)$ with the given boundary conditions, as shown in Fig. 19.2. To solve the equation, note that the constant of motion corresponding to energy is

$$\frac{1}{2} \left(\frac{d\phi}{d\tau} \right)^2 - V(\phi) = 0 \tag{19.20}$$

and the value of the action at the minimum is

$$S_0 = 2 \int_{-\infty}^{\infty} d\tau\, V(\phi(\tau)) \tag{19.21}$$

We use (19.20) to rewrite $V = \sqrt{V}\sqrt{V} = \sqrt{V/2}(d\phi/d\tau)$, and obtain

$$S_0 = \int_{\phi_1}^{\phi_2} d\phi \sqrt{2mV(\phi)} \tag{19.22}$$

Thus, in the saddle-point approximation we have

$$\langle \phi_2 | T | \phi_1 \rangle \approx e^{-S_0} = T_{\text{WKB}} \tag{19.23}$$

This show that the instanton is the Feynman path that dominates the tunneling amplitude in the semiclassical limit.

19.3 DEPINNING OF CHARGE DENSITY WAVES

We now apply the instanton description of tunneling to a physical problem. In linear-chain conductors, such as TTF–TCNQ (tetrathiafulvalene–tetracyanoquinodimethane), there exists a periodic charge distribution $\rho(x, t)$ called a *charge density wave* (CDW)[4]:

$$\rho(x, t) = \bar{\rho} + \rho_0 \cos(Qx + \phi(x, t)) \tag{19.24}$$

where $\bar{\rho}$ is a uniform background density, ρ_0 is the amplitude of the CDW, Q is the wave number, and $\phi(x, t)$ is the phase relative to an underlying periodic lattice. When the wave number Q is the same as that of the lattice, the latter presents a commensurate potential that pins the CDW, preventing it from sliding. In equilibrium, then, $\phi(x.t)$ has a value ϕ_0 everywhere. If we turn on an external electric field, the potential will become "tilted," as shown in Fig. 19.3, and the CDW can slide by tunneling from ϕ_0 to $\phi_1 = \phi_0 + 2\pi$. This process cannot take place throughout the chain simultaneously; the probability of that is nil. As a result of random fluctuations, thermal or quantum mechanical, a small one-dimensional "bubble" of phase ϕ_1 occurs somewhere, and then expands to engulf the whole chain [5].

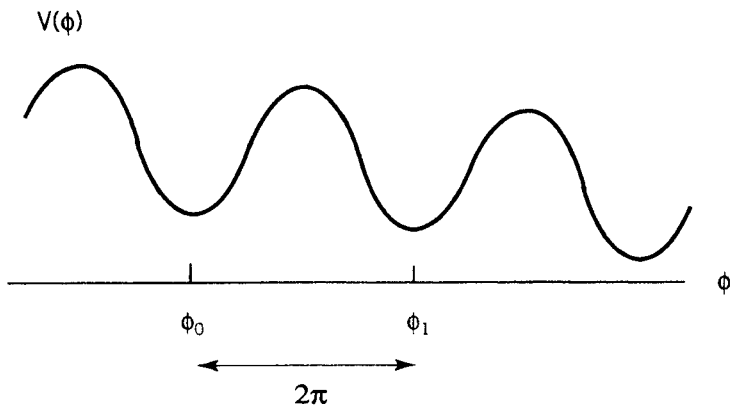


Figure 19.3 The potential seen by a charge density wave in an external electric field. In a small "bubble" in space, the phase angle ϕ can go from ϕ_0 to ϕ_1 by tunneling.

What happens in real time may be described as follows. In the absence of an electric field, the periodic lattice potential leads to a sine Gordon equation for $\phi(x, t)$, and there are soliton solutions. Random fluctuations can create a soliton–antisoliton pair somewhere along the chain, so that $\phi(x, t)$ looks momentarily like the configuration in Fig. 19.1*c*. In the presence of an electric field, there is incentive for the pair to grow in order to lower the energy, and the size of this excitation expands to eventually cover the entire chain.

We treat the creation of the bubble as a tunneling process—a quantum jump in Euclidean time. Outside this transition, the motion proceeds according to classical mechanics in real time. This description is a one-dimensional version of Coleman's theory [6] of a bubble of "true vacuum" in a background of "false vacuum."

The Lagrangian density in real time is given by

$$\mathcal{L}(x, t) = \frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\frac{\partial\phi}{\partial x}\right)^2 - V(\phi)$$

$$V(\phi) = W(1 - \cos\phi) - \epsilon\phi \tag{19.25}$$

where $-W\cos\phi$ is the commensurate potential due to the underlying lattice, and ϵ is the external electric field in suitable units. We have rescaled the variables to absorbed physical parameters such as charge and mass, so as to present a neat form for mathematical analysis. The fact that the electric field is coupled directly to the phase ϕ may be seen as follows. The current density in the system is $j \propto \partial\phi/\partial t$. Its interaction energy with an external electromagnetic field A contributes to the action the term

$$-\int dx\, dt\, A\frac{\partial\phi}{\partial t} = \int dx\, dt\frac{\partial A}{\partial t}\phi = -\int dx\, dt\, \epsilon\phi \tag{19.26}$$

where we have assumed that the external field is adiabatically turned on in the infinite past, and off in the infinite future.

We treat the electric field ϵ as a small perturbation. The equation of motion of the system is

$$\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} + \frac{dV(\phi)}{d\phi} = 0 \tag{19.27}$$

except for the quantum jump, which is described by a 2D instanton, a solution $\phi_E(x, \tau)$ of the equation in Euclidean time $\tau = it$:

$$\frac{\partial^2\phi_E}{\partial\tau^2} + \frac{\partial^2\phi_E}{\partial x^2} - \frac{dV(\phi_E)}{d\phi_E} = 0 \tag{19.28}$$

and seek a solution such that the entire chain is at $\phi_E(x, \tau) = \phi_0$ as $\tau \to -\infty$, and a finite-sized bubble of phase $\phi_1 = \phi_0 + 2\pi$ appears somewhere at $\tau = 0$. At this time,

the quantum jump is complete, and the bubble subsequently expands classically. The semiclassical picture is based on the fact that tunneling occurs with very small probability, and is treated as a rare interruption of the classical motion.

For the formal solution, we take advantage of the symmetry under Euclidean time reversal, and impose the boundary conditions

$$\lim_{\tau\to\pm\infty} \phi_E(x, \tau) = \phi_0$$

$$\left.\frac{\partial\phi_E(x, \tau)}{\partial\tau}\right|_{\tau=0} = 0 \tag{19.29}$$

but we use the solution only for $-\infty < \tau \leq 0$. The symmetry of the boundary condition makes it possible to have a solution that depends only on

$$s = \sqrt{x^2 + \tau^2} \tag{19.30}$$

The boundary conditions then become

$$\lim_{s\to\infty} \phi_E(s) = \phi_0$$

$$\phi_E'(0) = 0 \tag{19.31}$$

where the location of the quantum jump is taken to be $s = 0$. Because of translational invariance, the action should be independent of this location. The probability for a quantum jump per unit length of the chain is $\exp(-S)$, with

$$S = \int_{-\infty}^{\infty} d\tau \int_{-\infty}^{\infty} dx \left[\frac{1}{2}\left(\frac{\partial\phi_E}{\partial\tau}\right)^2 + \frac{1}{2}\left(\frac{\partial\phi_E}{\partial x}\right)^2 + W(1 - \cos\phi_E) - \epsilon(\phi_E - \phi_0)\right] \tag{19.32}$$

The solution we seek corresponds to a spherically symmetric configuration in the x–τ plane, as illustrated in Fig. 19.4. At $\tau = 0$ it looks like a soliton–antisoliton pair, and this is the bubble created through tunneling. The instanton should have a mean radius R, with $\phi_E = \phi_0$ outside the radius, and $\phi_E = \phi_1$ inside. The wall of the excitation is of thickness $1/W$, with energy W per unit circumference. Thus, the corresponding action is

$$S \approx 2\pi R\sqrt{W} - \pi R^2\epsilon(\phi_1 - \phi_0) \tag{19.33}$$

where $\phi_1 - \phi_0 = 2\pi$. Minimizing this with respect to R. we obtain the radius and corresponding action:

$$R_0 = \frac{\sqrt{W}}{2\pi\epsilon} \qquad S_0 \simeq \frac{W}{2\epsilon} \tag{19.34}$$

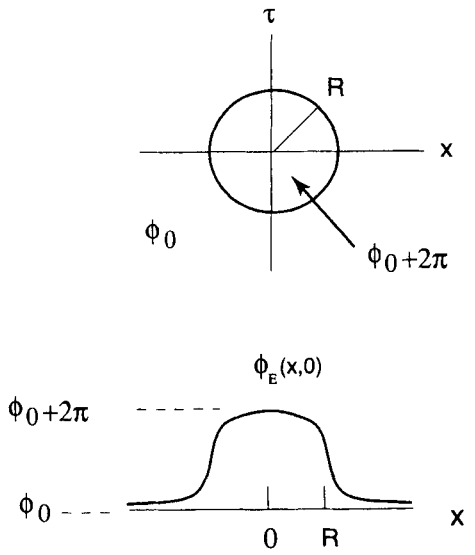


Figure 19.4 Upper panel represents the instanton that gives the tunneling amplitude for formation of a small bubble of new phase. Lower panel shows the field configuration at Euclidean time $\tau = 0$, which consists of a soliton–antisoliton pair created locally through tunneling. This bubble expands in real time to eventually engulf the whole x axis.

After the bubble is formed, it expands according to the equation of motion in real time. The behavior is described by the analytic continuation of the Euclidean solution, which becomes a function of $x^2 - t^2$. Thus the size of the bubble expands according to

$$R(t) = \sqrt{R_0^2 + t^2} \tag{19.35}$$

and the unit of time is set by whatever physical parameters we have used as scale.

What we have described is the sliding of the CDW by one lattice length due to tunneling. It is clear that this process repeats, and generates a current proportional to the transition probability:

$$I \simeq Ce^{-W/2\epsilon} \tag{19.36}$$

This current is extremely small, and has not been detected experimentally, because it is masked by currents arising from thermal fluctuations. In an incommensurate lattice, the CDW is pinned not by the lattice, but by impurities. The depinning due to tunneling across impurities can be described in a similar manner [7].

19.4 NONLINEAR SIGMA MODEL

We turn to a model with interesting topological properties. Consider a unit 3-vector field $\mathbf{n}(x)$ in a d-dimensional Euclidean space. The Euclidean action is

$$S = \tfrac{1}{4}\int d^d x\, \partial^\mu n_a\, \partial^\mu n_a \tag{19.37}$$

where $a = 1,2,3$ labels field components, and the Greek index $\mu = 1, \ldots, d$ is a Euclidean index for general d. For specific dimensions such as $d = 2$ or $d = 3$, we switch to Roman indices k. In physical applications, this model is variously known as the "classical Heisenberg model" or "$O(3)$ nonlinear sigma model" (for historical reasons that need not concern us). The nonlinearity arises from the constraint

$$n_a n_a = 1 \tag{19.38}$$

The designation $O(3)$ refers to the fact that $\mathbf{n}(x)$ may be identified with a point on the surface of a three-dimensional sphere, and therefore corresponds to an element of the rotation group $O(3)$.

Since rotations can also be represented by $SU(2)$, we can map the model to one involving a spinor field. Let z be a two-dimensional spinor:

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{19.39}$$

where z_i are complex-number fields, with the condition

$$z^\dagger z = z_1^* z_1 + z_2^* z_2 = 1 \tag{19.40}$$

We can represent $\mathbf{n}$ in the form

$$\mathbf{n} = z^\dagger \boldsymbol{\sigma} z \tag{19.41}$$

where σ_a ($a = 1,2,3$) are the 2×2 Pauli matrices. More explicitly, we put

$$\begin{aligned} n_1 &= 2\mathrm{Re}(z_1^* z_2) \\ n_2 &= 2\mathrm{Im}(z_1^* z_2) \\ n_3 &= z_1^* z_1 - z_2^* z_2 \end{aligned} \tag{19.42}$$

Thus

$$n_1^2 + n_2^2 + n_3^2 = (z^\dagger z)^2 = 1 \tag{19.43}$$

A straightforward calculation shows

$$\tfrac{1}{4}\,\partial^\mu n_a \partial^\mu n_a = \partial^\mu z^\dagger \partial^\mu z + (z^\dagger \partial^\mu z)(z^\dagger \partial^\mu z) \tag{19.44}$$

Thus the action can be written in the form

$$S = \int d^d x[\partial^\mu z^\dagger\ \partial^\mu z + (z^\dagger \partial^\mu z)(z^\dagger \partial^\mu z)] \tag{19.45}$$

We can define a vector potential by

$$A^\mu = -iz^\dagger \partial^\mu z \tag{19.46}$$

Then the action can be rewritten in the form (see Problem 19.2)

$$S = \int d^d x[(\partial^\mu + iA^\mu)z^\dagger][(\partial^\mu - iA^\mu)z] \tag{19.47}$$

which is invariant under the local gauge transformation

$$\begin{aligned} z(x) &\to e^{i\omega(x)} z(x) \\ A^\mu(x) &\to A^\mu(x) + \partial^\mu \omega(x) \end{aligned} \tag{19.48}$$

The field $\mathbf{n}(x)$ can be visualized as a point on the surface of a sphere, which constitutes the 2-sphere S^2. For a finite-action solution, it must approach a constant vector at infinity:

$$\mathbf{n}(x) \xrightarrow[|x|\to\infty]{} \mathbf{n}_0 \tag{19.49}$$

For definiteness, we take $\mathbf{n}_0 = (0, 0, 1)$. Spatial infinity is thus identified as one point, and the space is effectively compactified from Euclidean space R^d to the d–sphere S^d. A finite-energy solution therefore corresponds to a map

$$\mathbf{n}(x)\colon S^d \to S^2 \tag{19.50}$$

These maps fall into homotopy classes C_m characterized by a winding number m. The maps in each class are "homotopic," in the sense that they can be continuously deformed into one another [8]. The classes $\{C_0, C_1, \ldots\}$ form a group, the homotopy group. For a map $S^d \to X$, the homotopy group is denoted by $\pi_d(X)$, called the "dth homotopy group" of X. In particular, $\pi_1(X)$ is called the "fundamental group" of X.

For the $O(3)$ nonlinear sigma model, the relevant homotopy groups are $\pi_d(S^2)$:

$$\pi_1(S^2) = 0$$

$$\pi_2(S^2) = Z_\infty$$
$$\pi_3(S^2) = Z_\infty$$
$$\pi_4(S^2) = Z_2 \tag{19.51}$$

where 0 denotes the trivial group containing only the identity, Z_∞ denotes the set of integers $\{0, \pm1, \pm2, \ldots\}$, and Z_2 is the group $\{0, 1\}$ mod(2) under addition, or $\{1, -1\}$ under multiplication. We see that there are topological solitons for $d > 1$. They can be regarded either as a static soliton in d spatial dimensions, or an instanton in $d - 1$ spatial dimensions and one Euclidean time dimension. We discuss the different dimensions separately.

19.5 THE SKYRMION

The *Skyrmion* [9] is a static soliton of the $O(3)$ nonlinear sigma model in two spatial dimensions. Let us represent $\mathbf{n}(x)$ by a unit vector whose tip lies on a sphere in an internal space with axes labeled $a = 1,2,3$. There is thus an internal coordinate frame attached to each point x, and we take all these frames to have the same relative orientation. An element of the surface of the sphere is $d\mathbf{S} = (dS^1, dS^2, dS^3)$, with

$$dS^a = \frac{1}{2}\epsilon^{abc} dn^b\, dn^c = \frac{1}{2}\epsilon^{abc}\frac{\partial(n^b n^c)}{\partial(\xi^1\xi^2)} d\xi^1 d\xi^2 \tag{19.52}$$

where we have parametrized the surface of the sphere by two coordinates (ξ^1, ξ^2). The magnitude of the surface element is

$$dS = n^a dS^a = \frac{1}{2}\epsilon^{abc}\epsilon^{ij} n^a \frac{\partial n^b}{\partial \xi^i}\frac{\partial n^c}{\partial \xi^j} d\xi^1 d\xi^2 \tag{19.53}$$

As x ranges over the compactified spatial S^2, $\mathbf{n}(x)$ ranges over the internal S^2. The winding number of the map $\mathbf{n}(x)$ is the number of times the internal S^2 is covered, and is thus given by

$$Q = \frac{1}{4\pi}\int_x dS(x) \tag{19.54}$$

where $dS(x)$ is the surface element corresponding to $\mathbf{n}(x)$, and the preceding integral extends over the range such that x covers S^2 once. A convenient way to express this condition is to use (x^1, x^2) as the parameters (ξ^1, ξ^2). Therefore the winding number of the configuration $\mathbf{n}(x)$ is given by

$$Q = \frac{1}{8\pi}\int d^2x \epsilon^{ij}\epsilon^{abc} n_a(\partial^i n_b)(\partial^j n_c) \tag{19.55}$$

which has possible values 0, ±1, ±2, ... The configurations with $Q > 0$ are solitons, and those for $Q < 0$ are antisolitons.

We can rewrite the winding number in the form

$$Q = \frac{1}{8\pi}\int d^2x\, \epsilon^{ij}\mathbf{n}\cdot\partial^i\mathbf{n}\times\partial^j\mathbf{n} \tag{19.56}$$

Now note the identity

$$\tfrac{1}{4}\int d^2x|\partial^i\mathbf{n} + \epsilon^{ij}\partial^i\mathbf{n}\times\partial^j\mathbf{n}|^2 = \tfrac{1}{2}\int d^2x[(\partial^i\mathbf{n})^2 - \epsilon^{ij}\mathbf{n}\cdot\partial^i\mathbf{n}\times\partial^j\mathbf{n}] \tag{19.57}$$

Thus, the action can be written in the form

$$S = \tfrac{1}{4}\int d^2x(\partial^i\mathbf{n})^2 = 2\pi Q + \tfrac{1}{4}\int d^2x|\partial^i\mathbf{n} + \epsilon^{ij}\partial^i\mathbf{n}\times\partial^j\mathbf{n}|^2 \tag{19.58}$$

which shows

$$S \geq 2\pi Q \tag{19.59}$$

and the equality holds if and only if

$$\partial^i\mathbf{n} + \epsilon^{ij}\partial^i\mathbf{n}\times\partial^j\mathbf{n} = 0 \tag{19.60}$$

Solving this equations gives a soliton with finite action.

To find explicit solutions [10], use cylindrical coordinates for x:

$$\begin{aligned} x^1 &= r\cos\varphi \\ x^2 &= r\sin\varphi \end{aligned} \tag{19.61}$$

and parametrize $\mathbf{n}(x)$ through

$$n_a(x) = (\hat{\mathbf{x}}\sin f(r), \cos f(r)) \tag{19.62}$$

where $f(r)$ satisfies the boundary conditions

$$f(r) \to \begin{cases} \pi & (r \to 0) \\ 0 & (r \to \infty) \end{cases} \tag{19.63}$$

More explicitly,

$$\begin{aligned} n_1(x) &= \sin f(r)\cos\varphi \\ n_2(x) &= \sin f(r)\sin\varphi \end{aligned}$$

$$n_3(x) = \cos f(r) \tag{19.64}$$

The vector $\mathbf{n}(x)$ is $(0, 0, -1)$ at $x = 0$, and approaches $(0, 0, 1)$ as $|x| \to \infty$. We see that f and φ are just the polar angles of $\mathbf{n}$ in spherical coordinates. Now introduce the complex variable

$$w = e^{i\varphi} \tan \frac{f}{2} \tag{19.65}$$

Then[9]

$$n_1 + in_2 = \frac{2w}{1 + w^* w}$$

$$n_3 = \frac{1 - w^* w}{1 + w^* w} \tag{19.66}$$

and (19.60) becomes

$$\frac{\partial w}{\partial x^1} + i \frac{\partial w}{\partial x^2} = 0 \tag{19.67}$$

which are just the Cauchy–Riemann condition that w be an analytic function of z, with

$$z = x^1 + ix^2 \tag{19.68}$$

We rule out branch cuts, and obtain the general solution as a meromorphic function

$$w(z) = \prod_{i,j} \left(\frac{z - a_i}{\lambda} \right)^{m_i} \left(\frac{\lambda}{z - b_i} \right)^{n_j} \tag{19.69}$$

where λ is an arbitrary scale parameter and m_i and n_j are positive integers. In order that $f(r) \to 0$ at infinity, we must have

$$\sum_i m_i > \sum_j n_j \tag{19.70}$$

To find the winding number, note that for a given value w, z satisfies the polynomial equation

$$\prod_i \left(\frac{z - a_i}{\lambda} \right)^{m_i} - w \prod_j \left(\frac{\lambda}{z - b_i} \right)^{n_j} = 0 \tag{19.71}$$

which is of degree Σm_i, and therefore has Σm_i roots, generally different. That is, the

given value w occurs at Σm_i generally different points in space. Therefore the winding number is

$$Q = \sum_i m_i \tag{19.72}$$

The *antisoliton* is defined as the solution with negative Q, with the same boundary condition (19.49). It corresponds to the complex conjugate of $w(z)$, for then n_2 changes sign while n_1 and n_3 are unchanged, and Q changes sign.

For $Q = 1$, take $w(z) = z/\lambda$. This leads to

$$f(r) = 2 \tan^{-1} \frac{r}{\lambda} \tag{19.73}$$

The size of the soliton λ is arbitrary. Putting $\lambda = 1$, we have

$$\begin{aligned} n_1 &= \frac{r}{1+r^2} \cos \varphi \\ n_2 &= \frac{r}{1+r^2} \sin \varphi \\ n_3 &= \frac{1-r^2}{1+r^2} \end{aligned} \tag{19.74}$$

with corresponding spinor representation

$$\begin{aligned} z_1 &= \cos \frac{f}{2} \\ z_2 &= e^{i\varphi} \sin \frac{f}{2} \end{aligned} \tag{19.75}$$

To obtain the corresponding antisoliton, replace φ by $-\varphi$, or replace $n_2(x)$ by $-n_2(x)$. We visualize this soliton in the two views presented in Fig. 19.5. In Fig. 9.5*a* we sketch $\mathbf{n}(x)$ along a radius from the origin in the spatial plane. In Fig. 9.5*b*, we superimpose the internal S^2 on the compactified spatial S^2, showing a "hedgehog" configuration.

The solution can be viewed as an instanton, if we regard $x^1 = x$ as space, and $x^2 = \tau$ as imaginary time. As illustrated in Fig. 19.6, the instanton evolves along world lines shown in Fig. 19.6*a*, and the tip of the vector $\mathbf{n}(x, \tau)$ traces closed loops, as shown in Fig. 19.6*b*. When (x, τ) covers space–time once, the locus of $\mathbf{n}(x, \tau)$ sweeps over the 2-sphere once.

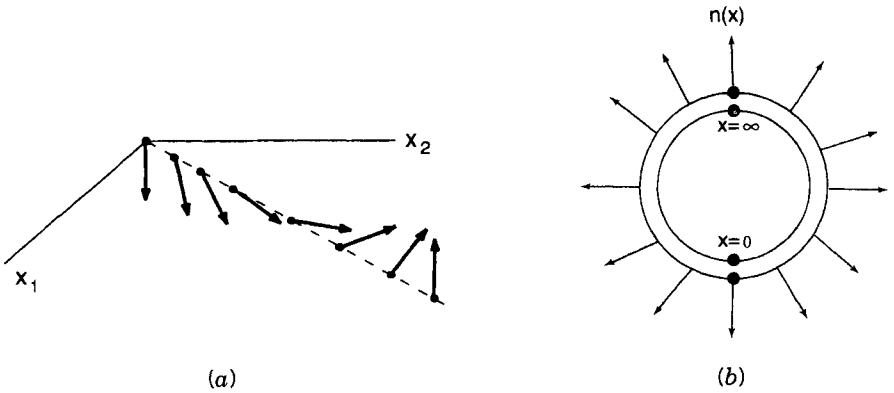


Figure 19.5 (*a*) Skyrmion as static soliton—the spin along a radius in the 2D plane turns 180° when we go from the origin to infinity; (*b*) when internal space is superimposed on the compactified real space, we get a "hedgehog."

19.6 THE HOPF INVARIANT

Viewing the Skyrmion as a static soliton in 2D, it is natural to ask how it propagated in time. To consider this question, we go to $d = 3$, where $\mathbf{n}(x)$ represents the Hopf map $S^3 \to S^2$.

Let us parametize S^2 by two parameters (ξ^1, ξ^2), and consider first a continuous map $R^3 \to S^2$, represented by the functions

$$\begin{aligned} \xi^1 &= f_1(x^1, x^2, x^3) \\ \xi^2 &= f_2(x^1, x^2, x^3) \end{aligned} \tag{19.76}$$

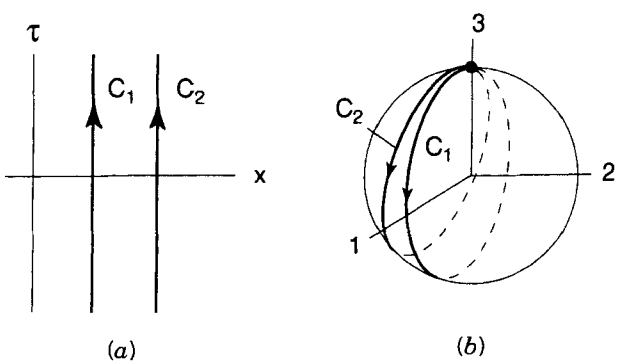


Figure 19.6 (*a*) Skyrmion viewed as instanton developing in imaginary time τ at fixed position x; (*b*) the spin traces out loops in internal space corresponding to the various world lines C_1, C_2.

For a given point on S^2,that is, for a given (ξ^1, ξ^2), this equation gives two relations among the coordinates $\{x^1, x^2, x^3\}$, and one combination of the coordinates remains arbitrary. This means that the inverse map of a point on S^2 is a curve on R^3, and the curve must be a closed loop, since a point has no boundary. Two different closed loops cannot intersect, for otherwise the intersection would be mapped to different points. To ensure that the closed loops do not run off to infinity, we require

$$f_i(x) \xrightarrow[|x|\to\infty]{} C \tag{19.77}$$

where C is a constant. This compactifies R^3 to S^3, and the map f is the Hopf map, which is illustrated schematically in Fig. 19.7. As shown in the figure, two loops in three-dimensional space are characterized by a topological invariant, the linkage number, giving the number of times one loop winds around the other. The linkage number is a property of the map, called the *Hopf invariant.* If we displace the image points P' and Q' continuously, the loops P and Q will change, but the linkage number will remain the same. It is now evident that $\pi_3(S^2) = Z$.

The Hopf map can be represented by the spinor representation introduced earlier in (19.41):

$$\mathbf{n}(x) = z^\dagger(x)\boldsymbol{\sigma} z(x) \tag{19.78}$$

We can write the spinor z in the form

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = Uz_0 \tag{19.79}$$

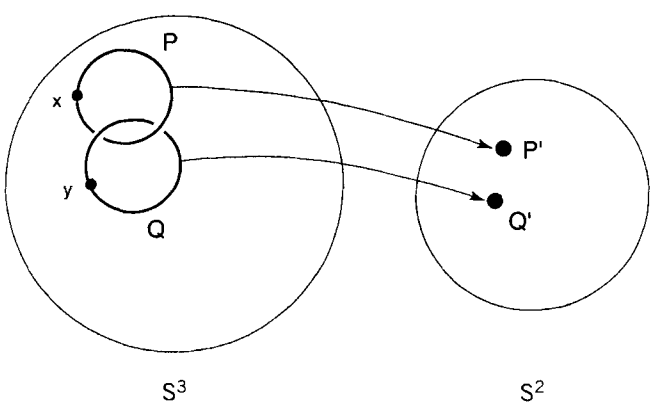


Figure 19.7 The Hopf map $S^3 \to S^2$ maps closed loops into points, because the two spaces have different dimensions. Any two closed-loop inverse images are characterized by the linkage number, which is the Hopf invariant of the map.

where $U \in SU(2)$ and z_0 is some fixed spinor. An element of $SU(2)$ is of the form

$$U = b_0 + \mathbf{b}\cdot\boldsymbol{\sigma} \tag{19.80}$$

with $b_0 + |\mathbf{b}|^2 = 1$. Therefore the manifold of $SU(2)$ is S^3, and (19.78) represents a map $S^3 \to S^2$. Since $\mathbf{n}$ is invariant under the gauge transformation $z \to e^{i\gamma}z$, the inverse map is a circle on S^3, as depicted in Fig. 19.7.

Since $z(x)$ is a map $S^3 \to S^3$, the representation (19.78) involves a two-step map $S^3 \to S^3 \to S^2$, and the winding number of the step $S^3 \to S^3$ is the Hopf invariant:

$$H = \frac{1}{2\pi^2}\int_x dS(x) \tag{19.81}$$

where $dS(x)$ is a volume element of the S^3 manifold of $z(x)$ and the integral is such that x ranges over its S^3 manifold once. The factor $2\pi^2$ is the volume of S^3. Parametrizing the manifold of z by 3 parameters $(\lambda_1, \lambda_2, \lambda_3)$, we can write

$$H = \frac{1}{2\pi^2}\int d^3x \frac{\partial(\lambda_1, \lambda_2, \lambda_3)}{\partial(x^1, x^2, x^3)} \tag{19.82}$$

The Jacobian in the integrand is $\epsilon^{ijk}\partial^i\lambda_1\partial^j\lambda_2\partial^k\lambda_3$, which is a 3-form constructed from z. There is only one such 3-form $\epsilon^{ijk}A^i\partial^jA^k$, with $A^k = -iz^\dagger\partial^k z$. Thus we have the Hopf invariant up to a normalization constant, which can be calculated from an explicit construction of $z(x)$. (See Problem 19.3.) The result is

$$H = \frac{1}{4\pi^2}\int d^3x\, \mathbf{A}\cdot\nabla\times\mathbf{A} \tag{19.83}$$

As we have seen in (19.47), $\mathbf{A}$ is like a vector potential. The corresponding magnetic field is a topological current density that is identically conserved:

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{19.84}$$

There is an interesting parallel with classical electromagnetism. Because of the gauge invariance (19.48), we can impose Coulomb gauge $\nabla\times\mathbf{A} = 0$. Putting

$$\nabla \times \mathbf{B} \equiv \mathbf{j} \tag{19.85}$$

we can write

$$\mathbf{A}(\mathbf{x}) = -\frac{1}{4\pi^2}\int d^3y \frac{\mathbf{j}(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \tag{19.86}$$

Thus the Hopf invariant can be rewritten as

$$H = \frac{1}{4\pi^2}\int d^3x \mathbf{B}\cdot\mathbf{A} = -\frac{1}{4\pi^2}\int d^3x\, d^3y \frac{\mathbf{B}(x)\cdot\mathbf{j}(y)}{4\pi|\mathbf{x}-\mathbf{y}|}$$

$$= -\frac{1}{(2\pi)^2}\frac{\epsilon^{ijk}}{4\pi}\int d^3x \int d^3y\, B^i(\mathbf{x})B^j(\mathbf{y})\partial^k \frac{1}{|\mathbf{x}-\mathbf{y}|} \tag{19.87}$$

This can be understood in terms of Gauss' formula for the linkage between two closed loops P and Q, such as those illustrated in Fig. 19.7:

$$L_{PQ} = \frac{\epsilon^{ijk}}{4\pi}\oint_P dx^i \oint_Q dy^j \partial^k \frac{1}{|\mathbf{x}-\mathbf{y}|} \tag{19.88}$$

To derive this formula, consider the solid angle subtended by loop Q at point $\mathbf{x}$, and count the number of 4π increments as $\mathbf{x}$ traverses loop P. The gradient of the solid angle in question is the magnetic field created by a unit current flowing in loop Q, which can be obtained via the Biot–Savart law. Thus (19.87) calculates the Hopf invariant of the configuration by calculating the Gaussian linkage between loops of the topological current.

Finally, we express the topological current (or magnetic field) $\mathbf{B}$ in terms of $\mathbf{n}(\mathbf{x})$:

$$B^i = \frac{1}{8\pi}\epsilon^{ijk}\epsilon^{abc} n_a \partial^j n_b\, \partial^k n_c \tag{19.89}$$

This is the only invariant we can construct out of $n_a(x)$ involving two derivatives. The constant in front can be determined by direct computation, but a faster way is as follows. The conservation law is

$$\frac{\partial B^1}{\partial x^1} + \frac{\partial B^2}{\partial x^2} + \frac{\partial B^3}{\partial x^3} \equiv 0 \tag{19.90}$$

The magnetic flux Q normal to the x^1–x^2 plane is a topological charge, which is constant of motion for Skyrmions, and should coincide with the winding number of $S^2 \to S^2$ calculated in (19.55):

$$Q = \int dx^1 dx^2 B^3 = \frac{1}{8\pi}\int dx^1 dx^2\, \epsilon^{abc} n_a (\partial^1 n_b\, \partial^2 n_c - \partial^2 n_b\, \partial^1 n_c) \tag{19.91}$$

19.7 FRACTIONAL SPIN

We now show that the spin of the 2D Skyrmion is an arbitrary real number [11]. We regard $\mathbf{n}(x) = \mathbf{n}(\mathbf{x}, \tau)$ as a 2D Skyrmion evolving in imaginary time τ, and imagine that it makes one complete rotation in space as it propagates from $\tau = -\infty$ to $\tau = \infty$.

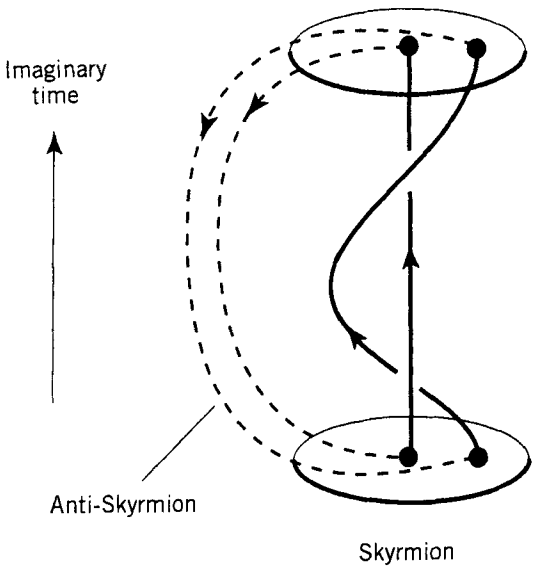


Figure 19.8 A Skyrmion makes complete rotations during its imaginary-time evolution. The number of rotations is the Hopf invariant of the space–time configuration.We introduce a spectator anti-Skyrmion, in order to have a vacuum on the space–time boundary.

In Fig. 19.8, we show two world lines of points on the Skyrmion, identified by given values of the spin. For example, the world line at the center of the Skyrmion has $\mathbf{n} = (0, 0, -1)$. The initial and final states of the Skyrmion are the same, but any two world lines are linked once. The number of complete rotations made by the Skyrmion is, in fact, given by the Hopf invariant of $\mathbf{n}(\mathbf{x}, \tau)$. To see this, imagine that at $\tau = -\infty$ a Skyrmion–anti-Skyrmion pair was created, and pulled apart. The Skyrmion propagates in the manner indicated, with the anti-Skyrmion as a passive spectator, and eventually the pair annihilates at $\tau = \infty$. World lines from the anti-Skyrmion are indicated by dotted lines in Fig. 19.8, which completes the Skyrmion world lines into closed loops. On the space–time boundary we now have the vacuum configuration $\mathbf{n} = (0, 0, 1)$. Hence the function $\mathbf{n}(\mathbf{x}, \tau)$ is a Hopf map $S^3 \rightarrow S^2$, and the linkage of any two closed world lines is the Hopf invariant.

Now describe the process in Minkowski space–time, by continuing $\tau = -it$. The Minkowski action is

$$S[\mathbf{n}] = \frac{1}{4}\int_{-\infty}^{\infty} dt \int d^2x \left[\left(\frac{\partial n_a}{\partial t} \right)^2 - |\nabla n_a|^2 \right] \tag{19.92}$$

We can add to this a term θH, where θ is an arbitrary real number, without changing the equation of motion. This is because H is an integer and cannot change in the

continuous variations made in applying the action principle. Thus, the amplitude for the process is

$$A(\theta) = \mathcal{N} \int D\mathbf{n}\, e^{iS[\mathbf{n}]+i\theta H} \tag{19.93}$$

where $\mathcal{N}$ is a normalization constant, such that the vacuum–vacuum transition amplitude is 1. For all paths contributing to the preceding integral, $H = 1$. Thus, relative to the vacuum–vacuum amplitude, we have

$$A(\theta) = e^{i\theta} \tag{19.94}$$

Equating this to $e^{2\pi i J}$, where J is the spin of the Skyrmion, we obtain an arbitrary real number:

$$J = \frac{\theta}{2\pi} \qquad \text{(Skyrmion)} \tag{19.95}$$

This result does not contradict basic principles of quantum mechanics, for the rotation group in 2D is the Abelian group $U(1)$, and thus the angular momentum may have continuous eigenvalues.

Continuing in this vein, we can view the time development of the Skyrmion, described by a configuration with nonzero Hopf invariant, as a 3D static soliton—the "Hopfion." To find the spin of the Hopfion, we go to $d = 4$, where $\mathbf{n}(x)$ becomes a map $S^4 \rightarrow S^2$, with $\pi_4(S^2) = Z_2$. The additive topological invariant has possible values

$$I = 0,\ 1 \bmod(2) \tag{19.96}$$

As before, we may add to the Minkowski action a term θI without affecting the equation of motion, but now we must require $e^{2i\theta} = 1$. Therefore the values for θ are quantized:

$$\theta = \pi n \tag{19.97}$$

where n is an integer. Repeating the argument in the last section, we find that the spin of the Hopfion must be integer or half-integer:

$$J = \frac{\theta}{2\pi} = \frac{n}{2} \qquad \text{(Hopfion)} \tag{19.98}$$

This is, of course, implied by the angular momentum commutation relations in 3D. What is unusual is that we can construct half-integer spin states from boson fields.

19.8 MONOPOLES, VORTICES, AND ANOMALIES

According to (19.91), the topological charge of a Skyrmion is the magnetic flux normal to the x^1–x^2 plane, in which the static Skyrmion resides. If we follow the time development in imaginary time x^3, the Skyrmion will sweep out a flux tube, which cannot terminate because of the conservation of the topological current (19.90). However, there exist monopole configuration in $d = 3$ that will destroy the conservation law, for example, the "hedgehog" configuration

$$n_a(x) = \frac{gx^a}{|x|} \tag{19.99}$$

which is singular at the position of the monopole. This means that the world line of the Skyrmion can be of finite length, terminated at both ends by monopole and anti-monopole, respectively. Such singular configurations may be ruled out in a classical theory, but cannot be ignored in quantum field theory, for the path integral extends over all fields, singular or not. In quantum theory, therefore, the monopoles render the topological current "anomalous."

To demonstrate the existence of the monopole, let us calculate the vector potential corresponding to a Skyrmion in the x^1–x^2 plane, with topological charge $Q = 1$. From the explicit solution (19.75), we find

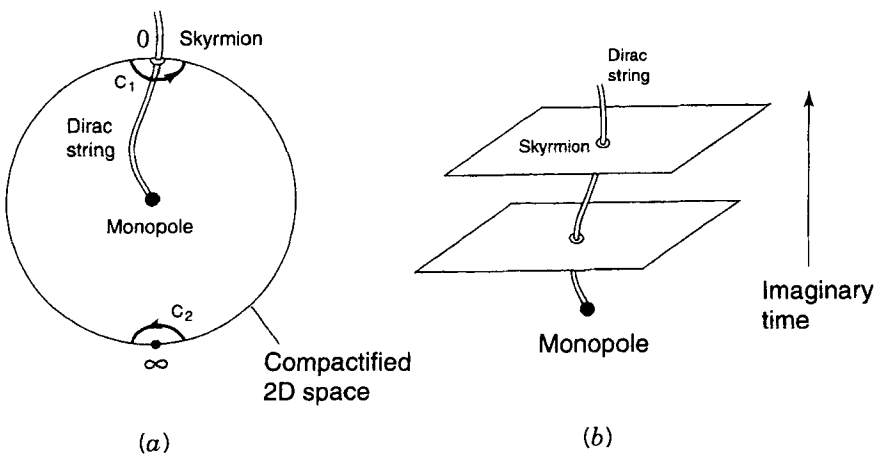


Figure 19.9 (*a*) A Skyrmion appears where the Dirace string of a 3D monopole pierces a sphere surrounding the monopole. (*b*) In another view, the world line of a Skyrmion in imaginary time is a Dirac string. The world line is generally terminated at both ends by monopole and antimonopole. The possibility of termination makes the topological Skyrmion current anomalous.

$$A^k(r, \varphi) = \partial^k \varphi \sin^2 \frac{f(r)}{2} \tag{19.100}$$

where $f(0) = \pi, f(\infty) = 0$. The magnetic flux through a loop C of radius r is given by

$$\Phi(r) = \oint_C dx^k A^k = 2\pi \sin^2 \frac{f(r)}{2} \tag{19.101}$$

which is 2π for $r = 0$, but approaches zero as $r \to \infty$. The total flux over the x^1–x^2 plane is therefore zero, in contradiction with the fact that $Q = 1$ by construction. The discrepancy is due to the failure of the relation between the flux and the topological winding number, because of the presence of a monopole in $d = 3$. As illustrated in Fig. 19.9*a*, where the compactified x^1–x^2 plane has a spherical surface, the flux through the closed loop C_1 is 2π, but that through C_2 is zero. This implies that there is a monopole inside the sphere, which terminates the flux tube of the Skyrmion. In this picture, the world line of the Skyrmion is the Dirac string of the monopole. When we open up the compactified sphere, the world line of the Skyrmion is as illustrated in Fig. 19.9*b*, which is terminated by the monopole.

Apart from the short-distance singularity, a single monopole has infinite action, and would not contribute to the path integral; however, a monopole–antimonopole pair can contribute. Thus the topological current in the $O(3)$ nonlinear sigma model in $d = 3$ should have a nonvanishing divergence given by the monopole density in the vacuum state. In analogy with the XY model in $d = 2$, the theory should have a phase in which monopole–antimonopole pairs are tightly bound, and one in which they are ionized. The topological current will be normal in the former phase and anomalous in the latter. Such a scenario is supported by numerical studies [12].

The mechanism that causes a current anomaly can be demonstrated analytically in the simpler XY model. In the presence of an external magnetic field W, the XY model is described by the Euclidean version of the sine–Gordon model (19.2), with action

$$S[\phi] = \int dx\, dy [\tfrac{1}{2}(\partial\phi)^2 + W(1 - \cos\phi)] \tag{19.102}$$

which describes the propagation of the 1D soliton in imaginary time. The topological current density is

$$j^k = \epsilon^{kj} \partial^j \phi \tag{19.103}$$

and we have $\partial^k j^k \equiv 0$ for nonsingular field configurations. However, the conservation is violated by a vortex configuration $\phi_0(x, y)$ that satisfies

$$\oint_C d\mathbf{s} \cdot \nabla \phi_0(x, y) = 2\pi v_0 \tag{19.104}$$

where C is a closed loop containing the origin of the x–y plane. As shown in Problem 18.1, a solution is

$$\phi_0(x, y) = v_0 \tan^{-1} \frac{x}{y} \tag{19.105}$$

For fixed $y > 0$, this is a one-dimensional soliton with boundary conditions

$$\phi_0(\infty, y) - \phi_0(-\infty, y) = \begin{cases} \pi v_0 & \text{if} \quad y > 0 \\ -\pi v_0 & \text{if} \quad y < 0 \end{cases} \tag{19.106}$$

As shown in Problem 18.1, we have for this configuration

$$\nabla \cdot \mathbf{j} = 2\pi v_0 \delta^2(\mathbf{x}) \tag{19.107}$$

Thus we expect the current to be anomalous in the high-temperature phase above the Kosterlitz–Thouless transition. This is verified in a more detailed analysis [13].

A physically important anomaly occurs in the the chiral current of a Dirac field, discussed in Problem 6.4. For a massless field, this current is conserved in a classical theory, but not when the theory is quantized. The divergence of the current turns out to be given by a magnetic charge density. The current is somewhat different from those considered above, in that it is not defined as a topological current. We refer the interested readers elsewhere for a full discussion [14].

PROBLEMS

19.1 To illustrate topological solitons in a theory different the sine–Gordon case, consider the nonlinear Schrödinger equation (Gross–Pitaevskii equation) introduced in Section 15.4, in one spatial dimension:

$$-\frac{\partial^2 \psi}{\partial x^2} + g|\psi|^2 \psi = i \frac{\partial \psi}{\partial t}$$

There exist topological soliton solutions, stabilized through the fact that the phase of $\psi(x, t)$ approach different values as $x \to \pm\infty$.

(a) Seek a solution of the form

$$\psi(x, t) = e^{i(kx - \omega t)} f(\xi) \qquad \xi \equiv x - Ut$$

where U is some constant, and $f(\xi)$ is real. Show that

$$f'' + i(2k - U)f' + (k^2 - \omega)f - gf^3 = 0$$

(b) Choose $k = U/2$. The equation then reduces to

$$f'' = -\frac{\partial V(f)}{\partial f}$$

$$V(f) = \frac{\alpha}{2} f^2 - \frac{g}{4} f^4$$

where $\alpha \equiv k^2 - \omega$. Regard the motion as that for a particle of unit mass in the potential $V(f)$, which for $\alpha > 0$ has a minimum about $f = 0$, and then falls off outside that. Consider motions confined to this central region. Obtain the orbit equation from conservation of "energy." The soliton corresponds to a zero-energy solution that traverses the central minimum when ξ goes from $-\infty$ to $+\infty$.

19.2 The $O(3)$ nonlinear sigma model has a remarkable local gauge invariance, revealed through the spinor representation (19.41), in which a vector potential (19.46) emerges. The form of the action (19.47), which exhibits the gauge invariance, can be established by explicit calculation, or through the following considerations.

(a) Consider the Lagrangian density

$$\mathcal{L}(x) = [(\partial^\mu + iA^\mu)z^\dagger]\,[(\partial^\mu - iA^\mu)z]$$

where $A^\mu(x)$ is regarded as an independent vector field. Show that the equations of motion give (19.46).

(b) Since there is no kinetic term for A^μ, it can be eliminated through the equations of motion. Show that $\mathcal{L}(x)$ is equivalent to

$$\mathcal{L}(x) = \partial^\mu z^\dagger \partial^\mu z + \tfrac{1}{4}(z^\dagger \overleftrightarrow{\partial_\mu} z)(z^\dagger \overleftrightarrow{\partial_\mu} z) = \partial^\mu z^\dagger\, \partial^\mu z + (z^\dagger \partial^\mu z)(z^\dagger \partial^\mu z)$$

where in the last step one uses the fact $z^\dagger z = 1$.

19.3 The spin variable $\mathbf{n}(x)$ in the $O(3)$ nonlinear sigma model in $d = 3$ represents a Hopf map $S^3 \to S^2$. Because of the boundary condition $\mathbf{n}(x) \to (0, 0, 1)$ at infinity, the space of $x \in R^3$ is compactified to S^3. In this problem we construct a configuration with Hopf invariant 1.

(a) Use spherical coordinates $x = (r, \theta, \varphi)$. Show that S^3, which is the surface of a four-dimensional sphere of unit radius, can be parametrized by the four coordinates

$$(b_0, \mathbf{b}) = (\cos\gamma, \sin\gamma\cos\theta, \sin\gamma\sin\theta\cos\varphi, \sin\gamma\sin\theta\sin\varphi)$$

$$(0 \le \gamma < \pi) \qquad (0 \le \theta < \pi) \qquad (0 \le \varphi < 2\pi)$$

where $\gamma(r)$ has the properties $\gamma(0) = \pi$, $\gamma(\infty) = 0$. The volume element is $\sin^2\gamma \sin\theta\, d\gamma\, d\theta\, d\varphi$, and the total volume of the space is $2\pi^2$.

(b) The map $\mathbf{n}(x)$ is constructed by identifying $(b_0, \mathbf{b})$ with its spinor representation (19.41):

$$z_1(x) = \cos\gamma + i\sin\gamma\cos\theta$$

$$z_2(x) = e^{i\varphi}\sin\gamma\sin\theta$$

Calculate $n_a(x)$, and show that $|z_1| = 1$, $z_2 = 0$ gives $\mathbf{n}_1 = (0, 0, 1)$ and $|z_2| = 1$, $z_1 = 0$ gives $\mathbf{n}_2 = (0, 0, -1)$. Sketch the loci $\mathbf{n}(x) = \mathbf{n}_1$ and $\mathbf{n}(x) = \mathbf{n}_2$ in R^3. Show that the former is the z axis and the latter is a circle in the equatorial plane. They are thus closed loops with linkage number 1. This shows that the Hopf invariant of the map is 1.

(c) Calculate the vector potential, and verify

$$A^k = \cos\theta\partial^k\gamma - \tfrac{1}{2}\sin 2\gamma \sin\theta\partial^k\theta + \sin^2\gamma\sin^2\theta\partial^k\varphi$$

$$\epsilon^{ijk}A^i\partial^j A^k = 2\sin^2\gamma\sin\theta\epsilon^{ijk}\partial^i\gamma\,\partial^j\theta\,\partial^k\varphi$$

(d) The Hopf invariant H is the winding number of the map $S^3 \to S^3$ represented by $z(x)$:

$$H = \frac{1}{2\pi^2}\int d^3x \sin^2\gamma \sin\theta \frac{\partial(\gamma\theta\varphi)}{\partial(x^1x^2x^3)} = \frac{1}{4\pi^2}\int d^3x\, \epsilon^{ijk}A^i\partial^j A^k$$

This verifies the normalization constant in (19.83).

REFERENCES

1. D. Finkelstein, *J. Math. Phys.* **7,** 1218 (1966).
2. J. S. Russell, *Proceedings of the British Association for the Advancement of Science*, London, 1845, p. 311.
3. G. 'tHooft, *Phys. Rev. Lett.* **37,** 8 (1976).
4. P. A. Lee, T. M. Rice, and P. W. Anderson, *Solid State Commun.* **14,** 703 (1974).
5. K. Maki, *Phys. Rev. Lett.* **39,** 46 (1977).
6. S. Coleman, *Phys. Rev.* **15,** 2929 (1977).
7. A. I. Larkin and P. A. Lee, *Phys. Rev.* **B17,** 1596 (1978).
8. S. T. Hu, *Homotopy Theory,* Academic Press, New York, 1959.
9. T. H. R. Skyrme, *Proc. Roy. Soc.* (London) **A247,** 260 (1958).
10. A. A. Belavin and A. M. Polyakov, *JETP Lett.* **22,** 245 (1975); A. M. Polyakov, *Gauge Fields and Strings,* Harwood Academic Publishers, Chur, Switzerland, 1987, Chapter 6.
11. F. Wilczek and A. Zee, *Phys. Rev. Lett.* **51,** 2250 (1983).
12. K. Huang, Y. Koike, and J. Polonyi, *Internatl. J. Mod. Phys.* **6,** 1267 (1991).
13. K. Huang and J. Polonyi, *Internatl. J. Mod. Phys.* **6,** 409 (1991).
14. K. Huang, *Quarks, Leptons, and Gauge Fields*, 2nd ed., World Scientific Publishers, Singapore, 1992, Chapter 11, Sections 12.5, 12.6.

APPENDIX A

Background Material

A.1 NOTATION

We generally use units in which $c = \hbar = 1$. When we have to distinquish between space and time, the dimensional of space is usually denoted by D, and that of space–time, Euclidean or Minkowskian, by d.

In relativistic systems, the metric tensor $g^{\mu\nu}$ in the $d = 4$ Minkowski space is diagonal, with $g^{00} = -g^{11} = -g^{22} = -g^{33} = 1$. The position and gradient 4-vectors are denoted as follows:

$$
\begin{aligned}
x^\mu &= (x^0, x^1, x^2, x^3) = (x^0, \mathbf{x}) = (t, x, y, z) \\
x_\mu &= g_{\mu\nu} x^\nu = (x^0, -\mathbf{x}) \\
x^2 &= x^\mu x_\mu = (x^0)^2 + x^k x_k = (x^0)^2 - \mathbf{x}^2 \\
\partial^\mu &= \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial x_0}, -\nabla \right) \\
\partial_\mu &= \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x_0}, \nabla \right) \\
\partial^2 &= \frac{\partial^2}{\partial x_0^2} - \nabla^2 \equiv \Box^2
\end{aligned}
\tag{A.1}
$$

Greek indices have the range 0, 1, 2, 3, while Roman indices have the range 1, 2, 3. Repeated indices are summed over their range. The dot product between two 4-vectors can be written in various forms:

$$
x \cdot y = x^\mu y_\mu = x^0 y^0 - x^k y^k = x^0 y^0 - \mathbf{x} \cdot \mathbf{y} \tag{A.2}
$$

The Kronecker δ is defined by

$$\delta_{nm} = \begin{cases} 1 & \text{if} \quad n = m \\ 0 & \text{if} \quad n \neq m \end{cases} \tag{A.3}$$

When n and m consist of complicated expressions, we write $\delta_K(n - m) = \delta_{nm}$, and the subscript "K" (Kronecker) is omitted if there is no chance of confusion.

Dirac δ function is defined by

$$\delta(t - t_0) = 0 \qquad (\text{if} \quad t \neq t_0)$$

$$\int_{-\infty}^{\infty} dt\, f(t)\delta(t - t_0) = f(t_0) \tag{A.4}$$

It has the properties

$$\delta(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{i\omega t}$$

$$\delta(at) = \frac{\delta(t)}{|a|}$$

$$\delta(t^2 - a^2) = \frac{1}{2|a|}[\delta(t - a) + \delta(t + a)] \tag{A.5}$$

A useful relation is

$$\frac{1}{t - i\eta} = P\frac{1}{t} + i\pi\delta(t) \qquad (\eta \to 0^+) \tag{A.6}$$

where P denotes principal value:

$$\int_{-\infty}^{\infty} dt\, P\frac{1}{t} = \lim_{\eta \to 0^+}\left[\int_{-\infty}^{-\eta} + \int_{\eta}^{\infty}\right]\frac{dt}{t} \tag{A.7}$$

We have occasion to use the representation

$$e^{i|\alpha|} = \frac{1}{i\pi}\int_{-\infty}^{\infty} dt \frac{e^{-i\alpha t}}{t^2 - 1 - i\eta} \qquad (\eta \to 0^+) \tag{A.8}$$

Other useful functions are

$$\theta(t) = \begin{cases} 1 & (\text{if} \quad t > 0) \\ 0 & (\text{if} \quad t < 0) \end{cases} = \frac{1}{2i\pi}\int_{-\infty}^{\infty} d\omega \frac{e^{i\omega t}}{t - i\eta} \qquad (\eta \to 0^+)$$

$$\epsilon(t) = \begin{cases} 1 & (\text{if} \quad t > 0) \\ -1 & (\text{if} \quad t < 0) \end{cases} = \frac{1}{i\pi}\int_{-\infty}^{\infty} d\omega\, e^{i\omega t} P\frac{1}{t} \tag{A.9}$$

A.2 CLASSICAL MECHANICS

A *classical particle* is an object characterized by a position vector q whose components $\{q^1, \cdots, q^D\}$ are functions of the time t. In the nonrelativistic regime, the time dependence is governed by Newton's equations:

$$m\ddot{q}^i = -\frac{\partial V(q)}{\partial q^i} \qquad (i = 1, \cdots, D) \tag{A.10}$$

where m is the mass and $V(q)$ is the potential. The Lagrangian function is defined by

$$L(q, \dot{q}) = \tfrac{1}{2}m\dot{q}^2 - V(q) \tag{A.11}$$

where $q^2 = q^i q^i$. Newton's equations can be represented in the form

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0 \tag{A.12}$$

If there are N particles, this equation still applies, provided we interpret q as the collection of all the coordinates.

A.3 QUANTUM MECHANICS

Classical mechanics must be supplanted by quantum mechanics, whenever any relevant physical quantity of dimension *energy* × *time* becomes comparable to or smaller than Planck's constant

$$h = 6.27 \times 10^{-27} \text{ ergs/s} \tag{A.13}$$

The state of a system in quantum mechanics is associated with the *direction* of a vector in a Hilbert space—a *ray* in Hilbert space. Two vectors differing only in normalization represent the same state. Physical observables are associated with Hermitian operators on the Hilbert space. The inner product between two state vectors Ψ and Φ are denoted either by the notation (Ψ, Φ) commonly used in mathematics or by the Dirac notation $\langle\Psi|\Phi\rangle$. The inner product between Ψ and $O\Phi$, where O is a Hermitian operator, is denoted by the equivalent notation

$$(\Psi, O\Phi) = \langle\Psi|O|\Phi\rangle \tag{A.14}$$

For a nonrelativistic particle in one dimension, the relevant operators are momentum p and position x, which are defined by the commutation relation

$$[p, x] = -i\hbar \tag{A.15}$$

where $\hbar = h/2\pi$. We summarize properties for a one-dimensional system. It is straightforward to generalize them to higher dimensions.

We assume that there exists a Hamiltonian operator $H(p, x)$, which is Hermitian and time-independent, and it generates time translations:

$$|t + dt\rangle = \left[1 - \frac{i}{\hbar} H dt\right] |t\rangle + O(dt^2) \tag{A.16}$$

where $|t\rangle$ is the state of the system at time t. For time translation over a finite interval Δt, we subdivide the interval into N successive infinitesimal intervals. In the limit $N \to \infty$ we have

$$|t + \Delta t\rangle = \underset{N\to\infty}{\longrightarrow} \left(1 - \frac{iH}{\hbar}\frac{\Delta t}{N}\right)^N = \exp\left(-\frac{i}{\hbar} H \Delta t\right)|t\rangle \tag{A.17}$$

For a nonrelativistic particle, we take

$$H = \frac{p^2}{2m} + V(x) \tag{A.18}$$

by correspondence with classical mechanics.

A useful basis for the Hilbert space consists of eigenstates $|x\rangle$ of the position operator, denoted here by x_{op}, with the properties

$$\begin{aligned} x_{\text{op}}|x\rangle &= x|x\rangle \\ \langle x|x'\rangle &= \delta(x - x') \\ \int_{-\infty}^{\infty} dx\, |x\rangle\langle x| &= 1 \end{aligned} \tag{A.19}$$

An alternative basis is the set of eigenstates of the momentum operator $|p\rangle$, with the properties

$$\begin{aligned} p_{\text{op}}|p\rangle &= p|p\rangle \\ \langle p|p'\rangle &= 2\pi\delta(p - p') \\ \int_{-\infty}^{\infty} \frac{dp}{2\pi} |p\rangle\langle p| &= 1 \end{aligned}$$

$$\langle x|p\rangle = \exp\left(\frac{i}{\hbar} px\right) \tag{A.20}$$

In the coordinate basis, the state is represented by the wave function

$$\psi(x, t) = \langle x|t\rangle \tag{A.21}$$

The momentum operator is represented by $-i\hbar\partial/\partial x$ and the Hamiltonian, by a differential operator:

$$H(p, x) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \tag{A.22}$$

The statement that $H(p, x)$ generates infinitesimal time translations becomes

$$\psi(x, t + dt) = \left[1 - \frac{i}{\hbar}H(p, x)dt\right]\psi(x, t) + O(dt^2) \tag{A.23}$$

where we have used the property $\langle x'|H|x\rangle = \delta(x - x')H(p, x)$. This leads to the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\psi = i\hbar\frac{\partial\psi}{\partial t} \tag{A.24}$$

It follow from this equation that there is a conservation law

$$\frac{\partial\rho}{\partial t} + \frac{\partial j}{\partial x} = 0 \tag{A.25}$$

with

$$\rho = \psi^*\psi$$

$$j = \frac{\hbar}{2mi}\left(\psi^*\frac{\partial\psi}{\partial x} - \frac{\partial\psi^*}{\partial x}\psi\right) \tag{A.26}$$

where ρ is the probability density for finding the particle at x at time t and j is the corresponding probability current density. For a real wave function, $j = 0$.

The operators p and x in infinite space have continuous spectra of eigenvalues. It is more convenient to work with countable sets of eigenvalues and eigenvectors. To do this, we enclose the system in a large but finite box of dimension L, and let $L \to \infty$ at the end of all calculations. The boundary condition imposed on the surface of the box is not crucial, because it usually affects the wave function only in a finite layer near the boundary, and has no effect on volume properties, such as the energy density. A convenient choice is the periodic boundary condition $\psi(x + L) = \psi(x)$. For a plane wave $\psi(x) = e^{ikx}$, this means that the wave number k is restricted to the discrete values

$$k = \frac{2\pi n}{L} \qquad (n = 0, \pm 1, \pm 2, \cdots) \tag{A.27}$$

A sum over states labeled by k approaches an integral in the large L limit:

$$\sum_k \xrightarrow[L\to\infty]{} \frac{L}{2\pi} \int_{-\infty}^{\infty} dk \tag{A.28}$$

APPENDIX B

Linear Response

We illustrate the calculation of a response function to a disturbance in the path-integral formalism, where no references are made to Hilbert-space states and operators. It is suited to macroscopic descriptions based on the Ginsburg–Landau free energy, which may be looked upon as a very general quantum field theory. We consider the example of superfluidity, and derive formulas used in Chapter 15.

Superfluidity is a transport property defined in terms of the response of a system to an imposed velocity field. We assume that the imposed field $w^i(x, t)$ is infinitesimal, and adiabatically turned on and off:

$$w^i(x, t) \xrightarrow[|t|\to\infty]{} 0 \tag{B.1}$$

In response, the momentum density of system changes by an amount $\delta g^i(x, t)$. The Fourier transforms of these quantities are denoted by

$$\begin{aligned} \tilde{w}^i(k, \omega) &= \int d^D x\, e^{i(\omega t - k\cdot x)} w^i(x, t) \\ \delta\tilde{g}^i(k, \omega) &= \int d^D x\, e^{i(\omega t - k\cdot x)} \delta g^i(x, t) \end{aligned} \tag{B.2}$$

We consider spatial dimensions $D = 2,3$. The linear response is defined by a response function χ:

$$\delta\langle \bar{g}^i(k, \omega)\rangle = \chi^{ij}(k, \omega)\tilde{w}^j(k, \omega) \tag{B.3}$$

where $\langle\ \rangle$ denotes thermal average. The inverse Fourier transform reads

$$\delta\langle g^i(x, t)\rangle = \int \frac{d^D k}{(2\pi)^d} e^{-i(\omega t - k\cdot x)} \chi^{ij}(k, \omega)\tilde{w}^j(k, \omega) \tag{B.4}$$

Consider the static response

$$\chi^{ij}(k) \equiv \lim_{\omega \to 0} \chi^{ij}(k, \omega) \tag{B.5}$$

For an infinite system with rotational invariance, this is a tensor dependent only on k^i. Its most general form is thus given by

$$\chi^{ij}(k) = \frac{k^i k^j}{k^2} A(k^2) + \left(\delta_{ij} - \frac{k^i k^j}{k^2}\right) B(k^2) \tag{B.6}$$

where A and B are two scalar functions associated respectively with the longitudinal and transverse responses. The longitudinal response for long wavelengths defines the total mass density ρ of the system:

$$\rho \equiv A(0) \tag{B.7}$$

while the transverse response defines the "normal fluid density":

$$\rho_n \equiv B(0) \tag{B.8}$$

The "superfluid density" is the difference

$$\rho_s \equiv A(0) - B(0) \tag{B.9}$$

These definitions can be justified as follows [1].

For concreteness, consider $D = 3$. The response in ordinary space is

$$\delta\langle g^i(\mathbf{x})\rangle = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \chi^{ij}(\mathbf{k}) \tilde{w}^i(\mathbf{k}) \tag{B.10}$$

Consider the limit of a spatially uniform velocity

$$w^i(\mathbf{x}, t) = w^i \tag{B.11}$$

The slow time dependence from adiabatic switching is left understood. In an sufficiently large system, $\delta\langle g^i(\mathbf{x})\rangle$ should be independent of position, so we can put $\mathbf{x} = 0$:

$$\delta\langle g^i(0)\rangle = \int \frac{d^3k}{(2\pi)^3} \chi^{ij}(\mathbf{k}) \tilde{w}^i(\mathbf{k}) \tag{B.12}$$

For a system enclosed in a rectangular box of size $L_1 \times L_2 \times L_3$, we have

$$\tilde{w}^i(\mathbf{k}) = \delta v^i \int d^3x\, e^{-i\mathbf{k}\cdot\mathbf{x}} = w^i F(k^1) F(k^2) F(k^3) \tag{B.13}$$

where

$$F(k^j) = \int_{-L_j/2}^{L_j/2} dx^j \exp(-ik^j x^j) \underset{L_j \to \infty}{\longrightarrow} 2\pi\, \delta(k^j) \tag{B.14}$$

Thus,

$$\delta\langle g^i(0)\rangle = w^i \int \frac{d^3k}{(2\pi)^3} F(k^1)F(k^2)F(k^3)\chi^{ij}(\mathbf{k}) \tag{B.15}$$

Choose δv^i to point along the x axis. We make different choices according to the geometry.

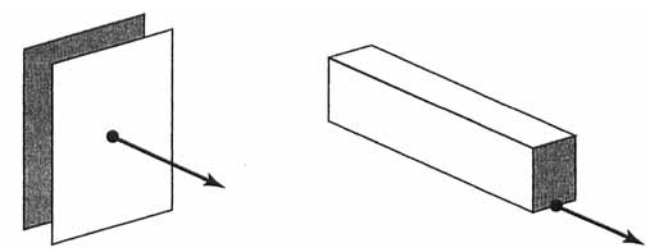
First put the system between flat disks with normal along the x axis, as shown in Fig. B.1. This is realized by first letting $L_2 \to \infty$, $L_3 \to \infty$, and then $L_1 \to \infty$. The imposed velocity field can be created by moving the disks in the x direction with infinitesimal velocity w^i. The entire system must move with the disks, and therefore the induced momentum density is the total density ρ times the velocity:

$$\begin{aligned} \rho w^1 &= w^1 \lim_{L_1\to\infty} \lim_{L_2\to\infty} \lim_{L_3\to\infty} \int \frac{d^3k}{(2\pi)^3} F(k^1)F(k^2)F(k^3)\chi^{11}(\mathbf{k}) \\ &= w^1 \lim_{k_1\to 0} \lim_{k_2\to 0} \lim_{k_3\to 0} \chi^{11}(\mathbf{k}) \end{aligned} \tag{B.16}$$

Therefore

$$\rho = \lim_{k^1\to 0} \lim_{k^2\to 0} \lim_{k^3\to 0} \chi^{11}(\mathbf{k}) \qquad \text{(disk geometry)} \tag{B.17}$$

Next put the system in a long pipe with axis along the x axis, moving in the x direction at velocity δv^i, as shown in Fig. B.1. This is realized by first letting $L_1 \to \infty$, and



(*a*) (*b*)

Figure B.1 Fluid placed in containers of different geometry, to illustrate the concept of normal fluid: (*a*) Disk geometry—the system is placed between large moving plates, and the whole system moves with the plates; (*b*) pipe geometry—the system is placed in a long moving pipe. The normal fluid component moves with the pipe, while the superfluid component remains at rest.

then $L_2 \to \infty$, $L_3 \to \infty$. The normal fluid is the part of the system that exhibits friction, and moves along with the pipe. Its density ρ_n is therefore given by

$$\rho_n = \lim_{k^3 \to 0} \lim_{k^2 \to 0} \lim_{k^1 \to 0} \chi^{11}(\mathbf{k}) \qquad \text{(pipe geometry)} \tag{B.18}$$

To calculate χ^{11}, we turn to (B.6), and note that

$$\frac{k^1 k^1}{k^2} = \frac{(k^1)^2}{(k^1)^2 + (k^2)^2 + (k^3)^2} \to \begin{cases} 1 & \text{(disk geometry)} \\ 0 & \text{(pipe geometry)} \end{cases} \tag{B.19}$$

Therefore

$$\begin{aligned} \rho &= A(0) \\ \rho_n &= B(0) \end{aligned} \tag{B.20}$$

Let us now calculate the superfluid density for a system described by a velocity field $\mathbf{v}(x)$, with Ginsburg-Landau free energy

$$E[\mathbf{v}] = \frac{\rho_0}{2} \int d^D x \, |\mathbf{v}(x)|^2 \tag{B.21}$$

The partition function of the system is

$$Z = \int D\mathbf{v} e^{-\beta E[\mathbf{v}]} \tag{B.22}$$

where β is the inverse temperature. We impose a prescribed velocity field $\mathbf{w}(x)$, understood to be switched on and off adiabatically. The energy becomes

$$E_w[\mathbf{v}] = \frac{\rho_0}{2} \int d^D x \, |\mathbf{v}(x) - \mathbf{w}(x)|^2 \tag{B.23}$$

The first-order change in energy is

$$\delta E = \int d^D x \, g^i(x) w^i(x) \tag{B.24}$$

where

$$g^i(x) = \rho_0 v^i(x) \tag{B.25}$$

which is defined as the momentum density. This averages to zero in the unperturbed system. With the external velocity field, the average momentum density becomes

$$\langle g^i(x)\rangle_w = \frac{\rho_0}{Z}\int D\mathbf{v} e^{-\beta E_w[\mathbf{v}]} v^i(x) \tag{B.26}$$

Taking the functional derivative with respect to $w^j(y)$, we obtain

$$\frac{\delta\langle g^i(x)\rangle_\omega}{\delta w^j(y)} = \frac{\beta\rho_0^2}{Z}\int D\mathbf{v}\, e^{-\beta E_w[\mathbf{v}]} v^i(x) v^j(x)$$

$$-\frac{\beta\rho_0^2}{Z}\int D\mathbf{v}\, e^{-\beta E_w[\mathbf{v}]} v^i(x)\int D\mathbf{v}\, e^{-\beta E_w[\mathbf{v}]} v^j(x)$$

$$= \beta\left[\langle g^i(x) g^j(y)\rangle_w - \langle g^i(x)\rangle_w \langle g^j(y)\rangle_w\right] \tag{B.27}$$

To first order in ω, we can neglect the second term, and integrate both sides to obtain the linear response

$$\delta\langle g^i(x)\rangle = \beta\int d^D y \langle g^i(x) g^j(y)\rangle w^j(y) \tag{B.28}$$

$$= \beta\rho_0^2\int d^D y \langle v^i(x-y) v^j(0)\rangle w^j(y) \tag{B.29}$$

where $\langle\ \rangle$ denotes thermal average without external field and we have assumed translational invariance. We Fourier-analyze $w^j(y)$, and on comparison of the result with (B.10), obtain the response function

$$\chi^{ij}(k) = \beta\rho_0^2\int d^D x e^{-ik\cdot x}\langle v^i(x) v^j(0)\rangle \tag{B.30}$$

Then we decompose the velocity field into longitudinal and transverse parts

$$\mathbf{v} = \mathbf{v}_L + \mathbf{v}_T \tag{B.31}$$

with $\nabla \cdot \mathbf{v}_L = 0$, and $\nabla \times \mathbf{v}_T = 0$. Then

$$\chi^{ij}(k) = \beta\rho_0^2\int d^D x e^{-ik\cdot x}[\langle v_L^i(x) v_L^j(0)\rangle + \langle v_T^i(x) v_T^j(0)\rangle] \tag{B.32}$$

We have used $\langle v_L^i(x) v_T^j(0)\rangle = 0$, which follows from the fact that there is no cross-term in the energy $E[\mathbf{v}]$.

To obtain the superfluid density, we have to decompose χ^{ij} into longitudinal and transverse parts. We write in matrix notation

$$\chi(k) = A(k^2) P_L + B(k^2) P_T \tag{B.33}$$

where

$$P_{\mathrm{L}}^{ij} \equiv \frac{k^i k^j}{k^2}$$

$$P_{\mathrm{T}}^{ij} \equiv \delta_{ij} - \frac{k^i k^j}{k^2} \tag{B.34}$$

Taking the matrix trace of χ, we have

$$\mathrm{Tr}\,\chi(k) = A(k^2) + (D-1)B(k^2) \tag{B.35}$$

Thus

$$B(k^2) = \frac{\mathrm{Tr}\,\chi(k) - A(k^2)}{D-1} \tag{B.36}$$

The superfluid density is

$$\rho_s = A(0) - B(0) = A(0) - \frac{\mathrm{Tr}\,\chi(0) - A(0)}{D-1}$$

$$= \frac{D}{D-1}A(\mathrm{k0}) - \frac{\mathrm{Tr}\,\chi(0)}{D-1} \tag{B.37}$$

We now calculate Tr $\chi(0)$:

$$\mathrm{Tr}\,\chi(0) = \beta\rho_0^2 \int d^D x[\langle v_{\mathrm{L}}^i(x) v_{\mathrm{L}}^i(0)\rangle + \langle v_{\mathrm{T}}^i(x) v_{\mathrm{T}}^i(0)\rangle] \tag{B.38}$$

The longitudinal contribution satisfies the f-sum rule:

$$\beta\rho_0^2 \int d^D x \langle v_{\mathrm{L}}^i(x) v_{\mathrm{L}}^i(0)\rangle = 1 \tag{B.39}$$

Thus

$$\mathrm{Tr}\,\chi(0) = \rho_0 + \beta\rho_0^2 \int d^D x\, \langle v_{\mathrm{T}}^i(x) v_{\mathrm{T}}^i(0)\rangle \tag{B.40}$$

This leads to the formula

$$\rho_s = \rho_0 - \frac{\beta\rho_0^2}{D-1} \int d^D x\, \langle v_{\mathrm{T}}^i(x) v_{\mathrm{T}}^i(0)\rangle \tag{B.41}$$

REFERENCES

1. P. C. Hohenberg and P. C. Martin, *Ann. Phys.* (NY) **34,** 291 (1965).
2. G. Baym, in *Mathematical Tools of Solid State and Superfluid Theory*, R. G. Clark, Ed., Oliver & Boyd, Edinburgh, 1969.
3. D. Forster, *Hydrodynamics, Fluctuations, Broken Symmetry, and Correlation Functions*, Benjamin, Reading, MA, 1978, Chapter 10.

Index